University of Groningen

Stable states and tipping points

Schreuder, Marieke

DOI:
[10.33612/diss.199938171](https://doi.org/10.33612/diss.199938171)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2022

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Schreuder, M. (2022). *Stable states and tipping points: a complex dynamic systems perspective on youth mental health*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen.
<https://doi.org/10.33612/diss.199938171>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Stable states and tipping points

a complex dynamic systems perspective on youth mental health

Marieke J. Schreuder

Cover: Charlotte Hendriks

Lay-out: Wendy Bour-van Telgen

Printed by: Ipskamp Printing

© 2022, Marieke Schreuder

No part of this thesis may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording or any information storage or retrieval system, without prior written permission of the copyright owner.

**rijksuniversiteit
 groningen**

**Stable states and tipping points:
 a complex dynamic systems perspective
 on youth mental health**

Proefschrift

ter verkrijging van de graad van doctor aan de
 Rijksuniversiteit Groningen
 op gezag van de
 rector magnificus prof. dr. C. Wijmenga
 en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

woensdag 23 februari 2022 om 14.30 uur

door

Marieke Josien Schreuder

geboren op 16 september 1992
 te Rotterdam

Promotores

Dr. C.A. Hartman

Prof. dr. M.C. Wichers

Copromotor

Dr. J.T.W. Wigman

Beoordelingscommissie

Prof. dr. P. Kuppens

Prof. dr. A. Lichtwarck-Aschoff

Prof. dr. U. Ebner-Priemer

Table of contents

Chapter 1	General introduction	9
Chapter 2	Lower emotional complexity as a prospective predictor of psychopathology in adolescents from the general population	25
Chapter 3	Early warning signals in psychopathology: what do they tell?	47
Chapter 4	Can we predict the direction of sudden shifts in symptoms? Transdiagnostic implications from a complex systems perspective on psychopathology	77
Chapter 5	Measuring psychopathology as it unfolds in daily life: addressing key assumptions of intensive longitudinal methods in the TRAILS TRANS-ID study	97
Chapter 6	For whom do intensive longitudinal studies work? Participation and compliance in a six-month daily diary study among individuals at risk for mental health problems	137
Chapter 7	Subthreshold psychopathology: stable or transient?	165
Chapter 8	Anticipating transitions in mental health in at-risk youth: a six-month daily diary study into early warning signals	197
Chapter 9	Anticipating manic and depressive shifts in patients with bipolar disorder using early warning signals	233
Chapter 10	Anticipating the direction of symptom progression using critical slowing down: a proof-of-concept study	265
Chapter 11	General summary and discussion	295

Nederlandse samenvatting	346
Dankwoord/acknowledgements	357
Curriculum vitae	361
List of publications	362
SHARE publications	364

1

General introduction

Almost all individuals experience mental health problems at some point in their lives¹. For some, such problems can amount to a clinical disorder. Most mental disorders have their first onset in adolescence, when individuals are between 12 and 24 years old^{1,2}. Adolescence is therefore considered a critical period in terms of mental health: approximately 20 to 25% of youth has had a mental disorder during the past year³, and by the age of 24, more than 60% of individuals from the general population ever met the diagnostic criteria for a mental disorder¹. It is also adolescents for whom the burden of mental disorders is highest⁴. This is because mental health problems may interfere with all experiences that come with growing up: completing an education, finding and keeping a job, establishing friendships and romantic relationships, et cetera³. Since impairments in these aspects of life are often difficult to overcome and have lasting consequences, mental disorders can be considered “the chronic diseases of the young”⁵.

Given the substantial prevalence and burden of mental disorders in young people, there is a clear need for effective early intervention and prevention. A first and necessary step to this end concerns the early detection of vulnerability: if we can identify those adolescents for whom mental health problems are likely to emerge and persist and offer help at an early stage, the prognosis of potentially impairing mental health problems likely brightens. However, despite decades of research into risk factors for mental disorders, early detection of mental disorders still poses a challenge⁶. That is, we are still unable to infer from someone’s genes, brain, personality, or life history whether, and at what specific moment in time, they will develop a mental disorder.

The challenge of early detection is fueled by at least two issues. First, the factors that may give rise to mental health problems are numerous – ranging from biological predispositions to environmental circumstances to psychological characteristics. It is therefore difficult, if not impossible, to establish how all these risk factors interact and cause mental disorders. Identifying the pathways that lead to psychopathology becomes even more difficult when taking into account individual differences. This introduces a second issue: what causes mental health problems for one person may not be relevant for another person. More generally, the causes of psychopathology are specific to individual persons⁷. As a consequence, averaging across individuals – for instance by comparing a group of individuals with major depressive disorder to a group of individuals without such a diagnosis – does not necessarily lead to insights into the mechanisms that apply to a specific individual. In sum, the complex and person-specific nature of psychopathology means that a simple,

one-size-fits-all solution to early detection is unlikely.

It follows that research into the early detection of psychopathology is in need of an approach that can account for the inherent complexity and person-specific nature of mental disorders. A complex dynamic systems approach may answer this call. Briefly, a complex dynamic systems approach proposes that although the mechanisms that cause mental disorders are not fully understood, it is still possible to predict the onset or remission of psychopathology at the individual level⁸. Specifically, monitoring moment-to-moment fluctuations in the system could allow for determining the system's vulnerability without requiring mechanistic insight in this vulnerability. This thesis was inspired by this complex dynamic systems approach to psychopathology. Before elaborating on the clinical promise of this approach, as well as the practical challenges that come with its empirical investigation, I will provide some background on complex dynamic systems theory and its translation to psychiatry.

Complex dynamic systems

Although research on complex dynamic systems spans several decades and scientific disciplines, there is still little consensus on what it means to be “complex”⁹. For the purposes of the present thesis, the notion put forward by Richardson and colleagues¹⁰ suffices: complex dynamic systems are systems that (i) consist of a large number of interacting elements, (ii) show emergent behavior that results from self-organization rather than central control, and (iii) evolve over time. An example of such a system is an ecosystem, where many species interact and together maintain a stable equilibrium. Other systems that have been proposed to resemble complex dynamic systems include financial markets, societies, neuronal firing patterns, and the climate⁸.

An interesting property of complex dynamic systems is that they may undergo critical transitions. A critical transition reflects a sudden shift between two stable states – for instance, from a greenhouse to an icehouse climate⁸. Critical transitions may happen without an obvious (overt) trigger, are difficult to revert, and have a large impact on the system. Given that the consequences of critical transitions can be detrimental, anticipating and potentially preventing transitions has inspired many research efforts. An important advance in this pursuit is the recognition that critical transitions are typically preceded by critical slowing down^{11,12}. Critical slowing down refers to the phenomenon that, as a system approaches a

transition, it becomes increasingly difficult for the system to recover from perturbations. That is, it takes longer and longer before the equilibrium state of the system is restored after minor perturbations. Critical slowing down cannot be observed directly, but can be inferred from the dynamics of the system. The dynamics that capture critical slowing down have therefore been labelled early warning signals (EWS): they anticipate otherwise unpredictable transitions, and thus, “warn” us^{11,12}. Thus, we might be able to foresee critical transitions in complex dynamic systems through monitoring the dynamics of the system.

Numerous EWS have been proposed in the past decades, and they have been listed elsewhere^{11,13}. The EWS that are most prominently covered in this thesis are rising autocorrelations^{14,15}. An autocorrelation reflects the extent to which a system’s current state (x_t) is predictive of its future state (x_{t+1}). Conceptually, autocorrelations reflect whether perturbations have long-lasting effects (high autocorrelations), or instead, are quickly negated (low autocorrelations). Rising autocorrelations thus reflect that the impact of perturbations is increasing – *i.e.*, that the system’s ability to recover its stable equilibrium gradually diminishes (critical slowing down). From this description, it follows that monitoring EWS requires intensive longitudinal data (*i.e.*, time series), where a system’s state (x) is assessed every minute, hour, day, or year, depending on the dynamics of the system of interest. Studies that used such data showed that rising autocorrelations reflect relatively robust EWS: they precede sudden shifts in the brightness of stars¹⁶, epileptic seizures¹⁷, critical transitions in the climate¹⁴, the collapse of ecosystems and civilizations^{18–21}, as well as the outbreak of the infectious diseases^{22,23}. From this elaboration, it can be seen that EWS are generic, in the sense that they anticipate critical transitions regardless of the type of system wherein such transitions occur. This raises the question whether EWS – and complex dynamic systems properties in general – also generalize to transitions in mental health.

Psychopathology from a complex dynamic systems perspective

Mental health, too, might be described in terms of complex dynamic systems principles^{6,24–27}. The system here refers to an individual’s network of interacting momentary mental states (*e.g.*, feeling happy, down, etc.). Together, these mental states give rise to an overall condition (*e.g.*, mental health), which can be considered a stable state (or equilibrium). Perturbations, such as a flat tire or sunny weather, may lead to temporary departures from this stable state, thereby creating dips or uplifts in the individual’s mood. Eventually, however, the stable

state is restored. It is perhaps easiest to think of stable states as reflecting either mental health or mental disorder – and therefore, the remainder of this introduction will focus on these two states. Keep in mind, however, that alternative stable states are also possible: individuals may for instance experience an enduring state characterized by mild symptoms of psychopathology that reflects neither mental health nor mental disorder²⁸ (see also **chapter 7**).

Like other complex dynamic systems, a system consisting of interacting mental states may undergo sudden transitions that may seem unpredictable, in the sense that they are not necessarily preceded by an impactful trigger (*e.g.*, a major life event) or overt warning signs (*e.g.*, gradually feeling more and more sad). Such transitions can involve, for instance, the onset of psychosis^{27,29}, relapse of depression³⁰, relapse of addiction³¹, shifts between manic and depressed episodes that characterize bipolar disorder³², and the progression of eating disorders³³. Transitions towards remission may also occur^{34,35}, but are not the focus of the current thesis and will therefore not be discussed further. If transitions in mental health follow the same generic principles as transitions in other complex dynamic systems, they may be preceded by critical slowing down (Fig. 1.1). In the context of mental health, critical slowing down can be understood intuitively: as an individual becomes more vulnerable for a mental disorder, it becomes increasingly difficult to recover from perturbations. Put differently, restoring one's stable state (mental health) upon perturbations takes more and more time, up to a point where recovery is no longer possible. At this "tipping point", even a minor perturbation can evoke a transition towards another stable state (mental disorder). Provided that mental health and psychopathology behave according to complex dynamic systems principles, foreseeing this tipping point might be possible through monitoring indicators of critical slowing down (*i.e.*, EWS). This suggests that monitoring how an individual's mental states fluctuate from moment to moment – and keeping track of EWS based on these dynamics – could forewarn upcoming mental health problems.

Fig. 1.1. *Critical slowing down for transitions in mental health.* A) A transition between two stable states, here classified as mental health and mental disorder. B) If we zoom in on the period before the transition, we can observe that (i) the system fluctuates around its stable state and (ii) as the system approaches the transition, its recovery from perturbations (red arrows) slows down. The latter process – critical slowing down – means that autocorrelation rises as the transition approaches. Far from the transition, the system does not linger (low autocorrelation), while closer to the transition, the system becomes more self-predictive (higher autocorrelation). This can only be observed in intensive longitudinal data (or, time series).

A complex dynamic systems approach provides a compelling answer to the challenges we face in pursuit of identifying individuals at risk for psychopathology. Firstly, it does not necessitate extensive knowledge about the complex mechanisms that give rise to symptoms of psychopathology in order to detect upcoming mental health problems. Rather, a complex dynamic systems approach proposes that symptom progression might be anticipated based on the dynamics of mental states (*i.e.*, critical slowing down). Secondly, a complex dynamic systems approach to psychopathology could allow for anticipating upcoming transitions *within individuals*. This suits the recent trend towards precision (or personalized) psychiatry³⁶: rather than group-level findings, a complex dynamic systems approach may yield person-specific findings that could directly translate to clinical practice. Specifically, it could allow for addressing a fundamental question in early intervention settings, namely: how likely is this person to develop a mental disorder in the near future? This illustrates a fundamental difference between the promise of EWS and existing methods for inferring individuals' risk for psychopathology: whereas existing risk calculators typically yield a static, group-level risk estimate (*e.g.*, “on average, young adult males with a current depression have a 0.7% chance of developing psychosis in the next year³⁷”), EWS may yield an individual-level, dynamic risk estimate (*e.g.*, “this individual's risk for developing psychosis has increased over the past month”). The latter insight has great clinical value, as it may signal at what moment in time

an intervention is needed.

Particularly youth at risk for mental health problems may have much to gain from the inferences that may be drawn from EWS. Vulnerability during this developmental period is often still pluripotent, meaning that it can foreshadow a broad range of psychopathological symptoms (*e.g.*, depression, anxiety, psychosis, substance use, etc.)^{27,38}. Therefore, we hypothesized that EWS could have an additional utility in at-risk youth: besides warning for an upcoming transition in mental health, EWS might be warning for the *specific types of symptoms* for which an individual is vulnerable. Identifying vulnerability with such precision could facilitate targeted and tailored early interventions. In sum, a complex dynamic systems approach to psychopathology promises a novel route to detecting impending mental disorders, potentially allowing us to intervene before a downward spiral has set.

Given the clinical promise of EWS, there is a clear relevance for empirically investigating complex dynamic systems principles in the context of psychopathology. This is challenging, as investigating the dynamics of mental states within individuals requires intensive longitudinal data, where individuals monitor their mental state for prolonged periods of time. Such data are more difficult to collect compared to filling in a single survey in cross-sectional designs, as they require considerable time from both participants and researchers. Further, in order to investigate transitions across domains of psychopathology, participants who are at increased risk for mental disorders are needed. Only when these at-risk individuals regularly monitor their mental states in a period during which transitions in their mental health will occur, is it possible to address the question whether the complex dynamic systems principles laid out above apply to psychopathology. Four years ago, it was still unknown whether the collection of such intensive longitudinal data in at-risk individuals would be feasible. It could be, for instance, that individuals are simply not willing or able to monitor their thoughts and feelings every day for several months. It could also be that certain types of individuals – *e.g.*, those with severe mental health problems, or conversely, those who never experienced a mental disorder – are less likely to comply with this type of design. This would reduce the usefulness of intensive longitudinal data – and EWS therein – to predict the course of psychopathology. Part of this thesis is therefore devoted to the practical implications of applying complex dynamic systems principles to mental health: (for whom) can we collect the required data?

This thesis

Group- versus individual-level studies

The current thesis starts with group-level studies – where we compared moment-to-moment emotions in youth with varying psychopathological symptom trajectories – and continues to within-individual studies – which allow for uncovering processes that happen within individuals. As the distinction between group-level (or nomothetic/between-individual) studies and individual-level (or idiographic/within-individual) studies will reappear throughout this thesis, it warrants a brief clarification. Group-level studies generally investigate whether individuals who belong to a specific group (*e.g.*, individuals diagnosed with a disorder) – or have a relatively high score on one parameter (*e.g.*, symptoms of psychopathology) – differ from others (*i.e.*, those without a diagnosis or with lower symptom severity). Such studies may yield, for instance, that individuals with depression have a higher variability in their day-to-day emotions compared to individuals without depression. These findings can create a profile of what we consider normative or typical, and allow for comparing an individual to a benchmark (*e.g.*, is your score on this questionnaire higher than average?). Complementary to group-level studies are individual-level studies. Such studies inform on processes that occur within individuals, for instance: does your emotional variability rise as you become more depressed? Such studies can accommodate the person-specific nature of psychopathology, but may come at the cost of limited generalizability across individuals. Ultimately, the utility of group- versus individual level studies is determined by the research question and inferential goals of a study.

Outline

Chapters 2 and 3 both address whether emotion dynamics predict future psychopathology in adolescents from the general population. These adolescents monitored their emotions ten times a day for six consecutive days, which allowed us to compute emotional complexity (**chapter 2**) and inertia (**chapter 3**). In **chapter 2**, we investigated whether emotional complexity predicts the course of psychopathological symptoms over one year. This study was inspired by the notion that complex dynamic systems with high diversity – such as ecosystems that consist of many different species – are generally resilient to sudden collapses⁸. As diversity declines, however, systems may become increasingly vulnerable to a collapse. If a similar principle holds in the context of mental health, individuals with more

diverse emotional experiences may be less likely to develop psychopathology compared to individuals with less diverse emotions. This would mean that individuals who can experience a wide range of emotions – for instance, feeling irritated, ashamed, or disappointed – are less vulnerable to future mental health problems compared to individuals who tend to lump all emotions into a single mental state – for instance, feeling bad. We assessed such individual differences in the complexity of emotional experiences using two measures, namely (i) emotional differentiation³⁹, which is well-known in emotion dynamics literature, and (ii) emotional diversity⁴⁰, which resembles measures of diversity in ecosystems. We hypothesized that adolescents with lower emotional complexity would be more likely to report worsening or persisting symptoms of psychopathology compared to adolescents with higher emotional complexity, and that this effect would remain present after accounting for individual differences in mean emotionality (*i.e.*, mean levels of negatively valenced emotions).

Besides declining diversity, vulnerability in complex dynamic systems can be monitored through measures of critical slowing down (*i.e.*, EWS). EWS could therefore be considered personalized risk markers for future psychopathology, revealing whether a specific individual at a specific moment in time is in need of treatment. Before empirically testing this idea, we aimed to replicate a group-level study that provided initial, indirect support for the hypothesis that critical slowing down may anticipate transitions in mental health⁴¹. Specifically, in **chapter 3**, we investigated whether between-individual differences in the autocorrelation of emotions (also called inertia) reported by adolescents from the general population predicted symptom changes one year later. Our reasoning was as follows: if rising autocorrelations precede a transition in mental health within individuals, this could translate to a positive group-level association between autocorrelations and future psychopathology. In other words, individuals with high levels of emotional inertia – meaning that emotions return to their baseline relatively slowly – were hypothesized to show greater symptom increases compared to individuals with lower emotional inertia. Additionally, we aimed to investigate whether the hypothesized association between inertia and future psychopathology is more pronounced in corresponding emotions and psychopathological domains (*e.g.*, feeling down and depression; feeling suspicious and psychosis, etc.). This allowed us to indirectly test the idea that EWS may not only indicate *whether* symptoms are likely to worsen, but also indicate the *type* of symptoms that will worsen.

Chapters 2 and 3 were inspired by a complex dynamic systems approach to psychopathology. It should be noted, however, that both emotional complexity as well

as inertia can also be understood from a more psychological perspective. Low emotional complexity, for instance, means that individuals may experience difficulty in labeling and regulating their emotions³⁹. Inertia, in turn, could signal an insensitivity to environmental cues, which may also complicate emotion regulation^{42,43}. Hence, the emotion dynamics that we covered in **chapters 2 and 3** are not new to mental health research. The idea that these dynamics may change *within* individuals – as proposed by a complex dynamic systems approach to mental health – is, however, relatively new. The remaining chapters of this thesis address why it is interesting to monitor such within-individual fluctuations (**chapter 4**), whether and for whom this is possible (**chapters 5 and 6**), and how within-individual fluctuations in emotions relate to (in)stability of mental health (**chapters 7-10**).

Firstly, **chapter 4** denotes several hypotheses that follow from a complex dynamic systems approach to mental health. Specifically, this chapter describes that EWS may foresee transitions towards reduced mental health as well as the type of psychopathological symptoms about to develop. This idea will be revisited in **chapters 9 and 10**. **Chapter 4** also describes the type of design that is necessary to investigate complex dynamic systems principles in the context of psychopathology. This chapter thus lays the groundwork for the study that we conducted (TRAILS TRANS-ID), which is described in **chapter 5**. Briefly, TRAILS TRANS-ID includes a six-month daily diary study, where at-risk young adults (N=134) monitored their mental state every evening for a period of six consecutive months (± 183 observations per person). In **chapter 5**, we additionally addressed three assumptions that are typically made, but seldomly verified, in intensive longitudinal designs. **Chapter 6** provides a more detailed picture of the participants recruited in TRAILS TRANS-ID. Specifically, we examined whether – and in what way – these individuals differ from (i) individuals from the general population and (ii) at-risk individuals who decided not to participate in TRAILS TRANS-ID. Together, **chapters 5 and 6** provide an extensive description of the intensive longitudinal study that we performed and reflect on the types of individuals for whom such designs are suitable.

The remaining **chapters 7-10** of the thesis each provide an illustration of what can be done with the data that intensive longitudinal designs yield. These chapters all directly address hypotheses that follow from a complex dynamic systems approach to psychopathology. First, in **chapter 7**, we examined the idea of alternate stable states in the context of psychopathology. It is typically assumed that such stable states reflect either ‘mental health’ or ‘mental disorder’, meaning that subthreshold psychopathology is merely a transient in-between. We investigated this assumption by considering the stability of psychopathological symptoms across levels of severity. Both parameters are based on the

intensive longitudinal data we collected in the TRAILS TRANS-ID study. In **chapter 8**, we investigated the transitions between stable states. Specifically, we investigated whether sudden drops in mental health – reflecting a transition from a stable state with higher well-being towards a state with lower well-being – are preceded by EWS. This study provides an intra-individual verification of what has indirectly been assessed in **chapter 3**. In order to examine whether EWS were indeed specific to transitions in mental health, we also looked for EWS in youth who did *not* show a drop in their mental health. This allowed us to establish the sensitivity, specificity, and predictive value with which EWS precede drops in mental health experienced by at-risk youth. In **chapter 9**, we extended the search for EWS to a clinical population, namely patients with bipolar disorder. Here, we analyzed data from 20 patients with bipolar disorder who provided five assessments of their mental state every day for a period of four months (± 491 observations per person). For these individuals, transitions in mental health do not manifest as drops in mental health, but rather, as relapses into manic or depressed episodes. We investigated whether EWS could accurately predict such relapses by examining their sensitivity, specificity, and predictive value. Together, **chapters 8 and 9** provide a first, explorative illustration of whether EWS can anticipate transitions in psychopathology both in an at-risk population and in a sample of bipolar patients. Finally, **chapter 10** describes whether the utility of critical slowing down – which is the theoretical foundation for EWS – extends beyond indicating the likelihood of a future transition. That is, we investigated whether critical slowing down might also inform on the future direction of transitions – *i.e.*, whether they involve remission or worsening of symptoms. To this end, we examined a dataset wherein principles of critical slowing down were previously confirmed^{30,44} as well as the data collected in the TRAILS TRANS-ID study. For both datasets, we used a method developed in ecological literature to examine whether the direction of critical slowing matches the direction of symptom progression.

In sum, this thesis aims to investigate emerging psychopathology from a novel angle, namely that of complex dynamic systems. Many of the principles discussed, including the potentially protective effects of diversity, critical transitions, critical slowing down, and EWS, have a long history in ecological literature^{8,15}. This led to an interdisciplinary effort, where I have tried to investigate how principles that were validated in (simulated) ecosystems may translate to the field of psychopathology. It is my hope that, after reading this thesis, it is clear how complex dynamic systems could reshape the field of psychiatry, how we should investigate such principles, and what barriers should be overcome before clinically implementing complex dynamic systems principles.

References

1. Caspi A, Houts RM, Ambler A, et al. Longitudinal assessment of mental health disorders and comorbidities across 4 decades among participants in the Dunedin birth cohort study. 2020;3(4). doi:10.1001/jamanetworkopen.2020.3221
2. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry*. 2005;62:593-602. doi:10.1001/archpsyc.62.6.593
3. Patel V, Flisher AJ, Hetrick S, McGorry P. Mental health of young people: a global public-health challenge. *The Lancet*. 2007;369. doi:10.1016/S0140
4. Whiteford HA, Ferrari AJ, Degenhardt L, Feigin V, Vos T. The global burden of mental, neurological and substance use disorders: An analysis from the global burden of disease study 2010. *PLoS ONE*. 2015;10(2). doi:10.1371/journal.pone.0116820
5. Insel T, Fenton FS. Psychiatric epidemiology it's not just about counting anymore. *Archives of General Psychiatry*. 2005;62:590-592.
6. Wichers M. The dynamic nature of depression: a new micro-level perspective of mental disorder that meets current challenges. *Psychological Medicine*. 2014;44(7):1349-1360. doi:10.1017/S0033291713001979
7. Molenaar PCM. A manifesto on psychology as idiographic science: bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research & Perspective*. 2004;2(4):201-218. doi:10.1207/s15366359mea0204_1
8. Scheffer M. *Critical Transitions in Nature and Society*. Princeton University Press; 2009.
9. Ladyman J, Lambert J, Wiesner K. What is a complex system? *European Journal for Philosophy of Science*. 2013;3(1):33-67. doi:10.1007/s13194-012-0056-8
10. Richardson MJ, Dale R, Marsh KL. Complex dynamical systems in social and personality psychology. In: *Handbook of Research Methods in Social and Personality Psychology*. ; 2014:253-282. doi:10.1017/cbo9780511996481.015
11. Scheffer M, Bascompte J, Brock WA, et al. Early-warning signals for critical transitions. *Nature*. 2009;461(7260):53. doi:10.1038/nature08227
12. Scheffer M, Carpenter SR, Lenton TM, et al. Anticipating critical transitions. *Science*. 2012;338(6105):344-348. doi:10.1126/science.1225244
13. Scheffer M, Carpenter SR, Dakos V, van Nes EH. Generic indicators of ecological resilience: inferring the chance of a critical transition. *Annual Review of Ecology, Evolution, and Systematics*. 2015;46:145-167. doi:10.1146/annurev-ecolsys-112414-054242
14. Dakos V, Scheffer M, van Nes EH, Brovkin V, Petoukhov V, Held H. Slowing down as an early warning signal for abrupt climate change. *PNAS*. 2008;105(38):14308-14312. doi:10.1073/pnas.0802430105
15. Wissel C. A universal law of the characteristic return time near thresholds. *Oecologia*. 1984;65:101-107.
16. George SV, Kachhara S, Misra R, Ambika G. Early warning signals indicate a critical transition in Betelgeuse. *Astronomy and Astrophysics*. 2020;640:1-7. doi:10.1051/0004-6361/202038785
17. Maturana MI, Meisel C, Dell K, et al. Critical slowing down as a biomarker for seizure susceptibility. *Nature Communications*. 2020;11(1):1-12. doi:10.1038/s41467-020-15908-3
18. Veraart AJ, Faassen EJ, Dakos V, van Nes EH, Lürling M, Scheffer M. Recovery rates reflect distance to a

- tipping point in a living system. *Nature*. 2012;481(7381):357-359. doi:10.1038/nature10723
19. Dai L, Vorselen D, Korolev KS, Gore J. Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science*. 2012;336(6085):1175-1177.
 20. Drake JM, Griffen BD. Early warning signals of extinction in deteriorating environments. *Nature*. 2010;467(7314):456-459. doi:10.1038/nature09389
 21. Scheffer M, van Nes EH, Bird D, Bocinsky RK, Kohler TA. Loss of resilience preceded transformations of pre-Hispanic Pueblo societies. 2021;118(18):2021. doi:10.1073/pnas.2024397118/-/DCSupplemental
 22. Kaur T, Sarkar S, Chowdhury S, Sinha SK, Jolly MK, Dutta PS. Anticipating the novel coronavirus disease (COVID-19) pandemic. *Frontiers in Public Health*. 2020;8(521):1-13. doi:10.1101/2020.04.08.20057430
 23. Harris MJ, Hay SI, Drake JM. Early warning signals of malaria resurgence in Kericho, Kenya. *Biology Letters*. 2020;16(3). doi:10.1098/rsbl.2019.0713
 24. Cramer AOJ, van Borkulo CD, Giltay EJ, et al. Major depression as a complex dynamic system. *PLoS one*. 2016;11(12):1-20. doi:10.1371/journal.pone.0167490
 25. Nelson B, McGorry PD, Wichers M, Wigman JT, Hartmann JA. Moving from static to dynamic models of the onset of mental disorder. *JAMA Psychiatry*. 2017;74(5):528-534. doi:10.1001/jamapsychiatry.2017.0001
 26. Odgers CL, Mulvey EP, Skeem JL, Gardner W, Lidz CW, Schubert C. Capturing the ebb and flow of psychiatric symptoms with dynamical systems models. *American Journal of Psychiatry*. 2009;166(5):575-582. doi:10.1176/appi.ajp.2008.08091398
 27. McGorry PD, Hartmann JA, Spooner R, Nelson B. Beyond the “at risk mental state” concept: transitioning to transdiagnostic psychiatry. *World Psychiatry*. 2018;17(June):133-142.
 28. McGorry PD, Hickie IB, Yung AR, Pantelis C, Jackson HJ. Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. *Australian and New Zealand Journal of Psychiatry*. 2006;40(8):616-622. doi:10.1111/j.1440-1614.2006.01860.x
 29. Tschacher W, Scheier C, Hashimoto Y. Dynamical analysis of schizophrenia courses. *Biological Psychiatry*. 1997;41:428-437.
 30. Wichers M, Groot PC, Psychosystems ESM group, EWS group. Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*. 2016;85:114-116. doi:10.1159/000441458
 31. Hufford MR, Witkiewitz K, Shields AL, Kodya S, Caruso JC. Relapse as a nonlinear dynamic system: Application to patients with alcohol use disorders. *Journal of Abnormal Psychology*. 2003;112(2):219-227. doi:10.1037/0021-843X.112.2.219
 32. Bayani A, Hadaeghi F, Jafari S, Murray G. Critical slowing down as an early warning of transitions in episodes of bipolar disorder: a simulation study based on a computational model of circadian activity rhythms. *Chronobiology International*. 2017;34(2):235-245. doi:10.1080/07420528.2016.1272608
 33. Zeeman EC. Catastrophe theory. *Scientific American*. 1976;234(4):65-83.
 34. Schiepek G, Tominschek I, Karch S, et al. A controlled single case study with repeated fMRI measurements during the treatment of a patient with obsessive-compulsive disorder: testing the nonlinear dynamics approach to psychotherapy. *World Journal of Biological Psychiatry*. 2009;10:658-668. doi:10.1080/15622970802311829
 35. Aderka IM, Nickerson A, Bøe HJ, Hofmann SG. Sudden gains during psychological treatments of anxiety and depression: a meta-analysis. *Journal of Consulting and Clinical Psychology*. 2012;80(1):93-101. doi:10.1037/a0026455

36. Os J van, Delespaul P, Wigman JTW, Myin-Germeys I, Wichers M. Psychiatry beyond labels: introducing contextual precision diagnosis across stages of psychopathology. *Psychological Medicine*. 2013;43(07):1563-1567. doi:10.1017/s0033291713000937
37. Fusar-Poli P, Rutigliano G, Stahl D, et al. Development and validation of a clinically based risk calculator for the transdiagnostic prediction of psychosis. *JAMA Psychiatry*. 2017;74(5):493-500. doi:10.1001/jamapsychiatry.2017.0284
38. van Os J. The dynamics of subthreshold psychopathology: implications for diagnosis and treatment. *American Journal of Psychiatry*. 2013;170:695-698.
39. Barrett LF, Gross J, Christensen TC, Benvenuto M. Knowing what you're feeling and knowing what to do about it: mapping the relation between emotion differentiation and emotion regulation. *Cognition and Emotion*. 2001;15(6):713-724. doi:10.1080/02699930143000239
40. Quidbach J, Gruber J, Mikolajczak M, Kogan A, Kotsou I, Norton MI. Emodiversity and the emotional ecosystem. *Journal of Experimental Psychology: General*. 2014;143(6):2057-2066. doi:10.1037/a0038025
41. van de Leemput IA, Wichers M, Cramer AOJ, et al. Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences*. 2014;111(1):87-92. doi:10.1073/pnas.1312114110
42. Suls J, Green P, Hillis S. Emotional reactivity to everyday problems, affective inertia, and neuroticism. *Personality and Social Psychology Bulletin*. 1998;24(2):127-136.
43. Bylsma LM, Rottenberg J. Uncovering the dynamics of emotion regulation and dysfunction in daily life with ecological momentary assessment. In: Nykliček I, Zeelenberg M, Vingerhoets A, eds. *Emotion Regulation and Well-Being*. Springer; 2011:225-244. doi:10.1007/978-1-4419-6953-8
44. Kossakowski JJ, Groot PC, Haslbeck JMB, Borsboom D, Wichers M. Data from 'Critical slowing down as a personalized early warning signal for depression.' *Journal of Open Psychology Data*. 2017;5. doi:10.5334/jopd.29

2

Lower emotional complexity as a prospective predictor of psychopathology in adolescents from the general population

published as: Schreuder, M.J., Wichers, M., Hartman, C.A., Menne-Lothmann, C., Decoster, J., van Winkel, R., Delespaul, P., de Hert, M., Derom, C., Thiery, E., Rutten, B.P.F., Jacobs, N., van Os, J., & Wigman, J.T.W. (2020). Lower emotional complexity as a prospective predictor of psychopathology in adolescents from the general population. *Emotion*.
doi: 10.1037/emo0000778

Abstract

Background. Emotional complexity (EC) involves the ability to distinguish between distinct emotions (differentiation) and the experience of a large range of emotions (diversity). Lower EC has been related to psychopathology in cross-sectional studies. This study aimed to investigate (i) if EC prospectively predicts psychopathology and (ii) if this effect is contingent on stressful life events. In order to further explore EC, the effects of differentiation and diversity were compared.

Methods. Adolescents from the general population (N=401) rated eight negatively valenced emotions ten times a day for six consecutive days. Further, they completed the Symptom Checklist-90 (baseline and one-year follow-up) and a questionnaire on past year's life events at follow-up. Logistic regression analyses tested whether EC – reflected by emotion differentiation (intra-class correlation coefficient; ICC) and diversity (diversity index; DI) – predicted prognosis (good: remitting/lacking symptoms, vs. bad: worsening/persisting symptoms).

Results. EC predicted prognoses, but only when based on the ICC ($OR_{EC,ICC} = 1.42, p = .02$). An EC_{ICC} one standard deviation above average increased the probability of good prognosis from .67 to .74. This effect was not related to stressful life events ($OR_{EC*events} = 1.03, p = .86$) and disappeared when emotion intensity (mean level) was taken into account ($OR_{EC} = 1.20, p = .20$).

Discussion. Predicting future prognosis does not necessitate complex measures of emotional experience (ICC, DI) but rather might be achieved through simpler indices (mean). The discrepant effects of the ICC and DI on prognosis suggest that impaired emotion representation (ICC) plays a more important role in vulnerability to mental ill-health than low diversity of emotions (DI).

Introduction

Emotional experience can be captured through various concepts, including – but not limited to – emotion differentiation (*i.e.*, the ability to distinguish between different emotions¹) and emotional diversity (*i.e.*, the ability to experience a wide range of emotions²). Together, these concepts describe the complexity of one's emotional experiences^{3–5}. High emotional complexity (EC) indicates that individuals experience a large number of diverse, fine-grained emotions. This is considered adaptive, because a detailed representation of emotions might facilitate dealing with them¹. Indeed, whereas high EC has been linked to well-being^{6,7} and social adjustment³, low EC may lead individuals to feel overwhelmed and to employ maladaptive coping strategies⁸. This might eventually contribute to the development of various psychopathological symptoms. For instance, individuals with depression⁹ or depressive symptoms^{2,10}, social anxiety disorder¹¹ and borderline personality disorder¹² have reported lower EC in negative affective states than have healthy controls.

Whereas the cross-sectional association between EC and psychopathological symptoms is well-established, the relation between EC and *future* symptoms has not yet been investigated. Yet, a better understanding of the prospective association between EC and psychopathology could provide meaningful clinical insights. For instance, if EC prospectively predicts the development of symptoms, interventions that improve EC could be speculated to lower one's vulnerability to psychopathology. This assumes that EC is malleable, which may indeed be the case. Specifically, it was recently suggested that repeatedly monitoring one's affective states, for instance through keeping an electronic diary, may increase emotion differentiation¹³. Before the promises of such interventions are to be tested, however, the association between EC and psychopathological symptoms needs to be determined in a longitudinal design¹⁴.

The few longitudinal studies into EC conducted so far have focused on moment-to-moment fluctuations in affective states, stressful experiences and psychopathology. These studies indicated that interactive effects between EC and stress might contribute to short-term increases in symptom severity^{14,15}. Specifically, poor EC may simultaneously follow from daily life stress exposure¹⁴ and exacerbate its short-term impact on symptom severity¹⁵. EC may thus affect momentary levels of symptom severity, which is particularly apparent in the context of daily life stress (*e.g.*, daily hassles). Whether these effects generalize to a more global level, involving long-term worsening of symptoms following more severe stressors

(*e.g.*, major life events, such as loss), has not yet been investigated. Yet, these long-term effects of EC at least partly determine its clinical relevance, and therefore warrant closer inspection.

Earlier studies into the role of EC in mental ill-health typically addressed only one aspect of EC – either emotion differentiation or emotional diversity. Emotion differentiation has been operationalized as the ICC¹. According to the ICC, low EC means that a change in one emotion (*e.g.*, feeling down) coincides with changes in other emotions (*e.g.*, feeling anxious, lonely, etc.). High EC, in contrast, occurs when emotions seem to be independent of one another (low ICC). A focus on emotion diversity instead of emotion differentiation as a source of EC proposes an alternative measure of EC, namely the diversity index (DI). According to this measure, not the covariance of emotions, but rather the number of emotions that individuals experience, is informative of mental ill-health. The DI suggests that it might be more adaptive when negative emotionality spreads across a large number of emotions (*i.e.*, feeling both down, anxious, and lonely) than when it is aggregated into one single emotion (*i.e.*, feeling down, without feeling anxious or lonely). This builds on the idea that emotional systems might behave similar to other systems, and thus, benefit from diversity^{2,16}. In conclusion, the ICC and DI emphasize distinct aspects of EC but have never been directly compared, which raises the question which facet of EC is most appropriate in the context of psychopathology.

The present study will investigate whether EC – reflected in both emotion differentiation (ICC) and emotion diversity (DI) – is predictive of future prognosis in terms of psychopathology. We focus exclusively on the complexity of emotions with negative (as opposed to positive) valence because the dynamics of negatively valenced emotions might be more informative of mental ill-health than the dynamics of positive emotions¹⁷. Further, the putative mechanism that underlies the link between EC and mental ill-health was only found for negative emotions¹. Our second aim was to investigate whether stressful life events moderate the association between EC and future prognosis. Finally, as mean levels of emotions have recently been shown to outperform metrics related to the dynamics of emotions – including EC – in terms of their predictive utility, it is important to verify whether the hypothesized associations remain present after accounting for emotion intensity¹⁸. Hence, our final aim was to examine whether the predictive utility of EC extends beyond that of emotion intensity.

Methods

Participants

Data were retrieved from the TwinssCan study¹⁹, which comprised a subset of a registered cohort of twins from the general population (*i.e.*, East Flanders Prospective Twin Survey)^{20,21}. The study included N=839 twins, most of whom were between 15 and 18 years old. The TwinssCan study was approved by the local Ethics Committee and all subjects provided written informed consent. For minors, parents provided additional written consent.

Experience sampling method

At baseline, participants provided repeated daily assessments of affective states, which were acquired through the experience sampling method (ESM). The ESM involved questionnaires which were administered through PsyMates® (electronic devices²²). The PsyMate was programmed to emit a beep-signal at ten semi-random time intervals within 90-minute blocks ranging from 7:30 AM to 10:30 PM for six consecutive days. The beep signal prompted participants to fill in a questionnaire concerning their current affective state (*e.g.*, “To what extent do you feel lonely?”). In total, 13 affective states (8 negatively valenced, 5 positively valenced) were rated on a 7-point Likert scale ranging from ‘not at all’ to ‘very much’ (see supplement 2.1 *List of ESM items*). For analyses, the data pertaining to negative affective states (feeling lonely, anxious, irritated, listless, suspicious, down, insecure, guilty) were selected. To ensure reliability and validity of these ratings, questionnaires that were completed more than 15 minutes after the beep were coded as missing²³.

Emotional complexity

EC was inferred from the differentiation (assessed by the within-person ICC) and diversity (assessed by the DI) of negatively valenced emotions. According to the ICC, optimal EC occurs when emotions do *not* fluctuate in a coordinated fashion. This measure emphasizes the covariance – rather than the absolute ratings – of emotions and requires repeated assessments of multiple emotions. In contrast, the DI proposes that optimal EC occurs when emotions are distributed evenly within time points. This measure emphasizes the number and relative abundance (*i.e.*, the proportional intensity) of emotions that individuals report, and can be computed based on a single assessment of multiple emotions².

Intraclass Correlation Coefficient (ICC)

The ICC quantifies the covariance (consistency ICC) or the equivalence (agreement ICC) in repeated, self-reported ratings of emotions²⁴. Whereas the consistency ICC regards emotions as differentiated when they do not change synchronously (regardless of their absolute score), the agreement ICC regards emotions as differentiated when they are given different absolute scores. In the context of emotion differentiation, the consistency ICC has been favored over the agreement ICC and we therefore investigated the former²⁵.

The ICC is computed by decomposing the variance in emotion ratings into the variance explained by time points (MS_{time}), variance explained by individual emotions, and residual variance that is explained by neither time points nor emotions ($MS_{residual}$; formula 1). Put differently, the ICC reflects the amount of variance in emotion ratings that can be attributed to the similarity between those ratings²⁴. In the absence of variability in emotion ratings across time (*e.g.*, because an individual never reports feeling lonely), the variance that is explained neither by time nor by emotions ($MS_{residual}$) can exceed the variance that is explained by time (MS_{time}). As a consequence, ICCs can be negative. Yet, negative ICC estimates are theoretically impossible and uninterpretable¹⁴ and are therefore commonly set to zero^{13,26}. Finally, to allow for between-person comparisons, ICCs were normalized by Fisher's Z transformation²⁷. ICC values were reversed for ease of interpretation, meaning that high values are indicative of high EC and vice versa^{6,14}.

$$ICC = \frac{MS_{time} - MS_{residual}}{MS_{time}}$$

Formula 1. Note that computation of the ICC requires variability along two dimensions, namely emotions (*e.g.*, feeling lonely, anxious, *etc.*) and time. For details concerning the computation of mean squares (MS), see Field²⁸.

Diversity Index (DI)

The DI reflects the number of emotions that are experienced (*i.e.*, richness of emotions) and relative abundance of each individual emotion (*i.e.*, evenness of emotions), and was averaged across time points to derive a single estimate for each individual^{2,5}. In formula 2, $p_{k,t}$ denotes the rating of emotion k at time point t divided by the total intensity of negative emotions at time point t (*i.e.*, the sum of emotion ratings at t). This ratio – indicative of the relative abundance of emotion k – was computed for each of n time points, and each of s emotions. The product of $p_{k,t}$ and its natural logarithm were then multiplied by -1, summed and averaged across time points⁵. High values indicate high EC.

$$DI = \frac{1}{n} \sum_{t=1}^n \sum_{k=1}^s (-1 * p_{k,t} * \ln(p_{k,t}))$$

Formula 2. Note that computation of the DI requires variability along one dimension, namely emotions.

Assessment of psychopathology

At baseline and at one year follow-up, participants completed the Symptom Checklist-90 (SCL-90)²⁹. The SCL-90 consists of 90 items that are rated on a 5-point Likert scale ranging from ‘not at all’ (score 0) to ‘very often/always’ (score 4; *e.g.*, “During the last week, I felt empty”). Future psychopathology – the outcome of interest – could be conceptualized as baseline-corrected symptom severity at follow-up. Yet, this would confuse persistently symptomatic with persistently a-symptomatic individuals: both groups may report small absolute changes in symptom severity, yet they have opposite prognoses. Hence, future psychopathology was operationalized as the dichotomous prognosis (good vs. bad) rather than a continuous estimate of symptom change. This allowed for interpreting a change in symptom severity in the context of baseline psychopathology. Prognosis was based on the total SCL-90 score, or global severity index (GSI), at baseline and follow-up (cf. Schauenburg and Strack³⁰). A change in GSI was considered substantial if it exceeded the reliable change index (RCI) reported by Schauenburg and Strack³⁰. Depending on whether the baseline GSI was below or above the cut-off of 0.57, the RCI was equal to 0.16 or 0.43, respectively³⁰. An increase in GSI that exceeded the RCI was defined as a bad prognosis. Additionally, a baseline GSI above cut-off (0.57) in absence of a reliable change was considered indicative of a bad prognosis. Good prognosis was defined as a decrease in GSI that exceeded the reliable change index or a baseline GSI below cut-off in absence of reliable change.

Assessment of stressful life events

At follow-up, participants completed a Life Events Questionnaire (LEQ), which was based on the Interview for Recent Life Events³¹. The LEQ lists 61 events that cover the following domains: work; education; finance; health; bereavement; migration; courtship, marriage and cohabitation; legal; family and social relationships. Participants reported whether these events occurred in the past year, and if so, rated their impact on a 5-point Likert scale ranging from ‘very pleasant’ to ‘very unpleasant’. Events that were rated as unpleasant (4) or very unpleasant (5) were counted, resulting in a score that reflected the number of stressful life events experienced between baseline and follow-up.

Analyses

Two separate multilevel logistic regression analyses were used to assess whether EC – reflected in the ICC or DI, respectively – predicted prognosis, accounting for similarities between individuals from the same family. Second, we explored the interaction effect of EC and stressful life events on prognosis. Analyses were conducted in R³² (version 1.1.453).

Results

Analyses included those individuals for whom (i) at least 30 (50%) affect ratings were available²³ and (ii) both baseline and follow-up ratings on symptom severity were available, resulting in N=401. Included participants were slightly older (mean age 17.8 years vs. 17.0 years, $t(739) = 3.32, p < .01$) and less likely to be male (34% vs. 46% males, $\chi^2(1) = 13.47, p < .01$) than excluded participants (N=438). Baseline symptom severity did not differ significantly between in- and excluded participants (mean GSI 0.49 vs. 0.47, $t(808) = 0.67, p = .50$). Attrition did not compromise power, which was estimated post-hoc. Specifically, we created 100 simulations of data with 401 individuals from 238 families, a random intercept variance of 0.10 and a single predictor with a small effect size (Cohen’s $d = 0.20$)³³. In all simulations, the effect of the predictor (*e.g.*, ED) was detected (power > .99).

Table 2.1. *Sample characteristics*

Descriptive	M (SD)	N (%)
Total sample size		401
Monozygotic twins; dizygotic twins; triplets; siblings		147 (37); 224 (56); 9 (2); 21 (5)
Male participants		135 (34)
Age in years	17.8 (4.0)	
Life events	2.65 (1.19)	
Completed ratings	45 (7.7)	
Baseline SCL-90 GSI	0.49 (0.39)	
Participants above SCL-90 GSI cutoff at baseline		114 (28)
Follow-up SCL-90 GSI	0.43 (0.35)	
Participants above SCL-90 GSI cutoff at follow-up (%)		99 (25)
Prognosis (good; bad)		267 (67); 134 (33)
Negative emotions	1.70 (0.51)	
Positive emotions	4.90 (0.71)	

Note. GSI = global severity index (cut-off: 0.57). SCL-90 = Symptom Checklist 90. SD = standard deviation.

Table 2.2. *Correlation between emotional complexity (differentiation, diversity), mean negative emotions, and life events*

	Diversity (DI)	Mean negative emotions	Life events
Differentiation (ICC)	.12 ($p = .02$)	-.32 ($p < .01$)	-.08 ($p = .12$)
Diversity (DI)		-.37 ($p < .01$)	<.01 ($p = .92$)
Mean negative emotions			.15 ($p < .01$)

Note. Data refer to Pearson correlation coefficients. DI = diversity index. ICC = intra-class correlation coefficient.

On average, individuals completed 45 (75%) affect ratings. In terms of psychopathology, most participants improved from baseline to follow-up, with 83 individuals (21%) who reported a reliable decrease in symptoms (*i.e.*, a reduction in GSI that exceeded the reliable change index) and 184 individuals (46%) with GSI scores below the cut-off at baseline whose symptoms did not reliably change (for details concerning subscale scores, see supplement 2.2 *Subscale scores SCL-90*). Hence, 267 individuals (67%) were considered to have a good prognosis (see Table 2.1). The remaining 134 individuals (33%) either reported a reliable increase in symptoms ($N=63$, 16%) or a baseline GSI score above cut-off without reliable change ($N=71$, 18%), and were thus considered to have a bad prognosis. Prognosis was unrelated to sex ($\chi^2(1) = 2.19$, $p = .14$) or age ($t(399) = 0.57$, $p = .57$).

The product-moment correlations between differentiation, diversity, intensity, and life-events were small to moderate (see Table 2.2). On average, the probability of a good prognosis was .67. In models with the ICC as a measure of EC (see Table 2.3, models 1-3), higher EC increased the probability of a good prognosis (model 1: $OR_{EC} = 1.42$, $p = .02$). There was no interaction effect between EC and stressful life events on prognosis (model 2: $OR_{EC*events} = 1.03$, $p = .86$). The effect of EC on prognosis disappeared after accounting for the mean levels of negative emotions (model 3: $OR_{EC} = 1.20$, $p = .20$)^a. Specifically, for an individual in an average family with an average EC, one standard deviation increase in mean negative emotions lowered the probability of a good prognosis from .67 to .57. In models where EC was assessed by the DI (Table 2.3, models 4-6), EC did not affect prognosis as a single predictor (model 4: $OR_{EC} = 1.15$, $p = .22$), or in interaction with life events (model 5: $OR_{EC*events} = 0.97$, $p = .79$), or after controlling for mean levels of negative emotions (model 6:

^a Removal of the 22 individuals (5%) for whom ICC values were negative, and hence set to zero^{13,14,26}, did not change the results (supplement 2.3 *Sensitivity analysis complete sample*).

$OR_{EC} = 0.96, p = .74$). Sensitivity analyses using continuous instead of dichotomized measures of symptom severity produced similar results, although the effect of EC on future symptoms was not statistically significant (supplement 2.4 *Sensitivity analysis continuous outcome*).

Table 2.3. *Results of multilevel logistic regression analyses with prognosis as outcome*

Model	OR [95% CI]	<i>p</i>	Probability of good prognosis	Conditional R^2	VIF
Model 1. EC (ICC)				.07	
EC	1.42 [1.06, 1.89]	.02	.74		
Model 2. EC (ICC) and life events				.06	
EC	1.40 [1.05, 1.87]	.02	.74		1.02
Life events	0.83 [0.67, 1.03]	.10	.63		1.07
EC X life events	1.03 [0.75, 1.42]	.86	.68		1.08
Model 3. EC (ICC) and mean negative emotions				.08	
EC	1.20 [0.91, 1.59]	.20	.71		1.08
Mean negative emotions	0.65 [0.52, 0.82]	<.01	.57		1.08
Model 4. EC (DI)				.04	
EC	1.15 [0.92, 1.43]	.22	.70		
Model 5. EC (DI) and life events				.04	
EC	1.14 [0.91, 1.43]	.25	.70		1.07
Life events	0.81 [0.65, 1.00]	.05	.62		1.00
EC X life events	0.97 [0.78, 1.21]	.79	.66		1.07
Model 6. EC (DI) and mean negative emotions				.07	
EC	0.96 [0.76, 1.22]	.74	.66		1.12
Mean negative emotions	0.61 [0.48, 0.77]	<.01	.55		1.12

Note. EC was assessed by the ICC (models 1-3) or DI (models 4-6). The OR reflects the increase in the odds of a good prognosis (relative to the odds of a bad prognosis) following a one standard deviation increase in the predictor for the average family, holding other predictors in the model constant. The inverse logit of the OR returns the probability of a good prognosis following a one standard deviation increase in the predictor for the average family. The conditional R^2 illustrates the amount of variance explained by the model. CI = confidence interval. DI = diversity index. EC = emotional complexity. ICC = intra-class correlation coefficient. OR = odds ratio. VIF = variance inflation factor (values > 5 are indicative of multicollinearity)³⁴.

Discussion

The present study investigated, for the first time, the prospective effects of EC on psychopathology at one-year follow-up while comparing two measures of EC (*i.e.*, emotion differentiation and emotion diversity). We found that low EC predicts an increased risk for future worsening or continuation of psychopathological symptoms when EC is based on the covariance of emotions (ICC), but not when EC is based on the diversity of emotions (DI). The significant effect of ICC was not contingent on the experience of stressful life events. Irrespective of its operationalization, lower EC coincided with a heightened intensity of negative affective states. This rendered the unique effect of EC (assessed by the ICC) on prognosis non-significant after controlling for the intensity of negative emotions. In conclusion, predicting future prognosis might not necessitate complex measures of emotional experience (the ICC and DI) but rather might be achieved through more parsimonious measures such as mean affect levels¹⁸. Nevertheless, the discrepant effects of the ICC and DI on prognosis do provide insight in the emotional processes that might contribute to individual differences in mental health. Specifically, intact emotion representation (ICC) might be more important for mental health than the richness and evenness of emotions (DI).

The extent to which emotions co-vary – assessed by the ICC – might reveal how emotions are represented and regulated³⁵. A high covariance between negative emotions suggests that individuals represent their emotions on a single dimension, ranging from pleasant to unpleasant. This suggests little explicit emotion knowledge, meaning that individuals with low EC (indexed by a high ICC) might have limited awareness not only of the distinction between different emotions, but also of the causes of emotional experiences, the expression of emotions, and the strategies for regulating emotions¹. Since such awareness may aid adequate coping with stressful situations, strongly co-varying emotions might predispose individuals to mental ill-health. In contrast, whether or not emotions are evenly distributed at single moments – assessed by the DI – appears to be uninformative of future psychopathology. This could mean that the analogy between ecosystems – which benefit from diversity – and emotional systems does not hold. In ecosystems, the domination of one species is potentially detrimental because species sustain each other and thereby preserve the system¹⁶. In emotional systems, in contrast, such symbiotic relations are implausible. Feeling down, for instance, does not rely on feeling anxious in the same way as a population of foxes relies on a population of rabbits. In conclusion, our findings suggest

that reduced emotion representation – indexed by strongly co-varying emotions – is more likely to underlie mental ill-health than low emotional diversity – indexed by the tendency to aggregate negative emotionality into a small number of emotions. Further, as the ICC and DI were largely unrelated, mentally representing a wide range of emotions (ICC) does not seem to result in reporting all these emotions to equal extent (DI). The independence between the ICC and the DI may suggest that EC primarily covers emotion differentiation rather than diversity.

In principle, EC – whether indexed by the ICC or DI – is independent of the absolute ratings on emotions. In practice, however, low EC co-occurs with heightened affect intensity (see Table 2.2). This has been reported in earlier studies as well^{6,18} – although contrasting findings have also been published^{10,15}. The association between emotion complexity and intensity might have several reasons. First, negative emotions reported by individuals from the general population might show floor effects, meaning that there is little variability over time¹⁸. This might lower the covariance between emotions (and hence, limit the ICC), leading to a negative association between EC (assessed by the reversed ICC) and emotion intensity. Second, high levels of negative affect may go together with a tendency to over-generalize affective states³⁶. For instance, depressed individuals typically describe their emotions in a non-specific, undifferentiated manner (*e.g.*, “I feel terrible”). Low EC and emotion intensity may thus be manifestations of the same construct, which could closely resemble neuroticism. Indeed, neuroticism not only relates to heightened negative emotion intensity³⁷ but also to low complexity of negative emotions^{37,38}, and seems to play an important role in individuals’ vulnerability for mental ill-health^{39–41}.

It is possible that adverse emotional experience impacts mental health particularly in times of stress. In line with this reasoning, we hypothesized that the adverse effects of low EC on prognosis would be most pronounced when adaptive emotion regulation is most essential, that is, in the face of stressful life events²⁵. Results did not support this hypothesis, suggesting that the interaction between EC and stress on psychopathology might be restricted to smaller (*i.e.*, moment-to-moment) timescales^{14,42}. However, the current design – where EC was measured at baseline and stressful life events were assessed in the ensuing year – might have limited our ability to test the interaction effect between EC, life events, and psychopathology. A design that measures EC and life events concurrently might be more suitable for investigating whether the effect of EC on mental health is contingent on stressful life events.

Strengths, limitations and future directions

The evaluation of two unrelated measures of EC (the ICC and DI) provided an improved understanding of the aspects of EC that are most informative of mental ill-health. Two other strengths of the current study are its prospective design, which allowed us to extend the existing cross-sectional findings on EC, and its considerable sample size ($N=401$), which allowed us to detect even small effects. As a result of this power, we found small but significant differences in age and sex distribution between lost ($N=438$) and retained participants. As both factors are unrelated to EC^{10,43} and the difference between groups was small (*i.e.*, retained participants: 10 months older, 12% less males), selection bias is unlikely to have affected our findings. A first limitation of the present study concerns the fact that – similar to earlier studies – we inferred EC from ratings on eight fixed emotions. This allowed for comparing ratings across individuals, but inevitably put constraints on the number of emotions individuals reported. In line with this, it has been suggested that designs where individuals can report the emotion(s) they experience without pre-defined labels composed by researchers might be more suitable for investigating EC²⁵. Such designs could particularly support the predictive utility of the DI, which has a limited range in designs with a fixed number of emotions⁴⁴. However, non-standardized formats for reporting emotions also challenge group-level analyses. As a second limitation, our sample was retrieved from the general population, suggesting that the worsening or persistence of symptoms (indicative of a bad prognosis) might not have the same clinical relevance as the symptoms reported in clinical samples^{9,11,13,43,45}. This might have compromised the association between EC and prognosis. Finally, we could not address whether and how EC changes over time. Investigating such changes might allow for disentangling emotional complexity from intensity. Further research might aim to delineate the boundary conditions for the association between EC and psychopathology and illustrate the dynamic associations between EC, affect intensity, and psychopathology.

Conclusions

We showed that the role of emotion complexity in mental ill-health is conditional on covariance of emotions, rather than the extent to which negative emotionality is evenly distributed across emotions. Further, both higher intensity and lower complexity of negative emotions mark a vulnerability for the persistence or worsening of psychopathological symptoms. However, intensity seems to suffice when predicting prognosis. Replication of the

current findings is needed in order to disentangle emotional intensity and complexity in the context of psychopathology. Finally, we found no evidence that stressful life events enhance the association between emotional complexity and future prognosis, although this requires further research.

Supplement 2

2.1 List of ESM items

The ESM questionnaire contained 13 items on affective states, of which 5 items were positively valenced and 8 items were negatively valenced. Only the negatively valenced items – marked with ● – were used for analysis.

- I feel cheerful
- I feel relaxed
- I feel satisfied
- I feel enthusiastic
- I feel energetic
- I feel insecure
- I feel lonely
- I feel anxious
- I feel irritated
- I feel listless
- I feel suspicious
- I feel down
- I feel guilty

Note that the ESM questionnaire also incorporated items on event appraisal (*e.g.*, “How pleasant was today’s most pleasant event?”), psychopathological symptoms (*e.g.*, “To what extent are you afraid of losing control?”), self-esteem (*e.g.*, “To what extent are you satisfied with yourself?”), activity and social company, and physical experiences (*e.g.*, “To what extent are you feeling hungry?”). These items are not listed here because they are not relevant for the current study.

2.2 Subscale scores SCL-90

Table S2.1. *SCL-90 subscale scores*

Subscale	Mean (standard deviation) baseline	Mean (standard deviation) follow-up
Sleeping problems	0.80 (0.79)	0.76 (0.81)
Inadequacy of thinking and acting	0.70 (0.54)	0.62 (0.51)
Depression	0.55 (0.53)	0.48 (0.46)
Interpersonal sensitivity	0.53 (0.47)	0.45 (0.42)
Somatic complaints	0.48 (0.44)	0.45 (0.47)
Hostility	0.41 (0.46)	0.36 (0.36)
Anxiety	0.41 (0.46)	0.34 (0.43)
Agoraphobia	0.19 (0.36)	0.14 (0.30)

2.3 Sensitivity analysis complete sample

Table S2.2. *Multilevel logistic regression analyses based on sample without individuals whose ICC value was negative (N=22), resulting in N=379*

	OR [95% CI]	<i>p</i>	Probability good prognosis	Conditional R ²	VIF
Model 1. EC (ICC)				.13	
EC	2.44 [1.75-3.39]	<.01	.86		
Model 2. EC (ICC) and life events				.08	
EC	2.70 [1.26-5.79]	.01	.87		1.20
Life events	0.95 [0.71-1.28]	.74	.70		1.87
EC X life events	1.57 [0.76-3.28]	.23	.79		1.96
Model 3. EC (ICC) and mean negative emotions				.10	
EC	1.60 [0.76-3.37]	.21	.78		1.09
Mean negative emotions	0.66 [0.52-0.84]	<.01	.59		1.09
Model 4. EC (DI)				.05	
EC	1.19 [0.94-1.51]	.15	.70		
Model 5. EC (DI) and life events				.05	
EC	1.19 [0.93-1.51]	.16	.70		1.04
Life events	0.81 [0.65-1.01]	.06	.61		1.00
EC X life events	0.97 [0.77-1.23]	.81	.65		1.04
Model 6. EC (DI) and mean negative emotions				.08	
EC	1.01 [0.78-1.29]	.97	.67		1.09
Mean negative emotions	0.62 [0.49-0.79]	<.01	.55		1.09

Note. EC was operationalized as the ICC (models 1-3) or DI (models 4-6). The OR reflects the increase in the odds of a good prognosis (relative to the odds of a bad prognosis) following a one standard deviation increase in the predictor for the average family, holding other predictors in the model constant. The inverse logit of the OR returns the probability of a good prognosis following a one standard deviation for the average family. The conditional R² illustrates the amount of variance explained by the model. CI = confidence interval. DI = diversity index. EC = differentiation of negative affective states. ICC = intra-class correlation coefficient. OR = odds ratio. VIF = variance inflation factor (values > 5 are indicative of multicollinearity³⁴).

2.4 Sensitivity analysis continuous outcome

For the analysis described in the main text, we converted a continuous measure (symptom severity at baseline and follow-up, quantified by the global severity index (GSI) of the SCL-90) into a dichotomous variable (good vs. bad prognosis). This was done in order to account for the fact that small absolute changes in symptom severity can signal both benign (in case of low baseline symptom severity) and adverse (in case of high baseline symptom severity) symptom trajectories. In other words, analyses based on continuous measures of symptom severity would have treated persistently symptomatic and persistently a-symptomatic individuals as identical.

To explore the effects of analyzing symptom severity as a dichotomous rather than continuous measure, we re-ran our analyses with a continuous outcome (Table S2.3). Specifically, multilevel models were run to assess whether follow-up GSI could be predicted by (i) baseline GSI, (ii) emotional complexity – reflected in emotion differentiation and diversity (model 1, 4), (iii) the interaction between emotional complexity and life events (model 2, 5) and (iv) emotional complexity, accounting for the effect of emotion intensity (model 3, 6). To prevent erroneously grouping persistently symptomatic and a-symptomatic individuals – who both report small changes in symptom severity, yet have different clinical profiles – we excluded individuals who reported persistent symptoms ($N=71$) from analyses, resulting in $N=330$. As denoted in Table S2.3, the effect of EC (ICC) on follow-up symptoms corrected for baseline symptoms was not statistically significant ($\beta = -0.02$, $p = .10$). This might be due to the fact that baseline and follow-up symptoms were moderately correlated ($r = .42$, $p < .01$), leaving relatively little variance in follow-up symptoms left to be explained by emotional complexity.

Table S2.3. *Multilevel analyses based on sample without individuals with persistent above-threshold symptoms (N=71), resulting in N=330*

	β [95% CI]	p	Conditional R^2	VIF
Model 1. EC (ICC)			.22	
EC	-0.02 [-0.05-0.00]	.10		1.04
Model 2. EC (ICC) and life events			.32	
EC	-0.02 [-0.05-0.00]	.11		1.16
Life events	0.05 [0.02-0.08]	<.01		1.05
EC X life events	-0.01 [-0.04-0.02]	.58		1.12
Model 3. EC (ICC) and mean negative emotions			.24	
EC	-0.01 [-0.04-0.01]	.36		1.11
Mean negative emotions	0.05 [0.01-0.08]	<.01		1.31
Model 4. EC (DI)			.21	
EC	0.00 [-0.03-0.03]	.96		1.04
Model 5. EC (DI) and life events			.30	
EC	0.00 [-0.03-0.03]	.92		1.14
Life events	0.05 [0.02-0.08]	<.01		1.05
EC X life events	0.01 [-0.02-0.04]	.41		1.11
Model 6. EC (DI) and mean negative emotions			.23	
EC	0.02 [-0.01-0.05]	.29		1.18
Mean negative emotions	0.06 [0.02-0.10]	<.01		1.39

Note. EC was operationalized as the ICC (models 1-3) or DI (models 4-6). P-values were estimated using Satterthwaites approximations. For readability, the effect of baseline GSI was omitted from the model results. CI = confidence interval. DI = diversity index. EC = emotional complexity. ICC = intra-class correlation coefficient. VIF = variance inflation factor (values > 5 are indicative of multicollinearity³⁴).

References

1. Barrett LF, Gross J, Christensen TC, Benvenuto M. Knowing what you're feeling and knowing what to do about it: mapping the relation between emotion differentiation and emotion regulation. *Cognition and Emotion*. 2001;15(6):713-724. doi:10.1080/02699930143000239
2. Quoidbach J, Gruber J, Mikolajczak M, Kogan A, Kotsou I, Norton MI. Emodiversity and the emotional ecosystem. *Journal of Experimental Psychology: General*. 2014;143(6):2057-2066. doi:10.1037/a0038025
3. Kang SM, Shaver PR. Individual differences in emotional complexity: their psychological implications. *Journal of Personality*. 2004;72(4):687-726. doi:10.1111/j.0022-3506.2004.00277.x
4. Lindquist KA, Barrett LF. Emotional complexity. In: Lewis M, Haviland-Jones JM, Barrett LF, eds. *Handbook of Emotions*. 2nd ed. Guilford Press; 2000:513-530.
5. Grossmann I, Huynh AC, Ellsworth PC. Emotional complexity: clarifying definitions and cultural correlates. *Journal of Personality and Social Psychology*. 2016;111(6):895-916.
6. Erbas Y, Ceulemans E, Blanke ES, Sels L, Fischer A, Kuppens P. Emotion differentiation dissected: between-category, within-category, and integral emotion differentiation, and their relation to well-being. *Cognition and Emotion*. 2019;33(2):258-271. doi:10.1080/02699931.2018.1465894
7. Grossmann I, Gerlach TM, Denissen JJA. Wise reasoning in the face of everyday life challenges. *Social Psychological and Personality Science*. 2016;7(7):611-622. doi:10.1177/1948550616652206
8. Kashdan TB, Barrett LF, McKnight PE. Unpacking emotion differentiation. *Current Directions in Psychological Science*. 2015;24(1):10-16. doi:10.1177/0963721414550708
9. Demiralp E, Thompson RJ, Mata J, et al. Feeling blue or turquoise? Emotional differentiation in major depressive disorder. *Psychological Science*. 2012;23(11):1410-1416. doi:10.1177/0956797612444903
10. Willroth EC, Flett JAM, Mauss IB. Depressive symptoms and deficits in stress-reactive negative, positive, and within-emotion-category differentiation: a daily diary study. *Journal of Personality*. 2019;88(2):1-11. doi:10.1111/jopy.12475
11. Kashdan TB, Farmer AS. Differentiating Emotions Across Contexts: Comparing Adults with and without Social Anxiety Disorder Using Random, Social Interaction, and Daily Experience Sampling. *Emotion*. 2014;14(3):629-638. doi:10.1158/0008-5472.CAN-10-4002.BONE
12. Suvak MK, Litz BT, Sloan DM, Zanarini MC, Barrett LF, Hofmann SG. Emotional granularity and borderline personality disorder. *Journal of Abnormal Psychology*. 2011;120(2):414-426. doi:10.1037/a0021808
13. Widdershoven RLA, Wichers M, Simons CJP, et al. Effect of self-monitoring through experience sampling on emotion differentiation in depression. *Journal of Affective Disorders*. 2019;244:71-77. doi:10.1016/j.jad.2018.10.092
14. Erbas Y, Ceulemans E, Kaloerinos E, et al. Why I don't always know what I'm feeling: the role of stress in within-person fluctuations in emotion differentiation. *Journal of Personality and Social Psychology*. 2018;115(2):179-191.
15. Starr LR, Hershenberg R, Li YI, Shaw ZA. When feelings lack precision: low positive and negative emotion differentiation and depressive symptoms in daily life. *Clinical Psychological Science*. 2017;5(4):613-631. doi:10.1177/2167702617694657
16. Scheffer M. *Critical Transitions in Nature and Society*. Princeton University Press; 2009.
17. Houben M, van den Noortgate W, Kuppens P. The relation between short-term emotion dynamics and

- psychological well-being: a meta-analysis. *Psychological Bulletin*. 2015;141(4):901-930. doi:10.1037/a0038822
18. Dejonckheere E, Mestdagh M, Houben M, et al. Complex affect dynamics add limited information to the prediction of psychological well-being. *Nature Human Behaviour*. 2019;3(5):478-491. doi:10.1038/s41562-019-0555-0
 19. Pries LK, Guloksuz S, Menne-Lothmann C, et al. White noise speech illusion and psychosis expression: An experimental investigation of psychosis liability. *PLoS ONE*. 2017;12(8):1-10. doi:10.1371/journal.pone.0183695
 20. Loos R, Derom C, Vlietinck R, Derom R. The East Flanders Prospective Twin Survey (Belgium): a population-based register. *Twin Research*. 1998;1(4):167-175. doi:10.1375/136905298320566131
 21. Derom C, Thiery E, Rutten BPF, et al. The East Flanders Prospective Twin Survey (EFPTS): 55 years later. *Twin Research and Human Genetics*. 2019;22:1-6. doi:10.1017/thg.2019.64
 22. Myin-Germeys I, Birchwood M, Kwapil T. From environment to therapy in psychosis: a real-world momentary assessment approach. *Schizophrenia Bulletin*. 2011;37(2):244-247. doi:10.1093/schbul/sbq164
 23. Delespaul PAEG. Assessing schizophrenia in daily life: the experience sampling method. Published online 1995.
 24. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*. 1979;86(2):420-428. doi:10.1037/0033-2909.86.2.420
 25. Ottenstein C, Lischetzke T. Development of a novel method of emotion differentiation that uses open-ended descriptions of momentary affective states. *Assessment*. 2019;27(8):1928-1945. doi:10.1177/1073191119839138
 26. Bartko JJ. On various intraclass correlation reliability coefficients. *Psychological Bulletin*. 1976;83(5):762-765. doi:10.1037/0033-2909.83.5.762
 27. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychological Methods*. 1996;1(1):30-46. doi:10.1037/1082-989X.1.1.30
 28. Field A. Comparing several means: ANOVA (GLM1). In: *Discovering Statistics Using SPSS*. Third. SAGE Publications; 2009:348-358.
 29. Arrindell WA, Ettema JHM. *Symptom Checklist (SCL-90)*. Handleiding Bij Een Multidimensionele Psychopathologie-Indicator. Swets & Zeitlinger; 1986.
 30. Schauenburg H, Strack M. Measuring psychotherapeutic change with the Symptom Checklist SCL 90 R. *Psychotherapy and Psychosomatics*. 1999;68(4):199-206. doi:10.1159/000012333
 31. Paykel ES. The interview for recent life events. *Psychological Medicine*. 1997;27:301-310.
 32. R Core Team. R: A language and environment for statistical computing. Published online 2020.
 33. Olvera Astivia OL, Gadermann A, Guhn M. The relationship between statistical power and predictor distribution in multilevel logistic regression: a simulation-based approach. *BMC Medical Research Methodology*. 2019;19(1):1-20. doi:10.1186/s12874-019-0742-8
 34. Menard S. An introduction to logistic regression diagnostics. In: *Applied Logistic Regression Analysis*. 2nd ed. SAGE Publications; 2002:68-91. doi:10.4135/9781412983433.n4
 35. Gross JJ, Jazaieri H. Emotion, emotion regulation, and psychopathology: an affective science perspective. *Clinical Psychological Science*. 2014;2(4):387-401. doi:10.1177/2167702614536164
 36. van der Gucht K, Dejonckheere E, Erbas Y, et al. An experience sampling study examining the potential

- p>impact of a mindfulness-based intervention on emotion differentiation.
- Emotion*
- . 2019;19(1):123-131. doi:10.1037/emo0000406
37. Erbas Y, Ceulemans E, Lee Pe M, Koval P, Kuppens P. Negative emotion differentiation: its personality and well-being correlates and a comparison of different assessment methods. *Cognition and Emotion*. 2014;28(7):1196-1213. doi:10.1080/02699931.2013.875890
 38. Carstensen LL, Mayr U, Pasupathi M, Nesselroade JR. Emotional experience in everyday life across the adult life span. *Journal of Personality and Social Psychology*. 2000;79(4):644-655. doi:10.1037/0022-3514.79.4.644
 39. Ormel J, Rosmalen J, Farmer A. Neuroticism: a non-informative marker of vulnerability to psychopathology. *Social Psychiatry and Psychiatric Epidemiology*. 2004;39(11):906-912. doi:10.1007/s00127-004-0873-y
 40. Jeronimus BF, Kotov R, Riese H, Ormel J. Neuroticism's prospective association with mental disorders halves after adjustment for baseline symptoms and psychiatric history, but the adjusted association hardly decays with time: a meta-analysis on 59 longitudinal prospective studies with 443 313 participants. *Psychological Medicine*. 2016;46(14):2883-2906. doi:10.1017/S0033291716001653
 41. Lahey BB. Public health significance of neuroticism. *American Psychologist*. 2009;64(4):241-256. doi:10.1037/a0015309
 42. Smidt KE, Suvak MK. A brief, but nuanced, review of emotional granularity and emotion differentiation research. *Current Opinion in Psychology*. 2015;3:48-51. doi:10.1016/j.copsyc.2015.02.007
 43. Kimhy D, Vakhrusheva J, Khan S, et al. Emotional granularity and social functioning in individuals with schizophrenia: an experience sampling study. *Journal of Psychiatric Research*. 2014;53:141-148. doi:10.1016/j.biotechadv.2011.08.021.Secreted
 44. Brown NJL, Coyne JC. Emodiversity: robust predictor of outcomes or statistical artifact? *Journal of Experimental Psychology: General*. 2017;146(9):1372-1377. doi:10.1037/xge0000330
 45. Erbas Y, Ceulemans E, Boonen J, Noens I, Kuppens P. Emotion differentiation in autism spectrum disorder. *Research in Autism Spectrum Disorders*. 2013;7(10):1221-1227. doi:10.1016/j.rasd.2013.07.007

3

Early warning signals in psychopathology: what do they tell?

published as: Schreuder, M.J., Hartman, C.A., George, S.V., Menne-Lothmann, C., Decoster, J., van Winkel, R., Delespaul, P., de Hert, M., Derom, C., Thiery, E., Rutten, B.P.F., Jacobs, N., van Os, J., Wigman, J.T.W., & Wichers, M. (2020). Early warning signals in psychopathology: what do they tell? *BMC Medicine*, 18(1). doi: 10.1186/s12916-020-01742-3

Abstract

Background. Despite the increasing understanding of factors that might underlie psychiatric disorders, prospectively detecting shifts from a healthy towards a symptomatic state has remained unattainable. A complex systems perspective on psychopathology implies that such symptom shifts may be foreseen by generic indicators of instability, or early warning signals (EWS). EWS include, for instance, increasing variability, covariance, and autocorrelation in momentary affective states – of which the latter was studied. The present study investigated if EWS predict (i) future worsening of symptoms as well as (ii) the type of symptoms that will develop, meaning that the association between EWS and future symptom shifts would be most pronounced for congruent affective states and psychopathological domains (*e.g.*, feeling down and depression).

Methods. A registered general population cohort of adolescents (mean age 18 years, 36% males) provided ten daily ratings of their affective states for six consecutive days. The resulting time series were used to compute EWS in feeling down, listless, anxious, not relaxed, insecure, suspicious, and unwell. At baseline and one-year follow-up, symptom severity was assessed by the Symptom Checklist-90 (SCL-90). We selected four subsamples of participants who reported an increase in one of the following SCL-90 domains: depression (N=180), anxiety (N=192), interpersonal sensitivity (N=184), or somatic complaints (N=166).

Results. Multilevel models showed that EWS in feeling suspicious anticipated increases in interpersonal sensitivity, as hypothesized. EWS were absent for other domains. Although the association between EWS and symptom increases was restricted to the interpersonal sensitivity domain, post-hoc analyses showed that symptom severity at baseline was related to heightened autocorrelations in congruent affective states for interpersonal sensitivity, depression, and anxiety. This pattern replicated in a second, independent dataset.

Conclusions. The presence of EWS prior to symptom shifts may depend on the dynamics of the psychopathological domain under consideration: for depression, EWS may manifest only several weeks prior to a shift, whereas for interpersonal sensitivity, EWS may already occur one year in advance. Intensive longitudinal designs where EWS and symptoms are assessed in real-time are required in order to determine at what timescale and for what type of domain EWS are most informative of future psychopathology.

Introduction

Psychiatric disorders affect approximately 1 in 4 individuals at some point in their lives¹. Given the burden associated with such disorders, preventing their onset and progression may substantially improve individuals' well-being. Traditionally, the quest for prevention has been pursued through understanding the risk factors and mechanisms that may give rise to psychopathology². At present, numerous risk factors have been proposed. Yet, the complexity of the mechanisms by which they contribute to psychopathology challenge accurate identification of who is when at risk for developing symptoms. This calls for a novel perspective on psychopathology.

A complex systems perspective may answer this call^{3–5}. According to this perspective, sudden shifts – including increases in psychiatric symptoms – may be understood regardless of our limited insight in the mechanisms that cause them. The underlying premise is that psychopathology can be described by distinct equilibrium states featured by absent/mild or severe symptoms, respectively. Shifts between these equilibria, marking the onset or remittance of symptoms, may occur in an abrupt, stepwise fashion rather than gradually^{6,7}. The likelihood of such shifts depends on the system's stability⁸. In a stable system, temporary departures from equilibrium caused by small environmental perturbations are met with regulatory processes that quickly restore the equilibrium. In other words, stressful events (perturbations) may cause brief increases in symptoms (departure from equilibrium), followed by remission of these symptoms (return to equilibrium). As instability increases, the system's resilience against perturbations declines, and consequently, a shift towards another equilibrium (*e.g.*, presence of symptoms) becomes more likely^{9,10}. Exposing a system's instability might thus allow us to foresee sudden shifts from absent/mild to more severe symptoms.

Empirical support for complex systems principles in psychopathology was provided by studies that confirmed that markers of instability – specifically, critical fluctuations – indeed predict symptom shifts^{11–13}. The present chapter does not discuss critical fluctuations^{8,11,14–16}, but rather focuses on critical slowing down as a marker of instability. Critical slowing down refers to an increasingly slowed return to equilibrium^{10,17}. Because of their ability to warn for an upcoming shift to an alternate equilibrium state, indicators of critical slowing down have been referred to as early warning signals (EWS)¹⁰. EWS include, for instance, increasing variability, covariance, and autocorrelation, as well as multivariate extensions of such

metrics¹⁸. Because autocorrelations are relatively robust EWS and have been established previously in observational, psychological data^{4,5,19}, the remainder of this chapter^b will focus on this metric^{9,10,17,19}. In the context of psychopathology, autocorrelations refer to the degree to which an individual's current affective state (*e.g.*, feeling down) is predictive of his/her future affective state^{4,20–22}. When autocorrelations are high, temporary rises in an affective state in response to a stressful event (*e.g.*, feeling down after failing a test) persist over time. This resistance to change in affective states prevents the system from quickly returning to its equilibrium state (*e.g.*, absence of symptoms, and hence, not feeling down)^{20,23–25}. A complex systems perspective on psychopathology thus proposes that elevated autocorrelations in momentary affective states might warn for an upcoming shift in symptoms. Such shifts can theoretically reflect either a decrease or an increase in symptoms⁴, although most research so far has focused on increases in symptoms.

Earlier studies have provided indirect support for EWS in psychopathology by showing that increased autocorrelations in affective states might cross-sectionally relate to several maladaptive characteristics, including low self-esteem^{24,25}, neuroticism²³, and the presence of psychiatric disorders^{20,21,24,26,27}. These studies raised two compelling questions. Specifically, it remained unclear (i) whether autocorrelations in affective states also prospectively predict psychopathology – as would be expected from complex systems principles, and (ii) whether the predictive utility of autocorrelations extends beyond that of mean affect levels²⁸.

Initial support for the prospective association between EWS – operationalized by elevated autocorrelations in affective states – and symptom shifts was provided by Van de Leemput and colleagues⁴, who demonstrated that individuals with higher EWS in affective states, such as feeling down and anxious, later reported higher depressive symptom severity. EWS were inferred from momentary ratings of affective states acquired through the experience sampling method (ESM). These results were in line with what would be expected from complex systems principles, but did not prove the presence of EWS prior to worsening symptoms within individuals^{29–31}. This specific question was addressed in a single-case study, which showed that a relapse in depression was preceded by rising autocorrelations in affective states⁵. These studies thus tentatively suggest that shifts in symptoms may indeed follow principles of complex systems. This implies that EWS could serve as a person-specific marker for vulnerability to future psychopathology. Before translating the existing evidence

^b Throughout the rest of this chapter, EWS will refer to autocorrelations.

to clinical practice, however, these prior findings require substantiation on a larger scale. A first step in this direction is replicating the first prospective group-level study⁴.

A second step that awaits empirical substantiation is to investigate whether EWS not only indicate the likelihood of shifts in the *severity* of symptoms, but also reveal the *type* of symptoms that will develop (*e.g.*, depression, anxiety, etc.)³². This hypothesis proposes that the association between EWS in affective states and future increases in symptoms is particularly pronounced for affective states congruent with the psychopathological domain involved in the shift. For instance, EWS in feeling down might precede an upcoming shift towards depression, whereas EWS in feeling anxious might precede increases in anxiety³². If this hypothesis holds, EWS could inform clinicians about an individual's vulnerability for specific disorders.

The current study was designed to further explore the role EWS in emerging psychopathology in two ways. First, building on the findings reported by Van de Leemput *et al.*⁴, we aimed to test whether EWS in affective states anticipate impending increases in depressive symptom severity. Second, we aimed to extend previous findings by examining whether EWS are predictive of the type of symptoms that will develop. We hypothesized that the association between symptom increases and EWS would be particularly pronounced for affective states congruent with the domain in which symptoms increased. The present chapter specifically focusses on symptom increases, as opposed to decreases, because of our interest in emerging psychopathology and the possibility of early detection thereof. In view of a recent meta-analysis, which suggested that associations between autocorrelations and psychopathology might be attributable to individual differences in mean affect levels²⁸, we also investigated the associations between affect intensity (*i.e.*, mean levels) and symptom increases. Similar to Van de Leemput *et al.*⁴, we analyzed a general population sample who provided ten daily affect ratings for six days using the ESM, resulting in times series that were used to compute EWS. Prior to the ESM and one year later, symptom severity was assessed. Participants who reported an increase in symptoms in one of the following domains were analyzed (hypothesized congruent affective states between brackets): depression (feeling down, listless), anxiety (feeling anxious, not relaxed), interpersonal sensitivity (feeling insecure, suspicious), and somatic complaints (feeling unwell).

Methods

Participants

Data were retrieved from the TwinssCan study³³, which comprised a subset of a registered cohort of twins from the general population (*i.e.*, East Flanders Prospective Twin Survey)³⁴. The study included N=839 Caucasian twins aged between 15 and 34 years old. As particularly adolescents were invited to participate, most participants (N=639) were between 15 and 18 years old. Complete baseline and follow-up SCL-90 ratings were available for 467 participants. We selected those participants who reported an increased symptom severity over time (*i.e.*, baseline rating < follow-up rating) in one of the following domains of the Symptom Checklist-90 (SCL-90)³⁵: depression, anxiety, somatic complaints, and interpersonal sensitivity. This criterion was adopted because we were interested in the development, rather than the remission, of symptoms. Symptom increases covered a broad range from minor impairments to clinically significant shifts – thereby allowing us to investigate symptom increases continuously⁴.

To support the reliability of autocorrelation estimates, participants were excluded if less than 20 pairs of consecutive affect ratings were available. This resulted in four samples (N=166, 184, 188, and 192). Since most participants reported increases in more than one domain, these samples were not unique in their composition. In total, the samples comprised 293 unique individuals, of whom 222 were included in more than one sample (see supplement 3.1. *Sample composition*, Table S3.1). The TwinssCan study was approved by the Local Ethics Committee and all subjects provided written informed consent. For minors, parents provided additional written consent.

Experience sampling method

Affective states were assessed through the ESM, which is a structured diary technique suitable for assessing affect in daily life. The ESM involved questions on momentary affect, context, and behavior, that were administered through PsyMates^{®36}. Participants received a PsyMate (an electronic device) and an instruction about its usage at the beginning of the study. The PsyMate was programmed to emit a beep-signal at ten semi-random time intervals within 90-minute blocks ranging from 7:30 AM to 10:30 PM for six consecutive days. The beep-signal alarmed participants to fill in a questionnaire concerning their current affective state. To support reliability and validity of these ratings, questionnaires that were completed more

than 15 minutes after the beep were coded as missing³⁷. Items were rated on a 7-point Likert scale, ranging from ‘not at all’ to ‘very much’. Analyses were limited to a selection of seven items. These items were selected based on their congruency with one of the following psychopathological domains of interest (items between brackets): depression (feeling down, listless), anxiety (feeling anxious, not relaxed), interpersonal sensitivity (feeling insecure, suspicious), somatic complaints (feeling unwell; Table 3.1).

Table 3.1. *Psychopathological domains and congruent affective states*

Psychopathological domain (SCL-90)	Depression	Anxiety	Somatic complaints	Interpersonal sensitivity
ESM item	I feel down I feel listless	I feel anxious I do not feel relaxed	I feel unwell	I feel insecure I feel suspicious

Note. ESM = experience sampling method. SCL-90 = Symptom Checklist-90.

Assessment of psychopathology

Prior to the ESM as well as one year later, participants completed the Dutch version of the SCL-90³⁵. This questionnaire consists of 90 items and has good psychometric properties³⁸. Items of the SCL-90 are rated on a 5-point Likert scale, ranging from ‘not at all’ to ‘very often/always’ (e.g., “During the last week, I felt empty”). Besides self-reported ratings of symptom severity in eight domains, the SCL-90 provides a global severity index (GSI) which can be used for screening purposes³⁸. A GSI of 0.57 or higher has been suggested to optimally differentiate those who are at risk for psychiatric disorders from those who are not³⁹.

We selected the domains depression, anxiety, somatic complaints, and suspicion and interpersonal sensitivity (the latter two together referred to as interpersonal sensitivity) for analyses. Difference scores within these domains (*i.e.*, follow-up score – baseline score) were computed to estimate changes in symptom severity. The domains depression and anxiety represent symptoms recognized in the corresponding clusters in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5)⁴⁰. Scores on the somatic complaints domain reflect non-specific physical discomfort. Finally, the interpersonal sensitivity domain incorporates feelings of interpersonal incompetence, suspicion, and paranoia. Other symptom clusters assessed by the SCL-90 were not selected because (i) the SCL-90 domain was not covered by ESM items (e.g., agoraphobia, sleeping problems, hostility, and insufficient thinking/feeling) or (ii) the psychopathological domain was non-specific (e.g., other complaints).

Analyses

We specified a multilevel model^c to evaluate whether autocorrelations in affective states were related to increases in symptom severity. A similar model has been used by earlier studies^{4,23,24,26,41–43}. In this model, the momentary affect rating (*e.g.*, feeling down) of individual *i* belonging to twin-pair *j* at time *t* ($affected_{tij}$) was predicted by the individual's mean-centered affect rating at time *t-1* ($affected_{t-1ij} - \mu_i$), the individual's increase in psychopathological symptoms from baseline to follow-up (P_i) and an interaction term between symptom increases and the lagged person-mean centered affect rating ($P_i \times (affected_{t-1ij} - \mu_i)$). The model can be regarded as a multilevel extension of a standard autoregressive model: at the within-person level, we estimate an autoregressive model and at the between-person level, we estimate the interaction between autoregressive coefficients and symptom increases.

In line with recommendations of De Haan-Rietdijk *et al.*⁴⁴ and earlier studies^{42,45–48}, affect ratings at time *t-1* were person-mean centered in order to ensure that parameters in the model were not affected by between-person differences in mean affect. Further, the first rating for each day was coded as missing in order to avoid autocorrelations that spanned a whole night rather than a few hours^{24,41}. We accounted for within-person and within-twin covariances in affect ratings by including a random intercept at the level of individuals and twins. Furthermore, a random effect for lagged affect ratings was estimated. In our multilevel model, β_{3ij} signifies the interaction effect of person-mean centered lagged affect ratings and symptom increases on actual affect ratings. This parameter thus illustrates the association between auto-correlating affective states and shifts in symptom severity.

Our first aim was to investigate whether EWS precede increases in depression. Therefore, analyses were restricted to increases in depressive symptoms, with affect items feeling down and feeling listless. Van de Leemput *et al.*⁴ used similar, yet less specific affective states, which were labeled according to their valence (positive/negative) and arousal (high/low). Our second aim was to explore whether autocorrelations in affective states are informative for the direction of shifts in symptom severity. This aim was addressed by

^c Multilevel model to investigate the association between autocorrelation in affective states and shifts in psychopathology:

$$affected_{tij} = \beta_{0ij} + \beta_{1ij}(affected_{t-1ij} - \mu_i) + \beta_{2ij}P_i + \beta_{3ij}(P_i * (affected_{t-1ij} - \mu_i)) + \varepsilon_{ij}$$

This equation translates to R code as follows: $affected_t \sim (affected_{t-1} - \mu_i) + P + (P \times (affected_{t-1} - \mu_i)) + (1 + affected_{t-1} - \mu_i)$, where *twind* denotes the factor that distinguishes between twin pairs and *ID* represents the individual.

applying the multilevel model^c to each of the four psychopathological domains and each of the seven affect items. The coefficients of the interaction effect between lagged affect ratings and increases in symptom severity on current affect ratings (β_{3ij}) were standardized by multiplying the coefficients with the ratio of the standard deviation in the linear predictor ($P_i * (affect_{t-1ij} - \mu_i)$) and the standard deviation of the outcome ($affect_{tij}$), respectively⁴⁹. The standardized coefficients were then compared in order to assess for which affect items the relation between autocorrelations and increases in symptom severity was most pronounced. Comparisons were made across models (*i.e.*, items) within domains and only for those items that were statistically significant⁵⁰. The statistical significance of coefficients was adjusted according to Hochberg's procedure⁵¹, so that within each domain the Type I error rate was .05.

The association between mean affect levels and symptom increases could not be retrieved directly from the multilevel model^c. Hence, we employed linear regression analyses in order to examine whether mean affect levels are equally predictive of future symptom increases. Similar to Van de Leemput *et al.*⁴ and solely for the purposes of visualization, we re-ran the multilevel model outlined in footnote c with categorized instead of continuous change scores of symptom severity (P_i). For each psychopathological domain, P_i scores were categorized based on tertile scores, resulting in three categories with approximately equal N (*i.e.*, low, medium, and high symptom severity). Analyses were conducted in R⁵² (version 4.0).

Results

Participants were grouped according to the SCL-90 domain in which they reported an increase in symptom severity. The resulting samples (N=166-192) were similar regarding age, sex distribution, and SCL-90 domain scores at baseline and follow-up (Table 3.2; for details, see supplement 3.1. *Sample composition*, Table S3.1). Based on the global severity index (GSI) of the SCL-90, the majority of participants scored below the clinical threshold (at baseline: 82-87%, at follow-up: 67-71%)³⁹. Symptom increases exceeded the reliable change index reported by Schauenberg and Strack³⁹ in 26-31% of individuals. On average, participants completed 45.4 (76%) affect ratings.

Table 3.2. *Sample characteristics*

	Depression (N=180)	Anxiety (N=192)	Somatic complaints (N=184)	Interpersonal sensitivity (N=166)
No. male (%)	63 (35)	66 (34)	67 (36)	64 (39)
Age, mean (<i>SD</i>), years	17.9 (4.30)	18.0 (4.40)	17.9 (4.00)	17.9 (4.30)
No. of completed affect ratings, mean (<i>SD</i>)	45.3 (7.70)	45.4 (7.50)	45.4 (7.50)	45.4 (7.20)
Baseline SCL-90 domain score, mean (<i>SD</i>)	1.32 (0.33)	1.26 (0.34)	1.31 (0.37)	1.34 (0.34)
Follow-up SCL-90 domain score, mean (<i>SD</i>)	1.60 (0.51)	1.47 (0.51)	1.58 (0.56)	1.60 (0.50)
Baseline SCL-90 GSI, mean (<i>SD</i>)	0.36 (0.28)	0.39 (0.30)	0.40 (0.32)	0.36 (0.30)
Baseline no. above clinical threshold (%) ^a	24 (13)	35 (18)	32 (17)	25 (15)
Follow-up SCL-90 GSI, mean (<i>SD</i>)	0.49 (0.38)	0.48 (0.38)	0.49 (0.40)	0.51 (0.40)
Follow-up no. above clinical threshold (%)	55 (31)	56 (29)	57 (31)	54 (33)
Symptom increases above RCI (%) ^b	56 (31)	50 (26)	48 (26)	52 (31)
<i>ESM items, mean (SD)</i>				
Down	1.64 (0.78)	1.69 (0.80)	1.74 (0.85)	1.70 (0.83)
Listless	1.74 (0.86)	1.81 (0.88)	1.79 (0.87)	1.79 (0.90)
Anxious	1.39 (0.59)	1.45 (0.64)	1.45 (0.66)	1.44 (0.64)
Not relaxed	2.84 (1.18)	2.90 (1.20)	2.90 (1.19)	2.84 (1.21)
Unwell	2.73 (0.83)	2.78 (0.84)	2.79 (0.83)	2.77 (0.84)
Insecure	1.60 (0.79)	1.69 (0.84)	1.71 (0.85)	1.63 (0.83)
Suspicious	1.38 (0.56)	1.43 (0.61)	1.42 (0.61)	1.39 (0.58)

Note. Samples were labeled according to the SCL-90 domain that was evaluated in subsequent analyses. Note that mean SCL-90 domain scores refer to the domain of interest (e.g., for depression, mean raw SCL-90 scores for the depression domain are presented). Means and standard deviations of ESM items were calculated within individuals within samples. ESM = experience sampling method. GSI = global severity index. RCI = reliable change index. SCL-90 = Symptom Checklist 90. SD = standard deviation. ^aFollowing the recommendations of Schauenburg and Strack³⁹, a GSI of 0.57 was adopted as clinical threshold. ^bThe RCI equals 0.43 (for baseline GSI scores > .57) or 0.16 (for baseline GSI scores < 0.57³⁹).

Our first aim was to investigate whether increases in depressive symptom severity are related to heightened autocorrelations in the affective states *feeling down* and *feeling listless*. For both *feeling down* and *feeling listless*, multilevel models did not reveal a significant interaction effect between symptom increases and person-mean centered affect ratings at time *t-1* on affect ratings at time *t* (Table 3.3; *feeling down*: $\beta = 0.02$, $SE_{\beta} = 0.02$, $p = .75$, marginal $R^2 =$

0.04, conditional $R^2 = 0.37^d$; *feeling listless*: $\beta = 0.01$, $SE_{\beta} = 0.03$, $p = .75$, marginal $R^2 = 0.04$, conditional $R^2 = 0.40$). In other words, autocorrelations in *feeling down* and *feeling listless* were not predictive of future symptom increases in adolescents from the general population.

Our second aim was to examine whether EWS, reflected in autocorrelations in different affective states, are predictive of the direction of increases in symptoms. After correcting for an inflated Type I error rate using Hochberg's procedure⁵¹, autocorrelations in *feeling suspicious* were predictive of larger symptom increases in the interpersonal sensitivity domain ($\beta = 0.12$, $SE_{\beta} = 0.03$, $p < .01$, marginal $R^2 = 0.04$, conditional $R^2 = 0.42$). The association between auto-correlating affective states and symptom increases was not statistically significant for other psychopathological domains (Table 3.3, Fig. 3.1). Other coefficients estimated in the multilevel model described in footnote c – *i.e.*, main effects of lagged person-mean centered affect ratings and symptom increases, respectively – as well as the unstandardized effects are denoted in the supplement 3.2 *Model results*, Tables S3.2 and S3.3. The variance in affect ratings ($affect_t$) that could be explained by previous affect ($affect_{t-1}$) and symptom increases – also referred to as the marginal R^2 – ranged from 0.02 to 0.06 (mean marginal $R^2 = 0.03$, median = 0.04). Accounting for individual differences in autocorrelations and symptom increases raised the explained variance to 0.29 to 0.49 (mean conditional $R^2 = 0.38$, median = 0.37).

Finally, we ran linear regression analyses to explore whether mean affect ratings were similarly predictive of symptom increases. Results were corrected for Type I error inflation. Affect intensity (*i.e.*, the average level of an affective state) was not predictive of future increases in symptoms after correcting for multiple testing (supplementary Table S3.4). Not correcting for multiple testing yielded similar results.

^d Marginal R^2 reflects the variance explained by the fixed effects; conditional R^2 reflects the variance explained by the full model (*i.e.*, fixed and random effects).

Table 3.3. *Standardized coefficients of interaction effects retrieved from multilevel models: EWS as predictors of symptom increases*

Affect item	Depression		Anxiety		Somatic complaints		Interpersonal sensitivity	
	(N=180)		(N=192)		(N=184)		(N=166)	
	β	SE	β	SE	β	SE	β	SE
Down	0.02	0.02	0.04 [†]	0.02	0.03	0.02	0.03	0.03
Listless	0.01	0.03	0.04	0.02	0.02	0.02	0.02	0.03
Anxious	0.02	0.03	0.05	0.03	0.01	0.03	0.04	0.03
Not relaxed	0.01	0.02	0.04 [†]	0.02	0.00	0.02	0.01	0.02
Unwell	-0.01	0.02	0.02	0.02	0.00	0.02	-0.03	0.02
Insecure	0.04	0.02	0.02	0.02	0.02	0.02	0.05 [†]	0.02
Suspicious	0.08 [†]	0.03	0.05	0.03	0.06	0.03	0.12*	0.03

Note. Coefficients refer to the standardized interaction effect of affect ratings at time $t-1$ and increases in psychopathological symptom severity on affect ratings at time t . This effect describes the relation between autocorrelations and symptom shifts. Note that other coefficients estimated by the multilevel model described in footnote c, as well as the unstandardized estimates, are reported the supplement, Tables S3.2 and S3.3. SE = standard error. *Significant at $\alpha=0.05$, with P-values adjusted according to Hochberg's procedure. [†]Significant at $\alpha=0.05$, without multiple testing correction.

Post-hoc analyses

The multilevel models adopted in the present study slightly deviate from the model described by Van de Leemput *et al.*⁴ (analyses concerning general population sample). Specifically, we decided to examine the association between autocorrelations and *increases* in symptom severity, instead of the association between autocorrelations and *symptom severity at follow-up, corrected for baseline symptom severity* (see also supplement 3.3. *Post-hoc analyses*). To examine to what extent the results of the present study did or did not replicate those of Van de Leemput *et al.*⁴, we tested both procedures in both datasets. When applying the model described by Van de Leemput *et al.*⁴ to the current dataset, our results matched those reported in this earlier study: higher autocorrelations in affective states predicted greater symptom severity at follow-up (supplementary Table S3.5). Similarly, when applying the model adopted in the present study to the dataset used by Van de Leemput *et al.*⁴, results confirmed those of the present study: autocorrelations in affective states were not related to symptom increases (supplementary Table S3.6). Hence, the results of the both studies replicated perfectly in a second, independent dataset.

The discrepancy between both models in terms of results might be due to the fact that the inclusion of symptom severity at baseline as a predictor (cf. Van de Leemput *et al.*⁴) corrects the *outcome* of the model (affect at time t) – but not another *predictor* in the model,

namely follow-up symptom severity – for baseline symptoms. Hence, including baseline symptoms does not mean that follow-up symptoms can be considered indicative of change scores. We therefore consider the model used in the present study – denoted in footnote c – to be better suited for investigating the association between autocorrelations and symptom increases than the model reported earlier⁴ (analyses concerning general population sample).

In both datasets, the correlation between symptom severity at baseline and follow-up was high (current dataset: $r = 0.79-0.86$, depending on the subsample, dataset Van de Leemput *et al.*⁴: $r = 0.68$). As a result, the association between autocorrelations and symptoms at follow-up (with baseline symptoms included as a predictor) that was found in both datasets might in fact reflect an association between autocorrelations and concurrent symptom severity. This indeed appeared to be the case both for the dataset provided by Van de Leemput *et al.*⁴ (supplementary Table S3.6) and for the current dataset (supplementary Table S3.7). In the latter, autocorrelations in *feeling suspicious* were most strongly related to baseline symptom severity (supplementary Table S3.7). This generic effect was followed by domain-specific effects, showing that the association between autocorrelations and baseline symptom severity was particularly pronounced for congruent combinations of affective states and psychopathological domains (Table 3.1, supplementary Table S3.7). Affect intensity (*i.e.*, the mean level of an affective state) was also predictive of higher symptom severity at baseline (supplementary Table S3.7), but these associations did not differentiate according to the type of affective state or psychopathological domain involved.

Fig. 3.1. Association between autocorrelations in affective states and increases in symptom severity in different psychopathological domains (i.e., depression, anxiety, somatic complaints and interpersonal sensitivity). SCL-90 scores were categorized solely for the purposes of visualization. After correcting for multiple testing, none of the trends were statistically significant (dotted lines), except for the association between increases in interpersonal sensitivity and EWS in feeling suspicious (solid line). Because visualization required the categorization of follow-up SCL-90 scores, the above figures approximate but not necessarily parallel the results in Table 3.3.

Discussion

This study found that increases in psychopathology after one year were generally not associated with baseline autocorrelations in negative affective states. The only exception was a significant association between elevated autocorrelations (EWS) in *feeling suspicious* and increases in interpersonal sensitivity. It follows that we found only marginal support for EWS as predictors of symptom shifts after one year, and no support for the hypothesis that EWS could signal the *type* of symptoms about to develop. Post-hoc analyses, which included a second dataset provided by Van de Leemput *et al.*⁴, revealed an identical pattern of results across two datasets: autocorrelations are predictive of *concurrent* depressive symptoms, but

not of *future* depressive symptoms. Thus, it seems that group-level designs (present study and Van de Leemput *et al.*⁴) do not lead to the same conclusion as individual-level studies that reported rising trends in autocorrelations in individuals who were about to relapse into depression^{5,53}. Methodological differences between both designs likely account for this, and provide useful guidance for further studies into EWS for psychopathological shifts.

Paradox between group- and individual-level studies

The majority of earlier studies concerning complex system principles in psychopathology involved either simulated data^{54,55} or intensive longitudinal data collected within individuals^{11,13}. Examples of the latter include two earlier case studies that adopted an experience sampling design, which allows for monitoring symptom shifts as well as affect dynamics (including autocorrelations; EWS) as they evolve within individuals. These studies found rising trends in the autocorrelation of negative mental states several weeks before a relapse in patients suffering from depression^{5,53}. The present findings suggest that these *within-individual* processes implied by complex systems principles – namely, accumulating instability prior to symptom shifts – hardly create differences *between individuals* – namely, higher autocorrelations in those who report higher symptom shifts³⁰. A likely explanation for this concerns the different timescales that were considered. Whereas within-individual designs investigated EWS over the course of several weeks^{5,53}, between-individual designs (including the present study) investigated EWS one to three years prior to symptom shifts^{4,21}. Taken together, these studies provide novel insight in the timescale at which critical slowing down manifests, and also suggest that this timescale may differ between psychopathological domains. For instance, depressive symptoms might be characterized by relatively fast dynamics, suggesting that EWS might only appear in the weeks prior to a symptom shift^{5,53}, whereas interpersonal sensitivity might show slower dynamics, meaning that EWS could already manifest one year in advance. Further research is needed in order to establish the timescale at which EWS unfold within individuals with varying psychopathological complaints. Such research could also address the size of symptom shifts for which EWS are relevant. In conclusion, the utility of complex systems principles to psychopathology may be restricted to specific circumstances, *e.g.*, characterized by specific time intervals, types of psychopathology, and sizes of shifts. It requires large-scale within-individual research, where EWS and symptom shifts are monitored in real-time in a considerable number of individuals, to delineate under what circumstances EWS could be informative of future symptom progression.

Autocorrelations in momentary affective states as signs of current symptomatology

Although associations between autocorrelations and *future symptom increases* were lacking for almost all affective states/psychopathological domains, post-hoc analyses showed that associations between autocorrelations in affective states and *symptom severity at baseline* were present across multiple domains. This pattern replicated in an independent dataset provided by van de Leemput *et al.*⁴. The extent to which momentary affective states carry over from one moment to the next thus appeared to be related to concurrent levels of psychopathological severity. Correspondingly, earlier studies found that emotional inertia, indexed by elevated autocorrelations in affective states, is cross-sectionally related to psychopathology^{20,24,26,27}. From an emotion-regulation perspective, emotional inertia signals what is labeled in this literature as rigidity^{21,27}. That is, individuals with high levels of inertia may ‘get stuck’ in negative affective states, and hence, be vulnerable for symptoms of psychopathology²⁷. However, inertia also coincides with heightened intensity of negative affect, raising the question to what extent their respective associations with psychopathology are shared. According to a recent meta-analysis, inertia and affect intensity show little overlap, with less than 6% of shared variance²⁸. Nevertheless, the association between inertia and depressive symptoms was weak and therefore dropped to ‘drastic non-significance’ after accounting for affect intensity²⁸. This might be explained by the small effect size of inertia on depressive symptoms. Hence, it seems that for predicting psychopathology, more parsimonious metrics of affective experience (affect intensity) suffice. Yet, inertia could still provide unique insight in the processes that give rise to psychopathology. Present findings suggest, for instance, that the association between inertia (autocorrelations) and psychopathology might show domain-specificity. Specifically, symptoms of depression, anxiety, and interpersonal sensitivity were more associated with autocorrelations in affective states that were a priori considered to cover that domain (Table 3.1; *e.g.*, for depression, symptom severity was most strongly associated to autocorrelations in feeling suspicious, down, and listless, whereas for anxiety, symptom severity was associated to autocorrelations in feeling suspicious and anxious). In contrast, mean levels were predictive of concurrent symptom severity regardless of the affective state or psychopathological domain assessed. Speculatively, this could mean that the intensity of negative affective states is informative of global symptom severity, whereas the dynamics of those affective states (reflected in their autocorrelations) might reveal the types of symptoms that are present. Further disentangling the unique effects of affect dynamics (including EWS) and affect intensity, respectively, in the context of psychopathology should be considered an important avenue for further research.

Strengths, limitations, and future directions

EWS have been suggested to hold great promise for clinical practice: if found to be predictive of future psychopathology, EWS could help us to foresee and possibly prevent illness progression. This means that clinical advancements would no longer be hindered by limited understanding of the complex mechanisms that underlie onset or progression of psychopathology. To examine the value of this promise, comprehensive and critical investigation of the utility of EWS to psychopathology is necessary. We therefore aimed to replicate and extend the first – and to date, only – study that provided empirical, group-level support for the association between EWS and symptom increases⁴. Across two datasets with a similar sample, experience sampling protocol and assessment of psychopathology, we found an identical pattern of results, which considerably strengthens our conclusions. However, neither the reliability of individual experience sampling items nor the reliability of autocorrelations could be verified, which highlights the need for psychometric reliability metrics suitable for experience sampling data. Another limitation is that the present group-level design could only indirectly investigate the relation between EWS and psychopathology and, like earlier studies^{4,11,14,21,24,26,41,43,56}, does not allow for within-individual inferences. A more direct evaluation of the hypotheses that follow from a complex systems perspective on psychopathology requires designs where individuals who are at increased risk for psychopathology prospectively monitor their mental state for a prolonged time period (*e.g.*, several months) during which a shift is likely to occur⁵. Such designs could verify whether, within individuals, sudden shifts in symptom severity are preceded by rising patterns in EWS. Real-time monitoring of EWS and symptom shifts allows for addressing critical research questions that will ultimately determine the usefulness of EWS to clinical practice. Such questions relate to the timing of EWS, the size and type of symptom shifts that can be anticipated, the sensitivity and specificity of EWS in the context of psychopathology, and the absolute thresholds in EWS that may inform clinical decision-making.

Conclusions

Empirical investigation of complex systems principles in psychopathology is still in its infancy. Hence, the present study should be considered a first step towards investigating whether complex systems principles apply to psychopathology. Across two datasets, we found that autocorrelations in affective states are related to concurrent symptoms of depression, but not to increases in those symptoms after one year. Yet, increases in interpersonal sensitivity

were preceded by heightened autocorrelations, which suggests that the association between early warning signals and psychopathology might depend on the domain wherein symptoms increase. Further research is needed to delineate when, for what type of symptoms, and for what size of shifts, early warning signals are predictive of future psychopathology. This requires monitoring early warning signals and symptoms in real-time, thereby allowing for a more direct evaluation of the inferences that follow from a complex systems perspective on psychopathology.

Supplement 3

3.1 Sample composition

Due to attrition, SCL-90 data at follow-up were only available for a subset of individuals, which reduced our sample size from $N=839$ to $N=467$. Compared to individuals with complete SCL-90 data, individuals with incomplete data ($N=372$) were younger ($t(837) = 3.27, p < .01, d = .23$) and more likely to be male ($\chi^2(1) = 22.71, p < .01$). Individuals with complete versus incomplete SCL-90 data did not differ in total SCL-90 score at baseline ($t(808) = 0.56, p = .57, d = .04$). Of the 467 individuals with complete baseline and follow-up SCL-90 ratings, we selected those individuals who reported increased symptom severity from baseline to follow-up. This resulted in four samples of $N=230$ (depression), $N=252$ (anxiety), $N=243$ (somatic complaints), and $N=220$ (interpersonal sensitivity). Within these samples, individuals with at least 20 consecutive affect ratings were selected, resulting in samples of $N=180$ (depression), $N=192$ (anxiety), $N=184$ (somatic complaints), and $N=166$ (interpersonal sensitivity). Because most participants reported increased symptom severity in more than one symptom cluster, these samples were not unique in their composition. That is, any two samples showed an overlap ranging from 110 to 124 individuals (Table S3.1). Of the 293 unique individuals that made up the samples, 64 were present in each sample. These participants did not differ from other participants with respect to age ($t(291) = 0.76, p = .45, d = .11$) or gender ($\chi^2(1) = 0.74, p = .39$), but did report higher total SCL-90 scores at baseline ($t(291) = 3.46, p < .01, d = .50$). Because participants were retrieved from a registered cohort of twins, they were not independent from each other. This dependency was accounted for in multilevel models. The sample of 293 unique individuals consisted of 170 (58.0%) twins, 116 (39.6%) genetically unrelated participants, 5 (1.7%) triplets and 2 (0.7%) siblings.

Table S3.1. *Overlap between samples*

	Depression ($N=180$)	Anxiety ($N=192$)	Somatic complaints ($N=184$)
Anxiety ($N=192$)	124		
Somatic complaints ($N=184$)	112	122	
Interpersonal sensitivity ($N=166$)	119	113	110

Note. Numbers denote the number of individuals that were present in both the sample denoted in the row and the sample denoted in the column.

3.2 Model results

Table S3.2. Standardized coefficients of main effects retrieved from multilevel models: EWS as predictors of symptom increases

Affect item	Depression (N=180)			Anxiety (N=192)			Somatic complaints (N=184)			Interpersonal sensitivity (N=166)		
	Lagged affect β_{1i}	Sympt. increase β_{2ij}	SE	Lagged affect β_{1i}	Sympt. increase β_{2ij}	SE	Lagged affect β_{1i}	Sympt. increase β_{2ij}	SE	Lagged affect β_{1i}	Sympt. increase β_{2ij}	SE
Down	0.14**	0.02	0.11	0.04	0.11**	0.04	0.12**	0.02	0.13**	0.04	0.13**	0.03
Listless	0.16**	0.02	0.10	0.05	0.13**	0.02	0.14**	0.02	0.10	0.04	0.16**	0.03
Anxious	0.10**	0.03	0.09	0.04	0.08**	0.02	0.12**	0.02	0.15**	0.04	0.09**	0.03
Not relaxed	0.15**	0.02	0.05	0.04	0.11**	0.02	0.16**	0.02	0.02	0.04	0.14**	0.02
Unwell	0.18**	0.02	0.14*	0.05	0.16**	0.02	0.20**	0.02	0.14*	0.05	0.19**	0.02
Insecure	0.11**	0.02	0.05	0.04	0.12**	0.02	0.12**	0.02	0.05	0.04	0.10**	0.02
Suspicious	0.06*	0.03	0.08	0.04	0.08**	0.02	0.09**	0.02	0.14**	0.04	0.10**	0.03

Note. Coefficients refer to the standardized effects of (1) person-mean centered affect ratings at time $t-1$ (β_{1i}) and (2) increases in psychopathological symptom severity (β_{2ij}) on affect ratings at time t , accounting for baseline symptom severity (model: $\text{affect}_{t,ij} = \beta_{0i} + \beta_{1i}(\text{affect}_{t-1,ij} - \mu) + \beta_{2ij}(P_i + \beta_{3ij}(P_i * (\text{affect}_{t-1,ij} - \mu)) + \epsilon_{ij})$). The interaction effects of affect ratings at time $t-1$ and increases in psychopathological symptom severity on affect ratings at time t (β_{3ij}) are reported in the main text (Table 3.3). SE = standardized standard error. Sympt. = symptom. *Significant at $\alpha=0.05$, **Significant at $\alpha=0.005$, with P-values adjusted according to Hochberg's procedure.

Table S3.3. Unstandardized coefficients of effects retrieved from multilevel models: EWS as predictors of symptom increases

Affect item	Depression (N=180)			Anxiety (N=192)			Somatic complaints (N=184)			Interpersonal sensitivity (N=166)		
	B _{-1ij}	SE	B _{-2ij}	SE	B _{-3ij}	SE	B _{-1ij}	SE	B _{-2ij}	SE	B _{-3ij}	SE
Down	0.17	0.03	0.36	0.14	0.05	0.05	0.14	0.02	0.60	0.17	0.12	0.06
Listless	0.20	0.03	0.33	0.16	0.02	0.06	0.16	0.02	0.49	0.18	0.11	0.06
Anxious	0.12	0.03	0.24	0.10	0.04	0.06	0.10	0.03	0.48	0.13	0.12	0.07
Not relaxed	0.18	0.02	0.20	0.17	0.03	0.05	0.14	0.02	0.25	0.19	0.13	0.06
Unwell	0.24	0.03	0.48	0.17	-0.02	0.06	0.21	0.02	0.55	0.19	0.08	0.07
Insecure	0.13	0.03	0.15	0.14	0.10	0.05	0.14	0.02	0.36	0.17	0.06	0.07
Suspicious	0.07	0.03	0.20	0.11	0.17	0.07	0.10	0.03	0.42	0.12	0.12	0.07

Note. Coefficients refer to the unstandardized effects of (1) person-mean centered affect ratings at time $t-1$ (B_{-1ij}), (2) increases in psychopathological symptom severity (B_{-2ij}), and (3) the interaction between person-mean centered affect ratings and increases in psychopathological symptom severity (B_{-3ij}) on affect ratings at time t (model: $\text{affect}_{t+1ij} = B_{-1ij} + B_{-2ij}(\text{affect}_{t-1ij} - \mu_i) + B_{-3ij}(\text{affect}_{t-1ij} - \mu_i) + \epsilon_{ij}$). SE: unstandardized standard error. Due to limited space, the statistical significance of coefficients could not be displayed. However, this significance is reported in Table S3.2 (B_{-2ij} and B_{-3ij}) and Table 3.2 (main text, B_{-3ij}).

Table S3.4. Affect intensity as a predictor of symptom increases, future symptom severity, and current symptom severity

Affect item	Depression (N=180)			Anxiety (N=192)			Somatic complaints (N=184)			Interpersonal sensitivity (N=166)		
	Symptom increases	Future symptom severity	Current symptom severity	Symptom increases	Future symptom severity	Current symptom severity	Symptom increases	Future symptom severity	Current symptom severity	Symptom increases	Future symptom severity	Current symptom severity
Down	0.03	0.25**	0.19**	0.03	0.24**	0.17**	0.05	0.23**	0.16**	0.04	0.24**	0.17**
Listless	0.02	0.17**	0.12**	0.01	0.19**	0.15**	0.03	0.19**	0.13**	0.04	0.17**	0.11**
Anxious	0.04	0.20**	0.13**	0.03	0.25**	0.18**	0.06	0.20**	0.12**	0.03	0.17**	0.12**
Not relaxed	0.01	0.09*	0.07*	0.00	0.13**	0.10**	0.01	0.05	0.04	0.01	0.09*	0.07*
Unwell	0.04	0.20**	0.14**	0.02	0.21**	0.15**	0.03	0.18**	0.12**	0.01	0.14**	0.11**
Insecure	0.01	0.14**	0.10**	0.00	0.20**	0.15**	0.01	0.11*	0.08*	0.02	0.14**	0.10**
Suspicious	0.03	0.13**	0.08**	0.04	0.22**	0.15**	0.07	0.21**	0.12**	0.03	0.13**	0.08*

Note. Numbers denote standardized coefficients retrieved from linear regressions where mean affect ratings predicted, respectively, symptom severity at follow-up while correcting for symptom severity at baseline (i.e., symptom increases), symptom severity at follow-up (i.e., future symptom severity), and symptom severity at baseline (i.e., current symptom severity). *Significant at $\alpha=0.05$, **Significant at $\alpha=0.005$, with P-values adjusted according to Hochberg's procedure; †Significant at $\alpha=0.05$, without multiple testing correction.

3.3 Post-hoc analyses

The primary aim of this study was to test whether EWS—operationalized as the autocorrelations in momentary affective states—precede increases in symptoms. We originally addressed this aim through a multilevel model similar to the model described by Van de Leemput *et al.*⁴ (analyses concerning the general population sample). In this model, outlined in formula 1, the affect rating of individual i belonging to family j at time t ($affect_{tij}$) is predicted by the same individual's mean-centered affect rating at time $t-1$ ($affect_{t-1ij} - \mu_i$), his/her symptom severity at baseline (Pb_i), his/her symptom severity at follow-up (Pf_i), and an interaction effect between symptom severity at follow-up and the lagged person mean-centered affect rating (β_{4ij}).

$$affect_{tij} = \beta_{0ij} + \beta_{1ij}(affect_{t-1ij} - \mu_i) + \beta_{2ij}Pb_i + \beta_{3ij}Pf_i + \beta_{4ij}(Pf_i \times (affect_{t-1ij} - \mu_i)) + \varepsilon_{tij}$$

Formula 1. Multilevel model that tests the association between the autocorrelation in momentary affective states and symptom severity at follow-up.

Thanks to an anonymous reviewer, we realized that the coefficient denoting the interaction effect in this model (β_{4ij}) does not test whether elevated autocorrelations precede *increases* in symptoms—as hypothesized from a complex systems perspective on psychopathology—but rather tests the relation between autocorrelations in affective states and *symptom severity at follow up*. We therefore adapted the model above to accommodate our questions and hypotheses (formula 2; see analysis section in the main text for a description of this model).

$$affect_{tij} = \beta_{0ij} + \beta_{1ij}(affect_{t-1ij} - \mu_i) + \beta_{2ij}P_i + \beta_{3ij}(P_i * (affect_{t-1ij} - \mu_i)) + \varepsilon_{tij}$$

Formula 2. Multilevel model that tests the association between the autocorrelation in momentary affective states and increases in symptoms. In this model, P_i denotes the increase in symptoms from baseline to follow-up.

In order to test how this adaptation of the model affected our results, both models were run. Results obtained from the second model (formula 2)—which addressed our primary aim—revealed that elevated autocorrelations preceded increases in interpersonal sensitivity, but not increases in other psychopathological domains (*i.e.*, depression, anxiety, somatic complaints; see Table 3.3 in main text). Results obtained from the first model (formula 1)—which mirrored the model reported by Van de Leemput *et al.*⁴—showed that autocorrelations in momentary affective states were positively related to the severity of symptoms of

depression, anxiety, and interpersonal sensitivity at follow up (see Table S3.5).

Post-hoc, we examined whether the discrepant results of both models replicated in the dataset analyzed by Van de Leemput *et al.*⁴. This dataset consists of 621 individuals from the general population who provided ratings of their momentary affective state ten times a day for six consecutive days. At baseline and at four follow-up assessments – which were scheduled three months apart – participants completed the Symptom Checklist-90 (SCL-90). Symptom severity at follow-up was calculated as the average score of these four follow-up assessments. In line with the analyses reported by Van de Leemput *et al.*⁴, affect ratings were aggregated according to their valence and arousal. The resulting scores will be referred to as feeling down (negative valence and low arousal) and feeling anxious (negative valence, high arousal). Further, analyses were restricted to the SCL-90 domain depression. Complete data were available for 533 participants. For a detailed description of the sample and methods, see Van de Leemput *et al.*⁴.

The results of these supplementary analyses confirmed our initial findings. Specifically, autocorrelations in feeling down and feeling anxious were predictive of higher depressive symptoms at follow-up, corrected for baseline (formula 1, cf. results of Van de Leemput *et al.*⁴) but not of higher *increases* in depressive symptoms from baseline to follow-up (formula 2; Table S3.6). The fact that this pattern emerged in two independent datasets led us to further explore the discrepancies between both models (formula 1 and 2, respectively). We hypothesized that the association between autocorrelations and symptoms at follow-up, corrected for baseline (formula 1), might in fact be accounted for by the association between autocorrelations and symptoms at baseline (which was not sufficiently controlled for in the model). In both datasets, this hypothesis was confirmed (Table S3.7).

Table S3.5. *Standardized coefficients of interaction effects retrieved from multilevel models: EWS as predictors of future symptoms*

Affect item	Depression (N=180)	Anxiety (N=192)	Somatic complaints (N=184)	Interpersonal sensitivity (N=166)
Down	0.10**	0.06 [†]	0.04	0.07 [†]
Listless	0.09*	0.06 [†]	0.02	0.05
Anxious	0.08 [†]	0.09*	0.04	0.06
Not relaxed	0.04	0.04 [†]	0.00	0.02
Unwell	0.04	0.01	0.01	0.01
Insecure	0.07*	0.05 [†]	0.02	0.08*
Suspicious	0.18**	0.10*	0.07 [†]	0.15**

Note. Coefficients refer to the standardized interaction effect of affect ratings at time *t*-1 and psychopathological symptom severity at follow-up on affect ratings at time *t*, accounting for baseline symptom severity (formula 1). This effect describes the relation between autocorrelations and future symptom severity. *Significant at $\alpha=0.05$, **Significant at $\alpha=0.005$, with P-values adjusted according to Hochberg's procedure; [†]Significant at $\alpha=0.05$, without multiple testing correction.

Table S3.6. *Standardized coefficients of interaction effects retrieved from two multilevel models in two datasets*

	Increases in symptoms of depression				Severity of symptoms of depression at follow-up				Severity of symptoms of depression at baseline			
	Original study (N=180)		Replication study (N=533)		Original study (N=180)		Replication study (N=533)		Original study (N=180)		Replication study (N=533)	
	β	B	β	B	β	B	β	B	β	B	β	B
Down	0.02	0.05	0.00	-0.01	0.10**	0.12**	0.13*	0.07*	0.12**	0.23**	0.13*	0.06*
Anxious	0.02	0.04	-0.02	-0.04	0.08*	0.09*	0.11*	0.06*	0.09*	0.16*	0.16**	0.08**

Note. Original study refers to the data used in the main text of this chapter, which were obtained from the TwinssCan study³⁴; replication study refers to the general population sample reported by Van de Leemput *et al.*⁴. The first columns ('increases in symptoms of depression') denote the interaction effect of increases in symptoms of depression and lagged person-mean centered affect ratings on actual affect ratings (formula 2). This effect describes the association between autocorrelations in momentary affective states and symptom increases. The middle columns ('severity of symptoms of depression at follow-up') denote the interaction effect of depressive symptom severity at follow-up and lagged person-mean centered affect ratings on actual affect ratings (formula 1). This effect describes the association between autocorrelations in momentary affective states and symptom severity at follow-up. Finally, the last columns ('severity of symptoms of depression at baseline') denote the interaction effect of depressive symptom severity at baseline and lagged person-mean centered affect ratings on actual affect ratings. To remain consistent with the results reported by Van de Leemput *et al.*⁴, both standardized (β) and non-standardized (B) coefficients are denoted. *Significant at $\alpha=0.05$; **Significant at $\alpha=0.005$.

Table S3.7. *Standardized coefficients of interaction effects retrieved from multilevel models: EWS as predictors of current symptoms*

Affect item	Depression (N=180)	Anxiety (N=192)	Somatic complaints (N=184)	Interpersonal sensitivity (N=166)
Down	0.12**	0.04	0.02	0.06 [†]
Listless	0.11**	0.05 [†]	0.01	0.06 [†]
Anxious	0.09*	0.08*	0.04	0.05
Not relaxed	0.04*	0.02	-0.01	0.02
Unwell	0.06*	0.00	0.01	0.03
Insecure	0.06*	0.05 [†]	0.01	0.06 [†]
Suspicious	0.15**	0.09*	0.05 [†]	0.10*

Note. Coefficients refer to the standardized interaction effect of affect ratings at time *t-1* and psychopathological symptom severity at baseline on affect ratings at time *t*. This effect describes the relation between autocorrelations and future symptom severity. *Significant at $\alpha=0.05$, **Significant at $\alpha=0.005$, with P-values adjusted according to Hochberg's procedure; [†]Significant at $\alpha=0.05$, without multiple testing correction.

References

1. Kessler RC, Aguilar-Gaxiola S, Alonso J, et al. The global burden of mental disorders: an update from the WHO World Mental Health (WMH) surveys. *Epidemiologia e Psichiatria Sociale*. 2009;18(1):23-33. doi:10.1017/S1121189X00001421
2. Nolen-Hoeksema S, Watkins ER. A heuristic for developing transdiagnostic models of psychopathology: explaining multifinality and divergent trajectories. *Perspectives on Psychological Science*. 2011;6(6):589-609. doi:10.1177/1745691611419672
3. Nelson B, McGorry PD, Wichers M, Wigman JTW, Hartmann JA. Moving from static to dynamic models of the onset of mental disorder a review. *JAMA Psychiatry*. 2017;74(5):528-534. doi:10.1001/jamapsychiatry.2017.0001
4. van de Leemput IA, Wichers M, Cramer AOJ, et al. Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences*. 2014;111(1):87-92. doi:10.1073/pnas.1312114110
5. Wichers M, Groot PC, Psychosystems ESM group, EWS group. Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*. 2016;85:114-116. doi:10.1159/000441458
6. Hayes AM, Laurenceau JP, Feldman G, Strauss JL, Cardaciotto LA. Change is not always linear: the study of nonlinear and discontinuous patterns of change in psychotherapy. *Clinical Psychology Review*. 2007;27(6):715-723. doi:10.1016/j.cpr.2007.01.008
7. Schiepek G, Eckert H, Aas B, Wallot S, Wallot A. *Integrative Psychotherapy - a Feedback-Driven Dynamic Systems Approach*. Hogrefe Publishing; 2015.
8. Olthof M, Hasselman F, Strunk G, et al. Critical fluctuations as an early-warning signal for sudden gains and losses in patients receiving psychotherapy for mood disorders. *Clinical Psychological Science*. 2019;8(1):1-11. doi:10.1177/2167702619865969
9. Scheffer M, Carpenter SR, Lenton TM, et al. Anticipating critical transitions. *Science*. 2012;338(6105):344-348. doi:10.1126/science.1225244
10. Scheffer M, Bascompte J, Brock WA, et al. Early-warning signals for critical transitions. *Nature*. 2009;461(7260):53. doi:10.1038/nature08227
11. Lichtwarck-Aschoff A, Hasselman F, Cox R, Pepler D, Granic I. A characteristic destabilization profile in parent-child interactions associated with treatment efficacy for aggressive children. *Nonlinear Dynamics, Psychology, and Life Sciences*. 2012;16(3):353-379.
12. Kowalik ZJ, Schiepek G, Kumpf K, Roberts LE, Elbert T. Psychotherapy as a chaotic process II. The application of nonlinear analysis methods on quasi time series of the client-therapist interaction: a nonstationary approach. *Psychotherapy Research*. 1997;7(3):197-218. doi:10.1080/10503309712331331973
13. Schiepek G, Tominschek I, Karch S, et al. A controlled single case study with repeated fMRI measurements during the treatment of a patient with obsessive-compulsive disorder: testing the nonlinear dynamics approach to psychotherapy. *World Journal of Biological Psychiatry*. 2009;10:658-668. doi:10.1080/15622970802311829
14. Schiepek GK, Tominschek I, Heinzel S. Self-organization in psychotherapy testing the synergetic model of change processes. *Frontiers in Psychology*. 2014;5:1-11. doi:10.3389/fpsyg.2014.01089

15. Heinzel S, Tominschek I, Schiepek G. Dynamic pattern in psychotherapy- discontinuous changes and critical instabilities during the treatment of obsessive compulsive disorder. *Nonlinear Dynamics, Psychology, and Life Sciences*. 2014;18(2):155-176.
16. Schiepek G, Strunk G. The identification of critical fluctuations and phase transitions in short term and coarse-grained time series-a method for the real-time monitoring of human change processes. *Biological Cybernetics*. 2010;102:197-207. doi:10.1007/s00422-009-0362-1
17. Scheffer M, Carpenter SR, Dakos V, van Nes EH. Generic indicators of ecological resilience: inferring the chance of a critical transition. *Annual Review of Ecology, Evolution, and Systematics*. 2015;46:145-167. doi:10.1146/annurev-ecolsys-112414-054242
18. Liu R, Chen P, Aihara K, Chen L. Identifying early-warning signals of critical transitions with strong noise by dynamical network markers. *Scientific Reports*. 2015;5(1):1-13. doi:10.1038/srep17501
19. Dakos V, van Nes EH, D'Odorico P, Scheffer M. Robustness of variance and autocorrelation as indicators of critical slowing down. *Ecology*. 2012;93(2):264-271. doi:10.1890/11-0889.1
20. Houben M, van den Noortgate W, Kuppens P. The relation between short-term emotion dynamics and psychological well-being: a meta-analysis. *Psychological Bulletin*. 2015;141(4):901-930. doi:10.1037/a0038822
21. Kuppens P, Sheeber LB, Yap MBH, Whittle S, Simmons JG, Allen NB. Emotional inertia prospectively predicts the onset of depressive disorder in adolescence. *Emotion*. 2012;12(2):283-289. doi:10.1037/a0025046
22. Trull TJ, Lane SP, Koval P, Ebner-Priemer UW. Affective dynamics in psychopathology. *Emotion Review*. 2015;7(4):355-361. doi:10.1177/1754073915590617
23. Suls J, Green P, Hillis S. Emotional reactivity to everyday problems, affective inertia, and neuroticism. *Personality and Social Psychology Bulletin*. 1998;24(2):127-136.
24. Kuppens P, Allen NB, Sheeber LB. Emotional inertia and psychological maladjustment. *Psychological Science*. 2010;21(7):984-991. doi:10.1177/0956797610372634
25. Heimpel SA, Wood J V., Marshall MA, Brown JD. Do people with low self-esteem really want to feel better? Self-esteem differences in motivation to repair negative moods. *Journal of Personality and Social Psychology*. 2002;82(1):128-147. doi:10.1037/0022-3514.82.1.128
26. Brose A, Schmiedek F, Koval P, Kuppens P. Emotional inertia contributes to depressive symptoms beyond perseverative thinking. *Cognition and Emotion*. 2015;29(3):527-538. doi:10.1080/02699931.2014.916252
27. Koval P, Kuppens P, Allen NB, Sheeber L. Getting stuck in depression: the roles of rumination and emotional inertia. *Cognition and Emotion*. 2012;26(8):1412-1427. doi:10.1080/02699931.2012.667392
28. Dejonckheere E, Mestdagh M, Houben M, et al. Complex affect dynamics add limited information to the prediction of psychological well-being. *Nature Human Behaviour*. 2019;3(5):478-491. doi:10.1038/s41562-019-0555-0
29. Bos EH, de Jonge P. "Critical slowing down in depression" is a great idea that still needs empirical proof. *Proceedings of the National Academy of Sciences*. 2014;111(10):1-1. doi:10.1073/pnas.1323672111
30. Wichers M, Borsboom D, Tuerlinckx F, et al. Reply to Bos and De Jonge: between-subject data do provide first empirical support for critical slowing down in depression. *Proceedings of the National Academy of Sciences*. 2014;111:1. doi:10.1073/pnas.1323835111
31. Fisher AJ, Medaglia JD, Jeronimus BF. Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*. 2018;115(27):E6106-E6115.

- doi:10.1073/pnas.1711978115
32. Wichers M, Schreuder MJ, Goekoop R, Groen RN. Can we predict the direction of sudden shifts in symptoms? Transdiagnostic implications from a complex systems perspective on psychopathology. *Psychological Medicine*. 2019;49(3):380-387. doi:10.1017/S0033291718002064
 33. Derom C, Thiery E, Rutten BPF, et al. The East Flanders Prospective Twin Survey (EFPTS): 55 years later. *Twin Research and Human Genetics*. 2019;22:1-6. doi:10.1017/thg.2019.64
 34. Loos R, Derom C, Vlietinck R, Derom R. The East Flanders Prospective Twin Survey (Belgium): a population-based register. *Twin Research*. 1998;1(4):167-175. doi:10.1375/136905298320566131
 35. Derogatis LR. *SCL-90-R Symptom Checklist-90-R Administration, Scoring, and Procedures Manual*. PsychCorp; 1977.
 36. Myin-Germeys I, Birchwood M, Kwapil T. From environment to therapy in psychosis: a real-world momentary assessment approach. *Schizophrenia Bulletin*. 2011;37(2):244-247. doi:10.1093/schbul/sbq164
 37. Delespaul PAEG. Assessing schizophrenia in daily life: the experience sampling method. Published online 1995.
 38. Holi M. Assessment of psychiatric symptoms using the SCL-90. Published online 2003.
 39. Schauenburg H, Strack M. Measuring psychotherapeutic change with the Symptom Checklist SCL 90 R. *Psychotherapy and Psychosomatics*. 1999;68(4):199-206. doi:10.1159/000012333
 40. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (5th Ed.)*; 2013.
 41. Koval P, Kuppens P, Allen NB, Sheeber L. Getting stuck in depression: The roles of rumination and emotional inertia. *Cognition and Emotion*. 2012;26(8):1412-1427. doi:10.1080/02699931.2012.667392
 42. Nelson J, Klumpparendt A, Doeblen P, Ehrling T. Everyday Emotional Dynamics in Major Depression. *Emotion*. 2018;20:179-191. doi:10.1037/emo0000541
 43. Curtiss J, Fulford D, Hofmann SG, Gershon A. Network dynamics of positive and negative affect in bipolar disorder. *Journal of Affective Disorders*. 2019;249:270-277. doi:10.1016/j.jad.2019.02.017
 44. de Haan-Rietdijk S, Kuppens P, Hamaker EL. What's in a day? A guide to decomposing the variance in intensive longitudinal data. *Frontiers in Psychology*. 2016;7(891). doi:10.3389/fpsyg.2016.00891
 45. van Roekel E, Verhagen M, Engels CRME, Kuppens P. Variation in the serotonin transporter polymorphism (5-HTTLPR) and inertia of negative and positive emotions in daily life. *Emotion*. 2018;18:229-236.
 46. Thompson RJ, Mata J, Jaeggi SM, Buschkuhl M, Jonides J, Gotlib IH. The everyday emotional experience of adults with major depressive disorder: examining emotional instability, inertia, and reactivity. *Journal of Abnormal Psychology*. 2012;29(4):997-1003. doi:10.1016/j.biotechadv.2011.08.021.Secretd
 47. Koval P, Brose A, Pe ML, et al. Emotional inertia and external events: The roles of exposure, reactivity, and recovery. *Emotion*. 2015;15(5):625-636. doi:10.1037/emo0000059
 48. Koval P, Sütterlin S, Kuppens P. Emotional inertia is associated with lower well-being when controlling for differences in emotional context. *Frontiers in Psychology*. 2016;6:1-11. doi:10.3389/fpsyg.2015.01997
 49. Hox JJ. *Multilevel Analysis: Techniques and Applications*. 2nd ed. Routledge; 2010.
 50. Gujarati DN. Extensions of the two-variable linear regression model. In: *Basic Econometrics*. McGraw-Hill/Irwin; 2003:164-296.
 51. Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*. 1998;75(4):800-802.
 52. R Core Team. R: A language and environment for statistical computing. Published online 2020.

53. Wichers M, Smit AC, Snippe E. Early warning signals based on momentary affect dynamics can expose nearby transitions in depression: a confirmatory single-subject time-series study. *Journal for Person-Oriented Research*. 2020;6(1):1-15. doi:10.17505/jpor.2020.22042
54. Cramer AOJ, van Borkulo CD, Giltay EJ, et al. Major depression as a complex dynamic system. *PLoS ONE*. 2016;11(12):1-20. doi:10.1371/journal.pone.0167490
55. Burger J, Veen D van der, Robinaugh DJ, et al. Bridging the gap between complexity science and clinical practice by formalizing idiographic theories: a computational model of functional analysis. *BMC Medicine*. 2020;18(99):1-18. doi:10.1186/s12916-020-01558-1
56. Sperry SH, Walsh MA, Kwapil TR. Emotion dynamics concurrently and prospectively predict mood psychopathology. *Journal of Affective Disorders*. 2020;261:67-75. doi:10.1016/j.jad.2019.09.076

4

Can we predict the direction of sudden shifts in symptoms? Transdiagnostic implications from a complex systems perspective on psychopathology

published as: Wichers, M., Schreuder, M.J., Goekoop, R., & Groen, R.N. (2019). Can we predict the direction of sudden shifts in symptoms? Transdiagnostic implications from a complex systems perspective on psychopathology. *Psychological Medicine*, 49(3), 380–387. doi: 10.1017/S0033291718002064

Abstract

Recently, there has been renewed interest in the application of assumptions from complex systems theory in the field of psychopathology. One assumption, with high clinical relevance, is that sudden transitions in symptoms may be anticipated by rising instability in the system, which can be detected with early warning signals (EWS). Empirical studies support the idea that this principle also applies to the field of psychopathology. The current chapter discusses whether assumptions from complex systems theory can additionally be informative with respect to the specific symptom dimension in which such a transition will occur (*e.g.*, whether a transition towards anxious, depressive or manic symptoms is most likely). From a complex systems perspective, both EWS measured in single symptom dynamics and network symptom dynamics at large are hypothesized to provide clues regarding the direction of the transition. Challenging research designs are needed to provide empirical validation of these hypotheses. These designs should be able to follow sudden transitions ‘live’ using frequent observations of symptoms within individuals and apply a transdiagnostic approach to psychopathology. If the proposed assumptions are supported by empirical studies, this will signify a large improvement in the possibility for personalised estimations of the course of psychiatric symptoms. Such information can be extremely useful for early intervention strategies aimed at preventing specific psychiatric problems.

Introduction

Sudden transitions in symptom levels

In the past years there has been renewed interest in the potential application of complex systems theory in the field of psychiatry¹⁻⁵. In short, complex systems theory entails that complex systems, ranging from ocean ecosystems to climate, financial markets or the evolutionary development of species, all have certain principles in common that predict their behavior. These relate, for example, to the resilience of a system to remain in its present stable state. High resilience refers to a high level of stability of the system (deep basin of attraction), meaning that the system can easily face perturbations without being tipped out of its current equilibrium⁶ (see Fig. 4.1).

Fig. 4.1. 2-D stability landscape. The ball represents the current state of a complex system and the line constitutes a surface showing the stability of the ball (or system state) in the current situation. In the left panel the state of the system is in a deep basin of attraction, meaning that it will take a considerable perturbation of the system before psychopathology can develop. In the right panel, however, the situation is different. Here, only a small perturbation may already be enough for a transition towards a state of psychopathology.

In complex systems this level of resilience may slowly diminish, even without noticeable signs. Once low, the system is highly instable and at this point even very minor contextual disturbances, also called perturbations, can push the system over a tipping point towards another basin of attraction⁶. This is why complex systems are characterised by sudden transitions, so-called phase transitions, that appear to emerge ‘out of the blue’. Similar transitions have been observed in psychiatry, as psychiatric symptoms sometimes (re)appear in a very abrupt way⁷. Already at the end of the previous century, the idea arose that

psychological phenomena might also behave according to principles of complex systems^{8–10}. If this is indeed the case, this is very relevant since the principles of complex systems teach us important things about the nature of psychopathology and how to understand and foresee sudden transitions in symptoms. One of the interesting consequences would be that the *early identification* of alterations in the level of instability of the system could reveal the proximity of the system's tipping point, or in other words, the likelihood that a sudden shift in symptoms occurs^{11,12}. Some recently conducted studies were able to translate this idea into simulation studies and empirical designs that attempted to test this assumption in the field of psychiatry^{13–16}. Most of the focus in these studies has been on the possibility of foreseeing shifts in the levels of symptoms (increasing or decreasing levels). Yet, it would be important to not only foresee a shift in symptom level, but to also foresee the type of symptoms in which the transition occurs. For example, we want to foresee, if someone is approaching a symptom transition, whether that is a transition characterised by increasing manic, anxious or depressive symptoms. In this chapter, we therefore want to explore whether we can extend the assumptions based on a complex system's perspective on psychopathology to foresee the *type* of these symptom shifts. Thus, rather than applying assumptions from the complex system perspective on foreseeing level shifts in symptoms on a single dimension of psychopathology (in which the constituent symptoms can move from being absent to being present), it would be interesting to explore whether we could extend the impact of these assumptions to a multidimensional psychopathological space, in which we can foresee *what types of symptoms* are likely to develop.

The ability to detect the direction of transitions in psychopathology in high-risk individuals is highly relevant. During subclinical stages of psychopathology, people often experience a combination of symptoms that cross a wide range of psychopathological dimensions^{17,18}. Therefore, it is often entirely unclear how and towards what type of symptoms psychopathology will develop in these high-risk individuals. This, however, is an urgent question as optimal clinical decision making in an early phase is important to reduce and prevent further development of psychopathology^{19,20}. This urgency is expressed in the amount of resources that are invested in the development of adequate clinical staging and profiling techniques^{21–23}. The complex systems perspective may contribute to these aims as it provides a complementary angle from which to understand the development of psychopathology and find solutions to improve personalised prediction.

This chapter will provide an overview of ideas and hypotheses that follow from

taking a complex systems perspective on psychopathology, focusing on foreseeing the type of symptoms that are most likely to show sudden transitions. First, we will describe in more detail what parts of complex systems theory, regarding early warning signals and transitions, have already been related to the field of psychopathology. Second, we will explain what additional predictions may follow from this theory with respect to foreseeing and differentiating the type of sudden shifts in symptoms. Also, we will discuss what research designs would be needed to empirically test these predictions and the relevance of these ideas for clinical practice.

Support for a complex systems perspective on psychopathology

As mentioned above, there are reasons to assume that psychopathology behaves according to principles of complex systems. First, sudden shifts in symptoms are observed in patients. Although psychopathology seems dimensional in nature in the sense that individuals can be anywhere on a continuum between having no symptoms and having severe symptoms, the road of symptom change within individuals can be much bumpier. Patients often report sudden relapses or sudden improvements in symptoms^{7,24,25}. This has been confirmed by statistical analyses of symptom patterns over time in depressed patients, which revealed that most patients show a bimodal distribution in symptoms. In other words, they experience either low levels or high levels of symptoms²⁶. This suggests that they experience sudden jumps in their symptom levels. Also, a recent (N=1) double-blind time-series experiment¹⁵ in which levels of symptoms were weekly and prospectively monitored over 239 days, confirmed the presence of a sudden jump in depressive symptoms in this person. At this change point, the level of symptoms suddenly went up and seemed to stabilize afterwards at a higher point on the continuum of depression. Although not all symptom transitions may occur in an abrupt fashion, these observations at least suggest that abrupt symptom changes are quite common, which is in line with expectations from complex system theory. Currently, extensive empirical research that has mapped symptom patterns in patients frequently and prospectively is lacking. More research is thus needed to confirm the assumption that symptom transitions often occur in an abrupt fashion.

Second, verbal descriptions of patients suggest that sudden and discontinuous changes in their symptom experience⁷ may occur in absence of an obvious, temporally proximal cause or reason. From a traditional approach, these unexpected symptom changes are difficult to explain as we know that external causes play an important role in symptom

development. Logic dictates that changes in symptoms are directly preceded by changes in specific factors (such as in the social environment, stressful events or therapy). From a complex system perspective, however, unexpected symptom change does make sense. This theory proposes that large shifts can also occur following minor, seemingly innocent stressors²⁷. These shifts are most likely when a system's resilience to remain in its current basin of attraction is very low (Fig. 4.1), meaning that the system is in an unstable situation. From a complex systems perspective, such an unstable situation may result from the impact of a distal cause, that happened some time ago, and that gradually led to the present loss of resilience. When resilience becomes very low, even minor, seemingly unimportant disturbances can tip over the system to an alternative state (see Fig. 4.1). This can explain why patients may experience long intervals between potential environmental causes and the onset of their symptoms. Also, it can explain why sudden shifts in symptoms may occur in absence of any obvious immediate trigger. Later we discuss how stability of the system can be empirically assessed.

Third, elements within complex systems are in a continuous and complex interplay with each other. In many complex systems, reinforcing feedbacks are present that, if strong enough, can push the system to another alternative state⁶. Such feedback loops are also likely to occur between mental states. Recent studies have confirmed the presence of feedback loops through network models, which showed that negative mental states, such as feeling down or irritated, are related to the occurrences of other negative mental states later in time. These effects may form vicious circles²⁸. Findings from most studies supported the hypothesis that reinforcing feedback loops were more pronounced in people with either higher levels of psychopathology or at risk of psychopathology, compared with individuals in the general population^{29–32}, although this was not confirmed by all studies^{33,34}. Moreover, a simulation study showed that networks with more strongly connected symptoms showed transitions to a depressed state more often compared with networks with weak connections¹⁶. Furthermore, exposure to external stress resulted in sudden shifts in symptoms only in strongly connected networks. Within such networks, removal of the stressor after the phase transition occurred did not cause the system to shift back to its original state¹⁶. The fact that recovery is not linearly related to the removal of the cause of the shift is called 'hysteresis'. Such non-linearity is typical for complex systems. This phenomenon also has face validity for the field of psychiatry as it may explain why people remain stuck in a clinical state even after removal of certain provocative factors that had prompted the mental complaints.

Finally, the most direct support for the idea that symptom changes behave according to the principles observed in complex system stems from empirical research showing that transitions in symptom levels can be anticipated by directly assessing changes in the stability of the system. From other fields of study (*e.g.*, ecology and computer science), it is known that these changes in stability can be observed using certain ‘early warning signals’ (EWS)^{35–37}. Such signals involve changes in the dynamics of important variables of a complex system, like increasing levels of autocorrelation (*i.e.*, the current state of an element of the system becomes a better predictor for its future state), variance (*i.e.*, elements of the system show greater amplitude changes in their intensity levels) or flickering (*i.e.*, sudden changes in intensity levels). These reflect increasing instability of the system and have been shown to closely precede critical phase transitions in various sorts of complex systems. Considering the above, we expect that EWS (which involve more complex aspects of time series dynamics than simple intensity changes of symptoms) may signal an increased likelihood for a phase transition and that such EWS can predict such transitions substantially earlier than simple changes in mean levels of these symptoms. This means that if psychopathology also behaves as a complex system, we may be able to find EWS that we can use to foresee important shifts in symptoms in an earlier phase and in a personalized manner. A few recent empirical studies already found support for this hypothesis^{13–15}. For example, in the time-series experiment in which a patient was followed over 239 days completing multiple affect measurements a day, EWS were observed in the sum score of all measured affect states. These EWS anticipated a subsequent phase transition in depressive symptoms¹⁵. Although these findings still await replication by a large-scale study in which EWS are followed over time within persons, support for the idea that psychopathology behaves as a complex system is accumulating. To conclude, many hypotheses have been formulated for the application of complex systems theory for the field of psychiatry and some empirical studies have been carried out suggesting we may be able to foresee transitions in levels of symptoms based on the system’s changes in stability. We now want to further explore what this theoretical framework can do for foreseeing the *type of symptom shifts* that individuals may express in the near future. Before we move to the above-mentioned theoretical explorations, however, it is important to first discuss the nature of psychopathology.

Redefining psychopathology as a multidimensional space

The idea that mental disorders are distinct and independent entities is not supported by empirical evidence³⁸. All evidence points to the fact that diagnostic classifications are not independent and show huge overlap with one another^{39,40}. Comorbidity is rule rather than exception⁴¹. Also, patterns of symptoms within diagnoses are quite heterogeneous⁴⁴. The human-made top-down boundaries and classifications established in diagnostic manuals may thus distort our view on the real structure of psychopathology. Clear boundaries between mental disorders seem absent and a patient's pattern of symptoms seems to spread across the various dimensions of psychopathology to form unique clusters of problems in each person. Despite such continuity, however, empirical studies show that some types of symptoms lump together more often whereas other combinations of symptoms hardly ever co-occur^{43–45}. A common explanation for the co-occurrence of certain symptoms or mental states (*e.g.*, feeling hopeless, guilty, down, fatigued, irritated) is that an underlying latent entity is responsible (*e.g.*, depression). However, this view alone cannot explain that the precise combinations of symptoms differ between people and can be mixed with symptoms from other clusters. A more recent view is the network perspective, which does not necessarily assume an underlying causal entity, but theorizes that symptoms can also trigger each other and thereby form clusters of co-occurring symptoms in a self-organized, or bottom-up fashion^{2,43,46}. We can then imagine a reality in which no absolute boundaries exist between diagnoses, and in which symptoms of psychopathology are all, to a varying extent, related to each other within one psychopathological space. However, some types of symptoms or mental states seem to lump together more often in this space whereas other combinations of symptoms hardly ever co-occur. From the network perspective this may make sense, as it can be easily imagined that one mental state, for example 'feeling down' easily triggers a certain other mental state, like 'worrying', but not so easily the mental state representing the feeling of being watched. Thus, certain mental states may be further apart from each other in space than other mental states. Similarly, certain groups of mental states may all be at a relatively close distance – meaning that they easily activate each other – making it likely that they are eventually 'switched on' together. Such groups may represent combinations of symptoms that have been frequently observed together and therefore received a group-identifying label such as 'depression' or 'psychosis'.

Thus, symptoms and dimensions of psychopathology show preferential connections, which puts constraints on the likelihood of certain psychopathological syndromes. Despite

such global constraints, however, individual patients are known to differ in the wiring patterns of their psychopathology networks, which gives rise to unique deviations from group-level connectivity. Thus, whereas in one person depressed feelings (such as sadness, guilt or shame) can easily trigger anxiety symptoms (such as worrying or inner tension) because they are close in the multidimensional space, this is not necessarily true in another person (see Fig. 4.2). Secondly, not only the connections between the clusters (lumps) of symptoms, but also the clusters themselves are likely to differ per individual with regard to their precise content: whereas for one person anxiety and depressive symptoms may form a single cluster or lump, this may not necessarily be the case for the next person. This implies that the structure of psychopathology and its supposed dimensions may differ per person and may not even be stable over time within a person. These theoretical ideas on the nature of psychopathology may explain the observed heterogeneity in symptom profiles (see Fig. 4.2).

4

Fig. 4.2. *Multidimensional space of psychopathology symptoms.* The circles represent different symptoms and the distance between them represents the ease with which they can trigger their neighbor symptom. The colored groups represent different clusters of symptoms. In scenario A it is more likely that depressive symptoms may eventually trigger the manic cluster than in scenario B, whereas depressive symptoms in scenario B will more easily activate anxiety symptoms.

Applying complex systems theory to foreseeing specific types of symptom transitions

As mentioned previously, the possibility to differentiate what type of symptom transitions are likely to occur in vulnerable individuals is highly relevant. The question is whether generic principles that apply to complex systems are informative also on this matter. If symptoms all

relate to each other within a multidimensional space as explained above, and if there are only relative boundaries between symptoms and symptom clusters, then instable attributes of the system at large may in theory inform us on the risk of transitions across the whole of psychopathology. This would mean that rather than working with a 2-dimensional (2D) stability landscape (Fig. 4.1), we suggest a 3D stability landscape (Fig. 4.3) in which system stability is depicted not only on the dimension from low to high stability for psychopathology in general, but also for different dimensions of psychopathology. In this 3D landscape we could then observe that the state of the system of an individual who is, for example, currently experiencing depressive symptoms, is close to a basin of attraction towards developing anxiety symptoms but not to developing manic symptoms (Fig. 4.3a). However, in the 3D landscape of another person this could be precisely opposite (Fig. 4.3b).

Fig. 4.3. *3-D stability landscape*. Landscape A shows a situation in which an individual in a currently depressed state is close to the basin of attraction towards developing anxiety symptoms but not to developing manic symptoms. In landscape B the situation is reversed. This individual who is currently in a depressed state is close to the attractor towards developing manic symptoms but not towards developing anxiety symptoms.

We know that the chances of individuals to develop various forms of psychopathology differ from one person to another. Therefore, it seems logical to expect that the same principles of system stability play a role in signaling phase transitions to a certain set of symptoms (*e.g.*, manic symptoms), rather than to another (*e.g.*, anxiety symptoms), depending on one's individual settings in the multidimensional space. The probability of making a transition to a certain group of symptoms is then reflected in the instability of a system's current state to move towards the corresponding basin of attraction of that particular group of symptoms in the psychopathological space. In a 3D instability landscape, we may assume that this instability is not equal around all basins of attraction. For example, at some point in time the state's position within the instability landscape can make it very easy for the state to roll into a large basin of attraction corresponding to anxiety symptoms (*i.e.*, move along the anxiety dimension), whereas it will not likely roll into the basin corresponding to manic symptoms (*i.e.*, move along the dimension of mania; see Fig. 4.3a).

Early warning signals and local points of instability

The above suggests the possibility that 'local points of instability' exist in the landscape which are related to specific symptom transitions. For example, we can assume that if we find rising EWS specifically, for example, in patterns of the mental state 'feeling down', that this is more likely to signal a transition towards depressed states than towards other symptom clusters. Similarly, rising EWS in other mental states, *e.g.*, 'feeling tense' or 'being extremely talkative', may be less pronounced prior to transitions towards depression relative to transitions towards anxious or manic states, respectively (Fig. 4.4). A first novel assumption therefore is the idea that EWS, such as rising autocorrelation or variance, may signal local points of instability in the 3D psychopathology landscape and can therefore predict what type of symptom transitions are most likely to occur. If this assumption is correct, then EWS patterns of various mental states may inform us not only on the possibility of a nearby transition, but also on the type of psychopathology that may develop in the near future.

Fig. 4.4. Association between early warning signals (EWS) and the proximity of tipping points. As a system approaches a tipping point, a transition towards an alternate state becomes more likely. A) A system that is likely to shift towards a depressed state is hypothesized to show pronounced EWS in related mental states, such as feeling down. B) In contrast, future transitions towards manic states might be preceded by EWS in other mental states, such as feeling talkative. Patterns in EWS might therefore reflect the direction of future transitions.

Fig. 4.5. Network connectivity as a prognostic indicator. Reported symptoms are a mix of clusters A and B. The dynamic networks reveal two different scenarios. In scenario 1 the dynamic network shows high connectivity in cluster A symptoms, whereas in scenario 2 the dynamic network shows high connectivity in cluster B symptoms. The network theory predicts that this leads to different likelihood of future symptomatology. Scenario 1 signals an increased likelihood to further develop symptoms of cluster A, whereas scenario 2 signals an increased likelihood to further develop symptoms of cluster B.

Network structure and local points of instability

A second hypothesis that may follow from the above-described multidimensional view on psychopathology is that *the structure of connections* between symptoms may also inform us on the likelihood for specific directions of symptom transitions. As explained above, symptoms are expected to trigger other symptoms, leading to clusters of symptoms that we label as 'syndromes' (e.g., depression, anxiety, psychosis). If no absolute boundaries exist then we can expect that symptoms can connect not only within fixed groups of symptoms but also across such groups. Symptoms that connect across boundaries are called bridge symptoms. These ideas have been formulated previously^{43,46} and can explain why people with one categorical psychiatric diagnosis (e.g., unipolar depression, when a depression cluster is active) are more likely to later fulfill criteria for another diagnosis (e.g., schizoaffective disorder, when a psychosis component joins in). Since bridge symptoms facilitate most of the communication between the clusters (syndromes), changes in the states of these bridge symptoms may be particularly good candidates for prediction of the direction of phase transitions in psychopathology networks. What follows from these ideas is that the specific patterns of connections between symptoms, like the EWS, may provide us with clues regarding the likelihood of a transition to a particular set of symptoms (a dimension in the landscape) (Fig. 4.5). For instance, if feeling down strongly triggers being paranoid in a particular individual, and if being paranoid is strongly connected to (and easily activates) many other symptom that make up the cluster of psychotic symptoms, then we can say that the state of feeling down is proximal and close to the transition towards psychosis for this person. In the 3D instability landscape this situation would be reflected by a gully in the hilly landscape running from the state of feeling down towards the basin of attraction corresponding to a psychotic state. In such a way the network structure of individuals' symptoms may be informative for local points of instability in the multidimensional space of psychopathology and may signal, depending on the current activation state of the network, what direction of transition is most likely. We thus hypothesize that both approaches, EWS and network structure, may inform on the direction of symptom transitions. An interesting question is whether both approaches are complementary in providing information and whether the combination of information further increases the precision of the estimated likelihood. This question needs to be resolved empirically.

Operationalization in novel research designs and clinical implications

Above we have tried to explain how a complex system view may lead to assumptions concerning the direction of symptom shifts. It is important, however, to not only create new theories on the nature of psychopathology, but also to pave the way for empirical verification of the proposed ideas. This requires a translation from the concept of system instability to the direct testing thereof, formulating specific predictions and utilizing novel research designs. We will discuss each of these challenges. The concept of system instability is often used as a nice metaphor. However, recent discoveries by Scheffer and colleagues^{11,12}, showing that system instability can be estimated by specific EWS, make system instability empirically testable. Whereas time-series of a CO₂ proxy were used to find EWS on climate change¹¹, time-series of emotions/mental states can be used to estimate EWS anticipating symptom transitions. Experience sampling, a diary technique to sample people's experiences in the flow of daily life⁴⁷, is one possibility to get access to such information. An operationalization hereof has been described in Wichers and colleagues¹⁵. Similar statistical models as used in this study to estimate system instability with EWS can be used to test the currently proposed novel hypotheses regarding predictions on the precise type of symptoms that are likely to make a transition (*i.e.*, via local points of instability visualized in a 3D rather than a 2D stability landscape). However, in a 3D stability landscape the situation is a bit more complex. For example, to derive an approximation of the stability landscapes as depicted in figure 4.3, one needs to examine EWS patterns in multiple mental states in these depressed individuals. Strongly rising levels in EWS patterns in certain mental states (*e.g.*, in anxiety) would then correspond with instable locations in the landscape where the state of the system can easily move towards that specific attractor (in that case the anxiety attractor). In that way, (multiple) local instabilities in a system become empirically testable.

However, these empirical tests have their statistical and methodological complexities. For example, the current research questions explicitly hypothesize changing EWS patterns over time. However, most statistical models assume stationarity. Fortunately, there are some solutions available. A first solution, also used in the N=1 experiment¹⁵, is to make use of moving window techniques. This means that autocorrelation, as one form of EWS, is estimated separately in a high number of overlapping time windows that in total cover the complete time-period of investigation. Another solution is to use the recently

developed time-varying vector autoregressive models⁴⁸ that allow for changing parameters over time.

Another complexity in the empirical testing of these ideas is that they require intensive measurement regimes, which should be timed in a period during which transitions are likely to occur. Also, to derive an approximation of the stability landscapes as depicted in figure 4.3, one needs to examine EWS in a broad range of mental states (*i.e.*, a depression cluster and its neighboring clusters, such as anxiety, mania and inhibition clusters). This requires the use of highly parsimonious yet informative questionnaires and data collection techniques. A question is whether such intensive designs are a feasible option in (high-risk) psychiatric populations. Recently, however, intensive time-series datasets (with a length of 3-6 months and with a sampling frequency of once to five times a day) have been successfully collected in patient groups with severe psychiatric problems. Compliance on measurements reached 76% on average and focus groups revealed that patients felt positive regarding the idea of monitoring themselves in such a way as it helps them to acquire more insight into their symptoms^{49,50}. Thus, although data collection is intensive, application hereof in clinical groups, and with the promise of using it for improving personalized prediction of course, may be well feasible. Nevertheless, we need to keep in mind that using complex system tools for this purpose only has added benefit if these outweigh simpler measures, such as plain mean levels of mental states or patients' own indications of risk, to predict symptom trajectories in psychopathology.

In short, psychiatry is strongly in need of personalized approaches. If we succeed, this novel approach may yield strongly improved personalized estimations on the course of psychiatric symptoms. Such information can be extremely useful for early intervention strategies aimed at preventing specific psychiatric problems.

References

1. Heinzel S, Tominschek I, Schiepek G. Dynamic patterns in psychotherapy--discontinuous changes and critical instabilities during the treatment of obsessive compulsive disorder. *Nonlinear Dynamics Psychol Life Sci.* 2014;18(2):155-176.
2. Borsboom D. A network theory of mental disorders. *World Psychiatry.* 2017;16(1):5-13.
3. Haken H, Tschacher W. How to modify psychopathological states? Hypotheses based on complex systems theory. *Nonlinear Dynamics, Psychology, and Life Sciences.* 2017;21(1):19-34.
4. Schiepek GK, Viol K, Aichhorn W, et al. Psychotherapy is chaotic- (not only) in a computational world. *Frontiers in Psychology.* 2017;8(379):1-23. doi:10.3389/fpsyg.2017.00379
5. Nelson B, McGorry PD, Wichers M, Wigman JTW, Hartmann JA. Moving from static to dynamic models of the onset of mental disorder. *JAMA Psychiatry.* 2017;74(5):528-534. doi:10.1001/jamapsychiatry.2017.0001
6. Scheffer M. *Critical Transitions in Nature and Society.* Princeton University Press; 2009.
7. Hayes AM, Laurenceau JP, Feldman G, Strauss JL, Cardaciotto LA. Change is not always linear: the study of nonlinear and discontinuous patterns of change in psychotherapy. *Clinical Psychology Review.* 2007;27(6):715-723. doi:10.1016/j.cpr.2007.01.008
8. van der Maas HL, Molenaar PC. Stagemwise cognitive development: an application of catastrophe theory. *Psychol Rev.* 1992;99(3):395-417.
9. Hayes AM, Strauss JL. Dynamic systems theory as a paradigm for the study of change in psychotherapy: an application to cognitive therapy for depression. *J Consult Clin Psychol.* 1998;66(6):939-947.
10. Schiepek G, Perltz V. Self-organization in clinical psychology. In: Meyers RA, ed. *Encyclopedia of Complexity and Systems Science.* Springer US; 2009:263-285. doi:10.1007/978-1-0716-0421-2_472
11. Scheffer M, Bascompte J, Brock WA, et al. Early-warning signals for critical transitions. *Nature.* 2009;461(7260):53. doi:10.1038/nature08227
12. Scheffer M, Carpenter SR, Lenton TM, et al. Anticipating critical transitions. *Science.* 2012;338(6105):344-348. doi:10.1126/science.1225244
13. Schiepek G, Tominschek I, Karch S, et al. A controlled single case study with repeated fMRI measurements during the treatment of a patient with obsessive-compulsive disorder: testing the nonlinear dynamics approach to psychotherapy. *World Journal of Biological Psychiatry.* 2009;10:658-668. doi:10.1080/15622970802311829
14. van de Leemput IA, Wichers M, Cramer AOJ, et al. Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences.* 2014;111(1):87-92. doi:10.1073/pnas.1312114110
15. Wichers M, Groot PC, Psychosystems ESM group, EWS group. Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics.* 2016;85:114-116. doi:10.1159/000441458
16. Cramer AOJ, van Borkulo CD, Giltay EJ, et al. Major depression as a complex dynamic system. *PLoS one.* 2016;11(12):1-20. doi:10.1371/journal.pone.0167490
17. Fusar-Poli P, Yung AR, McGorry P, van Os J. Lessons learned from the psychosis high-risk state: towards a general staging model of prodromal intervention. *Psychological Medicine.* 2014;44(1):17-24. doi:10.1017/S0033291713000184

18. McGorry P, Nelson B. Why we need a transdiagnostic staging approach to emerging psychopathology, early diagnosis, and treatment. *JAMA Psychiatry*. 2016;73(3):191-192. doi:10.1001/jamapsychiatry.2015.2868
19. Cross SP, Hermens DF, Scott EM, Ottavio A, McGorry PD, Hickie IB. A clinical staging model for early intervention youth mental health services. *Psychiatric Services*. 2014;65(7):939-943. doi:10.1176/appi
20. McGorry PD, Hickie IB, Yung AR, Pantelis C, Jackson HJ. Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. *Australian and New Zealand Journal of Psychiatry*. 2006;40(8):616-622. doi:10.1111/j.1440-1614.2006.01860.x
21. Berk M, Post R, Ratheesh A, et al. Staging in bipolar disorder: from theoretical framework to clinical utility. *World Psychiatry*. 2017;16(3):236-244. doi:10.1002/wps.20441
22. Wigman JTW, van Os J, Thiery E, et al. Psychiatric diagnosis revisited: towards a system of staging and profiling combining nomothetic and idiographic parameters of momentary mental states. *PLoS one*. 2013;8(3):1-8. doi:10.1371/journal.pone.0059559
23. McGorry P, Keshavan M, Goldstone S, et al. Biomarkers and clinical staging in psychiatry. *World Psychiatry*. 2014;13(3):211-223.
24. Tang TZ, DeRubeis RJ, Hollon SD, Amsterdam J, Shelton R. Sudden gains in cognitive therapy of depression and depression relapse/recurrence. *Journal of Consulting and Clinical Psychology*. 2007;75(3):404-408. doi:10.1037/0022-006X.75.3.404
25. Heinzl S, Tominschek I, Schiepek G. Dynamic pattern in psychotherapy- discontinuous changes and critical instabilities during the treatment of obsessive compulsive disorder. *Nonlinear Dynamics, Psychology, and Life Sciences*. 2014;18(2):155-176.
26. Hosenfeld B, Bos EH, Wardenaar KJ, et al. Major depressive disorder as a nonlinear dynamic system: bimodality in the frequency distribution of depressive symptoms over time. *BMC Psychiatry*. 2015;15(1):1-9. doi:10.1186/s12888-015-0596-5
27. Boeing G. Visual analysis of nonlinear dynamical systems: chaos, fractals, self similarity and the limits of prediction. *Systems*. 2016;4(4):1-18. doi:10.3390/systems4040037
28. Wichers M. The dynamic nature of depression: a new micro-level perspective of mental disorders that meets current challenges. *Psychological Medicine*. 2014;44(7):1349-1360.
29. Klippel A, Viechtbauer W, Reininghaus U, et al. The cascade of stress: a network approach to explore differential dynamics in populations varying in risk for psychosis. *Schizophrenia Bulletin*. 2018;44(2):328-337. doi:10.1093/schbul/sbx037
30. Pe ML, Kircanski K, Thompson RJ, et al. Emotion-network density in major depressive disorder. *Clinical Psychological Science*. 2015;3(2):292-300. doi:10.1177/2167702614540645
31. Wigman JTW, van Os J, Borsboom D, et al. Exploring the underlying structure of mental disorders: cross-diagnostic differences and similarities from a network perspective using both a top-down and a bottom-up approach. *Psychological Medicine*. 2015;45:2375-2387. doi:10.1017/S0033291715000331
32. Bringmann LF, Pe ML, Vissers N, et al. Assessing temporal emotion dynamics using networks. *Assessment*. 2016;23(4):425-435. doi:10.1177/1073191116645909
33. de Vos S, Wardenaar KJ, Bos EH, Wit EC, Bouwmans MEJ, de Jonge P. An investigation of emotion dynamics in major depressive disorder patients and healthy persons using sparse longitudinal networks. *PLoS one*. 2017;12(6):1-18. doi:10.1371/journal.pone.0178586
34. Groen RN, Snippe E, Bringmann LF, et al. Capturing the risk of persisting depressive symptoms: a dynamic network investigation of patients' daily symptom experiences. *Psychiatry Research*. 2019;271:640-648. doi:10.1016/j.psychres.2018.12.054

35. Tretyakov AY, Takayasu H, Takayasu M. Phase transition in a computer network model. *Physica A*. 1998;253:315-322.
36. Dakos V, Scheffer M, van Nes EH, Brovkin V, Petoukhov V, Held H. Slowing down as an early warning signal for abrupt climate change. *PNAS*. 2008;105(38):14308-14312. doi:10.1073/pnas.0802430105
37. van Nes EH, Scheffer M. Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift. *The American Naturalist*. 2007;169(6):738-747.
38. Carragher N, Krueger RF, Eaton NR, Slade T. Disorders without borders: current and future directions in the meta-structure of mental disorders. *Social Psychiatry and Psychiatric Epidemiology*. 2015;50(3):339-350. doi:10.1007/s00127-014-1004-z
39. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry*. 2005;62:593-602. doi:10.1001/archpsyc.62.6.593
40. Merikangas KR, He JP, Burstein M, et al. Lifetime prevalence of mental disorders in U.S. adolescents: results from the national comorbidity survey replication-adolescent supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry*. 2010;49(10):980-989. doi:10.1016/j.jaac.2010.05.017
41. Krueger RF, Eaton NR. Transdiagnostic factors of mental disorders. *World Psychiatry*. 2015;14(1):27-30. doi:10.1002/wps.20175
42. Wardenaar KJ, de Jonge P. Diagnostic heterogeneity in psychiatry: towards an empirical solution. *BMC Med*. 2013;11:201. doi:10.1186/1741-7015-11-201
43. Goekoop R, Goekoop JG. A network view on psychiatric disorders: network clusters of symptoms as elementary syndromes of psychopathology. *PLoS one*. 2014;9(11):1-47. doi:10.1371/journal.pone.0112734
44. Boschloo L, van Borkulo CD, Rhemtulla M, Keyes KM, Borsboom D, Schoevers RA. The network structure of symptoms of the diagnostic and statistical manual of mental disorders. *PLoS one*. 2015;10(9):1-12. doi:10.1371/journal.pone.0137621
45. Bekhuis E, Schoevers RA, van Borkulo CD, Rosmalen JGM, Boschloo L. The network structure of major depressive disorder, generalized anxiety disorder and somatic symptomatology. *Psychological Medicine*. 2016;46(14):2989-2998. doi:10.1017/S0033291716001550
46. Cramer AOJ, Waldorp LJ, van der Maas HLJ, Borsboom D. Comorbidity: a network perspective. *Behavioral and Brain Sciences*. 2010;33(2010):137-193. doi:10.1017/S0140525X09991567
47. Myin-Germeyns I, Oorschot M, Collip D, Lataster J, Delespaul P, van Os J. Experience sampling research in psychopathology: opening the black box of daily life. *Psychological Medicine*. 2009;39:1533-1547. doi:10.1017/S0033291708004947
48. Bringmann LF, Hamaker EL, Vigo DE, Aubert A, Borsboom D, Tuerlinckx F. Changing dynamics: time-varying autoregressive models using generalized additive modeling. *Psychological Methods*. 2017;22(3):409-425. doi:10.1037/met0000085
49. Bos FM, Snippe E, Bruggeman R, Wichers M, van der Krieke L. Insights of patients and clinicians on the promise of the experience sampling method for psychiatric care. *Psychiatric Services*. 2019;70(11):983-991. doi:10.1176/appi.ps.201900050
50. Schiepek G, Aichhorn W, Gruber M, Strunk G, Bachler E, Aas B. Real-time monitoring of psychotherapeutic processes: concept and compliance. *Frontiers in Psychology*. 2016;7(604):1-11. doi:10.3389/fpsyg.2016.00604

5

Measuring psychopathology as it unfolds in daily life: addressing key assumptions of intensive longitudinal methods in the TRAILS TRANS-ID study

published as: Schreuder, M.J.*, Groen, R.N.*, Wigman, J.T.W., Hartman, C.A., & Wichers, M. (2020). Measuring psychopathology as it unfolds in daily life: addressing key assumptions of intensive longitudinal methods in the TRAILS TRANS-ID study. *BMC Psychiatry*, 20(1), 1–14. doi: 10.1186/s12888-020-02674-1

*Contributed equally

Abstract

Background. Intensive longitudinal (IL) designs provide the potential to study symptoms as they evolve in real-time within individuals. This has promising clinical implications, potentially allowing conclusions at the level of specific individuals. The current study aimed to establish the feasibility of IL designs, as indicated by self-rated burden and attrition, in the context of psychiatry. Additionally, we evaluated three core assumptions about the instruments (diary items) used in IL designs. These assumptions are: diary items (i) reflect experiences that change over time within individuals (indicated by item variability), (ii) are interpreted consistently over time, and (iii) correspond to retrospective assessments of psychopathology.

Methods. TRAILS TRANS-ID is an add-on IL study in the clinical cohort of the TRAILS study. Daily diaries on psychopathological symptoms for six consecutive months were completed by 134 at risk young adults (mean age = 22.6 years, $SD = 0.6$ years). At baseline, immediately after the diary period, and one year after the diary period, participants completed a diagnostic interview.

Results. Excellent compliance (88.5% of the diaries completed), low participant burden (mean = 3.21, $SD = 1.42$, range = 1-10), and low attrition (8.2%) supported the feasibility of six-month IL designs. Diary items differed in their variability over time. Evaluation of the consistency of diary item interpretations showed that within-individual variability in scores could not be attributed to changing interpretations over time. Further, daily symptom reports reasonably correlated with retrospective assessments (over a six month period) of psychopathology obtained with the diagnostic interview, suggesting that both measures might complement each other.

Conclusion. The current study is the first to show that IL designs over extensive periods (*i.e.*, multiple months) in psychiatry are feasible, and meet three core assumptions to study change in psychopathology. This might allow for addressing novel and promising hypotheses in our field, and might substantially alter how we treat and study mental ill-health.

Introduction

Symptoms of mental ill-health are inherently time-varying: they may progress from mild complaints to severe disorders and vice versa¹. This may play out over longer periods, such as months or even years, but also can occur much more rapidly over the span of several days^{2,3}. Traditional longitudinal designs, which generally assess large groups of individuals a few times over long time-intervals, are capable of investigating slower changes at the group level such as developmental changes (*e.g.*, developmental differences between typically and non-typically developing children). However, these designs cannot provide insight at the individual level, *e.g.*, whether a particular individual is likely to develop anxiety symptoms in the near future, or conversely, is likely to remit. In order to address questions relevant to the individual patient, we need to examine how daily experiences of a specific individual change over the course of several months. Such designs – hereafter referred to as intensive longitudinal (IL) designs – are rigorous but necessary, as they will enable us to connect macro-level changes in individuals (exposed by traditional longitudinal designs) with micro-level processes, such as their vulnerability to mental ill-health at a particular moment in time.

IL designs accommodate several complex features of psychopathology that cannot be accounted for otherwise^{1,4,5}. A first feature is that symptom manifestation and illness progression seem highly person-specific. Individuals differ from other individuals: even when diagnosed with the same disorder (*e.g.*, major depression), individuals typically show marked differences in terms of the specific symptoms they experience and on what timescale they come and go⁶. Further, individuals differ from themselves over time: symptoms fluctuate within individuals both in severity and type (*e.g.*, anxiety symptoms may substitute depressive symptoms, or vice versa)^{7–9}. Fluctuations in severity may entail disorder onset, remission or relapse, and may also occur on a smaller scale, for instance between ‘mild’ and ‘moderate’ levels of severity^{10,11}. These person-specific features cannot be studied through traditional longitudinal designs, which is in part due to the relatively small number of repeated measurements (*i.e.*, usually less than ten) in such studies. In contrast, intensive longitudinal designs such as diary or ecological momentary assessment studies easily yield over 50 time points per individual. This high number and granularity of measurements allows for characterizing the within-individual process, and thus provides the possibility to gain insight in core questions of psychiatry, for instance why some individuals develop symptoms, whereas others do not, and why some individuals experience a mild symptom

course whereas others deteriorate over time.

When monitoring the development of psychopathology, a second feature is that assessments should be acquired during the period in which changes in symptomatology occur. For instance, in order to develop measures that enable detection of upcoming changes in symptoms, or gain insight in how switches in type of symptoms occur, the process of change needs to be observed while it unfolds. Instead of static snapshots with relatively long time intervals in between as in traditional longitudinal designs (*e.g.*, every 12 months), IL designs yield detailed insight into how symptoms change from day to day or moment to moment⁴. As symptom changes may involve both the type of symptoms that manifest and their severity, it is important to assess symptoms of various diagnostic categories simultaneously on a continuous scale. By allowing for variability between and within individuals in symptom profiles, intensive longitudinal designs may enable us to capture individual trajectories from subthreshold symptoms to full-blown disorders as well as transitions that occur between disorders.

As technology advances, the implementation of personalized designs has become increasingly feasible. For instance, IL data can be collected through electronic questionnaires which can be delivered via smartphones, making it possible to lengthen the measurement period. The latter is particularly important when studying how symptoms develop over time, as one needs to maximize the possibility of observing changes in symptoms. Despite increasing availability of methods to collect IL data, prospectively monitoring psychopathological symptom development through intensive designs with extended duration (> three months) has rarely been pursued before¹². It is therefore important to investigate the feasibility of these designs. Thus, the first aim of this chapter is to establish whether prospectively assessing symptoms through a six-month diary study is feasible. Specifically, participant burden – indicated by drop-out and non-compliance – will be evaluated. Second, because IL data show great promise in being able to capture symptom development, it is important to evaluate some of the core assumptions underlying the instruments used in these designs. In the current study we evaluate three such assumptions about daily diary items: (i) diary items are able to detect fluctuations in symptoms, (ii) the interpretation of diary items remains consistent over time, and (iii) diary items reflect the same constructs that are measured with traditional instruments (*e.g.*, retrospective diagnostic interviews).

These assumptions may call for some elaboration. The first assumption states that diary items are suitable for detecting fluctuations in respective symptoms. However, it is

important to evaluate which items may be particularly well-suited for monitoring changes in specific symptom domains or populations. For instance, diary items that are appropriate for specific clinical samples (*e.g.*, “To what extent did you hear voices today?” In those at high risk for psychosis) may not be suitable for samples from the general population or transdiagnostic samples with high risk for more common types of psychopathology, and vice versa. The second assumption involves the consistency of items over time. Because intensive longitudinal designs have long durations and require frequent introspection, it is possible that participants change their interpretation of diary items over the course of the study. For instance, feeling down might initially be perceived as intense sadness, and later be regarded as a mild sense of stress. Such a change makes interpreting the variability in feeling down over time difficult. Although varying item interpretations can also occur in short-lasting designs, *e.g.*, where symptoms are sampled for one or two weeks, the probability of changing interpretations increases with study duration. Hence, in the interest of the current as well as future intensive longitudinal designs, the time-invariance of diary item interpretations should be verified. Finally, given that both methods tap into psychopathology, a third assumption concerns the correspondence of symptom reports in the diary study to symptom reports in conventional assessments of psychopathology such as diagnostic interviews. In other words, we assume that the results of a diagnostic interview – which concern macro-level experiences based on retrospective assessments of psychopathological symptoms during several months – map to some extent onto micro-level experiences based on self-reports of symptoms in daily life. Analyzing this correspondence could potentially reveal which diary items are most suitable for detecting individual differences in specific diagnostic domains. In sum, the present chapter will investigate whether intensive longitudinal assessments of psychopathology are feasible and whether the instruments in those designs (diary items) (i) pick up variability, (ii) are consistently interpreted over time, and (iii) correlate with macro-scale assessments (retrospective diagnostic interview). This may pave the way for the use of intensive longitudinal designs to address important questions in the field of psychopathology.

The current chapter will address the above aims using data from a six-month daily diary study in young adults at risk for psychopathology. The study, TRAILS TRANS-ID, was designed to monitor the day-to-day fluctuations in psychopathological symptoms of participants from a clinical cohort as they entered young adulthood, which is considered a particularly vulnerable period in terms of mental ill-health (see TRAILS TRANS-ID infrastructure)^{13,14}. For six consecutive months, participants completed daily assessments concerning 58 symptoms

of psychopathology. The TRAILS TRANS-ID study has not been described before, and we will therefore provide a detailed description of its methods. Ultimately, the resulting data might allow for conclusions that could considerably advance our understanding of how, why, and when symptoms evolve.

Methods

TRAILS TRANS-ID infrastructure

TRAILS TRANS-ID connects two large, ongoing projects, namely TRAILS (TRacking Adolescents' Individual Lives Survey) and TRANS-ID (Transitions In Depression^e). TRANS-ID was designed to prospectively monitor sudden transitions in psychopathology. TRAILS TRANS-ID is one of the studies within the TRANS-ID project, and includes participants from TRAILS. TRAILS is an ongoing, prospective cohort study investigating mental health across pre-adolescence and young adulthood. TRAILS consists of a general population cohort (TRAILS-PC, which started in 2000) and a transdiagnostic clinical cohort (TRAILS CC, which started in 2004). Participants of TRAILS TRANS-ID were recruited in TRAILS CC. TRAILS CC was designed to enrich its general population counterpart by selectively sampling individuals who were at heightened risk for mental ill-health. Inclusion in TRAILS CC was based on at least one referral before the age of 11 years to the child psychiatric outpatient clinic of the University Medical Center Groningen. Because of this history, TRAILS CC participants are considered vulnerable for developing (additional) psychopathological complaints later in life (*e.g.*, in young adulthood). This is reflected in heightened mental ill-health in TRAILS CC compared to the general population cohort of TRAILS¹⁵. An extensive description of the sampling procedures for TRAILS CC has been published elsewhere¹⁶.

Study design

TRAILS TRANS-ID consisted of a six-month daily diary study that was designed to capture the ebb and flow of (subthreshold) symptoms of psychopathology in young adults at heightened risk for mental ill-health (see Study population). Immediately before (baseline), directly

^e Unlike TRANS-ID, TRAILS TRANS-ID does not solely focus on depression, but aims to examine psychopathology across diagnostic borders, including, but not limited to, mood disorders, anxiety disorders, subthreshold psychosis, thought problems, attention deficit hyperactivity disorder, autism spectrum disorder and aggression.

after (post assessment), and one year after the diary period (follow-up), a clinical interview took place. This interview included a diagnostic interview and two questionnaires on social functioning and life events, respectively. All measures were administered orally and are outlined in detail below. All assessments and other contact moments with participants (*e.g.*, phone calls) were performed by trained researchers (first two authors and research assistant) in a standardized way. Participants remuneration included up to a maximum of €200,- (for details, see supplement 5.1 *Financial compensation*) and a report of their own diary data.

Study population

All participants enrolled in TRAILS CC who participated in at least one of the previous measurement waves of TRAILS CC and had given consent to be approached for future assessments were eligible to participate in the TRAILS TRANS-ID study (see also TRAILS TRANS-ID infrastructure). At its first measurement wave, TRAILS CC included 543 children aged between 10 and 12 years old (mean age = 11.1, $SD = 0.50$, range = 10.1–12.4, 34.1% girls). TRAILS TRANS-ID took place when the TRAILS CC participants were between 21 and 24 years old. Of the 443 TRAILS CC participants eligible for participation in TRAILS TRANS-ID, 134 (30.2%) enrolled in the study (see details on recruitment below). The TRAILS CC cohort study was approved by the Dutch Central Committee on Research Involving Human Subjects (CCMO), in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. Prior to each measurement wave of TRAILS, informed consent was obtained from the adolescents (at earlier waves also from the parents). The TRAILS TRANS-ID study was approved by the standing Ethics Committee, in accordance with the 1964 Declaration of Helsinki. All participants gave written informed consent before enrolment in TRAILS TRANS-ID.

Procedure

Recruitment

From 18 October 2017 until 12 February 2018, individuals of the TRAILS CC cohort were invited to participate in TRAILS TRANS-ID. Individuals received an invitation letter by mail, which contained information about the aims and design of TRAILS TRANS-ID. Individuals could indicate whether they were interested in participation through a website, by phone, or by returning the reply card that was enclosed to invitation letter. Participants who did not respond received a reminder by mail (approximately one month after the first invitation).

Upon registering, participants were contacted by phone by a researcher who provided further information about the study. This contact moment was also used to verify whether participants owned a smartphone and whether they anticipated any circumstances that might interfere with the diary period (*e.g.*, long-term stay in remote areas). Participants who did not own a smartphone were provided one for the duration of the diary period. Of the 142 participants who were interested in participating in TRAILS TRANS-ID, 8 participants (5.6%) were not included for various reasons (*e.g.*, because they could not be reached or had changed their minds).

Baseline assessment

Prior to the clinical interview, participants were once more informed about the study procedures and asked for written informed consent. Thereafter, the clinical interview (see Instruments) was conducted and the diary procedure was explained. Participants completed a practice session in presence of a researcher in order to familiarize themselves with the questions. Participants were instructed to carefully consider their mental state and use the extreme ends only if they could not imagine experiencing a certain symptom more (corresponding to a score of 100) or less (corresponding to a score of 0). To verify whether participants' interpretation of diary items remained consistent during the diary period – as outlined in our second aim – half of the participants were interviewed about their interpretation of a randomly selected set of ten items after the clinical interview (see Qualitative assessment: validity of diary items).

Diary period

Participants started the diary period immediately after the baseline assessment (supplement, Fig. S5.1). This period comprised filling in an electronic diary every evening for six consecutive months (*i.e.*, 183 days). Participants could access the diary via a link sent in a text message to their mobile phone. The diary consisted of questions assessing experiences during the past day (for an overview of items, see supplement 5.3 *Descriptives mini-SCAN and diary items*, Table S5.3). Participants therefore always received text messages in the evening, at a fixed time according to the participant's wishes. Although timing of the text messages differed between participants, all participants had 24 hours between each measurement point. Participants were asked to complete the questionnaire as soon as possible after receiving the text message. If the questionnaire was not completed within 30 minutes, participants

received a reminder text message. After this reminder, participants had 2.5 hours to fill in the diary. The questionnaires were administered and stored via Roqua (www.roqua.nl).

During the diary period, the researchers contacted the participants at least seven times. The first two phone contacts took place in the first and second week of the study, respectively, to resolve any issues regarding, for instance, the timing of the text message or answer questions about the study procedures. Thereafter, participants were contacted monthly or when they skipped diaries for several consecutive days. These phone contacts served to support participants, to solve any issues that occurred (*e.g.*, technical problems), and to monitor life events that might affect symptom dynamics. The latter were assessed by asking participants whether they had experienced important (positively or negatively appraised) events regarding their health, work or education, social relations, and financial situation. These events were noted and, together with the participant, retrospectively dated. Researchers were also available by telephone and e-mail if participants needed assistance at other times than the pre-specified contact moments.

Post and follow-up assessment

After the diary period, participants were invited for a post assessment, which involved a clinical interview spanning the diary period (*i.e.*, the past six months) and an evaluation of the diary. The latter included rating the burden of completing the diary period on a scale from one (not at all burdensome) to ten (extremely burdensome). Participants were offered a report on their personal data including descriptive information which was explained orally and provided on paper. Those who were included in the qualitative assessment of diary items at baseline, were asked to take part in this assessment again at post (see Qualitative assessment: validity of diary items). Finally, one year after finishing the diary period, participants were invited for the follow-up assessment. This assessment involved a clinical interview (mini-SCAN, GVSG, life events questionnaire) that spanned the past year.

Diary measures

The diary questions, listed in Table S5.3 (supplement), were selected to cover symptoms of all common psychiatric disorders. Items were derived from earlier studies that adopted ambulatory assessments or diary protocols^{17–20} as well as from existing questionnaires such as the Positive and Negative Affect Scale (PANAS)²¹, the Adult Behavior Checklist (ABCL)²², and the Symptom Checklist (SCL-90)²³. Items were considered for inclusion in the diary if they

(i) covered multiple diagnostic categories, (ii) were sufficiently distinct from other items, (iii) were expected to vary from day to day, and (iv) were easy to understand. To accomplish the latter, we conducted a pilot study among 12 adolescents with a similar educational attainment level to that of the TRAILS CC cohort (see supplement 5.2 *Pilot study* for more information). The final selection consisted of 58 items, which were all rated on a visual analogue scale (VAS) that ranged from 0 to 100. Participants were given the opportunity to add an item to the diary that they personally wished to monitor (*e.g.*, “To what extent was I productive at work today?”). At the end of the diary, an optional textbox allowed participants to make notes about their day. Participants were encouraged to report unusual or important events, such as starting a new job or being ill, which might have influenced their responses to the diary that day. Participants’ notes remained confidential and were solely examined during the post assessment for the descriptive report on their personal data.

Qualitative assessment: validity of diary items

To evaluate the validity of diary items – as indicated by the consistency of participants’ interpretation – the first 60 participants who were included completed an interview about their interpretation of a randomly selected set of ten items both at the baseline and post assessment. We interviewed 60 individuals rather than the whole sample because this was considered to yield a reliable indication of interpretation consistency, without needlessly burdening participants. During the qualitative assessment, participants were asked what a certain item, such as feeling stressed, meant to them. Participants were asked about the same set at baseline – when participants practiced with the diary – and during the post assessment – when participants discussed their personal data with the interviewer. At both assessments, participants’ interpretations were recorded and later transcribed. Participants were encouraged to elaborate and to provide examples. In case participants did not understand an item during the first session, the item’s meaning was explained by the researcher. In this case, consistency in interpretation was not evaluated for this item.

Instruments

Demographic characteristics

Demographic characteristics of the sample, including socio-economic status and current occupation (education, employment), were retrieved from earlier measurement waves of the clinical cohort of TRAILS. Socio-economic status consisted of a standardized score

that combined information about household income and the education and occupation of participants' parents, and was assessed during the first measurement wave when participants were approximately 11 years old.

Diagnostic interview

The clinical interview that was administered at baseline, post assessment, and follow-up consisted of a diagnostic interview complemented by orally administered questionnaires on social functioning and life events (see *Social functioning* and *Life events*). The diagnostic interview consisted of the computer-assisted mini-SCAN which was administered by trained interviewers and assessed the presence of psychiatric disorders²⁴. The mini-SCAN is a validated semi-structured interview, which by means of screening questions and skips efficiently assesses whether individuals meet the diagnostic criteria of a broad range of psychiatric disorders. Since current purposes were primarily of scientific, as opposed to traditional diagnostic, nature, we slightly adapted the assessment procedure (see also supplement 5.4 *Details on scoring for the mini-SCAN interview*). First, similar to an earlier study²⁵, we rated items on a three-point scale (absent, subthreshold, clinical) rather than a dichotomous scale (absent, present). This allowed for recognizing symptoms that were endorsed by participants, but did not meet the criteria of clinical significance (*e.g.*, with respect to severity, duration, or frequency; see supplement 5.4 *Details on scoring for the mini-SCAN interview*). For instance, feeling down for multiple consecutive days, but less than two weeks, could classify as a subthreshold symptom of depression. The distinction between absent, subthreshold and clinical symptoms was considered particularly relevant since our participants were at risk for psychopathological symptoms. This at-risk status may frequently coincide with mild signs of psychopathology that occupy the grey area between mental health and psychopathology. Adapting the conventional rating procedure of the mini-SCAN allowed for capturing this grey area. A second adaptation concerned the removal of skips that are traditionally imposed by the mini-SCAN. To be able to compare ratings across but also within individuals, we assessed the following sections in each participant: stress, anxiety (including post-traumatic stress disorder (PTSD), generalized anxiety disorder (GAD), panic disorder, social anxiety disorder, specific phobia), obsessive-compulsive disorder (OCD), depression, mania, hallucinations, delusions, substance use, attention deficit hyperactivity disorder (ADHD), autism spectrum disorder (ASD). For the majority of sections, all symptoms were rated regardless of the presence of hallmark symptoms (*e.g.*, for depression these are

depressed mood, or loss of interest). Sections for which skips were preserved were anxiety, hallucinations, and delusions, which all contained symptoms that were conditional upon the hallmark symptom(s). For instance, symptoms of PTSD are conditional upon the experience of a traumatic event and were hence disregarded in the absence of such an event.

The mini-SCAN does not cover oppositional or antisocial behavior, yet such behavior may represent an important precursor or manifestation of psychopathology²⁶. Therefore, we complemented the mini-SCAN with the aggressive behavior subscale of the Adult Self Report (ASR, also assessed at earlier measurement waves of TRAILS²²), which was orally administered. This subscale consists of 15 items that are rated on a three-point scale (never or seldom, sometimes, frequently or always). Scores on the aggressive behavior subscale were considered indicative of symptoms of oppositional defiant disorder (ODD) and will henceforward be referred to as such.

Social functioning

Social functioning was assessed with the Groningen Social Behavior Questionnaire (in Dutch: Groningse Vragenlijst Sociaal Gedrag, GVSG), which has good psychometric properties²⁷. The GVSG consists of 45 items that assess functioning in several domains, including social relations (*i.e.*, with parents, children, friends, and romantic partners), work, school, household, and leisure-time. Items were rated on a 4-point Likert scale which ranged from ‘never’ to ‘always’. Domain scores that exceeded the upper limit cut-off provided by De Jong and Van der Lubbe²⁷ were categorized as ‘severe impairments in functioning’ and domain scores that exceeded the lower limit cut-off were categorized as ‘mild impairments in functioning’.

Life events

The life events questionnaire comprises an adapted version of the List of Threatening Experiences, which has good psychometric properties^{28,29}. The List of Threatening Experiences consists of 12 items describing major stressful life events, such as experiencing serious illness, losing a job, or ending a steady relationship. Eight items describing pleasant events, such as going on a holiday or graduating, were added to this list, resulting in a total of 20 items. The occurrence of these events (yes/no) was rated.

Statistical analyses

The data presented in the current chapter pertain to the baseline assessment, diary period,

and the post assessment. Follow-up interviews were still being conducted. Analyses were conducted in R³⁰ (version 1.463).

Aim 1: feasibility

Our first aim was to explore whether conducting a six-month daily diary study in at risk young adults is feasible. Feasibility was evaluated by examining descriptives (mean, standard deviation) concerning participants' drop-out and compliance, and their responses to the diary-evaluation questions in post assessment interview, which inquired about the burden and reasons for missing diary responses.

Aim 2: three core assumptions

Our second aim was to evaluate three common assumptions of IL assessments of psychopathology. The first assumption is that diary item capture variability in symptom expression. This assumption was evaluated by computing the between- and within-individual variability in symptom reports. Between-individual variability was operationalized as the deviance in mean item scores across individuals. Large values indicate that individuals differ considerably in their scores on the respective item, whereas small values suggest that individuals score relatively similar. Within-individual variability was examined through calculating the mean deviation in item scores within individuals. Large values indicate that the respective item varies considerably over time within individuals, whereas small values suggest that scores are relatively stable. As there are no criteria yet for the minimum amount of variability that can be considered informative, we compared different items descriptively.

The second assumption to evaluate was the consistency of participants' interpretations of diary items over time. This was addressed by comparing, per participant and per item, interpretations at baseline and at post assessment. Item interpretations were coded as consistent when participants used identical or similar words or examples to describe an item at baseline and post assessment, respectively. Inconsistent interpretations, in contrast, were scored in case the interpretations at baseline and post assessment showed marked discrepancies. All interpretations were coded by two raters separately. Interrater agreement was assessed through the B statistic^{31,32} instead of Cohen's Kappa because ratings of inconsistency were far less frequent than ratings of consistency. The B statistic quantifies the agreement between two raters, accounting for the agreement that would occur due to chance³¹.

To address the third assumption – concerning the correspondence between daily symptom reports and diagnostic measures assessed in a clinical interview – we examined the correlation between mean diary ratings and mean scores on distinct psychopathological domains (assessed by the diagnostic interview at post assessment). Because the data did not fulfill parametric assumptions, Spearman correlation coefficients were computed. We then determined correspondence by evaluating which diary items correlated most strongly with each domain.

Results

Sample characteristics

Of the 443 members of the TRAILS CC cohort who were invited to participate, 142 (32%) expressed their interest in TRAILS TRANS-ID. Of these individuals, 134 (94%, 30% of the invited number of participants) enrolled in the study (76 males (57%); mean age 22.6 years old, $SD = 0.6$). At the fourth measurement wave of TRAILS CC, when participants were approximately 21 years old, employment and education status were assessed. The majority of participants ($N=73$, 54%) were currently employed, of whom some with current education ($N=46$) and some without current education ($N=27$). Other participants either were employed in the past ($N=37$ (28%), of whom 23 with and 14 without current education), had never been employed ($N=14$ (10%), of whom 11 with and 3 without current education), or did not report employment or education ($N=10$, 7%). At the first measurement wave of TRAILS CC, when participants were approximately 11 years old, the average deviation quotient (which approximates the intelligence quotient) as measured by the Wechsler Intelligence Scale for Children³³ was 100.3 ($SD = 14.9$). Participants' socio-economic status, also assessed at the first measurement wave of TRAILS CC, was based on parental education, parental employment, and household income. The participants in the present study had a socio-economic status similar to that in the general Dutch population (low: $N=28$, 21%; medium: $N=69$, 52%; high: $N=37$, 28%). In terms of social functioning, 82 participants (67%) improved (baseline score < post assessment score) and 40 participants (33%) worsened (Table 5.1). Both at baseline and at the post-assessment, approximately 30% of participants met the DSM-5 criteria for a psychiatric disorder. Of the 40 participants (30%) with a diagnosis at baseline, 27 (20%) met criteria for a single diagnosis, 8 (6%) met criteria of two diagnoses, and 5 (4%) had three or more diagnoses. Seventeen participants (14%) who initially received a diagnosis did not meet

diagnostic criteria half a year later. A similar number of participants received a diagnosis at the post-assessment, but not at baseline (N=11, 9%). Mean scores and standard deviations for each mini-SCAN domain and diary item are listed in supplementary Tables S5.2 and S5.3.

Table 5.1. *Sample characteristics*

	Baseline (pre-diary)			Post-assessment		
	N=134			N = 122		
<i>Life events (mean, SD)</i>						
No. of negative life events	1.32 (1.41)			1.07 (1.10)		
No. of positive life events	2.04 (1.18)			2.17 (1.32)		
<i>GVSG impairments (N, %)*</i>	No	Mild	Severe	No	Mild	Severe
Parents	116 (87%)	10 (7%)	7 (5%)	95 (78%)	15 (12%)	12 (10%)
Partner	54 (40%)	15 (11%)	10 (7%)	42 (34%)	12 (10%)	13 (11%)
Children	9 (7%)	1 (1%)	0 (0%)	7 (6%)	2 (2%)	1 (1%)
Friends	96 (72%)	21 (16%)	15 (11%)	72 (59%)	36 (30%)	10 (8%)
Education	30 (22%)	20 (15%)	5 (4%)	31 (25%)	19 (16%)	4 (3%)
Occupation	57 (43%)	26 (19%)	14 (10%)	41 (34%)	38 (31%)	23 (19%)
Household	61 (46%)	54 (40%)	14 (10%)	52 (43%)	49 (40%)	18 (15%)
Spare time	63 (47%)	31 (23%)	40 (30%)	52 (43%)	34 (28%)	36 (30%)
<i>mini-SCAN (N, %)</i>						
Anxiety disorder	21 (16%)			12 (10%)		
Mood disorder	28 (21%)			23 (19%)		
Psychotic disorder	2 (1%)			5 (4%)		
Attention deficit and/or hyperactivity disorder	8 (6%)			8 (7%)		
Substance use disorder	3 (2%)			4 (3%)		

Note. Sample characteristics at follow-up are reported in the supplement (Table S5.5). Table S5.5 was not included in the published article, as the follow-up assessments were still being conducted at that time. GVSG = Groningse Vragenlijst Sociaal Gedrag, a measure of social functioning in eight domains. mini-SCAN = short version of the Schedules for Clinical Assessment in Neuropsychiatry. SD = standard deviation. *Not all domains were applicable to each participant (e.g., most participants did not have children). Hence, percentages do not necessarily add up to 100%. Severity of impairments was scored based on lower and upper cut-off scores reported by de Jong *et al.*²⁷.

Fig. 5.1. *Flowchart of recruitment and inclusion procedure for TRAILS TRANS-ID.* Percentages use the participants in the cell above as reference (e.g., 91% of those individuals who completed baseline assessments also completed post assessments). Practical reasons for not including participants were, for instance, long-term stay in remote areas with limited internet access.

Feasibility of the present design

Of the 134 included participants, one decided to quit the study immediately after the baseline interview and 11 participants (9%) dropped out during the diary period, resulting in 122 participants (91% from those included) who remained in the study (Fig. 5.1). The 11 participants who did not complete the study remained on average 83.45 days ($SD = 39.28$, percentage = 46%, range = 24-146) and completed 39.91 diaries ($SD = 23.22$, percentage = 46%, range = 9-76). They were similar to those who remained in the study with respect to age (mean 22.46 vs. 22.59 years old, $t(116) = 0.69$, $p = .49$), sex (58% vs. 57% male, $\chi^2(1) < 0.01$, $p > .99$), socio-economic status (mean standardized score 0.13 vs. 0, $t(132) = 0.59$, $p = .56$), IQ ($t(132) = 1.51$, $p = .13$), and diagnostic status at TRAILS TRANS-ID baseline (25% vs. 28% with diagnosis, odds ratio = 0.86, 95% CI = 0.14-3.73, $p > .99$). As denoted in Table 5.2, participants who remained in the study completed on average 88.4% of the diaries, and did not experience the diary period as very burdensome (mean burden: 3.21, on a scale from 1 (not at all burdensome) to 10 (extremely burdensome)). Missing a diary occurred most often because of interfering activities, such as social events, which participants were unwilling or unable to interrupt.

Table 5.2. *Completion and evaluation of diary protocol*

	Mean or frequency	Percentage
No. of completed diaries (<i>SD</i>)	162.30 (17.58)	88.4%
Burden on a scale from 1 (almost no burden) to 10 (extremely burdening) (<i>SD</i>)	3.21 (1.42)	
<i>Reason for missing a diary</i>		
I did not have my phone with me or was unaware of receiving a text message.	11	9%
Due to technical issues, I did not receive the text message or could not open the link.	9	7%
I was already asleep.	10	8%
I could not motivate myself to fill in in the diary.	13	11%
I was engaged in an activity I did not want to interrupt.	64	52%
I was in a location where it was impossible to fill in the diary.	7	6%
Other*	6	5%
Missing or not applicable**	2	2%

Note. *Other reasons for missing a diary, for instance, a broken smartphone. **Not applicable because participant did not miss any diaries. *SD* = standard deviation.

Assumption 1. Variability in symptom expression

The value of intensive longitudinal designs lies in their ability to capture both between- and within-individual variability in symptom reports. These are described in Table S5.3 (supplement). In general, items with lower mean ratings (averaged within, and subsequently across individuals) were less variable both between and within individuals. This is indicative of floor effects, for instance illustrated by the difference in mean and variability between a rarely endorsed item 44. *I had a fight* (mean = 7.50, $SD_{\text{within}} = 7.34$) and a commonly endorsed item 35. *I felt tired* (mean = 35.06, $SD_{\text{within}} = 18.78$). Other items where floor effects seem present are 53. *I used (soft)drugs today* (mean = 3.11, $SD_{\text{within}} = 3.57$) and 32. *I was easily startled* (mean = 9.67, $SD_{\text{within}} = 7.26$). In contrast, 22. *I could not bring myself to do anything* and 37. *It bothered me that things did not go as expected* showed high variability both between and within individuals and might thus be well-suited for intensive longitudinal designs (mean = 18.40 and 21.76, $SD_{\text{between}} = 11.49$ and 14.12, $SD_{\text{within}} = 15.36$ and 15.95, respectively). Compared to negatively valenced items, positively valenced items generally showed higher means (range = 46.88-61.41 vs. 3.11-35.06) and variability (range_{between} = 13.37-16.81 vs. 5.84-17.50; range_{within} = 12.79-18.34 vs. 3.57-18.79), and were hence less susceptible to floor effects. Across items, between- and within-individual variability were similarly large (overall $SD_{\text{between}} = 13.22$, overall $SD_{\text{within}} = 13.21$).

Assumption 2. Consistent interpretations of diary items over time

The majority of interpretations that could be evaluated was rated as consistent at baseline and post assessment by both raters (460, or 95%). Other interpretations were scored as inconsistent by one rater (15, or 3%) or by both raters (7, or 1%). Hence, fluctuations in symptoms as assessed by the diary could not be attributed to changing interpretations over time ($B = 0.96$, indicating excellent agreement). Due to attrition and technical issues, respectively, complete data on the interpretation of the diary items (*i.e.*, reports at both baseline and post assessment) were available for 55 of the 60 individuals who participated in the qualitative assessment. Because each participant was asked to describe ten items both at baseline and at post assessment, there were 550 sets of interpretations. Of these interpretations, 68 (12%) could not be rated due to insufficient information either at baseline or at post assessment. For instance, in some cases participants did not understand the item or could not come up with an interpretation other than the item (*e.g.*, “Feeling down means that I feel down”).

Assumption 3. Correspondence of diary items and clinical interview

Figure 5.2 shows the Spearman correlations between diary item scores (within-individual means) and mini-SCAN domains scores at post assessment (see Table S5.7 in the supplement for raw data). Overall, different diagnostic domains appeared to be related to distinct patterns in daily symptom reports (see Table S5.6 in the supplement). Internalizing problems – including the domains stress, anxiety, OCD, and depression – were strongly related to feeling tired (item 23), stressed (item 33), and worried (item 27). Anxiety was further covered by experiencing physical discomfort (item 51), whereas OCD and depression were more related to feeling down (item 20). Mania and psychosis were both related to feeling restless (item 30). Mania was additionally related to thought problems – reflected by 29. *I was easily distracted* and 34. *I felt overburdened* – whereas psychosis was more related to experiencing physical discomfort (item 51) and using (soft)drugs (item 53). ASD, ADHD, and substance abuse were all associated with items that specifically tapped into these domains. Specifically, ASD was primarily covered by feeling overstimulated (item 36) and being bothered because things did not go as expected (item 37), whereas ADHD was related to being impulsive (item 41) and being unable to sit still (item 40). The items that most strongly related to substance abuse were using (soft)drugs (item 53) and drinking alcohol (item 52). Finally, ODD was most strongly associated with a mix of items that covered internalizing domains – such as feeling

down (item 20) – and more ODD specific items such as feeling irritated (item 42) and having mood swings (item 47). In conclusion, daily reports of (subthreshold) symptoms aligned with retrospectively reported symptoms assessed in a diagnostic interview, and showed both specific and non-specific associations with the interview domains.

Fig. 5.2. *Spearman correlation coefficient between diary items and mini-SCAN domains.* The color and size of the dots reflect the direction and magnitude of the correlation. Correlations that were not statistically significant ($p > .05$) were omitted from the figure.

Discussion

In the current chapter, we addressed the feasibility of studying the development of psychopathological symptoms by means of six months' intensive longitudinal measurements and evaluated three assumptions that accompany similar designs using data from the TRAILS TRANS-ID study. We found that conducting a six-month diary study in young adults at risk for mental ill-health is feasible, demonstrated by excellent compliance, low perceived burden, and low attrition. Further, all diary items showed variability over time, but some more so than others. Currently, there are no objective criteria that determine what constitutes sufficient variability. Therefore, for now, these findings indicate that differences between items exist and that researchers need to consider this when compiling the diary instrument for their intensive longitudinal design. Importantly, for the vast majority of individuals, the meaning of diary items remained consistent over time, implying that daily diaries can validly capture changes in symptoms. The correspondence between retrospective diagnostic interviews and diary reports showed how psychopathological disorders manifest in daily life. In this high-risk, transdiagnostic sample, we found that diary items differ in the number and type of psychopathological domains they cover. Taken together, our findings show that day to day monitoring of psychopathological symptoms in at risk young adults is feasible and that our items are suitable for studying symptom development. In addition, our findings may inform future intensive longitudinal studies with respect to methodological considerations (*e.g.*, concerning the items that are most appropriate for monitoring particular psychopathological domains).

Six-month diary studies are feasible

We showed that relatively long intensive longitudinal designs with a large sample size are feasible. To date, the majority of personalized designs in psychiatry adopted high-frequency designs for short periods of time^{20,34–39}. By favoring sampling duration over sampling frequency, diary studies such as TRAILS TRANS-ID capture changes that occur over a larger timescale³⁸. This trade-off did not undermine the compliance of participants. Participants completed on average 89% of diaries- which is comparable to earlier studies that employed ambulatory assessments for shorter time intervals^{40,41}. Several factors might have contributed to this⁴². First, our participants have participated in scientific research for multiple years (*i.e.*, in TRAILS CC), which likely contributed to their commitment to also participate in this add-

on study. Second, we aimed to minimize the burden introduced by the diary protocol by tailoring the protocol to participants' daily lives (*e.g.*, by allowing participants to complete the diary on their mobile phone at the end of the day at a time convenient for them). Third, we contacted participants at least every month during the diary period which may have aided in their adherence to the study protocol⁴³. Regular contact also allowed us to remove barriers that prevented participants from completing the diaries, for instance optimization of the time at which the diary could be completed. Finally, participants were compensated for their efforts both personally and financially. The personal incentive consisted of a report on participants' own data, which was generally received with great interest. Participants further received a financial compensation that was contingent on the number of completed diaries. In conclusion, TRAILS TRANS-ID shows that the benefits of intensive longitudinal diary studies (*i.e.*, rich data pertaining to within-individual processes for a relatively long period of time) certainly outweigh their costs (*i.e.*, reported burden and attrition of participants, and time and financial investment of researchers).

Daily diaries fulfill prerequisites to measure psychopathology in daily life

Intensive longitudinal designs target within-individual changes. Therefore, the quality of measurement instruments should be evaluated on a within-individual level. Traditional longitudinal designs have addressed this, for instance, through structural equation modeling or dynamic factor modeling^{44,45}. These techniques are less suitable for IL designs because of differences in the data structure (*e.g.*, many more assessments per individual, missing and unbalanced data, non-stationary data). The present study suggests that our instruments are fit for their purposes. Specifically, daily diaries capture meaningful variability in symptoms, in the sense that this variability cannot be attributed to changes in the (subjectively perceived) content of items and reasonably corresponds to variability in diagnostic measures. First, a qualitative analysis of individuals' interpretation of diary items immediately before and after the six-month diary period showed that participants' interpretations were highly consistent over time. Participants were, however, sometimes unable to understand or verbalize certain items. This highlights the importance of verifying the comprehensibility of self-report measures, particularly in populations with average to limited verbal or introspective abilities⁴⁶. Pilot studies in samples with a background similar to that of the target population – such as the pilot study conducted prior to the current project – aid in accomplishing appropriate phrasing of items. Second, evaluation of the correspondence between diary

items and diagnostic interviews showed that retrospectively assessed diagnoses reasonably manifested in daily symptom reports (*e.g.*, ADHD and being unable to sit still). This shows that intensive longitudinal data and cross-sectional data might enrich each other by each providing partly overlapping and partly unique information concerning individuals' mental health. Together, these sources of information can provide a detailed picture of how, when, and for whom psychopathological symptoms develop.

Methodological issues

First, IL designs require a careful consideration of the timescale at which processes unfold. Whereas mood changes within minutes or hours, symptoms may evolve over longer timescales⁴⁷. This means that assessing dynamics in mood versus symptoms requires different designs. The within-individual variability in diary ratings found in the current study illustrates that the dynamics of psychopathological symptoms – or subthreshold expressions thereof – can indeed be captured by daily assessments. Second, it is possible that the individuals included in studies using IL designs are not representative of the general population, for instance because individuals who are willing to monitor their mental state for several consecutive weeks or months might be relatively more motivated⁴⁸. However, in the current study, financial compensation was given to participants which may have limited this issue. Further research is needed to explore who are most likely to enroll in and complete IL studies. Finally, the repeated assessments adopted in IL designs might evoke reactivity⁴³. That is, repeatedly assessing one's mental state might affect one's (appraisal of) experiences, thoughts, and feelings. Reactivity might be limited by providing participants with their own reports after instead of during the diary study – as was done in the current study. This is because it is less likely that participants alter their responses based on earlier responses when they are not confronted with earlier responses. In conclusion, designing and evaluating studies using intensive longitudinal methods requires appreciation of the timescale of interest as well as participants' response to these methods.

Future implications

The established feasibility of intensive longitudinal diary studies calls for a brief overview of the types of theories that could be investigated using these designs. Network theory^{49,50} and complex systems theory^{1,51,52}, for instance, both propose that the dynamics of symptoms – *i.e.*, the temporal variation in symptoms within individuals – might be predictive of long-term

prognosis. According to network theory, psychopathology can be conceptualized as a network of (causally) interacting symptoms. The more symptoms trigger each other, the more clusters of co-occurring symptoms arise. The strength and the structure of associations between symptoms may thus reveal individuals' vulnerability towards developing psychopathology. Second, psychopathology has been proposed to behave according to principles common to complex systems. This means that indicators of instability, referred to as early warning signals (EWS), may be informative of sudden transitions in symptom severity^{3,53,54}. These early warning signals are based on the temporal dynamics of symptoms, for instance reflected in the degree to which they carry over from one moment to the next. Because TRAILS TRANS-ID includes a broad set of symptoms spanning multiple disorder domains, these data may also be used to investigate whether EWS are informative of the type of symptom shifts in addition to shifts in symptom severity⁵¹. Both network and complex systems approaches to psychopathology raise hypotheses on the level of individuals. That is, changes in symptoms within an individual – operationalized either as network characteristics or as early warning signals – may be informative of the future development of that individual. Due to the individual-centered approach, the conclusions that follow from these hypotheses could directly translate to clinical practice. In order to examine the clinical utility of network and complex systems approaches, personalized designs such as TRAILS TRANS-ID are needed.

In addition to deriving individual-specific models, intensive longitudinal data collected in comparatively large samples such as in the TRAILS TRANS-ID study also allow for investigating what temporal processes may be shared by multiple individuals. Although some have argued that group-findings based on IL data do not translate well to individuals⁵⁵, it is also unlikely that every individual is unique and requires their own model describing psychological processes. Newly developed clustering techniques such as Group Iterative Multiple Model Estimation (GIMME⁵⁶) can be applied to IL data to detect (sub)groups of individuals that share temporal associations. Its embedding in the larger TRAILS cohort study yields the TRAILS-TRANS-ID study the unique possibility to investigate associations between current day to day symptom patterns and previously collected developmental and contextual variables as well as future outcomes.

Conclusion

TRAILS TRANS-ID has shown that prospectively monitoring symptoms as they develop over time in at risk young adults is within reach. We have demonstrated feasibility, within-person variability, consistency in meaning and correspondence with diagnostic outcomes. It is our hope that this pioneering study provides the groundwork for intensive longitudinal designs over extended periods that aim to investigate how psychopathology unfolds over time within individuals. The outcomes of such intensive longitudinal designs allow for investigating individual-level hypotheses derived from novel conceptualizations of psychopathology such as network theory and complex dynamic systems theory.

Supplement 5

Fig. S5.1. *Outline of the TRAILS TRANS-ID study design.* At baseline and post assessment, the following instruments were employed: diagnostic interview (short version of the Schedules for Clinical Assessment in Neuropsychiatry²⁴), aggressive behavior subscale of the Adult Self Report²², Groningse Vragenlijst Sociaal Gedrag²⁷, List of Threatening Experiences²⁸. The same instruments will be administered during the follow-up assessment, which was not completed yet at the time of writing this chapter. For a subset of 60 participants, the baseline and post assessments were complemented by a qualitative assessment that served to establish the validity of diary items.

5.1 Financial compensation

Participants received €20,- per interview (baseline, post assessment, and follow-up). Additionally, participants received compensation for the completion of daily diaries, which was contingent on the number of diaries completed (see table below). The maximum compensation amounted to €200,- (*i.e.*, €60,- for participating in all three clinical interviews, together with €140,- for completing ≥90% of the diaries).

Table S5.1. *Financial compensation for the diary period*

Number of daily diaries completed	Financial compensation
<128 (<70%)	€40,-
128-145 (70-79%)	€80,-
146-163 (80-89%)	€130,-
≥164 (≥90%)	€140,-

5.2 Pilot study

Participants for the pilot study were recruited in a local school for intermediate vocational education. Adolescents were informed about the purposes of the pilot study and provided written informed consent. In a semi-structured interview, individuals were asked to report their intuitive interpretation of diary items and to rate the extent to which this item described their current mood. Individuals were explicitly instructed to provide their personal view on items and were ensured that answers were not labeled ‘wrong’ or ‘right’. After completion of the interview, individuals received a voucher worth €20,-.

Interpretations were noted and discussed among the research team. Of the 63 items assessed in the pilot study, nine items were rephrased in order to better match participants’ vocabulary. Five items were deleted due to substantial perceived overlap with other items or due to frequently occurring misunderstandings. Finally, one item was added, resulting in a total of 58 items concerning mood (*e.g.*, “I felt happy”), (subthreshold) psychopathological symptoms (*e.g.*, “I felt stressed”), somatic experiences (*e.g.*, “I felt ill”), sleep quality (*e.g.*, “I slept well”), and event appraisal (*e.g.*, pleasantness of the most important event during the previous day).

5.3 Descriptives mini-SCAN and diary items

Table S5.2. Scores per mini-SCAN domain

Domain	Baseline	Range	Post assessment	Range
Stress	0.51 (0.48)	0-2	0.57 (0.55)	0-1.83
Anxiety	0.13 (0.23)	0-1.69	0.12 (0.20)	0-1.19
OCD	0.18 (0.36)	0-1.50	0.16 (0.33)	0-1.56
Depression	0.32 (0.35)	0-1.50	0.39 (0.41)	0-1.64
Mania	0.13 (0.24)	0-1.78	0.16 (0.25)	0-1.78
Psychosis	0.02 (0.07)	0-0.55	0.02 (0.08)	0-0.50
ASD	0.43 (0.39)	0-1.50	0.41 (0.40)	0-1.70
ADHD	0.66 (0.46)	0-1.79	0.67 (0.43)	0-1.70
Substance abuse	0.16 (0.19)	0-1.71	0.20 (0.19)	0-1.29
ODD	0.35 (0.30)	0-1.60	0.37 (0.32)	0-1.40

Note. Numbers denote mean scores per domain; numbers between brackets denote standard deviations. OCD = obsessive-compulsive disorder. ASD = autism spectrum disorder. ADHD = attention-deficit/hyperactivity disorder. ODD = oppositional defiant disorder.

Table S5.3. *Descriptives diary items*

Diary item	within-individual mean	between-individual SD	within-individual SD
1. Last night I slept well	55.09	13.37	16.29
2. I felt energetic	47.79	14.76	17.09
3. I was physically active	46.88	15.26	18.03
4. I was happy	57.71	15.83	14.23
5. I could enjoy things	58.57	14.37	14.84
6. I felt good	58.68	14.67	15.23
7. I felt relaxed	50.20	14.93	17.72
8. My day was worth living	61.41	14.98	15.07
9. I could handle the things that confronted me	61.35	14.08	15.20
10. Everything came easy	58.24	14.79	12.79
11. I felt like undertaking things	48.41	14.98	18.34
12. I am looking forward to tomorrow	55.79	16.58	16.01
13. My concentration was good	51.87	14.06	15.95
14. I felt appreciated	56.68	16.81	14.48
15. I felt like others liked me	58.16	15.30	14.42
16. I felt at ease with others	59.01	14.64	15.16
17. I was social	57.85	14.56	14.89
18. I felt lonely	14.63	13.31	11.24
19. I felt empty	14.08	12.48	10.95
20. I felt down	15.62	13.64	12.20
21. I felt guilty	12.64	10.74	10.17
22. I could not bring myself to do anything	18.40	11.49	15.36
23. I felt tired	35.06	15.86	18.79
24. I avoided things	16.00	12.68	12.47
25. I was afraid of making mistakes	16.61	12.67	12.34
26. I had difficulty making decisions	16.65	11.93	12.75
27. I worried	21.60	14.23	14.61
28. My thoughts would not leave me alone	23.68	17.14	14.45
29. I was easily distracted	19.35	13.48	14.01
30. I felt restless	19.55	13.71	13.66
31. I felt nervous	15.46	11.75	12.89
32. I was easily startled	9.67	9.18	7.26
33. I felt stressed	18.97	13.87	13.99
34. I felt overburdened	18.03	14.28	13.27
35. I was easily upset	12.96	10.76	10.99
36. I felt overstimulated	17.28	14.94	12.52
37. It bothered me that things did not go as expected	21.76	14.12	15.95
38. I was bored quickly	13.19	9.70	10.65
39. I felt impatient	14.28	10.69	11.05
40. I was unable to sit still	17.75	13.74	13.12

Diary item	within- individual mean	between- individual <i>SD</i>	within- individual <i>sd</i>
41. I was impulsive	20.76	15.44	12.99
42. I felt irritated	17.03	11.62	13.59
43. Others felt annoyed by me	12.70	9.62	9.65
44. I had a fight	7.50	8.11	7.34
45. I felt rebellious	12.15	10.59	9.07
46. Things did not bother me	21.18	16.01	13.82
47. I had moodswings	15.59	12.98	11.67
48. I felt strange	13.22	11.31	10.06
49. I felt suspicious	11.74	11.17	8.49
50. I experienced physical pain	17.51	14.05	13.88
51. I experienced physical discomfort	14.93	11.09	13.44
52. I drank alcohol today	7.92	8.07	10.24
53. I used (soft)drugs today	3.11	5.84	3.57
54. Number of pleasant events	35.64	17.48	14.47
55. Intensity of pleasant events	46.81	17.85	17.74
56. Number of unpleasant events	15.94	9.72	10.95
57. Intensity of unpleasant events	23.47	13.75	17.49
58. I'm dreading something that will happen soon	23.07	17.50	15.06

Note. Within-individual mean = averaged mean item score per individual. Between-individual *SD* = standard deviation of mean item scores. Within-individual *SD* = averaged standard deviation of item scores per individual.

5.5 Details on scoring for the mini-SCAN interview

Table S5.4 provides an overview of the scoring rules adopted for the mini-SCAN interview. The mini-SCAN interview is typically administered to infer whether individuals meet the diagnostic criteria for a psychiatric disorder, and therefore only assesses clinically significant symptoms. For the purposes of the present study, a more dimensional rating of mental health was adopted. That is, we rated symptoms that were insufficiently severe, enduring, or frequent to qualify as a sign of psychopathology – but nevertheless present and distressing – as subthreshold²⁵. To infer whether individuals met the criteria for a psychiatric disorder, the original, dichotomous rating scheme was used. Hence, only symptoms coded as clinical were taken into account when inferring whether participants met the criteria for a psychiatric disorder.

Table S5.4. *Scoring of the mini-SCAN interview*

Description	Mini-SCAN rating in TRAILS TRANS-ID	Original mini-SCAN rating*
Symptom causes significant burden due to impairments in functioning, duration and/or frequency of occurrence.	Clinical	Present
Symptom is endorsed, but does not cause significant burden (e.g., because of mild severity, short duration or infrequent occurrence).	Subthreshold	Absent
Symptom is absent.	Absent	Absent

*See Nienhuis *et al.*²⁴.

5.6 Other results

Table S5.5. *Sample characteristics at follow-up*

	Follow-up N=115		
<i>Life events (mean, SD)</i>			
No. of negative life events	1.10 (1.18)		
No. of positive life events	1.89 (1.20)		
<i>GVSG impairments (N, %)*</i>	No	Mild	Severe
Parents	95 (83%)	11 (10%)	8 (7%)
Partner	37 (32%)	15 (13%)	13 (11%)
Children	10 (9%)	2 (2%)	0 (0%)
Friends	73 (63%)	25 (22%)	11 (10%)
Education	15 (13%)	15 (13%)	1 (1%)
Occupation	32 (28%)	38 (33%)	26 (23%)
Household	44 (38%)	52 (45%)	16 (14%)
Spare time	37 (32%)	39 (34%)	39 (34%)
<i>mini-SCAN (N, %)</i>			
Anxiety disorder	21 (18%)		
Mood disorder	32 (28%)		
Psychotic disorder	4 (3%)		
Attention deficit and/or hyperactivity disorder	4 (3%)		
Substance use disorder	6 (5%)		

Note. This table was not included in the article published in BMC Psychiatry, as the data were not available at the time of publishing. GVSG = Groningse Vragenlijst Sociaal Gedrag, a measure of social functioning in eight domains. mini-SCAN = short version of the Schedules for Clinical Assessment in Neuropsychiatry.

Table S5.6. *Five diary items that correlated most strongly with each psychopathological domain assessed by the mini-SCAN (left: highest Spearman correlation coefficient)*

mini-SCAN domain	Diary item				
stress	28. My thoughts would not leave me alone	23. I felt tired	33. I felt stressed	27. I worried	30. I felt restless
anxiety	28. My thoughts would not leave me alone	27. I worried	57. Intensity of unpleasant events	51. I experienced physical discomfort	23. I felt tired
OCD	28. My thoughts would not leave me alone	23. I felt tired	33. I felt stressed	20. I felt down	27. I worried
depression	28. My thoughts would not leave me alone	20. I felt down	23. I felt tired	27. I worried	33. I felt stressed
mania	28. My thoughts would not leave me alone	29. I was easily distracted	30. I felt restless	34. I felt overburdened	20. I felt down
psychosis	53. I used (soft)drugs today	30. I felt restless	28. My thoughts would not leave me alone	51. I experienced physical discomfort	57. Intensity of unpleasant events
ASD	36. I felt overstimulated	37. It bothered me that things did not go as expected	28. My thoughts would not leave me alone	29. I was easily distracted	40. I was unable to sit still
ADHD	41. I was impulsive	40. I was unable to sit still	50. I experienced physical pain	46. Things did not bother me	39. I felt impatient
substance abuse	53. I used (soft)drugs today	52. I drank alcohol today	22. I could not bring myself to do anything	28. My thoughts would not leave me alone	20. I felt down
ODD	28. My thoughts would not leave me alone	57. Intensity of unpleasant events	29. I was easily distracted	30. I felt restless	23. I felt tired

Table S5.7. Correlation between mean ratings on diary items and mini-SCAN domains assessed during the post assessment interview

	Stress			Anxiety			OCD			Depression			Mania			Psychosis			ASD			ADHD			Substance abuse			ODD		
	r	p		r	p		r	p		r	p		r	p		r	p		r	p		r	p		r	p		r	p	
1. Last night I slept well	-0.21	0.02		-0.27	<0.01		-0.08	0.41		-0.37	<0.01		-0.17	0.06		-0.13	0.14		-0.06	0.55		-0.08	0.36		-0.11	0.21		-0.35	<0.01	
2. I felt energetic	-0.32	<0.01		-0.29	<0.01		-0.14	0.11		-0.49	<0.01		-0.21	0.02		-0.13	0.15		-0.11	0.23		-0.09	0.31		-0.07	0.46		-0.32	<0.01	
3. I was physically active	-0.06	0.50		-0.21	0.02		-0.03	0.77		-0.20	0.03		-0.02	0.80		-0.11	0.25		0.08	0.37		-0.02	0.83		-0.20	0.03		-0.13	0.16	
4. I was happy	-0.28	<0.01		-0.25	0.01		-0.19	0.03		-0.50	<0.01		-0.21	0.02		-0.11	0.23		-0.04	0.68		0.02	0.81		-0.17	0.06		-0.33	<0.01	
5. I could enjoy things	-0.24	0.01		-0.23	0.01		-0.17	0.06		-0.42	<0.01		-0.14	0.11		-0.14	0.11		-0.07	0.48		<0.01	0.98		-0.14	0.13		-0.33	<0.01	
6. I felt good	-0.30	<0.01		-0.26	<0.01		-0.14	0.13		-0.48	<0.01		-0.21	0.02		-0.10	0.30		-0.06	0.54		0.01	0.89		-0.12	0.20		-0.34	<0.01	
7. I felt relaxed	-0.34	<0.01		-0.32	<0.01		-0.19	0.04		-0.44	<0.01		-0.24	0.01		-0.13	0.14		-0.13	0.14		-0.04	0.69		-0.13	0.16		-0.41	<0.01	
8. My day was worth living	-0.22	0.02		-0.26	<0.01		-0.07	0.43		-0.43	<0.01		-0.14	0.13		-0.10	0.29		-0.02	0.85		0.01	0.93		-0.18	0.05		-0.30	<0.01	
9. I could handle the things that confronted me	-0.25	0.01		-0.20	0.03		-0.14	0.12		-0.39	<0.01		-0.15	0.11		-0.06	0.51		-0.16	0.08		-0.10	0.29		-0.15	0.09		-0.33	<0.01	
10. Everything came easy	-0.32	<0.01		-0.23	0.01		-0.16	0.08		-0.42	<0.01		-0.24	0.01		-0.07	0.42		-0.12	0.20		0.01	0.90		-0.11	0.25		-0.34	<0.01	
11. I felt like undertaking things	-0.27	<0.01		-0.28	<0.01		-0.11	0.23		-0.44	<0.01		-0.21	0.02		-0.21	0.02		0.01	0.94		-0.10	0.28		-0.20	0.03		-0.32	<0.01	
12. I am looking forward to tomorrow	-0.31	<0.01		-0.34	<0.01		-0.09	0.32		-0.46	<0.01		-0.17	0.06		-0.17	0.05		-0.07	0.47		-0.03	0.76		-0.18	0.04		-0.35	<0.01	
13. My concentration was good	-0.36	<0.01		-0.32	<0.01		-0.11	0.22		-0.41	<0.01		-0.29	<0.01		-0.14	0.11		-0.15	0.10		-0.16	0.08		-0.16	0.08		-0.40	<0.01	
14. I felt appreciated	-0.22	0.01		-0.25	0.01		-0.12	0.19		-0.42	<0.01		-0.21	0.02		-0.16	0.07		-0.06	0.49		0.03	0.72		-0.20	0.03		-0.32	<0.01	
15. I felt like others liked me	-0.16	0.07		-0.16	0.08		-0.12	0.20		-0.35	<0.01		-0.14	0.13		-0.10	0.25		-0.06	0.48		0.03	0.70		-0.17	0.06		-0.26	<0.01	
16. I felt at ease with others	-0.20	0.03		-0.18	0.04		-0.16	0.08		-0.40	<0.01		-0.17	0.06		-0.12	0.17		-0.11	0.22		-0.05	0.55		-0.18	0.05		-0.30	<0.01	
17. I was social	-0.14	0.11		-0.14	0.12		-0.16	0.07		-0.39	<0.01		-0.13	0.15		-0.14	0.13		-0.09	0.31		0.02	0.79		-0.15	0.11		-0.25	0.01	
18. I felt lonely	0.24	0.01		0.16	0.08		0.18	0.05		0.35	<0.01		0.25	0.01		0.07	0.41		0.07	0.44		-0.03	0.74		0.10	0.26		0.24	0.01	
19. I felt empty	0.26	<0.01		0.15	0.10		0.21	0.02		0.32	<0.01		0.26	<0.01		0.10	0.30		0.06	0.53		-0.02	0.81		0.07	0.43		0.24	0.01	

	Stress		Anxiety		OCD		Depression		Mania		Psychosis		ASD		ADHD		Substance abuse		ODD	
	r	p	r	p	r	p	r	p	r	p	r	p	r	p	r	p	r	p	r	p
20. I felt down	0.30	<.01	0.23	0.01	0.25	<.01	0.42	<.01	0.28	<.01	0.09	0.33	0.08	0.40	-0.03	0.78	0.11	0.24	0.28	<.01
21. I felt guilty	0.18	0.05	0.14	0.14	0.10	0.26	0.24	0.01	0.18	0.05	0.01	0.92	0.08	0.37	-0.04	0.64	<.01	0.97	0.19	0.04
22. I could not bring myself to do anything	0.29	<.01	0.24	0.01	0.21	0.02	0.35	<.01	0.28	<.01	0.12	0.20	0.16	0.09	0.10	0.29	0.16	0.08	0.29	<.01
23. I felt tired	0.45	<.01	0.31	<.01	0.27	<.01	0.42	<.01	0.23	0.01	0.08	0.36	0.20	0.03	0.02	0.86	-0.09	0.35	0.35	<.01
24. I avoided things	0.18	0.04	0.19	0.03	0.11	0.23	0.27	<.01	0.19	0.04	0.12	0.18	0.25	0.01	0.14	0.13	0.04	0.64	0.30	<.01
25. I was afraid of making mistakes	0.32	<.01	0.25	0.01	0.17	0.06	0.32	<.01	0.23	0.01	0.01	0.89	0.16	0.08	-0.02	0.79	-0.01	0.87	0.27	<.01
26. I had difficulty making decisions	0.17	0.06	0.14	0.12	0.13	0.15	0.18	0.05	0.18	0.05	0.08	0.41	0.23	0.01	0.08	0.39	0.06	0.51	0.25	<.01
27. I worried	0.40	<.01	0.33	<.01	0.25	0.01	0.40	<.01	0.25	<.01	0.08	0.40	0.12	0.20	-0.04	0.64	0.10	0.28	0.31	<.01
28. My thoughts would not leave me alone	0.51	<.01	0.35	<.01	0.35	<.01	0.46	<.01	0.37	<.01	0.16	0.07	0.26	<.01	0.05	0.60	0.12	0.19	0.42	<.01
29. I was easily distracted	0.37	<.01	0.24	0.01	0.20	0.03	0.37	<.01	0.35	<.01	0.10	0.28	0.26	<.01	0.10	0.26	0.06	0.49	0.35	<.01
30. I felt restless	0.38	<.01	0.26	<.01	0.23	0.01	0.37	<.01	0.33	<.01	0.18	0.05	0.23	0.01	0.05	0.62	0.07	0.42	0.35	<.01
31. I felt nervous	0.22	0.01	0.17	0.06	0.14	0.13	0.28	<.01	0.18	0.05	0.05	0.59	0.10	0.26	-0.06	0.52	0.03	0.75	0.17	0.06
32. I was easily startled	0.06	0.49	0.09	0.32	0.05	0.57	0.12	0.20	0.09	0.33	-0.06	0.50	0.05	0.57	-0.10	0.30	0.03	0.76	0.09	0.33
33. I felt stressed	0.40	<.01	0.28	<.01	0.26	<.01	0.38	<.01	0.28	<.01	0.07	0.44	0.20	0.03	<.01	0.96	0.04	0.66	0.33	<.01
34. I felt overburdened	0.37	<.01	0.20	0.02	0.20	0.03	0.31	<.01	0.28	<.01	0.01	0.94	0.20	0.03	0.10	0.30	-0.03	0.78	0.29	<.01
35. I was easily upset	0.23	0.01	0.16	0.08	0.14	0.11	0.25	0.01	0.22	0.01	-0.01	0.93	0.13	0.17	<.01	1.00	0.04	0.63	0.24	0.01
36. I felt overstimulated	0.31	<.01	0.15	0.10	0.19	0.04	0.23	0.01	0.28	<.01	0.03	0.75	0.30	<.01	0.15	0.10	0.01	0.89	0.26	<.01
37. It bothered me that things did not go as expected	0.31	<.01	0.25	0.01	0.17	0.06	0.28	<.01	0.21	0.02	0.09	0.30	0.30	<.01	0.09	0.33	<.01	0.98	0.34	<.01
38. I was bored quickly	0.05	0.58	0.08	0.38	0.06	0.53	0.14	0.12	0.10	0.27	0.10	0.26	0.01	0.95	<.01	0.98	0.10	0.28	0.12	0.20
39. I felt impatient	0.20	0.03	0.19	0.04	0.08	0.35	0.22	0.02	0.26	<.01	0.09	0.34	0.20	0.03	0.18	0.05	0.03	0.78	0.31	<.01
40. I was unable to sit still	0.14	0.12	0.05	0.56	0.04	0.62	0.14	0.11	0.25	<.01	0.06	0.51	0.26	<.01	0.30	<.01	0.03	0.71	0.22	0.01
41. I was impulsive	0.03	0.73	0.04	0.67	-0.01	0.93	0.08	0.37	0.18	0.05	0.09	0.35	0.20	0.03	0.30	<.01	0.06	0.52	0.20	0.03

	Stress		Anxiety		OCD		Depression		Mania		Psychosis		ASD		ADHD		Substance abuse		ODD	
	r	p	r	p	r	p	r	p	r	p	r	p	r	p	r	p	r	p	r	p
42. I felt irritated	0.26	<.01	0.20	0.03	0.05	0.59	0.22	0.01	0.18	0.05	0.02	0.82	0.18	0.05	0.09	0.34	0.02	0.84	0.34	<.01
43. Others felt annoyed by me	0.09	0.34	0.08	0.41	0.05	0.57	0.06	0.53	0.12	0.20	-0.05	0.60	0.13	0.15	0.14	0.12	0.04	0.66	0.19	0.04
44. I had a fight	-0.11	0.22	0.06	0.52	-0.05	0.61	-0.10	0.27	-0.10	0.26	-0.11	0.24	-0.02	0.83	0.06	0.48	0.07	0.43	0.06	0.51
45. I felt rebellious	0.10	0.25	0.07	0.47	0.02	0.84	0.08	0.37	0.18	0.05	-0.04	0.64	0.15	0.10	0.13	0.14	0.02	0.79	0.20	0.03
46. Things did not bother me	0.07	0.48	0.09	0.35	0.06	0.51	0.17	0.05	0.14	0.13	0.01	0.87	0.19	0.03	0.21	0.02	-0.02	0.85	0.18	0.05
47. I had moodswings	0.23	0.01	0.20	0.03	0.17	0.06	0.25	0.01	0.25	0.01	<.01	0.99	0.17	0.07	0.04	0.69	0.05	0.60	0.29	<.01
48. I felt strange	0.20	0.03	0.12	0.20	0.14	0.13	0.22	0.02	0.21	0.02	0.06	0.53	0.11	0.23	0.08	0.41	0.05	0.55	0.22	0.02
49. I felt suspicious	0.18	0.04	0.15	0.11	0.11	0.23	0.21	0.02	0.14	0.14	0.01	0.90	0.05	0.56	-0.04	0.66	0.05	0.62	0.21	0.02
50. I experienced physical pain	0.28	<.01	0.29	<.01	0.06	0.55	0.25	0.01	0.10	0.27	0.14	0.13	0.19	0.03	0.21	0.02	-0.01	0.92	0.22	0.01
51. I experienced physical discomfort	0.22	0.01	0.32	<.01	0.09	0.30	0.31	<.01	0.12	0.19	0.15	0.10	0.08	0.40	0.09	0.30	0.01	0.90	0.26	<.01
52. I drank alcohol today	-0.13	0.16	-0.01	0.90	0.03	0.76	-0.20	0.03	-0.02	0.82	0.01	0.94	-0.12	0.18	-0.01	0.93	0.49	<.01	-0.10	0.25
53. I used (soft)drugs today	0.05	0.58	0.15	0.11	0.24	0.01	0.15	0.11	0.11	0.24	0.25	0.01	-0.04	0.64	0.03	0.73	0.50	<.01	0.10	0.26
54. Number of pleasant events	-0.21	0.02	-0.23	0.01	-0.20	0.03	-0.23	0.01	-0.05	0.56	-0.11	0.22	0.05	0.55	-0.10	0.27	-0.06	0.52	-0.22	0.02
55. Intensity of pleasant events	0.03	0.73	<.01	1.00	-0.09	0.30	-0.10	0.29	0.10	0.26	0.02	0.81	0.13	0.14	0.06	0.52	-0.07	0.45	<.01	0.99
56. Number of unpleasant events	0.08	0.39	0.08	0.41	-0.02	0.83	0.17	0.06	0.05	0.57	-0.05	0.56	0.18	0.05	0.07	0.44	-0.06	0.52	0.18	0.04
57. Intensity of unpleasant events	0.33	<.01	0.33	<.01	0.12	0.18	0.35	<.01	0.23	0.01	0.14	0.11	0.21	0.02	0.11	0.21	-0.04	0.66	0.40	<.01
58. I'm dreading something that will happen soon	0.31	<.01	0.20	0.03	0.11	0.23	0.33	<.01	0.21	0.02	0.08	0.41	0.19	0.03	0.01	0.93	0.06	0.48	0.22	0.01

References

1. Nelson B, McGorry PD, Wichers M, Wigman JTW, Hartmann JA. Moving from static to dynamic models of the onset of mental disorder. *JAMA Psychiatry*. 2017;74(5):528-534. doi:10.1001/jamapsychiatry.2017.0001
2. Wichers M, Smit AC, Snippe E. Early warning signals based on momentary affect dynamics can expose nearby transitions in depression: a confirmatory single-subject time-series study. *Journal for Person-Oriented Research*. 2020;6(1):1-15. doi:10.17505/jpor.2020.22042
3. Wichers M, Groot PC, Psychosystems ESM group, EWS group. Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*. 2016;85:114-116. doi:10.1159/000441458
4. Os J van, Delespaul P, Wigman JTW, Myin-Germeys I, Wichers M. Psychiatry beyond labels: introducing contextual precision diagnosis across stages of psychopathology. *Psychological Medicine*. 2013;43(07):1563-1567. doi:10.1017/s0033291713000937
5. Hamaker EL, Wichers M. No time like the present: discovering the hidden dynamics in intensive longitudinal data. *Current Directions in Psychological Science*. 2017;26(1):10-15. doi:10.1177/0963721416666518
6. Fried EI, Nesse RM. Depression is not a consistent syndrome: an investigation of unique symptom patterns in the STAR*D study. *Journal of Affective Disorders*. 2015;172:96-102. doi:10.1016/j.jad.2014.10.010
7. McGorry PD, Hickie IB, Yung AR, Pantelis C, Jackson HJ. Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. *Australian and New Zealand Journal of Psychiatry*. 2006;40(8):616-622. doi:10.1111/j.1440-1614.2006.01860.x
8. McGorry P, van Os J. Redeeming diagnosis in psychiatry: timing versus specificity. *The Lancet*. 2013;381(9863):343-345. doi:10.1016/S0140-6736(12)61268-9
9. Fusar-Poli P, Yung AR, McGorry P, van Os J. Lessons learned from the psychosis high-risk state: towards a general staging model of prodromal intervention. *Psychological Medicine*. 2014;44(1):17-24. doi:10.1017/S0033291713000184
10. Hengartner MP, Lehmann SN. Why psychiatric research must abandon traditional diagnostic classification and adopt a fully dimensional scope: two solutions to a persistent problem. *Frontiers in Psychiatry*. 2017;8(JUN):7-11. doi:10.3389/fpsy.2017.00101
11. Kendell R, Jablensky A. Distinguishing between the validity and utility of psychiatric diagnoses. *American Journal of Psychiatry*. 2003;160:4-12.
12. Gil KM, Carson JW, Porter LS, et al. Daily stress and mood and their association with pain, health-care use, and school activity in adolescents with sickle cell disease. *Journal of Pediatric Psychology*. 2003;28(5):363-373. doi:10.1093/jpepsy/jsg026
13. Bijl RV, Ravelli A, van Zessen G. Prevalence of psychiatric disorder in the general population: results of the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social Psychiatry and Psychiatric Epidemiology*. 1998;33(12):587-595. doi:10.1007/s001270050098
14. Alonso J, Angermeyer MC, S. Bernert RB, et al. Disability and quality of life impact of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatrica Scandinavica*. 2004;109(420):21-27. doi:10.1111/j.1600-0047.2004.00329.x
15. Hartman CA, Hermanns VW, de Jong PJ, Ormel J. Self- or parent report of (co-occurring) internalizing and externalizing problems, and basal or reactivity measures of HPA-axis functioning: a systematic evaluation

- of the internalizing-hyperresponsivity versus externalizing-hyporesponsivity HPA-axis hypothesis. *Biological Psychology*. 2013;94(1):175-184. doi:10.1016/j.biopsycho.2013.05.009
16. Huisman M, Oldehinkel AJ, de Winter A, et al. Cohort Profile: The Dutch "TRacking Adolescents" Individual Lives' Survey'; TRAILS. *International Journal of Epidemiology*. 2008;37(6):1227-1235. doi:10.1093/ije/dym273
 17. Rosmalen JGM, Wenting AMG, Roest AM, de Jonge P, Bos EH. Revealing causal heterogeneity using time series analysis of ambulatory assessments. *Psychosomatic Medicine*. 2012;74(4):377-386. doi:10.1097/psy.0b013e3182545d47
 18. van Roekel E, Vrijen C, Heininga VE, Masselink M, Bos EH, Oldehinkel AJ. An exploratory randomized Controlled trial of personalized lifestyle advice and tandem skydives as a means to reduce anhedonia. *Behavior Therapy*. 2017;48(1):76-96. doi:10.1016/j.beth.2016.09.009
 19. Kramer I, Simons CJP, Hartmann JA, et al. A therapeutic application of the experience sampling method in the treatment of depression: a randomized controlled trial. *World Psychiatry*. 2014;13(1):68-77. doi:10.1002/wps.20090
 20. Krieke L van der, Jeronimus BF, Blaauw FJ, et al. HowNutsAreTheDutch (HoeGekIsNL): a crowdsourcing study of mental symptoms and strengths. *International Journal of Methods in Psychiatric Research*. 2016;25(2):123-144. doi:10.1002/mpr
 21. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*. 1988;54(6):1063-1070.
 22. Achenbach TM. *Manual for the Young Adult Self-Report and Young Adult Behavior Checklist.*; 1997.
 23. Derogatis LR. *SCL-90-R Symptom Checklist-90-R Administration, Scoring, and Procedures Manual*. PsychCorp; 1977.
 24. Nienhuis FJ, van de Willige G, Rijnders CAT, de Jonge P, Wiersma D. Validity of a short clinical interview for psychiatric diagnosis: the mini-SCAN. *British Journal of Psychiatry*. 2010;196(1):64-68. doi:10.1192/bjp.bp.109.066563
 25. Brugha TS, Nienhuis F, Bagchi D, Smith J, Meltzer H. The survey form of SCAN: the feasibility of using experienced lay survey interviewers to administer a semi-structured systematic clinical assessment of psychotic and non-psychotic disorders. *Psychological Medicine*. 1999;29(3):703-711. doi:10.1017/S003329179900834X
 26. Hofstra MB, van der Ende J, Verhulst FC. Continuity and change of psychopathology from childhood into adulthood: a 14-year follow-up study. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2000;39(7):850-858. doi:10.1097/00004583-200007000-00013
 27. de Jong A, van der Lubbe PM. *Groningse Vragenlijst Sociaal Gedrag: Zelfbeoordelvingsvragenlijsten Voor Het Vaststellen van Problemen in Het Interpersoonlijke Functioneren (Handleiding)*. Rob Giel Onderzoekscentrum/Rijksuniversiteit Groningen, disciplinegroep Psychiatrie; 2001.
 28. Brugha T, Bebbington P, Tennant C, Hurry J. The List of Threatening Experiences: a subset of 12 life event categories with considerable long-term contextual threat. *Psychological Medicine*. 1985;115:189-194.
 29. Motrico E, Moreno-Küstner B, De Dios Luna J, et al. Psychometric properties of the List of Threatening Experiences-LTE and its association with psychosocial factors and mental disorders according to different scoring methods. *Journal of Affective Disorders*. 2013;150:931-940. doi:10.1016/j.jad.2013.05.017
 30. R Core Team. R: A language and environment for statistical computing. Published online 2020.
 31. Shankar V, Bangdiwala SI. Observer agreement paradoxes in 2x2 tables: comparison of agreement

- measures. *BMC Medical Research Methodology*. 2014;14(1):1-9. doi:10.1186/1471-2288-14-100
32. Feinstein AR, Cicchetti D v. High agreement but low Kappa: II. Resolving the paradoxes. *Journal of Clinical Epidemiology*. 1990;43(6):551-558. doi:10.1016/0895-4356(90)90158-L
33. Silverstein AB. Validity of WISC-R short forms. *Journal of Clinical Psychology*. 1975;31:696-697.
34. Dworkin ER, Ullman SE, Stappenbeck C, Brill CD, Kaysen D. Proximal relationships between social support and PTSD symptom severity: a daily diary study of sexual assault survivors. *Depression and Anxiety*. 2019;35:43-49. doi:10.1002/da.22679.Proximal
35. Weiss NH, Bold KW, Contractor AA, Sullivan TP, Armeli S, Tennen H. Trauma exposure and heavy drinking and drug use among college students: identifying the roles of negative and positive affect lability in a daily diary study. *Addictive Behaviors*. 2018;79:131-137. doi:10.1016/j.addbeh.2017.12.015
36. Bouwmans MEJ, Bos EH, Hoenders HJR, Oldehinkel AJ, de Jonge P. Sleep quality predicts positive and negative affect but not vice versa. An electronic diary study in depressed and healthy individuals. *Journal of Affective Disorders*. 2017;207:260-267. doi:10.1016/j.jad.2016.09.046
37. Czyz EK, King CA, Nahum-Shani I. Ecological assessment of daily suicidal thoughts and attempts among suicidal teens after psychiatric hospitalization: Lessons about feasibility and acceptability. *Psychiatry Research*. 2018;267:566-574. doi:10.1016/j.psychres.2018.06.031
38. Telford C, McCarthy-Jones S, Corcoran R, Rowse G. Experience sampling methodology studies of depression: the state of the art. *Psychological Medicine*. 2012;42(6):1119-1129. doi:10.1017/S0033291711002200
39. Myin-Germeyns I, Kasanova Z, Vaessen T, et al. Experience sampling methodology in mental health research: new insights and technical developments. *World Psychiatry*. 2018;17:123-132.
40. Trull TJ, Ebner-Priemer U. Ambulatory assessment. *Annual Review of Clinical Psychology*. 2013;9:151-176. doi:10.1146/annurev-clinpsy-050212-185510
41. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annual Review of Clinical Psychology*. 2013;4:541-565. doi:10.1002/9781118384404.ch20
42. Hufford MR, Shiffman S. Assessment methods for patient-reported outcomes. *Disease Management & Health Outcomes*. 2003;11(2):77-86.
43. Palmier-Claus JE, Myin-Germeyns I, Barkus E, et al. Experience sampling research in individuals with mental illness: reflections and guidance. *Acta Psychiatrica Scandinavica*. 2011;123(1):12-20. doi:10.1111/j.1600-0447.2010.01596.x
44. van de Schoot R, Schmidt P, de Beuckelaer A, Lek K, Zondervan-Zwijnenburg M. Measurement invariance. *Frontiers in Psychology*. 2015;6:10-13. doi:10.3389/fpsyg.2015.01064
45. Fried EI, Borkulo CD van, Epskamp S, Schoevers RA, Tuerlinckx F, Borsboom D. Measuring depression over time... or not? Lack of unidimensionality and longitudinal measurement invariance in four common rating scales of depression. *Psychological Assessment*. 2016;28(11):1-14. doi:10.1037/pas0000275
46. McHugh RK, Behar E. Readability of self-report measures of depression and anxiety. *Journal of Consulting and Clinical Psychology*. 2009;77(6):1100-1112. doi:10.1037/a0017124
47. Wichers M. The dynamic nature of depression: a new micro-level perspective of mental disorder that meets current challenges. *Psychological Medicine*. 2014;44(7):1349-1360. doi:10.1017/S0033291713001979
48. Scollon CN, Kim-Prieto C, Scollon CN. Experience sampling: promises and pitfalls, strengths and weaknesses. *Journal of Happiness Studies*. 2003;4(1):5-34. doi:10.1023/a:1023605205115
49. Cramer AOJ, Waldorp LJ, van der Maas HLJ, Borsboom D. Comorbidity: a network perspective. *Behavioral and Brain Sciences*. 2010;33(2010):137-193. doi:10.1017/S0140525X09991567

50. Borsboom D, Cramer AOJ. Network analysis: an integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*. 2013;9(1):91-121. doi:10.1146/annurev-clinpsy-050212-185608
51. Wichers M, Schreuder MJ, Goekoop R, Groen RN. Can we predict the direction of sudden shifts in symptoms? Transdiagnostic implications from a complex systems perspective on psychopathology. *Psychological Medicine*. 2019;49(3):380-387. doi:10.1017/S0033291718002064
52. Cramer AOJ, van Borkulo CD, Giltay EJ, et al. Major depression as a complex dynamic system. *PLoS one*. 2016;11(12):1-20. doi:10.1371/journal.pone.0167490
53. Scheffer M, Bascompte J, Brock WA, et al. Early-warning signals for critical transitions. *Nature*. 2009;461(7260):53. doi:10.1038/nature08227
54. van de Leemput IA, Wichers M, Cramer AOJ, et al. Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences*. 2014;111(1):87-92. doi:10.1073/pnas.1312114110
55. Fisher AJ, Medaglia JD, Jeronimus BF. Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*. 2018;115(27):E6106-E6115. doi:10.1073/pnas.1711978115
56. Gates KM, Molenaar PCM. Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage*. 2012;63(1):310-319. doi:10.1016/j.neuroimage.2012.06.026

6

**For whom do intensive
longitudinal studies work?
Participation and compliance in
a six-month daily diary study
among individuals at risk for
mental health problems**

Abstract

Background. Intensive longitudinal (IL) measurement, which involves prolonged self-monitoring, may have important clinical applications but is also burdening. This raises the question who takes part in, and successfully completes IL measurements. This study investigated which demographic, personality, economic, social, psychological, or physical participant characteristics are associated with (i) participation and (ii) compliance in an IL study conducted in young adults at enhanced risk for psychopathology.

Methods. Young adults enrolled in the clinical cohort of the Tracking Adolescents' Individual Lives Survey (TRAILS) were invited to a six-month daily diary study. Participant characteristics came from five earlier TRAILS assessment waves collected from age 11 onwards. To evaluate participation, we compared diary study participants ($N=134$) to (i) non-participants ($N=309$) and (ii) a sex-matched subsample ($N=1926$) of individuals from the general population cohort of TRAILS. To evaluate compliance, we analyzed which characteristics were related to the proportion of completed diary entries.

Results. Participants (mean age 23.6 years, $SD = 0.67$) were largely similar to non-participants. In addition, compared to the general population, participants reported less advantageous scores on nearly all characteristics. Internalizing problems predicted higher compliance. Externalizing problems, antisocial behavior and daily smoking predicted lower compliance.

Conclusion. In young adults at enhanced risk for psychopathology, who were disadvantaged relative to the general population on nearly every characteristic evaluated, participation in a diary study is unbiased. Small biases in compliance occur, of which researchers should be aware. In conclusion, IL measurement is broadly applicable, which is a requirement for its usefulness in clinical practice.

Introduction

Intensive longitudinal designs (ILDs), such as daily diary studies, are increasingly conducted to study the dynamics of psychological processes. Key to these designs is that individuals are repeatedly measured, once a day or multiple times a day, for multiple weeks or months. This results in sufficient observations per person to assess within-person processes of change. Improved understanding of within-person change in psychological processes may offer (in the longer run) important opportunities for individuals in clinical practice^{1–4}. For example, repeatedly monitoring mental states and experiences may shed light on the circumstances that trigger symptoms, which may in turn inform therapeutic interventions. Ultimately, such feedback on psychological dynamics may lower symptom severity in persons with psychiatric problems⁵.

ILDs might be experienced as more intrusive and demanding in comparison to a one-time survey due to the daily character of the assessments. In addition, ILDs request a lengthier commitment of multiple weeks or months. Concerns have been expressed that ILDs, due to their higher burden, attract a certain type of individual: those who are more agreeable, conscientious, highly educated, and with more time on their hands^{6–8}. Such self-selection of a certain type of individual into these study designs⁹ raises the question if studies that employ ILDs are suitable for all patient groups.

Aspects of the ILD that are believed to influence self-selection in participation may likewise determine who, once started, does well in such a study. Good compliance requires, for instance, that individuals remember to participate (answer questions daily) and succeed in filling in the questions even when they do not feel like it^{10,11}. Possibly, individuals scoring higher on certain traits such as being conscientious or well-organized find it less of an effort to comply with the protocol than individuals who score lower on these traits. Again, this might mean that severely affected patient groups are underrepresented in a typical ILD, and hence, we do not know if such designs can be applied to a wide variety of patients in clinical practice.

Although the question for whom ILDs are useful is of high clinical relevance, knowledge on the characteristics related to self-selection and compliance is scarce. Both self-selection and compliance are difficult to evaluate because the required information is often unavailable. In the case of self-selection, we lack information on the individuals who did not respond to the invitation to participate in ILD studies. The risk of self-selection biases

may be particularly high if recruitment takes place through social media. In this situation, the target population is unknown since we do not know whom reached the invitation to begin with. In order to know who participates in ILT studies and to whom findings will generalize, characteristics of responders should be compared to population norms, but this is hardly feasible in practice. In the case of compliance characteristics, we lack information because researchers may limit the number of additional questionnaires on background characteristics (*e.g.*, personality or lifestyle), balancing the participation burden towards the repeated assessments. Hence, a comprehensive profile of who participates and does well in ILTs has so far not been established.

The present study aimed to investigate which demographic, economic, social, psychological, and physical characteristics were associated with participation and compliance in a six-month diary study that was conducted in a sample of Dutch young adults with heightened risk for psychopathology. Participants in the diary study were recruited within the ongoing Tracking Adolescents Individual Lives Survey clinical cohort study (TRAILS CC). The TRAILS CC respondents are at higher risk for psychological and psychiatric problems based on having been in contact with psychiatric services at any time before age 11 when TRAILS CC started. Around age 22, TRAILS CC participants were invited to participate in a diary study. In addition to TRAILS CC, TRAILS includes a representative population sample (TRAILS PC) of Dutch young adults for whom the same instruments were used at the same ages. We compared diary-participants (TRAILS CC participants who enrolled in the diary study) to both (i) diary-non-participants (TRAILS CC participants who did not enroll in the diary study), and to (ii) a general population reference group (TRAILS PC) on a wide range of characteristics that were measured in TRAILS during the five waves prior to the diary study. The first comparison allowed to investigate whether diary participants represented a relatively healthy and well-functioning subset of the at-risk cohort. In combination with analyses of diary participants' compliance, the second comparison allowed to determine whether ILTs also 'work' for individuals who score less favorable (*i.e.*, relative to the general population) on demographic, personality, economic, social, psychological and physical characteristics.

Method

Sample and design

This study is based on data collected as part of the TRacking Adolescents' Individual Lives Survey (TRAILS). TRAILS is an ongoing, prospective cohort study with bi- or triennial follow-up assessments investigating the development of mental health from pre-adolescence into adulthood in both a clinical (TRAILS CC) and a general population cohort (TRAILS PC). Both cohorts have been extensively described elsewhere^{12,13}. Briefly, children born between 1994 and 1995 who had been referred to a child psychiatric outpatient clinic in the Northern Netherlands any time before the age of 11 were eligible to participate in the clinical cohort (TRAILS CC). At 11 years old, participants either met criteria for a current psychiatric disorder (mostly externalizing domains), a lifetime diagnosis without a current disorder (mostly internalizing domains) or did not have a lifetime diagnosis¹⁴. Due to the referral before age 11, individuals participating in this cohort are considered to have heightened risk for psychiatric symptoms (see also Oldehinkel *et al.*¹³, who reported elevated prevalence rates of psychiatric disorders at age 16). Of the eligible individuals, 543 (43%) participated in the first measurement wave (mean age = 11.1 years, $SD = 0.5$, 34% girls) in 2004-2005. The general population cohort of the TRAILS (TRAILS PC) study preceded the clinical cohort, and its baseline assessment (T1) ran from March 2001 through July 2002. Children born between October 1st 1989 and September 30th 1991 in five municipalities in the North of the Netherlands were eligible for inclusion in TRAILS PC. Of the eligible individuals who met inclusion criteria and whose schools agreed to participate, 2230 (76%) participated (mean age = 11.1 years, $SD = 0.6$, 51% girls). Both cohorts include predominantly white individuals residing in both urban and rural areas. Each assessment wave of TRAILS was approved by Dutch Central Committee on Research Involving Human Subjects (CCMO), and adolescents (and at earlier waves also the parents) provided informed consent prior to participating in each wave.

Figure 6.1 shows how many participants completed each assessment wave. After the fifth assessment wave, when young adults were approximately 22 years old, all TRAILS CC participants who had not formally declined further participation in the TRAILS CC study were invited to participate in an add-on study called TRAILS TRANS-ID (Transitions In Depression). This study included a six-month daily diary period and aimed to investigate the day-to-day fluctuations in symptoms in individuals at increased risk for psychopathology. Of the 443

individuals who were invited for TRAILS TRANS-ID, 134 (30.2%) were included. Data collection procedures for TRAILS TRANS-ID are described in detail by Schreuder *et al.* (chapter 5)¹⁵. All participants provided written consent, and ethical approval was granted by the Medical Ethical Committee of the University Medical Center Groningen (2017/203).

The present study used data of the 443 individuals invited to TRAILS TRANS-ID collected during the five assessment waves of TRAILS CC prior to TRAILS TRANS-ID, and data on compliance obtained during the diary period of TRAILS TRANS-ID. Additionally, we used data of 1926 TRAILS PC participants collected during the first five waves of TRAILS PC. To account for the fact that the majority of TRAILS TRANS-ID participants (57%) is male, this sample concerns a sex-matched subset of the total number of TRAILS PC participants (N=2230). Matching was achieved by randomly selecting women from the sample, resulting in a subsample with 57% men to match the TRAILS TRANS-ID sex distribution (as opposed to 51% in the full TRAILS PC cohort).

Measures

We identified variables from five domains that could be associated with participation or compliance in the diary study. Supplementary Table S6.1 (published on the Open Science Framework, <https://osf.io/tnrqf/>) describes these variables, which cover the domains of personality and cognitive functioning, economic characteristics, social characteristics, psychological health, and physical health. All were assessed in TRAILS prior to the start of the TRANS-ID study. If variables were assessed at multiple waves, they were aggregated by calculating either the sum (*i.e.*, lifetime adverse events, chronic social stressors, and financial difficulties) or mean (*e.g.*, social problems and family functioning) across waves. The instruments or questions used to assess each characteristic and the timing at which these were assessed are summarized in Table S6.1.

Fig. 6.1. Flow-chart of the number of participants assessed per wave. CC and PC refer to the clinical and population cohort of TRAILS, respectively. The diary study central to TRAILS TRANS-ID is an add-on study in TRAILS CC. Participants who had not declined future participation in TRAILS CC were eligible for participation in the diary study. Of the 443 eligible participants, 132 (29.8%) had not participated in the fifth assessment wave of TRAILS. In sensitivity analyses, we excluded these participants. Numbers behind rows that denote the TRAILS assessment waves (T1-T5) refer to the percentage of the initial sample at T1 (TRAILS CC) or the sex-matched subsample (TRAILS PC) and mean age ($\pm$ standard deviation). We selected this sex-matched subsample in order to obtain equal sex distributions between TRAILS TRANS-ID participants ($N=134$) and individuals from the general population.

Analyses

The analyses of this study were preregistered on the Open Science Framework (<https://osf.io/ynu6b>). We complemented these analyses with an attrition analysis in which TRAILS CC participants who had not completed the fifth (most recent) assessment wave were excluded. Analyses were conducted in R¹⁶ (version 4.0.2).

To evaluate characteristics related to participation in a diary study and to compare them with reference values from the population cohort (*i.e.*, do diary participants from the at-risk group have lower scores on these characteristics), we conducted a series of regression analyses. In these analyses, individual characteristics were predicted by participant group (diary-participant vs. diary non-participant vs. general population). For continuously measured characteristics (including ordinal/interval variables with >5 categories), we used linear regression and for categorical variables, (multinomial) logistic regression. Continuous variables were standardized prior to analysis by subtracting the mean and dividing by the standard deviation. This yielded standardized regression coefficients, which facilitated comparisons of group differences across characteristics. For categorical variables, odds ratios were indicative of effect sizes. We applied a Holm-Bonferroni correction¹⁷ to maintain a family wise error rate of 0.05 within each global characteristic (*i.e.*, personality and cognitive functioning, economic, social, psychological health, and physical health). Finally, in order to ensure that potential group differences were not due to a general disinterest in continued participation in TRAILS rather than the daily diary study, we ran an attrition analysis in which we excluded individuals who had not participated in the fifth assessment wave (T5).

To evaluate whether personality and cognitive functioning, economic, social, physical health and psychological health characteristics were associated with compliance in a diary study, we conducted a series of univariate linear regressions. Here, the outcome variable compliance was expressed as the percentage completed daily diaries (*i.e.*, number of completed daily diary entries divided by the number of diaries sent to that participant). Compliance and continuous predictors were standardized by subtracting the mean and dividing by the standard deviation. A Holm-Bonferroni correction¹⁷ was applied to adjust for multiple testing.

Results

Sample description

Table 6.1 shows the demographic characteristics of the research sample.

Table 6.1. *Sample characteristics*

	Participants (N=134)	Non-participants (N=309)	General population (N=1926)
	<i>Mean (SD)/ N(%)</i>	<i>Mean (SD)/ N(%)</i>	<i>Mean (SD)/ N(%)</i>
Age in years at start TRAILS TRANS-ID	23.62 (0.67)	23.53 (0.71)	27.2 (0.55)*
Sex (% males)	76 (56.7)	212 (68.6)	1098 (57.0)
Non-Dutch ethnicity parents	0 (0)	7 (2.3)	264 (13.7)

*General population cohort of TRAILS started approximately 4 years before the clinical cohort of TRAILS (from which diary participants were sampled).

Relevance attrition analysis

Of the 134 individuals who participated in the diary study (referred to as participants), 9 (6.7%) had not participated in T5. Of the 309 individuals who declined participation in the diary study (referred to as non-participants in the remaining text), 123 (39.8%) had also not participated in T5. This difference indicates that non-participants are partly characterized by a more general disinterest in continued participation in TRAILS than specifically in the diary study as such, showing the relevance of the attrition analysis for drawing correct conclusions. Attrition was not specific to TRAILS CC (30% overall), as a similar proportion of participants in the general population cohort (TRAILS PC) had not participated in T5 (33% overall). Results of the attrition analyses, where we compared groups of individuals who completed T5 (*i.e.*, participants who completed T5 vs. non-participants who completed T5 and vs. TRAILS PC participants who completed T5), can be found in the supplementary Table S6.3. All discrepancies between the main and the attrition analyses are noted in the main text.

Diary study participants vs. non-participants

Women were more likely to participate in the diary study than men (43.3 vs. 31.4% female, $\chi^2(1) = 5.30$, $p = .02$). Figure 6.2 shows the comparison between participants and non-participants on the other characteristics (see Table 6.2 for a detailed representation) Participants had a higher IQ than non-participants ($M_p = 100.3$, $SD = 14.9$, $M_{NP} = 96.0$, $SD =$

16.0; $\beta = -.28$, $p_{cor} = .04$). Additionally, there was a moderate difference in number of reported adverse events ($M_p = 10.4$, $SD = 5.9$, $M_{NP} = 7.7$, $SD = 5.2$; $\beta = -.52$, $p_{cor} < .001$). However, after excluding individuals who had not participated in T5, both differences became smaller and non-significant (supplement, Table S6.3). This suggests that individuals with lower IQ-scores and individuals experiencing less negative life events may be less likely to continue participation in research in general and not necessarily be disinterested in a diary study per se.

Generally, effect sizes of other characteristics were small (average $\beta = 0.14$; average OR = 1.33). Notable exceptions were: moderate differences on childhood trauma ($\beta = -.25$), chronic social stressors ($\beta = -.22$) and several psychological health characteristics (*e.g.*, anti-social behavior, ADHD and externalizing problems; β 's between 0.20 and 0.25). Additionally, participants had 1.5 higher odds to report drug use, had 2.2 higher odds to have financial debts, and had 1.75 higher odds to be currently enrolled in education than non-participants. All these effects became smaller in the attrition analysis. This suggests that if these effects would have been significant in a larger sample, these differences would not be specific to the decision to participate in the diary study but related to study attrition in general. Hence, none of the assessed characteristics seemed specifically associated with participation in the diary study itself.

Diary study participants vs. general population

A reason for the absence of differences between participants and non-participants could be that the invited individuals, despite being sampled from an at-risk cohort, had non-deviating scores on the characteristics that were hypothesized to be related to participation in diary studies (*e.g.*, self-discipline, enrollment in education, etc.). To evaluate this, we compared the diary participants with the general population (see Table 6.2 and Figure 6.2).

Fig. 6.2. Differences between participants of an intensive diary study (TRAILS TRANS-ID) and (i) individuals who were invited, but declined to participate (upper plot), and (ii) individuals from the general population (lower plot). For readability, variables were merged if they were assessed in multiple informants (e.g., financial difficulties, social problems, physical health, and internalizing and externalizing problems) or if they were conceptually similar (e.g., cannabis, illicit drug use). If any of the lower-level variables was significantly different between groups, the averaged effect size is accompanied with a sign. Asterisks mark effects that were statistically significant after correcting for multiple testing (Holm's procedure); plus-signs mark effects that were statistically significant in sensitivity analyses. The illustrated data is also provided in a table (see Table 6.2 and supplement, Table S6.2).

With the exception of social economic status characteristics, which were not different between diary participants and the general population, diary participants scored higher on less favorable characteristics (*e.g.*, psychological ill-health or social problems) and scored lower on favorable characteristics (*e.g.*, family functioning or self-discipline). Effects sizes of the differences between diary participants and the sex- and age-matched general population sample were generally moderate (average $\beta = 0.39$; average OR = 2.4). The majority of significant differences found in main analysis remained in attrition analysis (except for differences in anti-social behavior and BMI; see supplement, Table S6.3). Diary participants thus consistently scored less favorably than the general population on characteristics that were hypothesized to influence who signs up for an ILD.

Table 6.2. *Characteristic differences between participants and non-participants, and between participants and a sex-matched general population sample*

	Participants vs. Non-participants					Participants vs. General population				
	B/OR	Beta/OR-1 [95% CI]	SE	<i>p</i>		B/OR	Beta/OR-1 [95% CI]	SE	<i>p</i>	
<i>Personality and cognitive functioning</i>										
Assertiveness	0.55	0.12 [-0.10, 0.33]	0.50	0.28		0.95	0.21 [0.02, 0.39]	0.43	0.03	
Extraversion	-0.11	-0.03 [-0.24, 0.19]	0.47	0.81		0.67	0.16 [-0.03, 0.34]	0.40	0.09	
Self-discipline	0.56	0.12 [-0.09, 0.33]	0.51	0.27		1.62	0.34 [0.16, 0.53]	0.44	<0.001	
Angry-hostility	-0.30	-0.07 [-0.28, 0.14]	0.44	0.49		-1.11	-0.27 [-0.45, -0.09]	0.37	0.003	
Impulsiveness	-0.06	-0.02 [-0.23, 0.19]	0.37	0.87		-0.52	-0.15 [-0.33, 0.03]	0.32	0.10	
Vulnerability	-0.54	-0.13 [-0.34, 0.07]	0.43	0.21		-1.64	-0.40 [-0.58, -0.23]	0.37	<0.001	
IQ	-4.32	-0.28 [-0.49, -0.08]	1.57	0.006		-3.09	-0.20 [-0.38, -0.03]	1.36	0.02	
<i>Economic characteristics</i>										
Financial difficulties (self)	-0.17	-0.22 [-0.43, -0.01]	0.08	0.04		-0.02	-0.03 [-0.21, 0.15]	0.09	0.73	
Educational attainment	-0.04	-0.04 [-0.27, 0.18]	0.12	0.72		0.11	0.11 [-0.07, 0.29]	0.09	0.24	
Financial difficulties (parent)	0.06	0.08 [-0.13, 0.28]	0.08	0.46		0	0 [-0.18, 0.17]	0.09	1.00	
Employment*	1.19	0.19 [-0.46, 1.62]	0.40	0.67		2.50	1.50 [0.26, 3.91]	0.35	0.01	
Social benefits*	1.21	0.21 [-0.28, 1.04]	0.26	0.47		0.25	-0.75 [-0.84, -0.60]	0.23	<0.001	
Participation*	0.77	-0.23 [-0.60, 0.47]	0.33	0.44		1.88	0.88 [0.05, 2.19]	0.28	0.024	
Current education*	0.57	-0.43 [-0.64, -0.09]	0.24	0.02		1.06	0.06 [-0.28, 0.55]	0.20	0.76	
Financial debt*	0.44	-0.56 [-0.81, -0.01]	0.41	0.05		1.58	0.58 [-0.06, 1.82]	0.28	0.10	
Educational att. mother	-0.07	-0.06 [-0.27, 0.14]	0.11	0.54		-0.18	-0.18 [-0.35, 0]	0.17	0.05	
Educational att. father	-0.17	-0.15 [-0.36, 0.07]	0.12	0.17		-0.07	-0.06 [-0.25, 0.12]	0.19	0.51	
Income*	0.96	-0.04 [-0.41, 0.57]	0.22	0.57		0.59	-0.41 [-0.61, -0.09]	0.22	0.99	
Occupation mother*	1.07	0.07 [-0.33, 0.70]	0.20	0.61		0.80	-0.20 [-0.47, 0.19]	0.20	0.86	
Occupation father*	0.79	-0.21 [-0.49, 0.21]	0.19	0.86		1.03	0.03 [-0.29, 0.49]	0.18	0.56	

	Participants vs. Non-participants				Participants vs. General population			
	B/OR	Beta/OR-1 [95% CI]	SE	p	B/OR	Beta/OR-1 [95% CI]	SE	p
<i>Social characteristics</i>								
Social problems (self)	-0.01	-0.05 [-0.25, 0.16]	0.03	0.65	-0.09	-0.34 [-0.51,-0.16]	0.02	<0.001
Social problems (parent)	-0.02	-0.06 [-0.25, 0.14]	0.03	0.57	-0.30	-0.89 [-1.06,-0.73]	0.03	<0.001
Family functioning	0.05	0.16 [-0.04, 0.36]	0.03	0.12	-0.09	-0.27 [-0.44,-0.10]	0.03	0.002
Adverse events	-2.76	-0.52 [-0.72,-0.31]	0.55	<0.001	-2.88	-0.54 [-0.71,-0.37]	0.47	<0.001
Chronic stress	-0.48	-0.22 [-0.41,-0.03]	0.21	0.02	-2.12	-0.99 [-1.15,-0.82]	0.18	<0.001
Trauma	-0.55	-0.25 [-0.45,-0.05]	0.23	0.02	-0.73	-0.33 [-0.51,-0.16]	0.20	<0.001
<i>Physical health characteristics</i>								
Physical health (self)	-0.01	-0.10 [-0.30, 0.10]	0.06	0.89	0.09	0.15 [-0.02, 0.33]	0.05	0.09
Physical health (parent)	-0.05	-0.23 [-0.43,-0.03]	0.06	0.35	0.21	0.39 [0.21, 0.56]	0.05	<0.001
BMI	-0.78	0.18 [-0.02, 0.39]	0.35	0.03	-1.08	-0.32 [-0.50,-0.15]	0.30	<0.001
Alcohol	0.92	-0.07 [-0.28, 0.13]	0.52	0.08	2.28	0.46 [0.28, 0.64]	0.44	<0.001
Absence (phys.)	-0.09	-0.07 [-0.28, 0.13]	0.13	0.47	-0.32	-0.27 [-0.44,-0.09]	0.11	0.003
Drugs*	0.67	-0.33 [-0.59, 0.12]	0.26	0.12	0.75	-0.25 [-0.50, 0.15]	0.21	0.17
Cannabis*	0.81	-0.19 [-0.46, 0.20]	0.17	0.86	1.07	0.07 [-0.24, 0.50]	0.17	0.64
Smoking*	1.13	0.13 [-0.24, 0.70]	0.18	0.73	1.06	0.06 [-0.25, 0.51]	0.17	0.63
<i>Psychological health characteristics</i>								
Internalizing problems (self)	-0.03	-0.17 [-0.37, 0.03]	0.02	0.10	-0.08	-0.43 [-0.61,-0.26]	0.02	<0.001
Internalizing problems (parent)	-0.01	-0.04 [-0.23, 0.15]	0.02	0.69	-0.16	-0.84 [-1.01,-0.67]	0.02	<0.001
Internalizing problems (teacher)	-0.04	-0.13 [-0.33, 0.07]	0.03	0.20	-0.18	-0.55 [-0.72,-0.37]	0.03	<0.001
Externalizing problems (self)	0.02	0.11 [-0.09, 0.31]	0.02	0.28	-0.04	-0.24 [-0.42,-0.07]	0.01	0.006
Externalizing problems (parent)	0.04	0.21 [0.02, 0.40]	0.02	0.03	-0.15	-0.72 [-0.89,-0.56]	0.02	<0.001
Externalizing problems (teacher)	0.07	0.18 [-0.03, 0.38]	0.04	0.09	0	-0.01 [-0.18, 0.17]	0.03	0.93
ADHD (self)	0.02	0.06 [-0.14, 0.26]	0.03	0.57	-0.14	-0.47 [-0.64,-0.30]	0.03	<0.001
ADHD (parent)	0.08	0.20 [0.01, 0.38]	0.04	0.04	-0.38	-0.91 [-1.07,-0.75]	0.03	<0.001
ADHD (teacher)	0.11	0.22 [0.03, 0.42]	0.05	0.03	-0.21	-0.42 [-0.60,-0.25]	0.04	<0.001
Autism	0.22	0.14 [-0.04, 0.32]	0.15	0.14	-1.69	-1.06 [-1.22,-0.91]	0.13	<0.001
Antisocial behavior	0.06	0.25 [0.05, 0.45]	0.03	0.02	0.06	0.24 [0.07, 0.42]	0.02	0.007
Absence (psych.)	-0.04	-0.07 [-0.27, 0.13]	0.07	0.51	-0.27	-0.41 [-0.58,-0.23]	0.06	<0.001
Medication*	1.21	0.21 [-0.24, 0.92]	0.24	0.42	0.1	-0.90 [-0.94,-0.85]	0.21	<0.001

*These variables were categorical (yes/no) or ordinal (e.g., smoking never, regularly, or daily) and therefore analyzed through (multinomial) logistic regression models. Beta refers to the increase (or decrease) in the odds of replying 'yes' (in case of binary variables) or to the increase (or decrease) in the odds of being in a higher category (in case of ordinal variables) if participants were not included in the diary study (left part of the table) or if participants were part of the general population (right part of the table). Bolded p-values were significant after the Holm-Bonferroni correction.

Associations between participant characteristics and compliance

On average, participants completed 84.9% ($SD = 15.3$) of the daily diaries^f. Table 6.3 shows the associations between participant characteristics and compliance. The majority of characteristics were not associated with participants' compliance. Characteristics that were negatively associated with compliance were teacher-reported externalizing symptoms ($\beta = -.28, p_{cor} < .05$), antisocial behavior ($\beta = -.29, p_{cor} < .01$), and daily smoking ($\beta = -.64, p_{cor} < .05$). In contrast, self-reported internalizing symptoms were positively associated with compliance ($\beta = .25, p_{cor} < .05$). A few other effects related to substance use (*i.e.*, cannabis) and psychological health (*i.e.*, ADHD and medication use) had comparable or larger associations with compliance ($>.30$ in case of cannabis use) but were not significant after correcting for multiple testing.

Table 6.3. *Linear regression results for compliance (i.e., % of completed daily diary assessments)*

	B	Beta [95% CI]	SE	p
Age	3.28	0.14 [-0.03, 0.32]	1.96	0.10
Sex				
Male (versus female)	-0.84	-0.05 [-0.40, 0.29]	2.69	0.78
<i>Personality and cognitive functioning</i>				
Assertiveness	-0.52	-0.18 [-0.36, 0]	0.27	0.06
Extraversion	-0.35	-0.11 [-0.29, 0.08]	0.30	0.25
Self-discipline	0.40	0.14 [-0.05, 0.32]	0.27	0.14
Angry-hostility	-0.05	-0.01 [-0.19, 0.16]	0.33	0.89
Impulsiveness	-0.45	-0.11 [-0.28, 0.07]	0.36	0.22
Vulnerability	0.07	0.02 [-0.15, 0.20]	0.32	0.82
IQ	0.00	0 [-0.17, 0.17]	0.09	1.00
<i>Economic characteristics</i>				
Financial difficulties (self)	-3.70	-0.17 [-0.34, 0]	1.87	0.05
Educational attainment	0.27	0.02 [-0.16, 0.20]	1.33	0.84
Financial difficulties (parent)	-1.75	-0.07 [-0.24, 0.10]	2.19	0.43
Employment*				
Employed (versus unemployed)	6.78	0.44 [-0.24, 1.13]	5.19	0.20
Social benefits				
Yes (versus no)	-0.65	-0.04 [-0.45, 0.36]	3.15	0.84
Participation				
Yes (versus no)	4.74	0.31 [-0.21, 0.83]	4.03	0.24
Current education				
Yes (versus no)	0.12	0.01 [-0.36, 0.37]	2.84	0.97
Financial debt				
Yes (versus no)	-2.28	-0.15 [-0.66, 0.36]	3.97	0.57

^f Note that this includes all participants who commenced the diary study ($N=133$), *i.e.*, also those who decided to quit during the diary study ($N=11$). Among the 122 participants who completed the diary study, compliance was 88.4% ($SD = 9.9$).

	B	Beta [95% CI]	SE	p
Educational att. mother	-0.57	-0.03 [-0.21, 0.14]	1.521	0.71
Educational att. father	-1.51	-0.11 [-0.29, 0.08]	1.333	0.26
Income*				
Middle (versus low, <1150)	4.85	0.32 [-0.47, 1.11]	6.11	0.43
High, >3500 (versus low, <1150)	4.85	0.32 [-0.60, 1.23]	7.09	0.49
Occupation mother*				
Middle (versus low)	-1.02	-0.07 [-0.67, 0.54]	4.65	0.83
High (versus low)	-5.32	-0.35 [-0.99, 0.29]	4.94	0.28
Occupation father*				
Middle (versus low)	4.09	0.27 [-0.23, 0.77]	3.86	0.29
High (versus low)	-1.69	-0.11 [-0.58, 0.36]	3.60	0.64
<i>Social characteristics</i>				
Social problems (self)	3.26	0.06 [-0.11, 0.23]	4.74	0.49
Social problems (parent)	-8.12	-0.20 [-0.37, -0.03]	3.40	0.018
Family functioning	4.10	0.09 [-0.08, 0.26]	3.99	0.31
Adverse events	-0.18	-0.07 [-0.24, 0.10]	0.23	0.42
Chronic stress	0.20	0.03 [-0.14, 0.20]	0.53	0.71
Trauma	-0.17	-0.02 [-0.19, 0.15]	0.69	0.81
<i>Physical health characteristics</i>				
Physical health (self)	-2.22	-0.08 [-0.25, 0.09]	2.39	0.35
Physical health (parent)	-1.02	-0.04 [-0.21, 0.14]	2.51	0.68
BMI	0.31	0.07 [-0.10, 0.24]	0.37	0.41
Alcohol	-0.34	-0.07 [-0.24, 0.11]	0.46	0.46
Absence (phys.)	0.25	0.02 [-0.15, 0.20]	1.02	0.81
Drugs*				
Yes (versus no)	-5.38	-0.35 [-0.75, 0.05]	3.12	0.09
Cannabis*				
Experimenter (versus never)	-6.86	-0.45 [-0.80, -0.10]	2.72	0.01
Regular user (versus never)	-12.10	-0.79 [-1.40, -0.18]	4.69	0.01
Smoking*				
Occasional (versus never)	-3.04	-0.20 [-0.63, 0.24]	3.36	0.37
Daily (versus never)	-9.83	-0.64 [-1.08, -0.21]	3.36	0.004
<i>Psychological health characteristics</i>				
Internalizing problems (self)	19.06	0.25 [0.08, 0.42]	6.49	0.004
Internalizing problems (parent)	9.72	0.13 [-0.04, 0.30]	6.31	0.13
Internalizing problems (teacher)	7.16	0.16 [-0.01, 0.33]	3.91	0.07
Externalizing problems (self)	-12.66	-0.13 [-0.30, 0.04]	8.62	0.14
Externalizing problems (parent)	-15.47	-0.24 [-0.41, -0.07]	5.53	0.006
Externalizing problems (teacher)	-12.65	-0.28 [-0.44, -0.11]	3.88	0.001
ADHD (self)	-1.79	-0.04 [-0.21, 0.14]	4.41	0.69
ADHD (parent)	-6.58	-0.20 [-0.37, -0.03]	2.88	0.024
ADHD (teacher)	-7.37	-0.24 [-0.41, -0.07]	2.61	0.006
Autism	-0.89	-0.11 [-0.28, 0.06]	0.71	0.21
Antisocial behavior	-27.52	-0.29 [-0.46, -0.13]	7.86	0.001
Absence (psych.)	-1.35	-0.07 [-0.25, 0.10]	1.62	0.40
Medication				
Yes (versus no)	-6.05	-0.40 [-0.79, 0]	3.09	0.05

*These variables were categorical with more than two levels (e.g., smoking seldomly, regularly, or often) and therefore analyzed using dummy variables. Bolded p-values were significant after using the Holm-Bonferroni procedure to correct for multiple testing.

Fig. 6.3. Average size (standardized beta) of the effect of different characteristics on compliance. Effects that were significant ($p < .05$) after correcting for multiple testing are marked with an asterisk. For continuous variables that were assessed across multiple informants (e.g., internalizing problems), effect sizes were averaged across informants. For categorical variables with two levels, the standardized beta describes the standardized difference in compliance between individuals who were part of the respective category (e.g., who were employed) versus those who were not (e.g., who were not employed). **For categorical variables with more than two levels (income, occupation parent, smoking, substance use), the standardized beta describes (i) the difference between the middle category (e.g., occasional smoking) and the reference category (no smoking) and (ii) the difference between the high category (e.g., daily smoking) and the reference category on top of the aforementioned difference. Only for parental occupation, these effects were in opposite direction (β middle vs. low education: 0.10, $p > 0.05$; β high vs. low education: -0.23, $p > 0.05$). For income, the difference between the highest vs. reference category and the middle vs. reference category is invisibly small (difference in β s: 0.005). For substance use, the effect size of the middle group reflects the averaged effect size of individuals who reported lifetime illicit drug use and individuals who reported experimental cannabis use. The highest level includes regular cannabis users.

Discussion

The current study investigated who participates in intensive longitudinal designs (ILDs), and, when enrolled in such a study, who does well in terms of compliance. In a cohort of at-risk individuals who were invited for a six-month daily diary study, we found that participation in ILDs is not reserved for relatively healthy, high-functioning individuals, as previously suggested⁸. Rather, individuals who are vulnerable for mental health problems participate and are compliant. First, comparing the diary participants and the cohort participants who decided not to participate in the diary study, we observed that the diary participants were not necessarily the better functioning part of the cohort. After ruling out differences (*e.g.*, higher IQ-scores and less behavioral problems for participants than non-participants) that were due to selective attrition effects in TRAILS CC rather than a specific disinterest in the diary study, there were no differences in characteristics (*e.g.*, personality, family socio-economic status, physical or mental health problems) that suggested participants functioned better than non-participants. This was not due to TRAILS CC participants functioning well, illustrated by the finding that participants scored less favorable on nearly every characteristic compared to the general population (*e.g.*, more social problems, lower physical and psychological health, etc.). Finally, with the exception of behavioral problems (*i.e.*, antisocial behavior and externalizing problems), none of these characteristics negatively affected compliance, and moreover compliance was excellent (*i.e.*, on average 84.9%). We conclude therefore that ILD studies are neither hindered by self-selection bias nor by low study compliance. It follows that ILDs can be successfully applied to study individuals with psychological and psychiatric problems.

As argued in the introduction, individuals for whom participation is no effort are thought to be overrepresented in ILDs⁸. These individuals are conscientious, self-disciplined, pro-social, and have relatively few health problems that could interfere with participation. Such characteristics are of particular importance for participation in ILDs, due to their higher burden and the lengthy commitment that is requested^{18,19}. However, there is also an opposing view on the profile of research participants that may typically take part in ILDs^{8,20}: individuals with reduced mental health. Individuals with mental health problems may be more motivated to contribute to studies into mental health, for instance because such studies may improve self-understanding or help others with similar problems^{2,21}. They may also have more time on their hands (*i.e.*, no job or not enrolled in education) and more interest in the remuneration provided to participants⁸. Both profiles imply a selection

bias, and for ILDs with clinical implications, the first profile ('happy and healthy') is most detrimental. We found no support for this profile: diary participants did not report more self-discipline nor were they healthier than non-participants. Importantly, compared to the general population, participants of the diary study scored less advantageous on practically all characteristics (including self-discipline and health), indicating that this first profile does not fit the average TRAILS TRANS-ID participant. In favor of the second profile, we observed that diary participants had experienced more stressors and traumatic events compared to non-participants, and reported somewhat higher behavioral problems. However, the attrition analysis, in which individuals that had not participated in the fifth assessment wave (T5) were excluded, revealed that this effect was likely driven by a disinterest in participating in research in general, as opposed to the diary study per se. Moreover, diary study participants did not display more psychological or physical health problems than those who decided not to participate. We thus conclude that intensive longitudinal designs are not necessarily biased in terms of the physical, psychological, and social characteristics of individuals that self-select into these kinds of studies, making it possible to study at-risk individuals for psychopathology. That said, we observed three characteristics associated with participation that could be paid attention to during recruitment of at-risk participants. Like previous ILDs we found that women were more likely to participate in the diary study than men^{22–24}. Moreover, although well below the general population mean, a slight underrepresentation of at-risk individuals scoring higher on anti-social behaviors or who have experienced less life events may occur. These biases are not specific to ILDs but rather apply to longitudinal designs in general²⁵. Hence, it is worth considering that results of ILDs in at-risk samples may not be representative of individuals with these characteristics (*i.e.*, being male, displaying behavioral problems, and reporting less adverse events).

Once enrolled in the diary study, internalizing problems were positively associated with compliance. This contrasts with a finding from a previous study, where depression severity was negatively associated with college students' study compliance²⁶. A recent meta-analysis on predictors of compliance in ILDs found no effect (*i.e.*, neither positive or negative) of depression severity on compliance²⁴. Hence, no conclusions can be drawn yet. Although intrinsic motivation or conscientiousness have been considered important determinants of compliance in diary studies^{8,27,28}, we found no association between self-discipline and compliance in this study. Likewise, we did not find that male participants were less compliant than female participants^{23,24,26}. Compliance was, however, lower in individuals with more

behavioral problems (*i.e.*, antisocial behavior and teacher-reported externalizing problems) and in those who smoked daily. Note that some of the other associations between behavioral problem domains and substance use with compliance had a similar effect size but were non-significant, likely due to reduced power given a low number of extreme values (*i.e.*, only 12 of 122 individuals were regular cannabis users). Thus, if substance use or behavioral problems are of interest to the study, it is good to be aware that the individuals displaying these behaviors may be less compliant, and therefore, they might be underrepresented.

Strengths and considerations

A major strength of this study is that extensive information was available for both individuals who participated and individuals who did not participate in the diary study, not only in terms of types of characteristics (*e.g.*, psychological, physical, economic, social, and demographic characteristics), but also in terms of repeated assessment of these characteristics, which yielded reliable estimates. An additional strength is that all characteristics were compared to normative scores from the general population (TRAILS PC), a study identical in set-up as TRAILS CC. This allowed us to confirm the at-risk status of our participants and thus rule out that the overall absence of strong biases between participants and non-participants was due to floor effects (*i.e.*, little mental health problems in both groups, despite their at-risk status). Additionally, the comparison with the general population showed that individuals (*i.e.*, the diary participants) who scored less favorable on characteristics associated with ILD study success had no trouble with participating in an ILD study. However, although participants showed higher adversity than the general population, we cannot guarantee that they are representative of the broader at-risk population. This is because diary participants were recruited within a long-lasting, ongoing cohort consisting of individuals at increased risk for mental health problems (TRAILS CC). The members of this cohort might be more motivated to participate in research than at-risk individuals in general^{29,30}, and might therefore differ from individuals in the full at-risk population. This could imply that associations between the investigated characteristics and the decision to participate, which would exist in the population, could not be observed in our sample. However, we think the impact of this study being embedded in TRAILS on current findings is limited, as the decision to enter the cohort study was made by parents rather than participants, who were aged 10-11 years at the time. The decision to participate in the diary study, in contrast, was made by the participants themselves. A final consideration is that due to sample size differences the comparison

between the general population sample (N=1926) and the diary participants (N=134) had more power than the comparison between diary participants and non-participants (N=309). Although this could have played a role in finding fewer significant differences between participants and non-participants than between participants and the general population sample, effect sizes of the differences were much smaller in the participant versus non-participant comparison, indicating that reduced power does not confound our conclusions. Moreover, the compliance analyses – which only involved diary participants and arguably had even less power than the participants versus non-participants comparison – resulted in several significant associations. Thus, we do not think that finding few differences between participants and non-participants is due to a lack of power.

Conclusions

Often, we do not know much about those individuals who were *not* included in our studies. This means we can only speculate about the selection bias that results from the fact that particular types of individuals might be more likely to participate. We had the unique possibility to compare participants to non-participants on a wide variety of traits, and found no selection bias towards highly conscientious, motivated, and healthy individuals, nor towards disadvantaged individuals. Instead, individuals who decided to participate in the diary study were largely similar to those who did not participate in terms of personality and psychological and physical health. However, researchers should be aware that females, those with higher IQ, and those who experienced more adverse events may be overrepresented in their study (due to self-selection), whereas individuals with more externalizing and less internalizing problems, more antisocial behavior, and daily smokers might be underrepresented (due to lower compliance). This is not specific to intensive longitudinal designs but applies to research participation in general. Individuals who took part in the diary study were disadvantaged relative to the general population in a broad variety of domains but showed excellent study compliance, confirming that diary studies are suitable for at-risk populations. Taken together, we found that diary participants were disadvantaged in several domains relative to the general population, but this did not affect their decision to participate in a six-month diary study nor their compliance. We conclude therefore that intensive longitudinal designs can be broadly applied and are not only suitable ‘for the happy few’.

Supplement 6

Table S6.2. Descriptive statistics and number of missing values for each characteristic in the three participant groups

	Participants		Non-participants		General population	
	Mean (SD)/ N(%)*	Missing N(%)	Mean (SD)/ N(%)*	Missing N(%)	Mean (SD)/ N(%)*	Missing N(%)
<i>Personality and cognitive functioning</i>						
Assertiveness	23.02 (5.16)	11 (8.21)	23.56 (4.85)	47 (15.21)	23.96 (4.52)	516 (26.79)
Excitement-seeking	27.23 (4.71)	11 (8.21)	27.11 (4.55)	47 (15.21)	27.90 (4.20)	517 (26.84)
Self-discipline	24.54 (5.20)	11 (8.21)	25.10 (4.99)	47 (15.21)	26.16 (4.55)	518 (26.9)
Angry-hostility	20.99 (4.16)	4 (2.99)	20.69 (4.21)	35 (11.33)	19.88 (4.07)	364 (18.9)
Impulsiveness	23.70 (3.73)	4 (2.99)	23.64 (3.65)	35 (11.33)	23.19 (3.42)	364 (18.9)
Vulnerability	20.69 (4.23)	4 (2.99)	20.15 (4.15)	35 (11.33)	19.04 (4)	364 (18.9)
IQ	100.34 (14.88)	0 (0)	96.02 (15.99)	0 (0)	97.25 (15.07)	8 (0.42)
<i>Economic characteristics</i>						
Financial difficulties (self)	2 (0.71)	4 (2.99)	1.83 (0.8)	64 (20.71)	1.97 (0.75)	432 (22.43)
Educational attainment	3.41 (1.02)	9 (6.72)	3.37 (1.02)	123 (39.81)	3.52 (1.01)	636 (33.02)
Financial difficulties (parent)	0.19 (0.61)	0 (0)	0.26 (0.75)	0 (0)	0.19 (0.82)	0 (0)
Employment		96 (71.64)		235 (76.10)		1501 (77.9)
No	17 (44.74)		30 (40.54)		104 (24.47)	
Yes	21 (55.26)		44 (59.46)		321 (75.53)	
Social benefits		9 (6.72)		123 (39.81)		636 (33.02)
No	94 (75.20)		133 (71.51)		1192 (92.4)	
Yes	31 (24.80)		53 (28.49)		98 (7.6)	
Participation		16 (11.94)		141 (45.63)		658 (34.16)
No	17 (14.41)		30 (17.86)		104 (8.2)	
Yes	101 (85.59)		138 (82.14)		1164 (91.8)	
Current education		9 (6.72)		123 (39.81)		636 (33.02)
No	45 (36)		92 (49.46)		447 (34.65)	
Yes	80 (64)		94 (50.54)		843 (65.35)	
Financial debt		10 (7.46)		131 (42.39)		660 (34.27)
No	108 (87.10)		167 (93.82)		1026 (81.0)	
Yes	16 (12.90)		11 (6.18)		240 (18.96)	
Educational att. mother	3.13 (0.89)	0 (0)	3.07 (0.97)	2 (0.65)	2.95 (1.07)	
Educational att. father	3.23 (1.09)	13 (9.7)	3.06 (0.99)	38 (12.30)	3.16 (1.16)	
Household income		8 (5.97)		14 (4.53)		121 (6.28)
Low < €1,150	7 (5.56)		14 (4.75)		219 (12.13)	
Middle	102 (80.95)		246 (83.39)		1405 (77.8)	
High > €3,500	17 (13.49)		35 (11.86)		181 (10.03)	
Occupation mother		40 (29.9)		90 (29.1)		490 (25.4)
Low	11 (11.7)		24 (11)		255 (17.8)	
Middle	52 (55.3)		125 (57.1)		740 (51.5)	
High	31 (33)		70 (32)		441 (30.7)	
Occupation father		31 (23.1)		75 (24.3)		389 (20.2)
Low	32 (31.1)		86 (36.8)		471 (30.6)	
Middle	30 (29.1)		58 (24.8)		461 (30)	
High	41 (39.8)		90 (38.5)		605 (39.4)	

	Participants		Non-Participants		General population	
	Mean (SD)/ N(%)*	Missing N(%)	Mean (SD)/ N(%)*	Missing N(%)	Mean (SD)/ N(%)*	Missing N(%)
<i>Social characteristics</i>						
Social problems (self)	0.34 (0.28)	0 (0)	0.32 (0.28)	0 (0)	0.24 (0.28)	3 (0.16)
Social problems (parent)	0.48 (0.39)	0 (0)	0.46 (0.45)	1 (0.32)	0.18 (0.28)	55 (2.86)
Family functioning	1.79 (0.34)	0 (0)	1.84 (0.35)	1 (0.32)	1.70 (0.32)	59 (3.06)
Adverse events	10.4 (5.88)	0 (0)	7.65 (5.22)	0 (0)	7.52 (5.27)	0 (0)
Chronic stress	3.41 (2.52)	0 (0)	2.94 (2.74)	0 (0)	1.29 (1.86)	0 (0)
Trauma	2.83 (1.92)	0 (0)	2.28 (2.47)	0 (0)	2.1 (2.16)	0 (0)
<i>Physical health characteristics</i>						
Physical health (self)	3.64 (0.56)	0 (0)	3.63 (0.63)	0 (0)	3.73 (0.59)	5 (0.26)
Physical health (parent)	3.93 (0.54)	0 (0)	3.87 (0.55)	1 (0.32)	4.14 (0.54)	59 (3.06)
BMI	21.26 (3.60)	0 (0)	20.48 (3.54)	1 (0.32)	20.18 (3.31)	11 (0.57)
Alcohol	2.38 (2.98)	3 (2.24)	3.30 (4.63)	16 (5.18)	4.66 (5.07)	123 (6.39)
Absence (phys.)	2.70 (1.31)	1 (0.75)	2.61 (1.25)	15 (4.85)	2.37 (1.18)	224 (11.63)
Other drugs		0 (0)		0 (0)		0 (0)
No	103 (76.87)		257 (83.17)		1572 (81.6)	
Yes	31 (23.13)		52 (16.83)		354 (18.38)	
Cannabis		3 (2.24)		27 (8.74)		299 (15.52)
Never	58 (44.27)		141 (50)		694 (42.66)	
Experimenter	61 (46.56)		118 (41.84)		777 (47.76)	
Regular	12 (9.16)		23 (8.16)		156 (9.59)	
Smoking		2 (1.49)		25 (8.09)		167 (8.67)
Never	78 (59.09)		156 (54.93)		1026 (58.3)	
Occasional	27 (20.45)		69 (24.30)		333 (18.93)	
Daily	27 (20.45)		59 (20.77)		400 (22.74)	
<i>Psychological health characteristics</i>						
Internalizing problems (self)	0.39 (0.20)	0 (0)	0.36 (0.20)	0 (0)	0.30 (0.19)	4 (0.21)
Internalizing problems (parent)	0.37 (0.21)	0 (0)	0.37 (0.23)	1 (0.32)	0.21 (0.17)	58 (3.01)
Internalizing problems (teacher)	0.56 (0.34)	2 (1.49)	0.52 (0.31)	5 (1.62)	0.38 (0.33)	51 (2.65)
Externalizing problems (self)	0.31 (0.15)	0 (0)	0.33 (0.18)	0 (0)	0.27 (0.16)	3 (0.16)
Externalizing problems (parent)	0.36 (0.24)	0 (0)	0.4 (0.25)	1 (0.32)	0.21 (0.18)	57 (2.96)
Externalizing problems (teacher)	0.25 (0.33)	2 (1.49)	0.32 (0.36)	5 (1.62)	0.25 (0.37)	49 (2.54)
ADHD (self)	0.71 (0.30)	0 (0)	0.73 (0.31)	0 (0)	0.57 (0.28)	3 (0.16)
ADHD (parent)	0.85 (0.46)	0 (0)	0.93 (0.43)	1 (0.32)	0.46 (0.37)	55 (2.86)
ADHD (teacher)	0.73 (0.50)	2 (1.49)	0.83 (0.48)	5 (1.62)	0.52 (0.47)	49 (2.54)
Autism	3.09 (1.88)	0 (0)	3.31 (1.85)	1 (0.32)	1.40 (1.29)	19 (0.99)
Antisocial behavior	0.19 (0.16)	0 (0)	0.25 (0.24)	0 (0)	0.25 (0.27)	1 (0.05)
Absence (psych.)	1.42 (0.83)	1 (0.75)	1.37 (0.88)	15 (4.85)	1.15 (0.58)	224 (11.63)
Medication		14 (10.45)		31 (10.03)		681 (35.36)
No	38 (31.67)		77 (27.7)		1026 (82.4)	
Yes	82 (68.33)		201 (72.3)		219 (17.59)	

Note. Absence (phys.) = absence due to physical complaints. Absence (psych.) = absence due to psychological complaints. ADHD = attention-deficit/hyperactivity disorder. SE = standard error.

Table S6.3. Sensitivity analysis including individuals who completed T5: characteristic differences between participants and non-participants, and between participants and a sex-matched general population sample

	Participants vs. Non-participants				Participants vs. General population			
	B/OR	Beta/1-OR [95% CI]	SE	p	B/OR	Beta/1-OR	SE	p
<i>Personality and cognitive functioning</i>								
Assertiveness	-0.09	-0.02 [-0.25, 0.21]	0.56	0.87	0.86	0.18 [-0.01, 0.37]	0.45	0.06
Extraversion	-0.64	-0.15 [-0.38, 0.09]	0.51	0.21	0.70	0.16 [-0.03, 0.35]	0.41	0.09
Self-discipline	0.63	0.13 [-0.10, 0.36]	0.56	0.27	1.80	0.38 [0.19, 0.57]	0.46	0.00
Angry-hostility	-0.74	-0.18 [-0.41, 0.04]	0.47	0.11	-1.45	-0.36 [-0.54, -0.17]	0.38	0.00
Impulsiveness	-0.38	-0.11 [-0.34, 0.12]	0.40	0.35	-0.60	-0.17 [-0.36, 0.01]	0.33	0.07
Vulnerability	-0.58	-0.14 [-0.37, 0.08]	0.46	0.21	-1.74	-0.43 [-0.61, -0.25]	0.37	0.00
IQ	-2.64	-0.18 [-0.40, 0.05]	1.72	0.13	-0.61	-0.04 [-0.22, 0.14]	1.39	0.66
<i>Economic characteristics</i>								
Financial difficulties (self)	-0.22	-0.31 [-0.53, -0.08]	0.08	0.01	-0.61	-0.04 [-0.22, 0.14]	1.39	0.66
Educational attainment	-0.04	-0.04 [-0.27, 0.18]	0.12	0.72	-0.03	-0.04 [-0.22, 0.15]	0.07	0.69
Financial difficulties (parent)	-0.01	-0.02 [-0.24, 0.21]	0.09	0.89	0.11	0.11 [-0.07, 0.29]	0.09	0.24
Employment*	1.19	0.19 [-0.46, 1.62]	0.40	0.67	2.50	1.50 [0.26, 3.91]	0.35	0.01
Social benefits*	1.21	0.21 [-0.28, 1.04]	0.26	0.47	0.25	-0.75 [-0.84, -0.60]	0.23	0.00
Participation*	0.77	-0.23 [-0.60, 0.47]	0.33	0.44	1.88	0.88 [0.05, 2.19]	0.28	0.02
Current education*	0.57	-0.43 [-0.64, -0.09]	0.24	0.02	1.06	0.06 [-0.28, 0.55]	0.20	0.76
Financial debt*	0.44	-0.56 [-0.81, -0.01]	0.41	0.05	1.58	0.58 [-0.06, 1.82]	0.28	0.10
Educational att. mother	-0.05	-0.05 [-0.27, 0.18]	0.12	0.69	-0.03	-0.03 [-0.21, 0.16]	0.09	0.78
Educational att. father	-0.12	-0.10 [-0.34, 0.14]	0.14	0.40	0.11	0.10 [-0.09, 0.29]	0.11	0.32
Income*	0.80	-0.2 [-0.55, 0.43]	0.24	0.78	0.76	-0.24 [-0.52, 0.23]	0.24	0.87
Occupation mother*	1.05	0.05 [-0.38, 0.78]	0.22	0.58	0.93	-0.07 [-0.40, 0.41]	0.22	0.63
Occupation father*	0.83	-0.17 [-0.49, 0.35]	0.20	0.77	1.14	0.14 [-0.23, 0.69]	0.19	0.75
<i>Social characteristics</i>								
Social problems (self)	0.00	0 [-0.23, 0.22]	0.03	0.97	-0.11	-0.46 [-0.64, -0.28]	0.02	0.00
Social problems (parent)	-0.04	-0.14 [-0.35, 0.07]	0.03	0.20	-0.31	-0.99 [-1.16, -0.82]	0.03	0.00
Family functioning	0.00	0.01 [-0.22, 0.23]	0.04	0.95	-0.10	-0.32 [-0.50, -0.13]	0.03	0.00
Adverse events	-1.11	-0.21 [-0.44, 0.01]	0.60	0.06	-1.49	-0.29 [-0.47, -0.10]	0.48	0.00
Chronic stress	-0.41	-0.18 [-0.40, 0.03]	0.25	0.10	-2.00	-0.89 [-1.07, -0.72]	0.20	0.00
Trauma	-0.29	-0.14 [-0.37, 0.09]	0.24	0.22	-0.24	-0.12 [-0.30, 0.07]	0.19	0.21
<i>Physical health characteristics</i>								
Physical health (self)	-0.03	-0.06 [-0.28, 0.17]	0.06	0.63	0.07	0.12 [-0.06, 0.30]	0.05	0.20
Physical health (parent)	-0.02	-0.04 [-0.27, 0.18]	0.06	0.70	0.20	0.39 [0.21, 0.57]	0.05	0.00
BMI	-0.02	0.00 [-0.23, 0.22]	0.37	0.97	-0.55	-0.17 [-0.35, 0.01]	0.30	0.07
Alcohol	0.41	0.09 [-0.13, 0.31]	0.53	0.43	2.40	0.52 [0.34, 0.70]	0.43	0.00
Absence (phys.)	-0.11	-0.09 [-0.32, 0.13]	0.14	0.42	-0.37	-0.32 [-0.50, -0.13]	0.11	0.00
Drugs*	1.18	0.18 [-0.30, 0.99]	0.26	0.54	1.15	0.15 [-0.24, 0.78]	0.22	0.53
Cannabis*	0.84	-0.16 [-0.46, 0.30]	0.18	0.78	1.10	0.10 [-0.23, 0.58]	0.17	0.71
Smoking*	1.09	0.09 [-0.30, 0.71]	0.18	0.65	1.06	0.06 [-0.26, 0.53]	0.18	0.62

	Participants vs. Non-participants					Participants vs. General population				
	B/OR	Beta/1-OR [95% CI]	SE	p		B/OR	Beta/1-OR	SE	p	
<i>Psychological health characteristics</i>										
Internalizing problems (self)	-0.03	-0.17 [-0.40, 0.05]	0.02	0.13		-0.09	-0.48 [-0.66,-0.3]	0.02	0.00	
Internalizing problems (parent)	-0.01	-0.05 [-0.27, 0.16]	0.02	0.62		-0.17	-0.89 [-1.06,-0.72]	0.02	0.00	
Internalizing problems (teacher)	-0.01	-0.04 [-0.26, 0.18]	0.04	0.74		-0.19	-0.59 [-0.77,-0.41]	0.03	0.00	
Externalizing problems (self)	-0.01	-0.06 [-0.29, 0.16]	0.02	0.57		-0.05	-0.36 [-0.54,-0.18]	0.01	0.00	
Externalizing problems (parent)	-0.01	-0.03 [-0.24, 0.19]	0.02	0.80		-0.16	-0.89 [-1.06,-0.72]	0.02	0.00	
Externalizing problems (teacher)	0.00	0.01 [-0.22, 0.24]	0.04	0.94		-0.06	-0.19 [-0.37, 0]	0.03	0.05	
ADHD (self)	-0.02	-0.07 [-0.29, 0.15]	0.03	0.53		-0.16	-0.57 [-0.75,-0.39]	0.03	0.00	
ADHD (parent)	0.01	0.02 [-0.19, 0.22]	0.04	0.88		-0.42	-1.09 [-1.26,-0.93]	0.03	0.00	
ADHD (teacher)	0.03	0.07 [-0.15, 0.30]	0.05	0.51		-0.27	-0.60 [-0.78,-0.42]	0.04	0.00	
Autism	0.01	0.00 [-0.19, 0.20]	0.15	0.97		-1.83	-1.24 [-1.40,-1.08]	0.12	0.00	
Antisocial behavior	0.02	0.10 [-0.13, 0.33]	0.02	0.38		0.01	0.04 [-0.14, 0.22]	0.02	0.66	
Absence (psych.)	-0.10	-0.16 [-0.39, 0.07]	0.07	0.17		-0.29	-0.49 [-0.68,-0.31]	0.06	0.00	
Medication*	1.18	0.18 [-0.31, 0.98]	0.27	0.54		0.09	-0.91 [-0.94,-0.86]	0.22	0.00	

Note. Beta refers to the increase (or decrease) in the odds of replying 'yes' (in case of binary variables) or to the increase (or decrease) in the odds of being in a higher category (in case of ordinal variables) if participants were not included in the diary study (left part of the table) or if participants were part of the general population (right part of the table). Bolded P-values were significant after the Holm-Bonferroni correction. *These variables were categorical (yes/no) or ordinal (e.g., smoking never, regularly, or daily) and therefore analyzed through (multinomial) logistic regression models. Educational att. = educational attainment. Absence (phys.) = absence due to physical complaints. Absence (psych.) = absence due to psychological complaints. ADHD = attention-deficit/hyperactivity disorder. SE = standard error.

References

1. van Os J, Delespaul P, Wigman J, Myin-Germeys I, Wichers M. Beyond DSM and ICD: introducing “precision diagnosis” for psychiatry using momentary assessment technology. *World Psychiatry*. 2013;12(2):113-117.
2. Bos FM, Snippe E, Bruggeman R, Wichers M, van der Krieke L. Insights of patients and clinicians on the promise of the experience sampling method for psychiatric care. *Psychiatric Services*. 2019;70(11):983-991. doi:10.1176/appi.ps.201900050
3. Wright AGC, Woods WC. Personalized models of psychopathology. *Annual Review of Clinical Psychology*. 2020;16:49-74. doi:10.1146/annurev-clinpsy-102419-125032
4. Piccirillo ML, Beck ED, Rodebaugh TL. A clinician’s primer for idiographic research: considerations and recommendations. *Behavior Therapy*. 2019;50(5):938-958. doi:10.1016/j.beth.2019.02.002
5. Kramer I, Simons CJP, Hartmann JA, et al. A therapeutic application of the experience sampling method in the treatment of depression: a randomized controlled trial. *World Psychiatry*. 2014;13(1):68-77. doi:10.1002/wps.20090
6. Krieke L van der, Jeronimus BF, Blaauw FJ, et al. HowNutsAreTheDutch (HoeGekIsNL): a crowdsourcing study of mental symptoms and strengths. *International Journal of Methods in Psychiatric Research*. 2016;25(2):123-144. doi:10.1002/mpr
7. Ohly S, Sonnentag S, Niessen C, Zapf D. Diary Studies in Organizational Research: An Introduction and Some Practical Recommendations. *Journal of Personnel Psychology*. 2010;9(2):79-93. doi:10.1027/1866-5888/a000009
8. Scollon CN, Kim-Prieto C, Scollon CN. Experience sampling: promises and pitfalls, strengths and weaknesses. *Journal of Happiness Studies*. 2003;4(1):5-34. doi:10.1023/a:1023605205115
9. Wilhelm P, Schoebi D, Perrez M. Frequency estimates of emotions in everyday life from a diary method’s perspective: A comment on Scherer et al.’s survey-study “emotions in everyday life.” *Social Science Information*. 2004;43(4):647-665. doi:10.1177/0539018404047712
10. Mulligan C, Schneider B, Wolfe R. Time use and population representation in the sloan study of adolescents. *National Bureau of Economic Research*. 2000;7:1-28. doi:10.3386/t0265
11. Larson RW, Moneta G, Richards MH, Wilson S. Continuity, stability, and change in daily emotional experience across adolescence. *Child Development*. 2002;73(4):1151-1165.
12. Huisman M, Oldehinkel AJ, de Winter A, et al. Cohort profile: the Dutch “Tracking Adolescents” Individual Lives’ Survey’; TRAILS. *International Journal of Epidemiology*. 2008;37(6):1227-1235. doi:10.1093/ije/dym273
13. Oldehinkel AJ, Rosmalen JGM, Buitelaar JK, et al. Cohort profile update: the tracking adolescents’ individual lives survey (TRAILS). *International Journal of Epidemiology*. 2015;44(1):76-76. doi:10.1093/ije/dyu225
14. Dietrich A, Ormel J, Buitelaar JK, Verhulst FC, Hoekstra PJ, Hartman CA. Cortisol in the morning and dimensions of anxiety, depression, and aggression in children from a general population and clinic-referred cohort: an integrated analysis. The TRAILS study. *Psychoneuroendocrinology*. 2013;38(8):1281-1298. doi:10.1016/j.psyneuen.2012.11.013
15. Schreuder MJ, Groen RN, Wigman JTW, Hartman CA, Wichers M. Measuring psychopathology as it unfolds in daily life: addressing key assumptions of intensive longitudinal methods in the TRAILS TRANS-ID study.

- BMC Psychiatry*. 2020;20(1). doi:10.1186/s12888-020-02674-1
16. R Core Team. R: A language and environment for statistical computing. Published online 2020.
 17. Holm S. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*. 1979;6(2):65-70.
 18. Gable SL, Reis HT, Elliot AJ. Behavioral activation and inhibition in everyday life. *Journal of Personality and Social Psychology*. 2000;78(6):1135-1149. doi:10.1037/0022-3514.78.6.1135
 19. Litt MD, Cooney NL, Morse P. Ecological momentary assessment (EMA) with treated alcoholics: methodological problems and potential solutions. *Health Psychology*. 1998;17(1):48-52.
 20. Courvoisier DS, Eid M, Lischetzke T. Compliance to a cell phone-based ecological momentary assessment study: the effect of time and personality characteristics. *Psychological Assessment*. 2012;24(3):713-720. doi:10.1037/a0026733
 21. Zullino D, Conus P, Borgeat F, Bonsack C. Readiness to participate in psychiatric research. *The Canadian Journal of Psychiatry*. 2003;48(7):480-484.
 22. Messiah A, Grondin O, Encrenaz G. Factors associated with missing data in an experience sampling investigation of substance use determinants. *Drug and Alcohol Dependence*. 2011;114(2-3):153-158. doi:10.1016/j.drugalcdep.2010.09.016
 23. Rintala A, Wampers M, Myin-Germeys I, Viechtbauer W. Response compliance and predictors thereof in studies using the experience sampling method. *Psychological Assessment*. 2019;31(2):226-235. doi:10.1016/j.psychres.2020.112896
 24. Vachon H, Viechtbauer W, Rintala A, Myin-Germeys I. Compliance and retention with the experience sampling method over the continuum of severe mental disorders: meta-analysis and recommendations. *Journal of Medical Internet Research*. 2019;21(12):1-20. doi:10.2196/14475
 25. Galea S, Tracy M. Participation rates in epidemiologic studies. *Annals of Epidemiology*. 2007;17(9):643-653. doi:10.1016/j.annepidem.2007.03.013
 26. Silvia PJ, Kwapił TR, Eddington KM, Brown LH. Missed beeps and missing data: dispositional and situational predictors of nonresponse in experience sampling research. *Social Science Computer Review*. 2013;31(4):471-481. doi:10.1177/0894439313479902
 27. Palmier-Claus JE, Myin-Germeys I, Barkus E, et al. Experience sampling research in individuals with mental illness: reflections and guidance. *Acta Psychiatrica Scandinavica*. 2011;123(1):12-20. doi:10.1111/j.1600-0447.2010.01596.x
 28. Piasecki TM, Hufford MR, Solhan M, Trull TJ. Assessing clients in their natural environments with electronic diaries: rationale, benefits, limitations, and barriers. *Psychological Assessment*. 2007;19(1):25-43. doi:10.1037/1040-3590.19.1.25
 29. Zunzunegui M v, Beland F, Gutierrez-Cuadra P. Loss to follow-up in a longitudinal study on aging in Spain. *Journal of clinical epidemiology*. 2001;54(5):501-510.
 30. Pietilä A-M, Rantakallio P, Läärä E. Background factors predicting non-response in a health survey of northern Finnish young men. *Scandinavian journal of social medicine*. 1995;23(2):129-136.

7

Subthreshold psychopathology: stable or transient?

adapted from: Schreuder, M.J., Wigman, J.T.W., Groen, R.N., Wichers, M., & Hartman, C.A. (2021). On the transience or stability of subthreshold psychopathology. *Scientific Reports*, 11(1), 1–7. doi: 10.1038/s41598-021-02711

Abstract

Symptoms of psychopathology lie on a continuum ranging from mental health to psychiatric disorders. Although much research has focused on progression along this continuum, for most individuals, subthreshold symptoms do *not* escalate in full-blown disorders. This study investigated how the stability of psychopathological symptoms (attractor strength) varies across severity levels (homebase). Data were retrieved from the TRAILS TRANS-ID study, where 122 at-risk young adults (mean age 23.6 years, 57% males) monitored their mental states daily for a period of six months (± 183 observations per participant). We estimated each individual's homebase and attractor strength using generalized additive mixed models. Regression analyses showed no association between homebases and attractor strengths (linear model: $B = 0.02$, $p = .47$, $R^2 < .01$; polynomial model: $B < 0.01$, $p = .61$, $R^2 < .01$). Sensitivity analyses where we (i) weighed estimates according to their uncertainty and (ii) removed individuals with a DSM-5 diagnosis from the analyses did not change this finding. This suggests that stability is similar across severity levels, implying that subthreshold psychopathology may resemble a stable state rather than a transient intermediate between mental health and psychiatric disorder. Our study thus provides additional support for a dimensional view on psychopathology, which implies that symptoms differ in degree rather than kind.

Introduction

Psychopathology is increasingly recognized as a dimensional phenomenon^{1–5}. From such a dimensional perspective, psychiatric disorders reflect the extreme end of a severity continuum ranging from the absence of symptoms to the presence of severe symptoms. Along this continuum lie subthreshold symptoms, which fall short of the diagnostic criteria for a clinical disorder but may still cause burden and functional impairments^{3,6,7}.

A dimensional view on psychopathology implies that the differences between subthreshold symptoms and their full-threshold counterparts are quantitative rather than qualitative. This is supported by studies showing that subthreshold symptoms and full-blown psychiatric disorders have a similar etiology, structure (based on symptom interrelations⁸), and treatment response (*i.e.*, phenomenological continuity⁹). For instance, mild psychiatric traits and disorders share similar genetic risk factors, illustrated by the finding that 80% of the covariance between subthreshold symptoms and psychiatric disorders is attributable to genetic overlap¹⁰. Similarly, the brain regions associated with subthreshold and clinical manifestations of psychopathology are largely overlapping¹¹. Environmental risk factors, such as childhood abuse and stressful life events, have also been related to both sub- and full-threshold expressions of psychopathology^{6,12}. Finally, like psychiatric disorders, subthreshold symptoms are associated with distress and declined functioning^{6,7,12,13}, which can improve following psychological treatment¹⁴. In sum, there is substantive evidence that the distinction between subthreshold symptoms and psychiatric disorders seems to be a matter of *degree* – *e.g.*, reflected in the number of symptoms and affected individuals – rather than *kind*^{15,16}.

Subthreshold symptoms are commonly considered clinically relevant not only because of the above-mentioned similarities to psychiatric disorders, but also because of their prognostic significance¹. That is, individuals with subthreshold symptoms are two to five times more likely to develop a psychiatric disorder compared to individuals without such symptoms^{1,17,18}. This implies that, for some individuals, subthreshold symptoms reflect a temporary phase between having no symptoms and having a psychiatric disorder. Yet, longitudinal cohort studies have shown that for the majority of individuals, subthreshold symptoms do *not* escalate into full-blown disorders. Specifically, the proportion of individuals with subthreshold symptoms that meet the criteria for a psychiatric disorder when assessed several years later ranges between 14–35% (depression^{1,2,19}), 14–15% (anxiety^{1,20}), 32% (bipolar disorder²¹), 25% (psychosis²²), and 36–38% (substance abuse¹). For other individuals,

subthreshold symptoms may either remit or persist. Such persistence contradicts the common notion that subthreshold symptoms are transient. Instead, subthreshold symptoms could – at least for some individuals – reflect stable states rather than transient transitional phenomena. This introduces the possibility of yet another qualitative similarity between subthreshold symptoms and psychiatric disorders: both might be stable phenotypes.

So far, the stability of psychopathological symptoms has mostly been investigated across very short timescales (*e.g.*, hour-to-hour) and relatively long timescales (*e.g.*, year-to-year). The present study aims to investigate the day-to-day stability of psychopathological symptoms across six months using a complex systems perspective^{23–26}. According to this perspective, symptoms might manifest as stable states, for instance labelled as mental health, subthreshold psychopathology, or psychiatric disorder^{27–29}. These stable states – commonly referred to as attractors – can be thought of as the set point to which systems tend to return again and again upon perturbations (*i.e.*, stressful or pleasant events). In healthy individuals, for instance, events may lead to temporary dips or uplifts in mood, but eventually, a state of mental health (*i.e.*, their attractor) is restored. Attractors result from regulatory processes, reflected in interactions between elements of the system (*e.g.*, feedback loops between mental states²⁴). In the presence of strong regulatory processes, systems are resistant to change. This translates to a strong tendency to remain in an attractor (*e.g.*, one with low symptom severity). As regulatory processes weaken, transitions from one attractor to another become more likely. Hence, the stability of an attractor can be inferred from regulatory processes, known as attractor strength. If subthreshold symptoms indeed reflect a stable attractor that behaves similar to the attractors with low and high symptom severity, the strengths of these attractors should be similar. This would mean that there is no clear association between the symptom severity of attractors (referred to as homebases) and attractor strengths (Fig. 7.1a). If, on the other hand, subthreshold symptoms reflect more transient phenomena (*i.e.*, temporary states between low and severe symptoms), there should be a quadratic relation between homebases and attractor strengths (Fig. 7.1b). We investigated this hypothesis in an intensive longitudinal study where 122 at-risk young adults monitored transdiagnostic (subthreshold) symptoms daily for a period of six months. Since subthreshold symptoms are considered diffuse, representing a mix of symptoms from different psychopathological domains, we focused on attractors of overall symptom severity, rather than attractors of specific symptom domains³⁰.

Fig. 7.1. *Illustration of the association between homebases and attractor strengths under two different scenarios.* The homebase corresponds to the severity of symptoms that characterize an attractor. A) If the subthreshold attractor is comparable to the healthy and disordered attractors in strength, there is no clear association between homebases and attractor strengths. B) If the subthreshold attractor is transient, there is a quadratic relation between homebases and attractor strengths.

Materials and methods

Participants

Participants were recruited from the clinical cohort of an ongoing study, named TRacking Adolescents' Individual Lives Survey (TRAILS³¹). At the time of inclusion in the clinical cohort of TRAILS (TRAILS CC), participants were between 10 and 12 years old and had been referred to mental health care services. Because of this history, they were considered at increased risk to develop mental health problems. Since their inclusion, participants completed bi- or tri-annual follow-up assessments. When TRAILS CC participants were approximately 22.6 years old (range = 21-24), they were invited to take part in an add-on diary study (TRAILS TRANS-ID). Of the 443 eligible participants, 134 (30.2%) were included in TRAILS TRANS-ID. The present study included the 122 individuals who completed the diary period. A more elaborate description of these participants, as well as other methodological details, has been published elsewhere (chapter 5)³². All participants provided written informed consent. This study was approved by the medical ethics committee of the University Medical Center Groningen (reference no. 2017/203). All procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation

and with the Helsinki Declaration of 1975, as revised in 2008.

Procedure

Participants completed daily questionnaires every evening for a period of six consecutive months. Each questionnaire consisted of 58 items pertaining to the past day (*e.g.*, “Today, I felt tired”) that were rated on a visual analogue scale (VAS) ranging from 0 to 100. These questionnaires, or diaries, were sent via text messages to participants’ mobile phones. Prior to and immediately after the diary period (*i.e.*, at baseline and post), a semi-structured diagnostic interview was orally administered (mini-SCAN). This interview was used to assess whether individuals met the diagnostic criteria for a DSM-5 disorder (for details on the procedure, please see Schreuder *et al.*³³ or chapter 5). The post assessment covered the entire diary period, and therefore, this assessment was used for sensitivity analyses (see data analysis and supplement 7.7 *Sensitivity analyses 2: DSM-5 diagnoses*).

Data analysis

The diary procedure yielded a maximum of 183 measurements of 58 mental states per individual, for 122 individuals (*i.e.*, >1.2 million observations in total). The data pertaining to the 35 negative mental states assessed in our study – listed in supplement 7.1 *List of mental states* – were selected for analyses. Together, these mental states were considered reflective of individuals’ overall symptom severity. We estimated overall symptom severity (sx) for individual i at time t by computing the mean rating across the individual’s negative mental states (ms) at time t , so that $sx_{i,t} = \sum ms_{i,t} / 35$. Subsequently, a generalized additive mixed model (GAMM) was fitted^{34,35}. Specifically, symptom severity of individual i at time t was predicted by this individual’s (i) intercept, (ii) autoregressive parameter, and (iii) non-linear trend in symptom severity over time (for details, see supplement 7.2 *GAMM details*). This model yielded an estimated homebase and attractor strength for each individual separately, while taking into account each individual’s change in symptoms over time. The homebase is given by the person-specific intercept (which is conceptually similar, but not equal to, the person’s mean), and reflects the symptom severity that characterizes an individual’s attractor³⁶. As such, relatively low homebases can be considered adaptive, whereas higher homebases may be maladaptive. The attractor strength reflects the regulatory forces that maintain the attractor, and is given by person-specific estimates of the inversed autoregressive parameter (*i.e.*, the effect of symptom severity at $t-1$ on symptom severity at t)³⁶. This operationalization

of homebases and attractor strengths has also been adopted in earlier studies^{36,37}, and can be considered a discrete-time translation of the parameters described in the DynAffect model³⁸ and the PersDyn model³⁹, which are formalized in continuous time.

The relation between homebases and attractor strengths was tested with regression analyses. Specifically, we compared models where attractor strength was predicted by homebase vs. squared homebase (*i.e.*, polynomial regression). This allowed for differentiating between the scenarios displayed in Fig. 7.1. We repeated these regression analyses in two sensitivity analyses. First, we fitted weighted regressions to account for the uncertainty in the estimates and attractor strengths. The weights in these models were proportional to the sum of the range of the confidence intervals around the homebases and attractor strengths. Second, we checked the effect of (co-morbid) full-blown disorders by omitting individuals who met the criteria for at least one DSM-5 diagnosis from the analyses. This was done to allow for the possibility that mental states, and the stability thereof, might have a different meaning for individuals with versus without psychiatric disorders⁴⁰. By re-running analyses in individuals without disorders, we could verify whether findings followed from between-individual differences in *e.g.*, the “threshold” for reporting a certain mental state. Individuals with a DSM-5 diagnosis were selected based on the mini-SCAN assessed at post, which covered the presence of psychiatric disorders during the entire diary period. All analyses were performed in R⁴¹ (version 4.0.2) using the package *mgcv* (version 1.8.33).

Results

Participants (N=122, 56.6% male) were on average 23.64 years old ($SD = 0.67$, range = 22.26–24.81) and had on average completed 163.39 diary assessments (88.6%, $SD = 17.12$, range = 116–190). At baseline, 40 individuals (32.79%) met the criteria for at least one DSM-5 diagnosis. After the diary period, 34 individuals (27.87%) had a DSM-5 disorder, of whom 23 (67.65%) were also diagnosed at baseline. Most prevalent were mood disorders (N=28 and 23 at baseline and post, respectively), followed by anxiety disorders (N=21 and 12) and ADHD (N=8 and 8).

The fitted values and the distribution of residuals indicated that assumptions of the GAMM were not violated (see supplement 7.3 *Assumptions of GAMM* and supplement 7.4 *Fitted values*). The GAMM had an adjusted R^2 of 77% and yielded homebases that varied between 2.85 and 46.51, with a mean of 17.81 ($SD = 9.80$). Attractor strengths varied

between 1.52 and 28.83 (mean = 4.16, $SD = 3.35$). Individuals who met criteria for a DSM-5 diagnosis at post had a higher homebase (mean = 21.57) compared to non-diagnosed individuals (mean = 16.36, $t(120) = 2.70$, $p < .01$, Cohen's $d = 0.55$), but did not differ in terms of attractor strength (mean = 4.26 vs. 4.12, respectively; $t(120) = 0.21$, $p = .84$, Cohen's $d = 0.04$).

Regression analyses indicated that there was no clear association between homebase and attractor strength (linear model: $B = 0.02$, $p = .47$, $R^2 < .01$; polynomial model: $B < 0.01$, $p = .61$, $R^2 < .01$; Fig. 7.2). This finding did not change after taking into account the uncertainty in the estimates (supplement 7.6 *Sensitivity analyses 1: uncertainty of homebase and attractor strength estimates*) nor after removing individuals with a DSM-5 diagnosis from the analyses (supplement 7.7 *Sensitivity analyses 2: DSM-5 diagnoses*).

Fig. 7.2. *Association between the homebase and attractor strength of symptoms of psychopathology.* Homebases and attractor strengths were estimated from a generalized additive mixed model using six months of daily diary data from 122 young adults. Individuals who received any DSM-5 diagnosis after the diary period are printed in blue. The black line shows the association between homebases and attractor strengths based on a linear model; the grey line shows the fit of a polynomial model. Neither model indicated an association between homebases and attractor strengths. For illustrative purposes, three outliers (individuals with an attractor strength of >10) were omitted from this figure. Including these individuals did not change the results (see the supplement, Fig. S7.2).

Discussion

Symptoms of psychopathology have been proposed to lie on a severity continuum, where the absence of symptoms and psychiatric disorders mark the extreme ends. This has been supported by the notion that subthreshold and full-threshold psychopathological symptoms show a comparable etiology^{6,9–12} and treatment outcome¹⁴. This study investigated the stability of psychopathological symptoms, *i.e.*, their attraction strength, along the severity continuum. We found that the stability of symptoms assessed daily over a period of six months is independent of the severity of symptoms. This provides additional support for a dimensional view on psychopathology, which implies that subthreshold and full-threshold psychopathological symptoms differ in degree (*i.e.*, severity) rather than in kind (*e.g.*, stability). In conclusion, just like some individuals may experience constant mental health or psychopathology, others may get stuck in subthreshold psychopathology. Subthreshold symptoms may thus resemble stable states, rather than transient conditions that mark the progression from relatively healthy towards disordered states (or vice versa).

A dimensional view on psychopathology does not necessarily preclude the existence of discrete, stable states along the severity continuum^{15,42}. Present findings show that such states not only lie on the extreme ends of the continuum – reflecting mental health and mental disorder – but may just as well occupy the regions in between these extremes – reflecting subthreshold psychopathology. It follows that the clinical relevance of subthreshold symptoms does not just lie in their associated burden^{3,6,7} and their tendency to precede full-threshold symptoms^{1,17,18} (their prognostic significance), but also in their stability. Stability here refers to a property of an attractor in a complex dynamic system, namely attractor strength. In the context of psychopathology, an attractor can be considered a set of mental states to which a system tends to return upon perturbations (*e.g.*, pleasant/stressful events^{23,27,29}). An attractor has a certain *homebase*, which may describe mild vs. more severe psychopathological symptoms, and *strength*, which reflects the regulatory processes that maintain the attractor. Relatively strong (stable) attractors with low homebase can be considered adaptive, as they illustrate a healthy system that is resilient to external perturbations. In contrast, attractors with higher symptom severity may be maladaptive, illustrated by the current finding that individuals with a DSM-5 disorder had higher homebases than those without a disorder.

We have shown that maladaptive attractors do not differ from more adaptive attractors in terms of strength. Yet, previous work has reported that individuals with a

psychiatric disorder may have weaker attractors compared to healthy controls, implying a negative association between attractor strengths and homebases^{36,37}. Similarly, studies that used alternative measures of stability (*i.e.*, adjusted square of successive differences^{43–45} and probability of acute change^{44,45}) found higher instability in patients compared to controls. However, this difference was likely driven by the standard deviation, meaning that patients and controls may differ primarily in the dispersion of mental states as opposed to the stability of mental states^{43,46}. An explanation for the discrepancy between earlier and present findings could be that the at-risk youth in the present study are more impaired than the healthy controls and/or less impaired than the patient samples in former studies^{36,37}, which in turn may have restricted the ranges in attractor strengths and homebases in the current study. However, the large variability in symptom severity in the present sample suggests that a restricted range of psychopathological symptom severity is unlikely to underlie current findings. Instead, the differences across studies concerning attractor strengths potentially follow from considerable differences in sampling frequency (*i.e.*, assessments with a 1-day interval vs. 15-minute/1-hour interval) and duration (*i.e.*, six months vs. one to two days): although individuals with psychiatric disorders may have a lower hour-to-hour stability of emotions compared to non-affected individuals^{36,37}, their day-to-day stability of mental states may be, as indicated by the current findings, similar. Hence, whereas stability on a short timescale could be adaptive – for instance because it signals adequate emotional responsivity to environmental cues⁴⁷ – the meaning of stability on a longer timescale depends on the homebase that is maintained. Specifically, stability on a longer timescale can be either adaptive (when it maintains mental health) or maladaptive (when it maintains sub- or full-threshold psychopathological symptoms). In conclusion, the meaning of psychological dynamics – such as the stability of mental states – crucially depends on the timescale under consideration.

Besides the timescale of assessments, the dynamics of psychopathological symptoms might be affected by the type of symptoms under consideration. It could for instance be hypothesized that certain symptom domains (*e.g.*, anxiety) are more stable than others (*e.g.*, psychosis, mania⁴⁸). At present, little is known about such between-domain differences: it has been reported that panic disorder and major depression show higher homebases (but similar attractor strengths) compared to borderline, post-traumatic stress and eating disorders^{37,49}, whereas negative psychotic symptoms may have a stronger attraction (but similar homebase) compared to positive symptoms of psychosis⁵⁰. However,

small sample sizes and methodological heterogeneity preclude firm conclusions. To investigate dissociations between homebases and attractor strengths across clinical stages and psychopathological domains, future studies should aim to include individuals with a wide range of symptoms of varying severity. The current study did so by including youth who experienced a widely varying degree of (mental health) problems and a wide variety of mental states.

Unlike the majority of earlier studies on subthreshold symptoms, the present study considered attractors on a continuum of symptom severity, and did not classify individuals into subgroups based on pre-set cut-offs. This is particularly advantageous given the considerable heterogeneity in definitions of “subthreshold” psychopathology that plagues research on this topic^{7,13,51}. Arguably, our decision to not categorize came at the cost of an unclear clinical significance of the homebase estimates, which were based on daily ratings of negative mental states. However, the fact that individuals with a DSM-5 diagnosis had significantly higher homebases than those without a diagnosis supports our inferences. Another potential limitation of the current study is that the aggregation of symptoms into global psychopathology might have obscured domain-specific associations between the homebase and strength of attractors. However, our operationalization was in line with the notion that subthreshold psychopathology may not be domain-specific³⁰, and therefore, fitted with our aim to study the dynamics of symptoms of varying severity. Finally, our estimates of attractor strengths require that the timescale of assessments (daily) matches the timescale of the process of interest (*i.e.*, strength of attraction, or the speed with which a homebase is restored). This issue is not specific to the current study, but rather applicable to all intensive longitudinal studies: within-person dynamics (including homebases and attractor strengths) can only be estimated with sufficient sampling frequency⁵². Although the present timescale (daily) is in line with our interest in long-term stability of symptoms – as opposed to momentary fluctuations in emotions^{43,46} – further work on the role of timescales in studies on symptom dynamics is hopefully awaited (for a recent example, see Sperry and Kwapił⁴⁴).

The lack of an association between homebases and attractor strengths found in the present study implies that individuals can get stuck anywhere on the severity continuum. Attractors do not, however, eternally persist: they may change over time, and such changes may involve a shift from subthreshold to full-threshold psychopathological symptoms or vice versa. Future research is needed to establish what triggers such shifts. After a shift towards a maladaptive attractor (one with a high homebase) has occurred, it is imperative to understand

what maintains the attractor. A complex systems perspective on psychopathology implies that attractors emerge from interactions between mental states – meaning that individuals with stronger attractors would be expected to show greater connectivity between mental states²⁴. An alternative avenue for further research concerns the comparison of attractors of different domains of psychopathology, which could expose how specific domains progress and persist, and may inform treatment.

Supplement 7

7.1 List of mental states

We used the data of 35 mental states, listed below.

1. I felt lonely
2. I felt empty
3. I felt down
4. I felt guilty
5. I could not bring myself to do anything
6. I felt tired
7. I avoided things
8. I was afraid of making mistakes
9. I had difficulty making decisions
10. I worried
11. My thoughts would not leave me alone
12. I was easily distracted
13. I felt restless
14. I felt nervous
15. I was easily startled
16. I felt stressed
17. I felt overburdened
18. I was easily upset
19. I felt overstimulated
20. It bothered me that things did not go as expected
21. I was bored quickly
22. I felt impatient
23. I was unable to sit still
24. I was impulsive
25. I felt irritated
26. Others felt annoyed by me
27. I had a fight
28. I felt rebellious
29. Things did not bother me

30. I had moodswings
31. I felt strange
32. I felt suspicious
33. I experienced physical pain
34. I experienced physical discomfort
35. I'm dreading something that will happen soon

7.2 GAMM details

The GAMM allows for estimating smooth non-linear relationships that may differ between individuals^{53–55}. This non-linearity is important to model because, in data that span six months, non-linear trends in symptom severity may be common. GAMMs estimate such non-linear relations using smooths, which comprise a set of basis functions (in this case, thin-plate regression splines³⁴). The wiggleness of each smooth term is determined by the number of basis functions, which was set to 10, and a penalization parameter, which was selected by generalized cross-validation. To avoid overfitting, we let the model use the same penalization parameter for all individuals. Hence, we assumed equal wiggleness (*i.e.*, an equal amount of non-linearity) across individuals. Note that this does not pose restrictions on the shape of parameters, such as an individual's trend in symptoms over time⁵⁴.

We fitted a model where symptom severity at time t was predicted by (1) an intercept, (2) person-mean centered symptom severity at $t-1$, and (3) a smooth function of centered time. Between-individual variation in these three parameters was estimated through including random effects (1 and 2) and through a factor-smooth interaction (3). The latter allowed for estimating non-linear trends in symptoms over time that differed between individuals⁵⁴. GAMM assumes that parameters do not change in a stepwise fashion, which was checked visually using plots of the fitted and observed values. Other assumptions of the GAMM, *i.e.*, that residuals are normally distributed with a mean of 0, that residuals are homoscedastic across levels of the linear predictor, and that the number of basis functions is sufficient, were checked visually using residuals diagnostics plots and a diagnostic test provided in the *mgcv* package⁵⁵ (see assumptions of GAMM).

Homebase and attractor strength estimation

In accordance with the DynAffect model³⁸, homebases were operationalized as the person-specific intercept (random effect 1 above) and attractor strengths were given by the inverse

of the person-specific autoregressive coefficient (random effect 2 above). Both were independent of individuals' trend in symptoms over time (effect 3 above). Confidence intervals around the homebase and attractor strength were computed by bootstrapping. The association between homebases and attractor strengths was examined based on regression models.

7.3 Assumptions of GAMM

Figure S7.1 was obtained using the *gam.check* function from *mgcv*. The left plots (S7.1a, S7.1c) show that the deviance residuals were approximately normally distributed. The departure from the straight line in the upper left plot likely results from the fact that symptom assessments were bounded between 0 and 100. Plot S7.1b suggests that the assumption of homoscedastic residuals was not violated. Finally, we checked whether the number of basis functions (10) was too low using the *k.check* function. The output showed edf-values ranging between 2 and 4 (all P-values > .05), and hence, the number of basis functions was considered sufficient.

Fig. S7.1. *Diagnostic information about the fit of the generalized additive mixed model.* A) QQ-plot of deviance residuals. B) Residuals against the linear predictor, to evaluate the assumption of homoscedastic residuals. C) Histogram of residuals, to evaluate the assumption of normally distributed residuals. D) Outcome against fitted values.

7.4 Fitted values

These plots on the next pages show the observed values (grey dots) and fitted values (blue lines). Each plot corresponds to a single individual, resulting in 122 plots. For each plot, the x-axis depicts time (6 months, ranging from observation 1 to observation 183) and the y-axis denotes symptom severity (the outcome of the generalized additive mixed models we fitted).

7.5 Figure 2 including all individuals

Figure 7.2 in the main text excluded three individuals whose attractor strength exceeded 10 (corresponding to an autoregressive coefficient of 0.10). The below figure (Fig. S7.2) does include these individuals. Note that the outlier estimates were very uncertain, with 95% confidence intervals ranging between -239.23 and 352.35. The mean absolute range of the 95% confidence intervals was 336.02 in the three individuals with an attractor strength > 10, and 10.27 in all other individuals (N=118). Given this uncertainty, we considered it justified to remove the estimates from the figure in the main chapter.

Fig. S7.2. Association between the homebase and attractor strength of symptoms of psychopathology. Here, we included all individuals (*i.e.*, also outliers).

7.6 Sensitivity analyses 1: uncertainty of homebase and attractor strength estimates

Homebases and attractor strengths were estimated with varying uncertainty, which was quantified by 95% confidence intervals obtained through bootstrapping. We used the inverse of the span of these confidence intervals to compute weights. These weights were then added to the regression analyses. Similar to our main analyses, the results did not indicate an association between homebases and attractor strengths (linear model: $B < 0.01$, $p = .69$, $R^2 < 0.01$; polynomial model: $B < 0.01$, $p = .50$, $R^2 < 0.01$; see figure S7.3 below).

Fig. S7.3. Association between the homebase and attractor strength of symptoms of psychopathology. Here, we weighed observations according to their uncertainty.

7.7 Sensitivity analyses 2: DSM-5 diagnoses

To examine whether our results were affected by the inclusion of individuals with psychiatric disorders, we re-ran the simple and polynomial regression analyses with attractor strength as outcome and (squared) homebase as predictor and excluded 34 individuals who met criteria for a DSM-5 diagnosis during the diary period from the analyses. This resulted in a subsample of 88 non-diagnosed individuals. Similar to our findings, the regression models indicated no association between homebases and attractor strengths (linear model: $B = 0.03$, $p = .46$, $R^2 = 0.01$; polynomial model: $B < 0.01$, $p = .53$, $R^2 < 0.01$; see figure S7.4 below).

Fig. S7.4. Association between the homebase and attractor strength of symptoms of psychopathology. Here, we only included individuals who did not meet diagnostic criteria.

References

1. Shankman SA, Lewinsohn PM, Klein DN, Small JW, Seeley JR, Altman SE. Subthreshold conditions as precursors for full syndrome disorders: a 15-year longitudinal study of multiple diagnostic classes. *Journal of Child Psychology and Psychiatry*. 2009;50(12):1485-1494. doi:10.1111/j.1469-7610.2009.02117.x.Subthreshold
2. Fergusson DM, Horwood LJ, Ridder EM, Beautrais AL. Subthreshold depression in adolescence and mental health outcomes in adulthood. *Archives of General Psychiatry*. 2005;62:66-72.
3. Keenan K, Hipwell A, Feng XIN, et al. Subthreshold symptoms of depression in preadolescent girls are stable and predictive of depressive disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2008;47(12):1433-1442. doi:10.1097/CHI.0b013e3181886eab
4. McGorry PD, Hartmann JA, Spooner R, Nelson B. Beyond the “at risk mental state” concept: transitioning to transdiagnostic psychiatry. *World Psychiatry*. 2018;17(June):133-142.
5. McGorry PD, Hickie IB, Yung AR, Pantelis C, Jackson HJ. Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. *Australian and New Zealand Journal of Psychiatry*. 2006;40(8):616-622. doi:10.1111/j.1440-1614.2006.01860.x
6. Wesselhoeft R, Sørensen MJ, Heiervang ER, Bilenberg N. Subthreshold depression in children and adolescents - a systematic review. *Journal of Affective Disorders*. 2013;151(1):7-22. doi:10.1016/j.jad.2013.06.010
7. Haller H, Cramer H, Lauche R, Gass F, Dobos GJ. The prevalence and burden of subthreshold generalized anxiety disorder: a systematic review. *BMC Psychiatry*. 2014;14(128). doi:10.1186/1471-244X-14-128
8. Groen RN, Wichers M, Wigman JTW, Hartman CA. Specificity of psychopathology across levels of severity: a transdiagnostic network analysis. *Scientific Reports*. 2019;9(1):1-10. doi:10.1038/s41598-019-54801-y
9. Markon KE, Chmielewski M, Miller CJ. The reliability and validity of discrete and continuous measures of psychopathology: a quantitative review. *Psychological Bulletin*. 2011;137(5):856-879. doi:10.1037/a0023678
10. Taylor MJ, Martin J, Lu Y, et al. Association of genetic risk factors for psychiatric disorders and traits of these disorders in a Swedish population twin sample. *JAMA Psychiatry*. 2019;76(3):280-289. doi:10.1001/jamapsychiatry.2018.3652
11. Besteher B, Gaser C, Nenadić I. Brain structure and subclinical symptoms: a dimensional perspective of psychopathology in the depression and anxiety spectrum. *Neuropsychobiology*. 2020;79(4-5):270-283. doi:10.1159/000501024
12. Meeks TW, Vahia I v., Lavretsky H, Kulkarni G, Jeste D v. A tune in “a minor” can “b major”: a review of epidemiology, illness course, and public health implications of subthreshold depression in older adults. *Journal of Affective Disorders*. 2011;129(1-3):126-142. doi:10.1016/j.jad.2010.09.015
13. Balázs J, Keresztény Á. Subthreshold attention deficit hyperactivity in children and adolescents: a systematic review. *European Child and Adolescent Psychiatry*. 2014;23(6):393-408. doi:10.1007/s00787-013-0514-7
14. Cuijpers P, Smit F, van Straten A. Psychological treatments of subthreshold depression: a meta-analytic review. *Acta Psychiatrica Scandinavica*. 2007;115(6):434-441. doi:10.1111/j.1600-0447.2007.00998.x
15. Meehl PE. Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of*

- Personality*. 1992;60(1):117-174. doi:10.1111/j.1467-6494.1992.tb00269.x
16. Flett GL, Vredenburg K, Krames L. The continuity of depression in clinical and nonclinical samples. *Psychological Bulletin*. 1997;121(3):395-416. doi:10.1037/0033-2909.121.3.395
17. Lee YY, Stockings EA, Harris MG, et al. The risk of developing major depression among individuals with subthreshold depression: a systematic review and meta-analysis of longitudinal cohort studies. *Psychological Medicine*. 2018;49:92-102. doi:10.1017/s0033291718000557
18. Bertha EA, Balazs J. Subthreshold depression in adolescence: a systematic review. *European Journal of Child and Adolescent Psychiatry*. 2013;22:589-603. doi:10.1007/s00787-013-0411-0
19. Tuithof M, ten Have M, van Dorsselaer S, Kleinjan M, Beekman A, de Graaf R. Course of subthreshold depression into a depressive disorder and its risk factors. *Journal of Affective Disorders*. 2018;241:206-215. doi:10.1016/j.jad.2018.08.010
20. Bosman RC, ten Have M, de Graaf R, Muntingh AD, van Balkom AJ, Batelaan NM. Prevalence and course of subthreshold anxiety disorder in the general population: a three-year follow-up study. *Journal of Affective Disorders*. 2019;247:105-113. doi:10.1016/j.jad.2019.01.018
21. Papachristou E, Oldehinkel AJ, Ormel J, et al. The predictive value of childhood subthreshold manic symptoms for adolescent and adult psychiatric outcomes. *Journal of Affective Disorders*. 2017;212:86-92. doi:10.1016/j.jad.2017.01.038
22. Kaymaz N, Drukker M, Lieb R, et al. Do subthreshold psychotic experiences predict clinical outcomes in unselected non-help-seeking population-based samples? A systematic review and meta-analysis, enriched with new results. *Psychological Medicine*. 2012;42(11):2239-2253. doi:10.1017/S0033291711002911
23. Wichers M, Schreuder MJ, Goekoop R, Groen RN. Can we predict the direction of sudden shifts in symptoms? Transdiagnostic implications from a complex systems perspective on psychopathology. *Psychological Medicine*. 2019;49(3):380-387. doi:10.1017/S0033291718002064
24. Cramer AOJ, van Borkulo CD, Giltay EJ, et al. Major depression as a complex dynamic system. *PLoS one*. 2016;11(12):1-20. doi:10.1371/journal.pone.0167490
25. Schiepek G. Complexity and nonlinear dynamics in psychotherapy. *European Review*. 2009;17(2):331-356. doi:10.1017/S1062798709000763
26. Olthof M, Hasselman F, Strunk G, et al. Critical fluctuations as an early-warning signal for sudden gains and losses in patients receiving psychotherapy for mood disorders. *Clinical Psychological Science*. 2019;8(1):1-11. doi:10.1177/2167702619865969
27. Jeronimus BF. Dynamic system perspectives on anxiety and depression. In: Kunnen ES, de Ruiter NMP, Jeronimus BF, van der Gaag MA, eds. *Psychosocial Development in Adolescence: Insights from the Dynamic Systems Approach*. Routledge Psychology; 2019. doi:10.1128/MCB.01405-08
28. Shapiro Y, Scott JR. Dynamical systems therapy (DST): complex adaptive systems in psychiatry and psychotherapy. In: Mittleton-Kelly E, Paraskevas A, Day C, eds. *Handbook of Research Methods in Complexity Science*. Edward Elgar Publishing Limited; 2018:567-590.
29. Hayes AM, Andrews LA. A complex systems approach to the study of change in psychotherapy. *BMC Medicine*. 2020;18(1):1-13. doi:10.1186/s12916-020-01662-2
30. van Os J. The dynamics of subthreshold psychopathology: implications for diagnosis and treatment. *American Journal of Psychiatry*. 2013;170:695-698.
31. Huisman M, Oldehinkel AJ, de Winter A, et al. Cohort Profile: The Dutch "TRacking Adolescents" Individual Lives' Survey'; TRAILS. *International Journal of Epidemiology*. 2008;37(6):1227-1235. doi:10.1093/ije/dym273

32. Schreuder MJ, Groen RN, Wigman JTW, Hartman CA, Wichers M. Measuring psychopathology as it unfolds in daily life: addressing key assumptions of intensive longitudinal methods in the TRAILS TRANS-ID study. *BMC Psychiatry*. 2020;20(1). doi:10.1186/s12888-020-02674-1
33. Schreuder MJ, Groen RN, Wigman JTW, Hartman CA, Wichers M. Measuring psychopathology as it unfolds in daily life : addressing key assumptions of intensive longitudinal methods in the TRAILS TRANS-ID study. *BMC Psychiatry*. 2020;20(1):1-14.
34. Wood SN. *Generalized Additive Models: An Introduction with R.*; 2006.
35. Hastie T, Tibshirani R. Generalized additive models. *Statistical Science*. 1986;1(3):297-318. doi:10.2307/2246134
36. Ebner-Priemer UW, Houben M, Santangelo P, et al. Unraveling affective dysregulation in borderline personality disorder: a theoretical model and empirical evidence. *Journal of Abnormal Psychology*. 2015;124(1):186-198. doi:10.1037/abn0000021
37. Santangelo PS, Limberger MF, Stiglmayr C, et al. Analyzing subcomponents of affective dysregulation in borderline personality disorder in comparison to other clinical groups using multiple e-diary datasets. *Borderline Personality Disorder and Emotion Dysregulation*. 2016;3(1):1-13. doi:10.1186/s40479-016-0039-z
38. Kuppens P, Oravecz Z, Tuerlinckx F. Feelings change: accounting for individual differences in the temporal dynamics of affect. *Journal of Personality and Social Psychology*. 2010;99(6):1042-1060. doi:10.1037/a0020962
39. Sosnowska J, Kuppens P, de Fruyt F, Hofmans J. A dynamic systems approach to personality: the Personality Dynamics (PersDyn) model. *Personality and Individual Differences*. 2019;144(February):11-18. doi:10.1016/j.paid.2019.02.013
40. Helmchen H, Linden M. Subthreshold disorders in psychiatry: clinical reality, methodological artifact, and the double-threshold problem. *Comprehensive Psychiatry*. 2000;41(2):1-7. doi:10.1016/S0010-440X(00)80001-2
41. R Core Team. R: A language and environment for statistical computing. Published online 2020.
42. Linscott RJ, van Os J. Systematic reviews of categorical versus continuum models in psychosis: evidence for discontinuous subpopulations underlying a psychometric continuum. Implications for DSM-V, DSM-VI, and DSM-VII. *Annual Review of Clinical Psychology*. 2010;6:391-419. doi:10.1146/annurev.clinpsy.032408.153506
43. Koval P, Pe ML, Meers K, Kuppens P. Affect dynamics in relation to depressive symptoms: variable, unstable or inert? *Emotion*. 2013;13(6):1132-1142. doi:10.1037/a0033579
44. Sperry SH, Kwapil TR. Bipolar spectrum psychopathology is associated with altered emotion dynamics across multiple timescales. *Emotion*. Published online 2020:1-14. doi:10.1037/emo0000759
45. Trull TJ, Lane SP, Koval P, Ebner-Priemer UW. Affective dynamics in psychopathology. *Emotion Review*. 2015;7(4):355-361. doi:10.1177/1754073915590617
46. Bos EH, de Jonge P, Cox RFA. Affective variability in depression: revisiting the inertia–instability paradox. *British Journal of Psychology*. 2018;110(4):814-827. doi:10.1111/bjop.12372
47. Kashdan TB, Rottenberg J. Psychological flexibility as a fundamental aspect of health. *Clinical Psychology Review*. 2010;30(4):865-878. doi:10.1016/j.cpr.2010.03.001
48. Sperry SH, Walsh MA, Kwapil TR. Emotion dynamics concurrently and prospectively predict mood psychopathology. *Journal of Affective Disorders*. 2020;261:67-75. doi:10.1016/j.jad.2019.09.076

49. Heller AS, Davidson RJ, Fox AS. Parsing affective dynamics to identify risk for mood and anxiety disorders. *Emotion*. 2019;19(2):283-292. doi:10.1037/emo0000440
50. Westermann S, Grezellschak S, Oravec Z, Moritz S, Lüdtkke T, Jansen A. Untangling the complex relationships between symptoms of schizophrenia and emotion dynamics in daily life: findings from an experience sampling pilot study. *Psychiatry Research*. 2017;257:514-518. doi:10.1016/j.psychres.2017.08.023
51. Cuijpers P, Smit F. Subthreshold depression as a risk indicator for major depressive disorder: a systematic review of prospective studies. *Acta Psychiatrica Scandinavica*. 2004;109(5):325-331. doi:10.1111/j.1600-0447.2004.00301.x
52. Haslbeck JMB, Ryan O. Recovering within-person dynamics from psychological time series. *Multivariate Behavioural Research*. 2021;56(2):1-32. doi:10.1080/00273171.2021.1896353
53. Wieling M. Analyzing dynamic phonetic data using generalized additive mixed modeling: a tutorial focusing on articulatory differences between L1 and L2 speakers of English. :1-51.
54. Pedersen EJ, Miller DL, Simpson GL, Ross N. Hierarchical generalized additive models in ecology: an introduction with mgcv. *PeerJ*. 2019;2019(5). doi:10.7717/peerj.6876
55. Wood S. *Generalized Additive Models: An Introduction with R*. 2nd ed. Taylor & Francis Group; 2017.

8

Anticipating transitions in mental health in at-risk youth: a six-month daily diary study into early warning signals

Abstract

If psychopathology behaves like a complex dynamic system, transitions in mental health, such as the onset or sudden worsening of symptoms, may be preceded by early warning signals (EWS). EWS could thus reflect personalized warning signals for impending psychopathology. We empirically investigated this hypothesis in at-risk youth ($N=122$, mean age 23.6 years, $SD = 0.7$, 57% males) from the clinical cohort of Tracking Adolescents' Individual Lives Survey (TRAILS CC), who provided daily mental state assessments for six months. We analyzed whether EWS (rising autocorrelations in mental states) preceded drops in mental health. EWS had low sensitivity (5%) and high specificity (96%). Thus, in the present sample, the proposed generic nature and clinical utility of EWS could not be substantiated. This calls for a more nuanced view on the application of complex dynamic systems principles to psychopathology and raises several hypotheses about when and for whom such principles may hold.

Introduction

Mental health varies over time: periods without symptoms of psychopathology may be followed by periods with symptoms, and vice versa¹. Alternations between such periods can occur suddenly and are difficult to foresee. For instance, relapse of depression or onset of psychosis can occur abruptly, in absence of an obvious trigger². This challenges early detection of emerging psychopathological disorders and hampers prevention and early intervention efforts.

A complex systems perspective on psychopathology provides a novel view on the early detection of progressing symptoms. According to this perspective, mental health can be described by alternative states (*e.g.*, mental health vs. mental ill-health). When these states are stable, they are resilient to external perturbations (Fig. S8.1). This means that stressful situations (perturbations) trigger only brief negative mental states⁸ or symptoms of psychopathology. As the instability of mental health states accumulates, the system's resilience to perturbations declines (a phenomenon referred to as critical slowing down)^{3,4}. This means that the negative mental states evoked by perturbations increasingly linger over time. Eventually, the system may reach its tipping point: symptoms no longer recede, and a transition towards mental ill-health takes place. The critical aspect is that, at this point, even small, seemingly innocent perturbations could trigger a transition – and therefore, it may be experienced as sudden and unpredictable. Yet, in a wide range of complex systems, such transitions can be anticipated by monitoring the gradual rise of instability within the system⁵⁻⁸.

The stability of complex systems can be inferred from certain dynamics in time series data. Specifically, rising instability coincides with rising autocorrelations, variances, and cross-correlations⁹. These measures are collectively referred to as early warning signals (EWS) and have been shown to anticipate a variety of transitions, including sudden crashes of financial markets¹⁰, shifts between sleep stages¹¹, epileptic seizures¹², the extinction of species¹³, the outbreak of a global pandemic¹⁴, and climate changes¹⁵. If psychopathology behaves similar to these other complex systems, transitions between mental health states

⁸ Throughout this chapter, 'negative mental states' refers to (subthreshold) symptoms of psychopathology, such as feeling down, anxious, or irritated. Mental health states, in contrast, refers to the global equilibria between which sudden transitions may occur (*e.g.*, mental health vs. depression).

could be anticipated by increasing lingering of mental states (autocorrelation) as well as rising amplitudes or frequencies of fluctuations (variation) and increasingly strong cross-correlations in mental states. This means that EWS could prospectively indicate individuals' risk for the onset or worsening of psychopathology.

The suggestion that EWS may be informative of psychopathology fits well with the finding that individuals who report higher symptom severity also tend to experience stronger autocorrelations^{16–19}, variance^{19,20}, and cross-correlations in negative mental states^{19,21}. The prospective associations between these respective dynamics and psychopathology are still ambiguous, which might be due to methodological heterogeneity across studies (*e.g.*, the time interval between assessments of mental states and assessments of symptom severity; characteristics of sample and design)^{20,22–24}. Notably, the aforementioned studies all investigated between-person processes (*i.e.*, Do individuals with relatively large symptom transitions also report relatively high levels of EWS?). Complex systems principles, in contrast, concern within-person processes (*i.e.*, Are symptom transitions preceded by rising EWS *within* individuals?). To date, the only confirmation of the latter question came from two case studies^{25,26}. Thus, although earlier studies were mostly in line with what would be expected based on complex systems principles, extensive empirical support for EWS as personalized risk markers for future transitions in psychopathology is still lacking.

In order to evaluate complex systems principles in psychopathology at a within-person level, mental health should be monitored in real-time, as it unfolds within individuals. This is practically challenging, as it requires individuals to register mental states frequently (*e.g.*, daily) during a period where transitions in mental health are likely to occur^{25,26}. The two earlier case studies met this requirement, and found that EWS – indexed by autocorrelations in negative mental states – preceded a relapse in depression^{25,26}. This suggests that EWS might reveal psychopathological vulnerability for a specific individual at a specific moment in time. These initial findings call for a more systematic investigation into the sensitivity and specificity of EWS – which will ultimately help to determine the utility of EWS to clinical practice. The current study therefore investigated whether EWS anticipate transitions in mental health in a large sample of young adults at increased risk for psychopathology. In line with earlier studies^{25,26}, we focused on rising autocorrelations in negative mental states as EWS. Our first aim was to establish the sensitivity and specificity of EWS. Based on daily diary data covering a six-month period from 122 individuals, we investigated how often transitions in mental health were preceded by EWS in negative mental states (sensitivity).

We also analyzed how often EWS emerged in individuals who did not undergo transitions in mental health (specificity). Second, we aimed to investigate the probability of a transition in mental health upon detection of EWS (predictive values^h). Clinically useful warning signals should not only be sensitive, specific, and highly predictive, but also occur early and be hard to miss (*i.e.*, strong). Therefore, our third aim was to compare the timing and strength of EWS across the negative mental states wherein they manifested. Since EWS for transitions in mental health have not been investigated in a sample with a size and clinical profile similar to ours before, this study had an exploratory character. In acknowledgement of the wealth of analytical options for data analysis, however, analyses were preregistered (see <https://www.osf.io/fumrx>).

Methods

Procedure

Data were retrieved from the TRAILS TRANS-ID study, which has been described in detail elsewhere²⁷. Briefly, the study included 134 participants from the clinical cohort of the TRacking Adolescents' Individual Lives Survey (TRAILS)²⁸. TRAILS is an ongoing, prospective cohort study with bi- or triennial assessments investigating the development of mental health from pre-adolescence into adulthood. Participants enrolled in the clinical cohort of TRAILS (TRAILS CC) had been referred to a child psychiatric outpatient clinic in the Northern Netherlands any time before the age of 11 years. Because of this history, they are considered at increased risk for psychopathology.

TRAILS TRANS-ID (Transitions In Depression) was conducted as an add-on study to the regular assessment waves of TRAILS CC. Of the 443 participants who were invited, 134 agreed to participate in TRAILS TRANS-ID. At the time of enrollment in TRAILS TRANS-ID, participants were approximately 23 years old. TRAILS TRANS-ID included a six-month daily diary study and aimed to investigate the day-to-day fluctuations in mental states in individuals at increased risk for psychopathology. The TRAILS TRANS-ID study was approved by the local Medical Ethical Committee (reference no. 2017/203). All participants provided written informed consent. All procedures contributing to this work comply with the ethical

^h Predictive value analyses were not preregistered, but these values flow logically from the already-calculated prevalence of transitions in combination with the sensitivity and specificity of EWS.

standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Diary study

The diary study involved the completion of a questionnaire (diary) every evening for a period of six months. Participants received these questionnaires through a link sent in a text message to their mobile phones. Questionnaires assessed thoughts and feelings during the past day (*e.g.*, “To what extent did you feel down today?”), and were rated on a visual analogue scale (VAS) ranging from 0 (not at all) to 100 (very much). Each questionnaire consisted of 58 items, of which 17 assessed positive mental states (*e.g.*, happy, relaxed, at ease with others), 35 assessed negative mental states (*e.g.*, stressed, tired, irritated), and 6 concerned event appraisal and drug- and alcohol consumption. A more elaborate description of the diary study, as well as a list of all assessed items, has been reported elsewhere²⁷ (chapter 5).

Diagnostic interview

Immediately before and after the diary study, the short version of the Schedules for Clinical Assessment in Neuropsychiatry (mini-SCAN) was administered²⁹. The mini-SCAN is a semi-structured diagnostic interview that assesses whether individuals meet the DSM-5 criteria for a psychiatric disorder (specifically: mood, anxiety, psychotic, substance use, and attention deficit hyperactivity disorder). Further, the mini-SCAN included a screener for autism spectrum disorder. The mini-SCAN was complemented by the aggressive behavior subscale of the Adult Self Report³⁰ (ASR), which covers oppositional or antisocial behavior. Both the mini-SCAN and the ASR subscale were orally administered.

Social functioning

Next to the diagnostic interview, we assessed social functioning using the Groningen Social Behavior Questionnaire (in Dutch: Groningse Vragenlijst Sociaal Gedrag, GVSG³¹). The GVSG assesses functioning on nine different domains: parents, partner, young children, old children, friends, education, occupation, household, leisure-time. Within each domain, five questions are rated on a 4-point Likert-scale ranging from ‘never’ to ‘always’. Functioning in each of these domains was only assessed if the domain was applicable to the participant (*e.g.*, in case of education, only if the participant was enrolled in education). The domain old children, which reflected participants’ relationship with their children of age > 15 years old,

was not applicable to any participant and is therefore not described further. Based on cut-off scores³¹, we categorized scores as reflecting severe impairments, mild impairments, or no impairments in functioning.

Analyses

Analyses consisted of the following steps, that are each described in detail below. First, we defined time series of each individual's mental health based on a combination of time series of several positive mental states. Second, we identified whether, and at what moment in time, individuals experienced a transition towards reduced mental health in this composite time series (see: *Transitions in mental health*). Finally, for each individual and each of the 35 negative mental states separately, we estimated time-varying trends in the autoregressive coefficient (see: *Early warning signals*). Significant rises in the autoregressive coefficient ending within two weeks prior to a transition in mental health were considered to reflect an EWS. The sensitivity, specificity, and predictive values of EWS were evaluated by examining the relative prevalence of EWS (i) in individuals who experienced a transition towards reduced mental health (sensitivity; positive predictive value) and (ii) in individuals who did not experience such a transition (specificity; negative predictive value). All analyses were conducted in R³² (version 4.0.2).

Transitions in mental health

For each individual, we determined how mental health changed over time. Mental health was defined as the mean score of four items that reflected positive mental states (*i.e.*, I felt good, I felt happy, I could enjoy things, My day was worth living). These items were selected because they showed the highest loadings in a single-factor model of all 17 positive mental states in a subset of ten randomly selected individuals (see supplement 8.2 *Definition of mental health* for additional details). We analyzed transitions in positive rather than negative mental states because (i) positive mental states were expected to show greater fluctuations than negative mental states¹⁹, (ii) positive mental states were expected to show stronger and more consistent cross-correlations than negative mental states on a between-person level³³, and (iii) this minimized the overlap between outcome (transitions in mental health) and predictor (EWS in negative mental states).

Transitions in mental health were identified by change point analyses. Change point analyses iteratively fit left- and right-sided Gaussian Kernel linear regressions (bandwidth

15 observations) for each mental health estimate in the time series for each individual separately³⁴. In other words, we estimated mental health at each time point based on (i) all weighted previous observationsⁱ and (ii) all weighted subsequent observations. The difference between the estimates based on previous and subsequent observations at a specific point in time is expected to increase if a sudden transition in mental health happened at that timepoint. This procedure was repeated for bootstrapped data. If the findings from empirical data were substantially different from the findings from bootstrapped data, the time point with the largest discrepancy was considered a change point or transition point (for a detailed description and illustration, see Smit *et al.*³⁵). This procedure was then repeated for the largest remaining segment (either before or after the transition), in order to establish whether a second transition took place. We determined maximally two transitions because, given time series data of 180 observations, a larger number of transitions seemed both theoretically implausible and impractical for later EWS estimation.

Early warning signals

Early warning signals were operationalized as rising trends in the autoregressive coefficient of negative mental states. For each individual and each negative mental state, a time-varying autoregressive (tv-AR) model was adopted in order to estimate trends in the autoregressive parameter³⁶. Tv-AR models are an extension of the AR(1) model, where a mental state at time t is predicted by (i) an intercept and (ii) the mental state at $t-1$. Unlike AR(1) models, which assume stationary time series, tv-AR models allow for gradual changes in both the intercept and the autoregressive coefficient of time series. Sudden transitions are modelled by adding a dummy predictor that encodes the timing of the transition. In tv-AR models, changes in parameters are estimated using nonparametric smooth functions based on thin-plate regression splines (cf. a general additive model framework)^{37–39}. Smoothness is determined by (i) the number of regression splines³⁶, which was set to 10, and (ii) a penalization parameter, which was selected by generalized cross-validation. Time-varying AR models assume that the predefined maximum number of regression splines is sufficient to capture the dynamics of the time series. This assumption was checked based on the effective degrees of freedom

ⁱ Note that ‘observation’ refers to each individual’s mean score on four items reflective of positive mental states (*i.e.*, I felt good, I felt happy, I could enjoy things, My day was worth living). Additional details can be found in the supplement.

and their associated P-values. Further, tv-AR models assume that residuals are mutually independent and normally distributed with constant variance. This was examined using residuals diagnostics plots³⁹.

To estimate whether trends in the autoregressive coefficient were significantly in- or decreasing, we computed the first derivative of the smooth function describing the autoregressive coefficient. This was done using the method of central finite differences described by Simpson³⁷. Trends in the autocorrelation were considered significant if the 95% confidence interval around the first derivative did not include zero. A similar approach has been adopted elsewhere^{40,41}. A multiple testing correction was not considered necessary because we compared EWS between two groups (with vs. without a transition in mental health; see below: *Sensitivity, specificity, and predictive values of early warning signals*), meaning that an overestimation of EWS (due to type I errors) would be canceled out.

The *strength* of EWS was operationalized as the average slope of the trend in the autoregressive coefficient. Higher slopes were considered indicative of stronger EWS. The *duration* of EWS was operationalized as the time interval between the earliest moment at which a significant rise in the autoregressive coefficient was detected and the transition. A longer duration suggests that EWS emerged relatively early (*i.e.*, long before the transition).

Sensitivity, specificity, and predictive values of early warning signals

First, we determined the sensitivity of EWS to detect transitions in mental health. We only inspected EWS prior to transitions that marked a shift from a period with a relatively high to a period with lower mental health. EWS were characterized based on their strength, duration, and the item wherein they manifested. We addressed the specificity of EWS by evaluating EWS in a subsample of individuals without any transition in mental health. This subsample was selected based on propensity score matching with age, sex, diary assessments (no. completed, mean, variability), and baseline psychopathology as covariates of interest (see supplement 8.3 *Matching of individuals with and without transitions in mental health* for details). For the subsample of individuals who did not experience any transition in mental health, we defined a reference transition point that was equal to the average transition point of individuals who did experience a transition in mental health. If this reference transition would often be preceded by EWS, the specificity of EWS would be low. To provide a complete overview of sensitivity and specificity, we computed both metrics (i) for each individual mental state and (ii) across negative mental states. The former reflects the proportion of

transitions preceded by EWS in a particular mental state (sensitivity) or the proportion of non-transitions (*i.e.*, reference transition points in individuals without transitions) that were *not* preceded by EWS (specificity). The latter was defined as the probability of EWS in any mental state, given that there was a transition (sensitivity) or the probability of *no* EWS, given that there was *no* transition (specificity). Finally, we computed the predictive values of EWS across negative mental states, which indicate the probability of a transition, given that EWS are detected (positive predictive value) and the probability that no transition will occur, given that EWS are not detected (negative predictive value). Predictive values take into account the overall probability of transitions (*i.e.*, the proportion of people that experienced a transition). Together, sensitivity, specificity, and predictive values may inform the utility of EWS for clinical practice. Formulas for computing sensitivity, specificity, and predictive values are provided in supplement 8.4 *Formulas for sensitivity, specificity, and predictive values*.

Results

Participant characteristics

Of the N=134 individuals who enrolled in TRAILS TRANS-ID, N=122 (91%) completed the diary study. Their characteristics are denoted in Table 8.1. All participants were of Dutch ethnicity, and most were enrolled in education (N=72) and/or had a job (N=65). Participants' socio-economic status, as assessed at enrollment in TRAILS CC, matched the socio-economic status of the general population in the Netherlands. At baseline, 40 (32.79%) individuals met the criteria for at least one psychiatric disorder. After the diary period, 34 (27.87%) met diagnostic criteria. Most of these individuals (N=23) were also diagnosed at baseline, whereas others (N=11) experienced the onset of a disorder during the diary period. Mood disorders (N=28 at baseline and N=23 at post), anxiety disorders (N=21 and 12), and ADHD (N=8 and 8) were most common. Approximately half of the participants (N=62 at baseline and N=68 at post) had severe impairments in at least one domain of social functioning. Impairments in functioning most often concerned problems in enjoying leisure time (*e.g.*, feeling bored), problems in occupation (*e.g.*, not performing well at work), and problems in engaging with friends (at baseline) or with a romantic partner (at post; Table 8.1).

Table 8.1. *Sample characteristics*

		Full sample		Subsample with a transition towards reduced mental health	
	N	122		20	
<i>Demographics</i>	sex (no. of males)	69 (57%)		10 (50%)	
	age ($\pm$ SD)	23.64 ($\pm$ 0.67)		23.38 ($\pm$ 0.75)	
<i>Socio-economic status</i>	high	25 (20%)		3 (15%)	
	middle	66 (54%)		13 (65%)	
	low	31 (25%)		4 (20%)	
	financial difficulties	29 (25%)		7 (37%)	
<i>Diary assessments</i>	completed (%)	88.45		84.74	
	mean	16.85		16.16	
	variability	12.34		11.66	
		<i>Baseline</i>	<i>Post</i>	<i>Baseline</i>	<i>Post</i>
<i>Psychopathology</i>	mood disorder	28	23	3	6
	anxiety disorder	21	12	1	2
	psychotic disorder	2	5	1	2
	substance use disorder	3	4	1	1
	ADHD	8	8	0	1
<i>Social functioning</i>	parents	5 (4.13%)	12 (9.84%)	1 (5.00%)	2 (10.00%)
	partner	6 (8.57%)	13 (19.40%)	2 (16.67%)	4 (33.33%)
	children	0 (0%)	1 (10.00%)	0 (0%)	0 (0%)
	friends	14 (11.67%)	10 (8.47%)	2 (10.00%)	3 (15.79%)
	education	4 (7.84%)	4 (7.41%)	1 (11.11%)	0 (0%)
	occupation	13 (14.94%)	23 (22.55%)	1 (6.67%)	4 (25.00%)
	household	12 (10.26%)	18 (15.00%)	2 (10.53%)	2 (10.00%)
	leisure-time	36 (29.51%)	36 (29.51%)	9 (45.00%)	8 (40.00%)

Note. The full sample includes all individuals who completed the diary study; the subsample of individuals with a transition towards reduced mental health concerns the focus of the present chapter, which investigates early warning signals for such transitions. Financial difficulties denotes the number of participants who experienced some or substantial problems in getting by financially. For diary assessments, mean refers to the average rating across negative mental states during the diary period. Psychopathology refers to the number of individuals that met the DSM-5 criteria for a psychiatric disorder at baseline (*i.e.*, immediately before the diary period) and post (*i.e.*, immediately after the diary period). Social functioning refers to the number and percentage of individuals who showed severely impaired functioning on the domains assessed by the GVSG. Percentages were based on the number of individuals for whom a certain domain was relevant. ADHD = attention-deficit/hyperactivity disorder. DSM-5 = Diagnostic and Statistical Manual of Mental Disorders. GVSG = Groningen Social Behavior Questionnaire.

Transitions in mental health

Change point analyses indicated that 44 individuals (36.07%) experienced a transition in mental health. Of these individuals, 20 (45.45%, 16.39% of the full sample) showed a drop in mental health (average drop = 6.70, $SD = 5.43$, range = 0.15-17.20). On average, transitions happened halfway through the diary period (day 94, $SD = 34.11$, range = 29-159). Individuals who experienced a transition towards reduced mental health were not different from the full sample in terms of sex ($\chi^2(1) = 0.15$, $p = .70$), age ($t(140) = 1.58$, $p = .12$), or the number of completed diary ratings ($t(140) = 1.51$, $p = .13$). Further, the mean and variability of negative mental states assessed during the diary period did not differ between individuals with a transition and the full sample ($t(140) = 0.28$, $p = .78$; $t(140) = 0.53$, $p = .60$). Of those individuals who experienced a drop in mental health, 6 and 7 met diagnostic criteria at baseline and post, respectively. The likelihood to meet diagnostic criteria at baseline or post was not different between individuals who experienced a drop in mental health and the full sample ($\chi^2(1) < 0.01$, $p > .99$; $\chi^2(1) = 0.15$, $p = .70$). Finally, those who experienced a drop in mental health did not differ from the full sample in their likelihood to experience severe problems in functioning in at least one domain at baseline (50.00% vs. 50.82%, $\chi^2(1) < 0.01$, $p > .99$) or post (55.00% vs. 55.74%, $\chi^2(1) < 0.01$, $p > .99$), nor in the number of domains wherein they experienced problems at baseline (mean 0.90 vs. 0.74, $t(140) = 0.73$, $p = .46$) or post (mean 1.15 vs. 0.96, $t(140) = 0.73$, $p = .46$).

Sensitivity and specificity of early warning signals

For 14 individuals (70.00% of the 20 individuals with transitions), a sudden transition in mental health was preceded by EWS in at least one out of 35 negative mental states (see Fig. 8.1 for an example). Of these individuals, 11 showed EWS in more than one mental state (mean no. of mental states with EWS per individual: 2.50, $SD = 1.02$, range = 1-4). The majority of mental states wherein EWS were detected were specific to a single individual, and hence, the sensitivity of EWS in individual mental states ranged between 0 and 15%. Put differently, individual mental states never (0%) or seldomly (5-15%) showed EWS prior to a drop in mental health (supplement, Table S8.3 and Fig. S8.2). The overall sensitivity of EWS across all negative mental states was 5.00% (95% CI = 3.51-6.89%; Table 8.2).

To determine the specificity of EWS, we selected a subsample ($N=20$) of individuals without a transition in mental health that matched the sample with a transition in terms of age, sex, diary assessments (no. completed, mean, variability), and baseline psychopathology.

For 13 (65.00%) of these individuals, EWS were found. Most of them showed EWS in a single negative mental state (mean no. of mental states with EWS per individual: 2.15, $SD = 1.46$, range = 1-6). Again, the negative mental states wherein EWS manifested were mostly unique to individuals, resulting in a specificity of 90 to 100% (supplement 8.5 *Sensitivity of EWS*, Fig. S8.2). In other words, EWS seldomly or never represented a false alarm, supported by an overall specificity of EWS across all negative mental states of 96.00% (95% CI = 94.27-97.33%; Table 8.2).

Fig. 8.1. *Illustration of an early warning signal prior to a transition towards reduced mental health.* A) Change point analyses identified a transition towards reduced mental health on day 129 (blue line). Black horizontal lines represent the average mental health score pre- and post-transition. B) Prior to this transition, marked by the blue vertical line, the autoregressive coefficient in *I experienced physical pain* significantly increased. This may suggest that the sudden drop in mental health was preceded by a period of gradually accumulating instability (critical slowing down).

Table 8.2. *Overall presence of early warning signals in 35 negative mental states*

	EWS	no EWS	Total
Transition	35	665	700
No transition	28	672	700
Total	63	1.337	

Note. Numbers refer to the number of EWS found in individuals with a transition ($N=20$) and a matched subsample of individuals without a transition ($N=20$) towards reduced mental health. As we examined EWS in 35 negative mental states, the total number of signals that could theoretically be detected for each for each subgroup (with vs. without transition) was $35 \times 20 = 700$. The prevalence of transitions towards reduced mental health was 16.39% (20/122). The numbers in this table were used to calculate the overall sensitivity and specificity of EWS as well as their predictive values.

Predictive value of early warning signals

Whereas sensitivity and specificity are informative of the probability that EWS will occur prior to a transition, predictive values indicate the probability of a transition following EWS (for a more elaborate description, see supplement 8.4 *Formulas for sensitivity, specificity, and predictive values*). EWS can be considered clinically useful if (i) their presence increases the probability of a transition (high positive predictive value, or PPV) and (ii) their absence increases the probability that no transition will occur (high negative predictive value, or NPV). In this sample, the probability of a transition towards reduced mental health was 16.39% (20/122). The presence of EWS increased this probability to 19.68% (95% CI = 13.10-28.48%; PPV). In absence of EWS, however, the probability of a drop in mental health remained 16.25% (NPV = 83.75%, 95% CI = 83.44-84.06%).

Timing and strength of early warning signals

The type of mental state wherein EWS manifested, as well as the timing and strength of EWS, varied considerably between individuals (Fig. 8.2, supplementary Fig. S8.3). The timing of EWS ranged between 0.29 weeks (2 days) to 21.71 weeks (almost 5 months). On average, EWS appeared 9.43 weeks prior to a transition in mental states. Except for EWS in I was bored quickly, I felt restless and I had a fight, EWS in all mental states took place at least one week prior to the transition. The earliest signs occurred in I was easily distracted, I was impulsive, and I worried. The strongest warning signals occurred in I experienced physical pain, Things did not bother me, and I felt restless. The timing and strength of EWS found in individuals without a transition in mental health are reported in supplement 8.7 *EWS in a subsample of individuals without transitions towards reduced in mental health*.

Fig. 8.2. Timing (A) and strength (B) of early warning signals across negative mental states (rows) and individuals (columns). Only negative mental states in which EWS were detected are shown. As illustrated, the timing and strength of EWS, as well as the negative mental state wherein EWS manifest, vary considerably between individuals. For seven individuals, no EWS were found (empty columns).

Discussion

Early warning signals – and complex systems principles in general – have been suggested to hold great promise for mental health research and clinical practice. In particular, rising autocorrelations in negative mental states have been introduced as potential personalized warning signs that could enable foreseeing the sudden onset or remission of psychiatric disorders^{25,26}. This would mean that by monitoring and analyzing individuals' negative mental states, clinicians could know whether a specific individual, at a specific moment in time, is at high risk for developing psychopathology. The present study concerns one of the first empirical investigations of EWS – conceptualized as rising trends in the autoregressive coefficient of negative mental states – as predictors of transitions in mental health. Using daily diary data collected for a period of six months in 122 at-risk young adults, we identified 20 individuals who experienced a drop in mental health. The majority of mental states did not show EWS prior to these transitions, illustrated by an overall sensitivity of 5%. *If* EWS occurred, however, they more often occurred in individuals with drops in mental health (35 EWS) than in individuals without such drops (28 EWS). In conclusion, in the present sample, EWS were too rare to be informative of future mental health. Thus, it seems that anticipating drops in mental health in the current sample by means of EWS remains beyond reach.

Present findings contradict two earlier single-case studies, which reported that EWS anticipate relapses in depression^{25,26}. The discrepancies between present and earlier findings might be attributed to differences in the severity of psychopathological symptoms – and relatedly, the magnitude of transitions – that individuals experienced. Specifically, earlier case studies included individuals diagnosed with major depression who showed a sudden increase in weekly symptom reports of depression. In contrast, the present study concerned sudden drops in day-to-day mental health reported by individuals who not necessarily had clinical diagnoses. Hence, the presently studied transitions might have been smaller than, and qualitatively different from, the relapses experienced by individuals with a history of major depression. This is a critical point, because support for EWS prior to transitions has mostly been found in the context of large – potentially irreversible – shifts in systems, such as the extinction of a species (ecosystems) or market crashes (financial systems). Including individuals without clinically significant symptoms, for whom transitions in mental health were likely smaller than those previously studied, may thus have lowered our power to detect EWS. Nonetheless, investigating the potential of EWS in this population is essential,

because precisely these individuals, who are at risk but have not yet developed a disorder, might profit most from the early detection of vulnerability that EWS could offer. Further, absolute thresholds for the magnitude of transitions that are preceded by EWS do not exist. Hence, we do not consider the nature of the present sample a limitation but an important addition to the literature.

Putting early warning signals into perspective

Whereas research into EWS as predictors for transitions in mental health is still in its infancy, other fields yielded valuable insight in EWS that can help to put current findings into perspective. Although numerous studies from various fields showed that transitions, for instance observed in ecological or financial systems, are preceded by rising autocorrelations^{15,42,43}, other studies did not support this^{44–46}. This led to the conclusion that EWS might be less generic than initially proposed⁴⁷. Rather, the presence and detectability of EWS depends upon numerous requirements. For instance, EWS are not expected when (i) instability accumulates too quickly to be captured by EWS; (ii) the transition is caused by a strong external trigger instead of gradually accumulated instability within the system^{48,49}; (iii) the instability of the system manifests in other variables than the variables wherein EWS are examined⁴⁷; (iv) the measurement frequency does not match the timescale at which the dynamics of interest operate^{10,50}; (v) the alternative state following from the transition is only temporary, and does not fulfill the (mathematical) requirements of an equilibrium^{48,51}; or (vi) the system is exposed to non-stochastic fluctuations^{46,52,53}. All of these conditions could play a role in the detection of EWS prior to transitions in mental health. Such transitions are still poorly understood, and therefore, the extent to which mental health meets the aforementioned requirements remains speculative. We will discuss this in-depth below and refer back to the above-mentioned requirements. First, transitions in mental health may follow from rapidly accumulating vulnerability – which cannot be captured by EWS (requirement i) – or from major external events (*e.g.*, traumatic events; requirement ii), rather than from gradually accumulating instability. Provided that there is a gradual build-up of instability, not all negative mental states might be equally affected. That is, EWS might manifest only in certain mental states (requirement iii), which may vary across individuals (see Fig. 8.2, supplement Fig. S8.3). The present study addressed this by examining EWS in a broad variety of mental states. Yet, little is known about the variables wherein vulnerability to mental ill-health actually manifests, and it could therefore be that other types of variables

(*e.g.*, physiological) are more suitable for detecting EWS. The detection of EWS further depends on the frequency of assessments: if the time interval between observations is too long, the correlation between adjacent observations (autocorrelation) is no longer informative of critical slowing down (requirement iv). This highlights yet another unresolved question: at what exact timescale does psychopathology unfold? It has been reasoned that mood states have faster dynamics than (subthreshold) symptoms or the mental states presently studied⁵⁴. We therefore adopted daily, instead of *e.g.*, hourly, measurements to infer EWS. Still, EWS could manifest on a smaller timescale. If that is the case, however, the eventual clinical utility of EWS would be limited: it seems unfeasible to ask at-risk individuals to monitor their mental states every hour for a considerable period of time (*e.g.*, six months). Hence, the daily assessments of the present study reflect a good balance between what is theoretically plausible and practically feasible. Two final prerequisites for detecting EWS relate to the states between which transitions may occur. In some systems, these states represent temporally stable equilibria with clear boundaries (*e.g.*, greenhouse vs. icehouse climate states). In mental health, however, alternative states may be less clear-cut. Rather, the progression from mental health to ill-health may involve temporary ‘in-between’ stages⁵⁵. Humans may shift back and forth between such stages, and hence, transitions in mental health might involve transient states instead of distinct stable equilibria (requirement v). Even if mental (ill-)health can be described as an equilibrium, the events to which people are exposed – which lead to temporary departures from equilibrium, *i.e.*, dips and uplifts in their mood – may not occur stochastically (requirement vi). That is, humans can exert some control over the events they experience, and therefore, these events do not happen completely at random. As individuals become more vulnerable to symptoms of anxiety, for example, they may increasingly avoid certain triggers. This complicates the detection of critical slowing down (and hence, EWS): we cannot assess slowed recovery from perturbations if there are few perturbations to recover from. In summary, little is known about the extent to which mental health meets the requirements for anticipating sudden transitions by means of EWS. This knowledge gap is to be addressed in the future.

One of the most compelling properties of EWS is that they may anticipate system behavior – such as sudden transitions – even though little is known about the mechanisms that underlie system behavior (*e.g.*, biological, social, psychological). This is highly attractive in the context of mental health, where changes within individuals – such as the onset or progression of symptoms – are still poorly understood from a mechanistic point of view

(*e.g.*, in terms of the biological, social, and psychological causes). Yet, although EWS do not necessitate such traditional mechanistic understanding, they *do* require an alternative level of understanding mental health. That is, from the list of requirements outlined above, it follows that EWS require insight in the dynamics that govern psychopathology. This entails delineation of the type of transitions that take place, the characteristics of the states between which such transitions may occur, and the timescale at which dynamics such as EWS unfold. This could ultimately help with figuring out which, if any, types of changes in mental health may be formally recognized as transitions that are preceded by critical slowing down.

Taken together, this study found little support for the notion that rising autocorrelations in negative mental states could be used as reliable, personalized risk markers for impending transitions towards reduced mental health. Yet, it is too early to reject this idea, and follow-up studies with different populations, types of transitions, and warning signals (*e.g.*, increasing trends in variances or cross-correlations; multivariate warning signals) are needed. Further, it may be fruitful broaden our perspective on what EWS could offer: besides anticipating transitions, EWS could be used to map resilience. Multivariate extensions of EWS could for instance indicate the type of symptoms for which a specific individual is most vulnerable (or least resilient)^{7,47,56}. Such insight might then inform preventive intervention efforts. In sum, the hypotheses and clinical promises that follow from a complex systems approach to mental health range far beyond what could be addressed in the present study, although the current findings indicate new angles for future research.

Strengths, limitations, and future directions

The unique data analyzed in the present study – which covered a maximum of 183 daily diary assessments in 122 at risk young adults – offered the possibility to provide one of the first large scale intra-individual studies into EWS (*i.e.*, rising autocorrelations in negative mental states) as predictors of transitions towards reduced mental health. By analyzing EWS not only in individuals with transitions in mental health, but also in individuals *without* such transitions, we could examine both the sensitivity and specificity of EWS. The latter is seldomly reported in empirical literature, yet is highly relevant from a clinical point of view (for a more elaborate discussion on this topic, see Boettiger *et al.*⁵⁷).

An important consideration in evaluating the merits and generalizability of the current research concerns our operationalization of transitions in mental health. In order to minimize the overlap between predictor (EWS) and outcome (transitions), we derived

drops in mental health from aggregated scores on positive mental states. Although it seems intuitive that sudden drops in positive mental states signal a reduction in mental health, it has also been argued that positive mental states and symptoms of psychopathology should be considered separate dimensions of mental health⁵⁸. This potential independence between positive and negative mental states may have limited the presence of EWS. This limitation touches upon a broader issue, namely that we know fairly little about the transitions that individuals experienced (*e.g.*, in terms of the clinical meaning of these transitions; the driving causes of transitions; the timescale at which they unfolded; their supposed non-linearity, etc.). Yet, such characteristics are crucial for the presence of EWS⁴⁷. A more elaborate exploration of transitions in mental health – for instance observed as a sudden onset or relapse of psychopathology – therefore concerns an important avenue for further research. As an example, recent work showed that the majority of remission trajectories occur in a stepwise fashion, as opposed to (log)linearly⁵⁹. Additionally, further studies are needed to define the periods before and after transitions in mental health. When, and for whom, do such periods resemble stable equilibria as opposed to transient, intermittent states? Finally, future studies could look beyond the traditional application of EWS (that is, as timely warning signs), and examine whether EWS can inform on the type of symptoms for which individuals may be most vulnerable⁵⁶.

Conclusion

In young adults at risk for psychopathology, rising autocorrelations in negative mental states – which have been considered personalized warning signs – hardly predict transitions towards reduced mental health. Unlike earlier studies, which tested EWS on group-level¹⁶, this study investigated EWS within individuals. There were considerable individual differences in the prevalence, timing, and strength of EWS, suggesting that a complex systems approach to mental health may fit some individuals better than others. This highlights the need to investigate under what circumstances complex systems principles are useful for understanding the progression of mental health problems (*e.g.*, type of complaints, characteristics of the transition, etc.). In conclusion, EWS do not provide a simple solution to predicting who will experience a sudden drop in mental health at what moment in time. It is too early, however, to throw out the child with the bathwater. Complex systems offer us a novel perspective on vulnerability and resilience to mental disorders, which has only just begun to be discovered.

Supplement 8

8.1 Illustration of critical slowing down

Fig. S8.1. An illustration of critical slowing down for transitions in mental health based on simulated data. Red arrows denote perturbations, or stressful events, that may temporarily lower an individual's mental health. The figure illustrates that (i) mental health may fluctuate over time; (ii) transitions in mental health may occur suddenly; (iii) as a transition towards mental ill-health approaches, resilience to perturbations declines, meaning that perturbations have increasingly lasting effects. This means that, after a perturbation, it takes longer and longer before the system returns to its original state, as reflected by rising trends in the autocorrelation (EWS).

8.2 Definition of mental health

For each individual, we defined mental health at time t as the individual's mean score on four positive mental states rated at time t . Rather than averaging all 17 positive mental states, we chose to select a limited number of positive mental states as a proxy for mental health. This was done in order to preserve construct validity, or in other words: to make sure that mental health reflected a similar construct for each individual. To determine which positive mental states most consistently reflected mental health across individuals, we performed factor analyses restricted to a single factor. Subsequently, we computed the mean factor loading for each mental state across individuals. As denoted in Table S8.1, mental health was best explained by the following mental states: I could enjoy things, I was happy, My day was worth living, and I felt good. Across individuals, these mental states explained approximately 50% of the variance in mental health. Other items explained considerably less variance in mental health, and were hence not included in its estimation.

Table S8.1. *Average explained variance of 16 mental states in a single-factor model*

Mental state	Explained variance
I slept well	0.13
I was physically active	0.19
I am looking forward to tomorrow	0.19
I felt like others liked me	0.21
I was social	0.23
My concentration was good	0.24
I felt at ease with others	0.24
I felt appreciated	0.27
I felt relaxed	0.27
Everything came easy	0.28
I felt like undertaking things	0.29
I could handle the things that confronted me	0.32
I felt energetic	0.36
I could enjoy things	0.48
I was happy	0.49
My day was worth living	0.51
I felt good	0.51

8.3 Matching of individuals with and without transitions in mental health

In order to address the specificity of EWS for transitions in mental health, we compared EWS in individuals with transitions towards reduced mental versus a matched subsample of individuals without such transitions. Matching was performed in order to minimize the probability that differences between these groups of individuals could be due to differences in, for instance, their age or sex. Propensity score matching was used to select a subsample of individuals without a transition in mental health that were similar to those with a transition in terms of age, sex, diary assessments (no. completed, mean, variability), and psychopathology⁶⁰. This involved calculating a propensity score for each participant (*i.e.*, the probability of belonging to the group of individuals with a transition based on the aforementioned characteristics), and subsequently matching individuals with similar propensity scores using a nearest neighbor algorithm. Table S8.2 denotes the result of this procedure. None of the characteristics were significantly different between individuals with versus without a transition in mental health.

Table S8.2. *Matching of individuals with vs. without a transition in mental health*

	Characteristic	No transition in mental health	Transition in mental health
<i>Descriptives</i>	N	20	20
	sex (no. of males)	8	10
	age ($\pm$ SD)	23.27 ($\pm$ 0.56)	23.38 ($\pm$ 0.75)
<i>Diary assessments</i>	completed (%)	88.39	84.74
	mean	16.90	16.16
	variability	12.70	11.66
<i>Psychopathology</i>	stress	3.60	3.30
	depression	7.25	7.90
	anxiety	3.05	4.95
	OCD	1.10	0.95
	psychosis	1.15	0.70
	mania	1.70	2.20
	ASD	2.75	3.60
	ADHD	28.45	28.75
	substance abuse	1.85	2.50
	ODD	5.30	5.50

Note. For diary assessments, mean refers to the average rating across negative mental states during the diary period. Variability refers to the average variability of negative mental states during the diary period. Psychopathology estimates refer to individuals' scores on domains assessed in a semi-structured diagnostic interview that preceded the diary period. ADHD = attention-deficit/hyperactivity disorder. ASD = autism spectrum disorder. OCD = obsessive-compulsive disorder. ODD = oppositional defiant disorder. SD = standard deviation.

8.4 Formulas for sensitivity, specificity, and predictive values

On the level of mental states, sensitivity is calculated as the proportion of transitions preceded by EWS in that mental state. Similarly, specificity is calculated as the proportion of non-transitions *not* preceded by EWS in that mental state.

$$\text{sensitivity}_{\text{mental state } k} = \frac{\text{EWS in mental state } k}{\text{total no. of transitions}}$$

$$\text{specificity}_{\text{mental state } k} = 1 - \frac{\text{EWS in mental state } k}{\text{total no. of non-transitions}}$$

Across mental states, sensitivity reflects the ratio of true positives (*i.e.*, EWS in the presence of a transition) and the total number of EWS that could have been present. The overall

specificity reflects the ratio of true negatives (*i.e.*, no EWS in absence of a transition) and the total number of EWS that could have been present. In the present study, the total number of EWS that could have been present equals 700 (20 transitions, each with 35 potential EWS).

$$sensitivity_{overall} = \frac{true\ positives}{true\ positives + false\ negatives}$$

$$specificity_{overall} = \frac{true\ negatives}{false\ positives + true\ negatives}$$

Finally, the positive and negative predictive values (PPV and NPV, respectively), reflect the probability of a transition given EWS (PPV) or the probability of *no* transition given that EWS are not detected (NPV). Note that these values take into account the prevalence of a transition (in this study, 16.39%).

$$PPV = \frac{sensitivity \times (1 - prevalence)}{(1 - specificity) \times prevalence + sensitivity \times (1 - prevalence)}$$

$$NPV = \frac{specificity \times (1 - prevalence)}{(1 - sensitivity) \times prevalence + specificity \times (1 - prevalence)}$$

8.5 Sensitivity of EWS

Table S8.3. *Sensitivity and specificity of early warning signals in individual negative mental states*

item	sensitivity	specificity
I had a fight	15%	100%
I was impulsive	10%	100%
I worried	10%	100%
My thoughts would not leave me alone	10%	100%
Things did not bother me	10%	100%
I felt empty	15%	95%
I experienced physical pain	10%	95%
I felt nervous	5%	100%
I felt overburdened	10%	95%
I felt stressed	10%	95%
I felt suspicious	5%	100%
I felt tired	5%	100%
I had difficulty making decisions	5%	100%
I had moodswings	10%	95%
I was easily distracted	10%	95%
I could not bring myself to do anything	5%	95%
I experienced physical discomfort	0%	100%
I felt down	5%	95%
I felt rebellious	0%	100%
I felt restless	5%	95%
I felt strange	0%	100%
I was easily upset	0%	100%
It bothered me that things did not go as expected	0%	100%
Others felt annoyed by me	0%	100%
I was afraid of making mistakes	5%	90%
I was bored quickly	5%	90%
I was unable to sit still	5%	90%
I felt guilty	0%	95%
I felt impatient	0%	95%
I felt irritated	0%	95%
I was easily startled	0%	95%
I'm dreading something that will happen soon	0%	95%
I felt lonely	0%	90%
I felt overstimulated	5%	85%
I avoided things	0%	85%

Note. Sensitivity refers to the percentage of individuals for whom the negative mental state in the first column demonstrated a rising trend autocorrelation (*i.e.*, an EWS) prior to a transition towards reduced mental health. Specificity refers to the percentage of individuals without a transition for whom the corresponding negative mental state did *not* show an EWS.

Fig. S8.2. Frequency of early warning signals across negative mental states for individuals with a drop in mental health (true positives) and individuals without a change in mental health (false positives).

8.6 EWS in individuals with transitions towards reduced mental health: additional figure

Fig. S8.3. *Timing and strength of early warning signals in distinct negative affective states prior to a transition towards reduced mental health.* The strength of EWS was operationalized as the mean gradient of the trend in the autoregressive coefficient. The gradient of lines corresponds to the mean strength of EWS; the thickness of lines corresponds to the number of transitions that were preceded by EWS in the corresponding negative mental state. For readability, the 22 mental states in which EWS were found were split in three separate graphs.

8.7 EWS in a subsample of individuals without transitions towards reduced mental health

To determine the specificity and predictive values of EWS as risk markers for upcoming transitions towards reduced mental health, we selected a subsample of individuals without a drop in mental health (N=20) and compared their EWS (false positives) to the EWS in individuals with a drop in mental health (true positives). The timing and duration of EWS in this subsample without transitions were not of interest to the main chapter. For the interested reader, we included the figures that illustrate these characteristics below. Timing of EWS ranged between 0.57 and 13.29 weeks (0.29 – 21.71 weeks in individuals with a transition). The strongest EWS occurred in *I felt lonely*, *I experienced physical pain*, and *I avoided things* (*I experienced physical pain*, *Things did not bother me*, and *I felt restless* in individuals with a transition).

Fig. S8.4. Timing (A) and strength (B) of early warning signals across negative mental states (rows) and individuals who did not undergo a transition in mental health (columns). Only negative mental states in which EWS were detected are shown. For seven individuals, no EWS were detected (empty columns).

Fig. S8.5. *Strength and timing of early warning signals in distinct negative affective states in individuals without a transition in mental health.* The strength of EWS was operationalized as the mean gradient of the trend in the autoregressive coefficient. The gradient of lines corresponds to the mean strength of EWS; the thickness of lines corresponds to the number of transitions that were preceded by EWS in the corresponding negative mental state. For readability, the 20 mental states in which EWS were found were split in three separate graphs.

References

1. Hofmann SG, Curtiss J. A complex network approach to clinical science. *European Journal of Clinical Investigation*. 2018;48(8). doi:10.1111/eci.12986
2. Hayes AM, Laurenceau JP, Feldman G, Strauss JL, Cardaciotto LA. Change is not always linear: the study of nonlinear and discontinuous patterns of change in psychotherapy. *Clinical Psychology Review*. 2007;27(6):715-723. doi:10.1016/j.cpr.2007.01.008
3. Scholz JP, Kelso JAS, Schöner G. Nonequilibrium phase transitions in coordinated biological motion: Critical slowing down and switching time. *Physics Letters A*. 1987;123(8):390-394. doi:10.1016/0375-9601(87)90038-7
4. Scheffer M. *Critical Transitions in Nature and Society*. Princeton University Press; 2009.
5. Jeronimus BF. Dynamic system perspectives on anxiety and depression. In: Kunnen ES, de Ruiter NMP, Jeronimus BF, van der Gaag MA, eds. *Psychosocial Development in Adolescence: Insights from the Dynamic Systems Approach*. Routledge Psychology; 2019. doi:10.1128/MCB.01405-08
6. Cramer AOJ, van Borkulo CD, Giltay EJ, et al. Major depression as a complex dynamic system. *PLoS one*. 2016;11(12):1-20. doi:10.1371/journal.pone.0167490
7. Wichers M, Schreuder MJ, Goekoop R, Groen RN. Can we predict the direction of sudden shifts in symptoms? Transdiagnostic implications from a complex systems perspective on psychopathology. *Psychological Medicine*. 2019;49(3):380-387. doi:10.1017/S0033291718002064
8. Olthof M, Hasselman F, Strunk G, et al. Critical fluctuations as an early-warning signal for sudden gains and losses in patients receiving psychotherapy for mood disorders. *Clinical Psychological Science*. 2019;8(1):1-11. doi:10.1177/2167702619865969
9. Scheffer M, Bascompte J, Brock WA, et al. Early-warning signals for critical transitions. *Nature*. 2009;461(7260):53. doi:10.1038/nature08227
10. Wen H, Ciamarra MP, Cheong SA. How one might miss early warning signals of critical transitions in time series data: a systematic study of two major currency pairs. *PLoS one*. 2018;13(3):1-22. doi:10.1371/journal.pone.0191439
11. de Mooij SMM, Blanken TF, Grasman RPPP, Ramautar JR, van Someren EJW, van der Maas HLJ. Dynamics of sleep: exploring critical transitions and early warning signals. *Computer Methods and Programs in Biomedicine*. 2020;193. doi:10.1016/j.cmpb.2020.105448
12. Maturana MI, Meisel C, Dell K, et al. Critical slowing down as a biomarker for seizure susceptibility. *Nature Communications*. 2020;11(1):1-12. doi:10.1038/s41467-020-15908-3
13. Clements CF, Ozgul A. Including trait-based early warning signals helps predict population collapse. *Nature Communications*. 2016;7. doi:10.1038/ncomms10984
14. Kaur T, Sarkar S, Chowdhury S, Sinha SK, Jolly MK, Dutta PS. Anticipating the novel coronavirus disease (COVID-19) pandemic. *Frontiers in Public Health*. 2020;8(521):1-13. doi:10.1101/2020.04.08.20057430
15. Dakos V, Scheffer M, van Nes EH, Brovkin V, Petoukhov V, Held H. Slowing down as an early warning signal for abrupt climate change. *PNAS*. 2008;105(38):14308-14312. doi:10.1073/pnas.0802430105
16. van de Leemput IA, Wichers M, Cramer AOJ, et al. Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences*. 2014;111(1):87-92. doi:10.1073/pnas.1312114110

17. McGorry PD, Hartmann JA, Spooner R, Nelson B. Beyond the “at risk mental state” concept: transitioning to transdiagnostic psychiatry. *World Psychiatry*. 2018;17(June):133-142.
18. Koval P, Kuppens P, Allen NB, Sheeber L. Getting stuck in depression: the roles of rumination and emotional inertia. *Cognition and Emotion*. 2012;26(8):1412-1427. doi:10.1080/02699931.2012.667392
19. Houben M, Van Den Noortgate W, Kuppens P. The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. *Psychological Bulletin*. 2015;141(4):901-930. doi:10.1037/a0038822
20. Sperry SH, Walsh MA, Kwapił TR. Emotion dynamics concurrently and prospectively predict mood psychopathology. *Journal of Affective Disorders*. 2020;261:67-75. doi:10.1016/j.jad.2019.09.076
21. Nelson J, Klumppendt A, Doebler P, Ehring T. Everyday emotional dynamics in major depression. *Emotion*. 2018;20:179-191. doi:10.1037/emo0000541
22. Brose A, Schmiedek F, Koval P, Kuppens P. Emotional inertia contributes to depressive symptoms beyond perseverative thinking. *Cognition and Emotion*. 2015;29(3):527-538. doi:10.1080/02699931.2014.916252
23. Kuppens P, Sheeber LB, Yap MBH, Whittle S, Simmons JG, Allen NB. Emotional inertia prospectively predicts the onset of depressive disorder in adolescence. *Emotion*. 2012;12(2):283-289. doi:10.1037/a0025046
24. Curtiss J, Fulford D, Hofmann SG, Gershon A. Network dynamics of positive and negative affect in bipolar disorder. *Journal of Affective Disorders*. 2019;249:270-277. doi:10.1016/j.jad.2019.02.017
25. Wichers M, Smit AC, Snippe E. Early warning signals based on momentary affect dynamics can expose nearby transitions in depression: a confirmatory single-subject time-series study. *Journal for Person-Oriented Research*. 2020;6(1):1-15. doi:10.17505/jpor.2020.22042
26. Wichers M, Groot PC, Psychosystems ESM group, EWS group. Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*. 2016;85:114-116. doi:10.1159/000441458
27. Schreuder MJ, Groen RN, Wigman JTW, Hartman CA, Wichers M. Measuring psychopathology as it unfolds in daily life: addressing key assumptions of intensive longitudinal methods in the TRAILS TRANS-ID study. *BMC Psychiatry*. 2020;20(1). doi:10.1186/s12888-020-02674-1
28. Huismán M, Oldehinkel AJ, de Winter A, et al. Cohort Profile: The Dutch “TRacking Adolescents’ Individual Lives’ Survey”; TRAILS. *International Journal of Epidemiology*. 2008;37(6):1227-1235. doi:10.1093/ije/dym273
29. Nienhuis FJ, van de Willige G, Rijnders CAT, de Jonge P, Wiersma D. Validity of a short clinical interview for psychiatric diagnosis: the mini-SCAN. *British Journal of Psychiatry*. 2010;196(1):64-68. doi:10.1192/bjp.bp.109.066563
30. Achenbach TM. *Manual for the Young Adult Self-Report and Young Adult Behavior Checklist*.; 1997.
31. de Jong A, van der Lubbe PM. *Groningse Vragenlijst Sociaal Gedrag: Zelfbeoordelingsvragenlijsten Voor Het Vaststellen van Problemen in Het Interpersoonlijke Functioneren (Handleiding)*. Rob Giel Onderzoekscentrum/Rijksuniversiteit Groningen, disciplinegroep Psychiatrie; 2001.
32. R Core Team. R: A language and environment for statistical computing. Published online 2020.
33. Mattsson M, Hailikari T, Parpala A. All happy emotions are alike but every unhappy emotion is unhappy in its own way: a network perspective to academic emotions. *Frontiers in Psychology*. 2020;11(742):1-11. doi:10.3389/fpsyg.2020.00742
34. Muller H-G. Change-points in nonparametric regression analysis. *The Annals of Statistics*. 1992;20(2):737-761.

35. Smit AC, Snippe E, Kunkels YK, Riese H, Helmich MA, Wichers M. Transitions In Depression (TRANS-ID) Tapering. Published 2020. Accessed November 6, 2020. www.osf.io/h75p9
36. Bringmann LF, Hamaker EL, Vigo DE, Aubert A, Borsboom D, Tuerlinckx F. Changing dynamics: time-varying autoregressive models using generalized additive modeling. *Psychological Methods*. 2017;22(3):409-425. doi:10.1037/met0000085
37. Simpson GL. Modelling palaeoecological time series using generalised additive models. *Frontiers in Ecology and Evolution*. 2018;6:1-21. doi:10.3389/fevo.2018.00149
38. Hastie T, Tibshirani R. Generalized additive models. *Statistical Science*. 1986;1(3):297-318. doi:10.2307/2246134
39. Wood S. *Generalized Additive Models: An Introduction with R*. 2nd ed. Taylor & Francis Group; 2017.
40. Burthe SJ, Henrys PA, Mackay EB, et al. Do early warning indicators consistently predict nonlinear change in long-term ecological data? *Journal of Applied Ecology*. 2016;53(3):666-676. doi:10.1111/1365-2664.12519
41. Sommer S, van Benthem KJ, Fontaneto D, Özgül A. Are generic early-warning signals reliable indicators of population collapse in rotifers? *Hydrobiologia*. 2017;796(1):111-120. doi:10.1007/s10750-016-2948-7
42. Veraart AJ, Faassen EJ, Dakos V, van Nes EH, Lürling M, Scheffer M. Recovery rates reflect distance to a tipping point in a living system. *Nature*. 2012;481(7381):357-359. doi:10.1038/nature10723
43. Drake JM, Griffen BD. Early warning signals of extinction in deteriorating environments. *Nature*. 2010;467(7314):456-459. doi:10.1038/nature09389
44. Diks C, Hommes C, Wang J. Critical slowing down as an early warning signal for financial crises? *Empirical Economics*. 2018;57(4):1201-1228. doi:10.1007/s00181-018-1527-3
45. Gsell AS, Scharfenberger U, Özkundakci D, et al. Evaluating early-warning indicators of critical transitions in natural aquatic ecosystems. *Proceedings of the National Academy of Sciences of the United States of America*. 2016;113(50):1-9. doi:10.1073/pnas.1608242113
46. Guttal V, Raghavendra S, Goel N, Hoarau Q. Lack of critical slowing down suggests that financial meltdowns are not critical transitions, yet rising variability could signal systemic risk. *PLoS one*. 2016;11(1):1-20. doi:10.1371/journal.pone.0144198
47. Boerlijst MC, Oudman T, de Roos AM. Catastrophic collapse can occur without early warning: examples of silent catastrophes in structured ecological models. *PLoS one*. 2013;8(4):1-6. doi:10.1371/journal.pone.0062033
48. Dakos V, Carpenter SR, van Nes EH, Scheffer M. Resilience indicators: prospects and limitations for early warnings of regime shifts. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2015;370(1659):1-10. doi:10.1098/rstb.2013.0263
49. Boettiger C, Hastings A. Quantifying limits to detection of early warning for critical transitions. *Journal of The Royal Society Interface*. 2012;9(75):2527-2539. doi:10.1098/rsif.2012.0125
50. Haslbeck JMB, Ryan O. Recovering within-person dynamics from psychological time series. *Multivariate Behavioural Research*. 2021;56(2):1-32. doi:10.1080/00273171.2021.1896353
51. Boettiger C, Ross N, Hastings A. Early warning signals: the charted and uncharted territories. *Theoretical Ecology*. 2013;6(3):255-264. doi:10.1007/s12080-013-0192-6
52. Scheffer M, Carpenter SR, Lenton TM, et al. Anticipating critical transitions. *Science*. 2012;338(6105):344-348. doi:10.1126/science.1225244
53. Dutta PS, Sharma Y, Abbott KC. Robustness of early warning signals for catastrophic and non-catastrophic

- transitions. *Oikos*. 2018;127(9):1251-1263. doi:10.1111/oik.05172
54. Wichers M, Wigman JTW, Myin-Germeys I. Micro-level affect dynamics in psychopathology viewed from complex dynamical system theory. *Emotion Review*. 2015;7(4):362-367. doi:10.1177/1754073915590623
55. McGorry PD, Hickie IB, Yung AR, Pantelis C, Jackson HJ. Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. *Australian and New Zealand Journal of Psychiatry*. 2006;40(8):616-622. doi:10.1111/j.1440-1614.2006.01860.x
56. Weinans E, Lever JJ, Bathiany S, et al. Finding the direction of lowest resilience in multivariate complex systems. *Journal of the Royal Society Interface*. 2019;16(159):1-9. doi:10.1098/rsif.2019.0629
57. Boettiger C, Hastings A. Early warning signals and the prosecutor's fallacy. *Proceedings of the Royal Society B: Biological Sciences*. 2012;279(1748):4734-4739. doi:10.1098/rspb.2012.2085
58. Keyes CLM. Complete mental health: an agenda for the 21st century. In: Keyes CLM, Haidt J, eds. *Flourishing: Positive Psychology and the Life Well-Lived*. American Psychological Association; 2004:293-312. doi:10.1037/10594-013
59. Scheffer M, Carpenter SR, Dakos V, van Nes EH. Generic indicators of ecological resilience: inferring the chance of a critical transition. *Annual Review of Ecology, Evolution, and Systematics*. 2015;46:145-167. doi:10.1146/annurev-ecolsys-112414-054242
60. Ho DE, Imai K, King G, Stuart EA. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*. 2007;15(3):199-236. doi:10.1093/pan/mdl013

9

Anticipating manic and depressive shifts in patients with bipolar disorder using early warning signals

published as: Bos, F.M.*, Schreuder, M.J.*, George, S.V., Doornbos, B., Bruggeman, R., van der Krieke, L., Haarman, B.C., Wichers, M., Snippe, E., & Bos, F. (2021). Anticipating manic and depressive shifts in patients with bipolar disorder using early warning signals. Preprint. doi: [10.31234/osf.io/v5b7t](https://doi.org/10.31234/osf.io/v5b7t)

*Contributed equally

Abstract

Background. In bipolar disorder treatment, accurate prediction of manic and depressive episodes is paramount but remains difficult. A novel idiographic approach to prediction is to monitor generic early warning signals (EWS), which may manifest in symptom dynamics. EWS could thus form personalized alerts in clinical care. The present study investigated whether EWS can anticipate mood shifts in individual patients with bipolar disorder.

Methods. Twenty bipolar type I/II patients participated in ecological momentary assessment (EMA), completing five questionnaires a day for four months (mean = 491 observations per person). Weekly completed symptom questionnaires on depressive (Quick Inventory for Depressive Symptomatology Self-Report) and manic (Altman Self-Rating Mania Scale) symptoms were used to determine transitions. EWS (rises in autocorrelation at lag-1 and standard deviation) were calculated in moving windows over 17 affective and symptomatic EMA items. Positive and negative predictive values were calculated to determine clinical utility.

Results. Eleven (of the twenty) patients reported 1-2 manic or depressive transitions. The presence of EWS increased the probability of impending transitions towards depression and mania from 32-36% to 46-48% (autocorrelation) and 29-41% (standard deviation). However, the absence of EWS could not be taken as a sign that no transition would occur in the near future. The momentary states that indicated nearby transitions most accurately (predictive values: 65-100%) were cheerfulness, focusing ability, full of ideas, worry, racing thoughts, agitation, energy, and tiredness. Large individual differences in the utility of EWS were found.

Conclusions. EWS may improve anticipating manic and depressive transitions in bipolar disorder.

Introduction

A major challenge in psychiatry is to timely identify impending psychopathological episodes for individual patients. Until now, research has mostly focused on group-level retrospective risk factors¹, which unfortunately say little about *which* individual patient will relapse *when*. However, rapid technological advances have enabled patients to easily monitor their mood and symptoms in real-time, opening the door to prospective and personalized anticipation of clinically relevant symptom changes in the near future². Such early identification of episodes might be especially relevant for patients with bipolar disorder (BD), who experience frequent and disruptive depressive and manic episodes, and whose treatment is strongly focused on episode recognition³. Now that intensive longitudinal monitoring through smartphones has become increasingly feasible⁴, the field is in need of tools to utilize these data to anticipate future increases in psychopathology. Techniques derived from complex dynamical systems theory may provide such tools. In complex dynamical systems, abrupt transitions to alternative states are anticipated by increasing instability of the system. This instability is reflected by a process termed critical slowing down⁵, which means that, as a system approaches a transition, it gets increasingly slow in recovering from minor perturbations⁶. Critical slowing down manifests in patterns in the dynamics of time series data, including rising autocorrelation (*i.e.*, the system's current state increasingly predicts its next state) and rising variance (*i.e.*, the system's current state shows increasing fluctuations). Because critical slowing down may occur prior to transitions in a system, rising autocorrelation and variance have been termed early warning signals (EWS). EWS have been shown to precede transitions in a wide variety of systems, such as climate change⁷, starlight shifts⁸, and animal extinction⁹.

Researchers increasingly recognize that psychopathology might also behave as a complex dynamical system^{10,11}. If this is the case, transitions to psychopathological episodes might be anticipated by EWS in momentary affective and symptomatic states^{12,13}. These states can be assessed through ecological momentary assessment (EMA), with which patients frequently monitor their affect and symptoms in daily life on their smartphones. In psychopathology, rising autocorrelations in momentary states (*e.g.*, affect, stress, worry) indicate higher carryover of one's affective state from one moment to the next. This means that the effects of perturbations (*e.g.*, stressful events) linger longer¹⁴. Rising trends in variance mean that perturbations have an increasingly strong impact on one's momentary state. Preliminary research suggests that EWS in momentary states can indeed anticipate

prospective shifts from a healthy to a depressed state^{15–17}. This raises the question whether EWS could also signal impending mood shifts in BD patients.

For BD patients, a smart personalized prediction tool based on smartphone monitoring could revolutionize treatment. Patients and clinicians describe this as one of the greatest promises of e-mental health^{18,19}. EMA might fulfill a dual purpose in this regard: the monitoring itself might already increase awareness of mood shifts, whereas the gathered data can be analyzed to provide a personalized EWS-based alert system²⁰. Indeed, previous research has demonstrated that daily and weekly mood fluctuations in BD patients follow principles of complex dynamical systems^{21–23}, and that EWS can be detected prior to mood transitions in simulated actigraphy data²⁴. A logical next step would be to examine in empirical data whether EWS in momentary states indeed anticipate mood shifts in BD patients.

Therefore, the present empirical and exploratory study is the first to investigate whether EWS precede mood shifts in BD patients and whether they might have clinical utility. To that end, we employed an exploratory replicated single-subject design in twenty BD patients who participated in EMA for four months. First, we investigated whether EWS improved the detection of impending manic as well as depressive transitions, and whether the absence of EWS could be taken to indicate a lack of impending transitions. Second, we studied which momentary states constituted the best EWS for manic and depressive transitions, and finally, we examined individual differences in the utility of EWS. Taken together, these findings might provide a comprehensive investigation into whether EWS may indeed be used to signal upcoming transitions in BD.

Methods

Participants

For the present prospective observational cohort study, twenty patients with BD type I or II were included (see Table 9.1). To be included, patients had to: (i) be ≥ 18 years, (ii) be diagnosed with and currently in treatment for BD type I/II, and (iii) demonstrate high occurrence of manic and/or depressive episodes (≥ 2 in the previous year). Clinicians of two Dutch tertiary care institutions invited patients for the study until the intended cap of twenty participants was reached. Interested patients were invited to the research facility, where the study was explained in detail and informed consent was signed.

Twenty-eight patients were referred, two of whom were unreachable. Six patients

declined participation during the first telephone call, expecting that study participation and the focus on mood would be too burdensome. This left twenty patients that started and finished the study. Although a seemingly small sample, it is relatively large for idiographic studies. Here, power depends not on the number of individuals, but rather on the number of assessments per individual²⁵. As such, a sample of twenty patients allows us to examine the relative robustness of EWS in improving the detection of transitions. The study was approved by the University Medical Center Groningen medical ethics committee (no. 201501161).

Study design

Participants engaged in EMA and completed weekly questionnaires for four months. Patients received five assessments per day: every three hours (time-contingent schedule), patients received a text message with a link to the EMA questionnaire on their smartphone, which took approximately 1-2 minutes to complete. Assessments were securely administered and stored via Roqua (www.roqua.nl) in patients' personal health records. Patients chose their own start and end time and had one hour to complete each assessment. No reminder prompts were given. Researchers contacted patients after the first three days of monitoring and if compliance was low, if regular contact was preferred, or if the weekly questionnaires indicated above-threshold symptoms. Participants were not offered financial compensation.

Measurements

Manic and depressive transitions

Weekly, patients completed the Dutch versions of the Altman Self-Rating Mania Scale (ASRM²⁶) and the Quick Inventory for Depressive Symptomatology Self-Report (QIDS-SR²⁷). A transition was defined as a clinically relevant abrupt (*i.e.*, within one week) increase in manic (ASRM ≥ 6 ²⁸) or depressive (QIDS-SR ≥ 6 ²⁷) symptoms, without such increases in the two weeks prior to the transition (see supplement for all transitions per individual). Our analytic approach further required at least three weeks (≥ 105 observations) of EMAs prior to the transition.

EMA items

The questionnaire (see Table S9.1) consisted of 26 items pertaining to momentary mood, symptoms, sleep, and activities. These items were based on previous EMA research^{29–31} and interviews with three patients and a psychiatrist on relevant constructs for people with BD.

For the calculation of EWS, we selected the 17 EMA items that assessed affect or symptoms on a continuous scale (*i.e.*, ranging from 0 (“not at all”) to 100 (“very much”)).

Compliance

Two patients were part of a pilot and therefore completed three months of EMA monitoring. The eighteen other patients completed on average 18 weeks of monitoring (range = 16-32). Average compliance, calculated as the number of completed assessments divided by the number of assessments participants received, was 76% (491 assessments, $SD = 137.8$).

Data analysis

For each EMA momentary state, two EWS indicators were examined: rises in the autocorrelation at lag-1 and rises in the standard deviation as indicator of the variance. These EWS were estimated using a moving windows approach^{15,32}. Briefly, this involves iterative computation of EWS within segments (or windows) of the time series, for each individual and transition separately. With each iteration, the window slides one time point ahead until the transition point. Within every window, we computed the autocorrelation and standard deviation of each momentary state. This yielded a new time series for each EWS indicator, allowing us to examine whether rises in the indicator actually preceded an abrupt manic or depressive transition. Pre-processing and analysis of the data were performed in the statistical programming language R³³. Results were visualized using the R-packages *ggplot2*³⁴ and *gridExtra*³⁵.

Pre-processing steps

Outliers were winsorized to minimize their influence on the results. To ensure there were no trends in the mean over time (*i.e.*, stationarity)³², and reduce the risk of false positives³⁶, the data were detrended by applying a Gaussian kernel smoothing function over the whole pre-transition period^{7,37}. Missing data were not imputed because this might result in spurious correlations³².

Window size

The window size is related to the timescale at which the system dynamics evolve and may affect results³². Therefore, sensitivity analyses were conducted to investigate the effects of windows of one, two or three weeks. We ensured that each window included an equal

number of weekend days, thereby negating their effects on the results²⁴. Sensitivity analyses indicated small differences for different window sizes (see supplement Table S9.2), but results were largely robust. Therefore, based on suggestions on how bipolar dynamics develop in patients with high mood instability³⁸, we opted for a window size of two weeks (*i.e.*, 70 observations), yielding on average 168 windows ($SD = 99$, range = 35-415) per transition.

Calculation of EWS

First, within each window, the autoregressive coefficient was calculated at lag-1 over the residuals obtained after detrending³². The autocorrelation thus indicates how well a momentary state (*e.g.*, feeling cheerful) predicts itself three hours later. Overnight lags were prevented by deleting the lagged observation of each day's first observation. Second, variance was estimated from the standard deviation¹ over the residuals within each window³². A rising standard deviation indicates that someone's momentary state varies more widely around the mean over time.

Significance testing

The above approach resulted in a new time series data set for every individual, with separate estimates of the autocorrelation and standard deviation. To test whether each EWS indicator significantly increased, we calculated the Kendall correlation coefficient (τ_K) over this data set³², in the two weeks prior to the transition (using the R-package *Kendall*³⁹). A significant positive correlation indicates that the EWS indicator significantly rose prior to a transition. The value of Kendall's tau was taken to reflect the strength of the EWS. Since the moving windows use overlapping data, we corrected for this dependency between nearby windows by applying the Hamed-Rao correction⁴⁰. Furthermore, we corrected for multiple testing by using the false discovery rate (FDR) approach as proposed by Benjamini and Hochberg⁴¹. This ensures that, across all tests performed within a single patient, the probability of false positives is 5%.

¹ The standard deviation is the square root of the variance. This transformation does not affect the rank correlation coefficient Kendall's tau.

Predictive value

To gain insight into the predictive utility of EWS for detecting mood transitions, positive and negative predictive values (PPV and NPV respectively) were calculated⁴². The PPV and NPV are based on the sensitivity (proportion of true positives, where EWS preceded transitions) and specificity (proportions of true negatives, where the absence of EWS indicates the absences of transitions). Predictive values indicate to what extent (i) the presence of EWS increases the likelihood of detecting a transition (PPV) and (ii) their absence increases the likelihood of detecting that *no* transition will take place (NPV). PPV and NPV were calculated for both EWS indicators averaged across momentary states, as well as for each momentary state separately. Predictive values follow the below formulas, using the sensitivity and specificity of a particular momentary state (i) for detecting transitions (t) towards either mania or depression:

$$PPV_{i,t} = \frac{sensitivity_{i,t} * prevalence_t}{sensitivity_{i,t} * prevalence_t + (1 - specificity_i) * (1 - prevalence_t)}$$

$$NPV_{i,t} = \frac{specificity_i * (1 - prevalence_t)}{(1 - sensitivity_{i,t}) * prevalence_t + specificity_i * (1 - prevalence_t)}$$

Higher predictive values indicate higher accuracy of EWS in detecting transitions. A PPV of 100% indicates that an EWS has no false positives, whereas a NPV of 100% indicates an EWS has no false negatives. The PPV and NPV demonstrate whether EWS improve the accuracy of detecting transitions above the general prevalence of transitions in our sample, and should thus be higher than the average transition prevalence (*i.e.*, the proportion of manic (32%), depressive (36%), or no transitions (68% for mania and 64% for depression)). See supplement for a step-by-step calculation of predictive values and supplement Table S9.3 for sensitivity and specificity for all EWS.

To estimate false positives, EWS were also computed for individuals without a transition. For them, we selected a period of approximately 245 observations (*i.e.*, mean number of observations in the pre-transition period of transitioning patients), without any changes of ≥ 4 on the ASRM and QIDS-SR. This was possible for seven out of nine non-transitioning patients.

Results

Sample characteristics

Demographic and clinical characteristics are depicted in Table 9.1. Of the twenty patients, eleven reported an abrupt manic or depressive mood shift, four of whom reported two transitions. Notably, during depressed transitions, all patients reported subthreshold mania scores. During manic transitions, all but two patients reported above threshold scores for depression.

Table 9.1. *Demographic and clinical characteristics*

	Full sample (N=20)	Patients with transition (N=11)	Patients without transition (N=9)
<i>Gender (N)</i>			
Male	4	3	1
Female	16	8	8
<i>Age (N)</i>			
20-35 years	9	4	5
36-50 years	8	5	3
51-65 years	3	2	1
<i>Education level (N)</i>			
Higher education	9	6	3
Secondary education	5	2	3
Secondary vocational education	3	1	2
Pre-vocational education	3	2	1
<i>Years since bipolar disorder diagnosis (M, SD)</i>	6.4 (6.3)	5.0 (5.8)	8.2 (6.8)
<i>Years in treatment (M, SD)</i>	10.6 (8.8)	10.1 (8.5)	11.27 (9.7)
<i>Bipolar disorder diagnosis (N)</i>			
Bipolar disorder type I	9	6	5
Bipolar disorder type II	11	5	4
<i>Comorbid diagnoses (N)</i>			
No comorbid Axis I/II disorder	12	5	7
Attention Deficit/Hyperactivity Disorder	1	1	0
Autism Spectrum Disorder	1	1	0
Sleep disorder	1	1	0
Alcohol/drug dependence	1	0	1
Personality disorder	6	5	1
<i>Medication use (N)</i>			
None	2	1	1
Amphetamine	1	1	0
Anti-epileptic	10	8	2

	Full sample (N=20)	Patients with transition (N=11)	Patients without transition (N=9)
Atypical antipsychotic	10	5	5
Benzodiazepine	9	6	3
Thyreomimetica	2	0	2
Lithium	5	0	5
Monoamine oxidase inhibitor	3	3	0
Selective serotonin reuptake inhibitor	4	2	2
Tricyclic antidepressant	1	1	0
<i>Transitions¹ (N)</i>			
To a manic episode		7	
To a depressive episode		8	
<i>Symptom increase in week of transition (M, SD)</i>			
Transition to a manic episode (ASRM)		6.7 (1.5)	
Transition to a depressive episode (QIDS-SR)		11.3 (5.8)	
<i>Episode duration after transition in weeks (M, SD)</i>			
Manic episode		1.9 (0.9)	
Depressive episode		2.6 (3.1)	
<i>Compliance to EMA (%)</i>	75.9	74.6	77.6

Note: ¹Four patients reported two transitions, the remaining seven patients reported one transition. ASRM = Altman Self-Rating Mania Scale. M = mean. N = number. QIDS-SR = Quick Inventory of Depressive Symptomatology Self-Report. SD = standard deviation.

A clinical illustration

We will first present a clinical case example to illustrate EWS in an individual patient (ID6): a 27-year-old woman diagnosed with BD type II. She reported both a manic and a depressive transition on the weekly ASRM and QIDS-SR (Figure 9.1a). Her weekly symptom scores had been stable in the previous weeks. Note that this would imply similar stability in her EMA momentary states until the transitions.

Figure 9.1b shows her EMA observations for the state ‘*extremely well*’. Higher scores indicate she is feeling more euphoric. It is difficult to distill clear patterns from these data; we see a lot of moment-to-moment variation, and more missed assessments after the manic transition. Figure 9.1c shows significant EWS in the autocorrelation for ‘*extremely well*’ prior to both transitions. This means that, within the two weeks before reporting a manic and depressive shift, her euphoria increasingly lingered over time. Figure 9.1d shows an EWS in the standard deviation for the depressive transition, but not for the manic transition, indicating that her euphoria varied more widely prior to the depressive transition.

Fig. 9.1. An illustration of early warning signals in one individual (ID6) in the item "I feel extremely well". A) Weekly manic (Altman Self-Rating Scale, ASRM, blue) and depressive (Quick Inventory of Depressive Symptomatology, QIDS-SR, orange) symptom scores. At week 8 and 15, the patient reports an abrupt transition to a manic and depressive episode, respectively. B) EMA observations for the item "I feel extremely well". Higher scores indicate that the patient is feeling more down. C) Significant early warning signals (EWS) in the autocorrelation for "I feel extremely well" for the transition to the manic (Kendall's Tau = .54, *corrected* $p < .001$) as well as the depressive episode (Kendall's Tau = .68, *corrected* $p < .001$). D) An EWS in the standard deviation prior to the depressive transition (Kendall's Tau = .75, *corrected* $p < .001$), but not prior to the manic transition (Kendall's Tau = .50, *corrected* $p = .07$).

Predictive value of early warning signals

All transitions, both depressive and manic, were preceded by at least one EWS (*i.e.*, a significant rise in the autocorrelation or standard deviation) in at least one of the EMA momentary states. Regarding the autocorrelation, on average, most EWS were found prior to manic transitions (mean = 5.6) versus depressive transitions (mean = 4.5) and non-transitions (mean = 4.2). The standard deviation also yielded the most EWS prior to manic transitions (mean = 6.7), versus depressive transitions (mean = 2.8) and non-transitions (mean = 5.3). In general, average PPVs and NPVs indicate that EWS were better in signaling the presence of transitions than their absence (see Fig. 9.2).

Regarding depressive transitions, the average PPV indicates that if at least one EWS in the autocorrelation was detected, the probability of anticipating a depressive transition increased from 36% (prevalence) to 46%. However, if EWS were absent, the probability of correctly inferring *no* depressive transition did not improve (NPV = 63% versus 64% prevalence). The standard deviation was not an accurate EWS for depressive transitions, with a PPV of 29% and a NPV of 58%.

For manic transitions, the autocorrelation was again slightly more accurate in signaling transitions. Here, EWS improved the probability of correctly inferring a manic transition from 32% (prevalence) to 48% (autocorrelation) or 41% (standard deviation). However, the NPV demonstrates that the probability of inferring *no* manic transition in the absence of EWS improved only slightly from 68% (prevalence) to 74% (autocorrelation) or 75% (standard deviation).

Predictive value of specific EMA momentary states

Figure 9.2 demonstrates the PPVs and NPVs for all momentary states separately. Four momentary states had a PPV of 100% for both depressive and manic transitions: *'cheerful'* (autocorrelation), *'ability to focus/switch'* (autocorrelation), *'full of ideas'* (autocorrelation), and *'worry'* (standard deviation), indicating they were never found for patients without a transition. The NPVs of these momentary states ranged from 70-83%. These momentary states might thus signal an impending transition without specifying its nature (depressive or manic).

Other momentary states were found to specifically improve the detection of either depressive or manic transitions, as indicated by an average PPV and NPV above 70%. For depressive transitions, the only momentary state yielding such EWS was *'tired'* (autocorrelation). For manic transitions, the best EWS were found in *'racing thoughts'* (autocorrelation), *'agitated'* (standard deviation), and *'full of energy'* (standard deviation).

'Socializing' and *'down'* never improved the detection of either manic or depressive transitions. A particularly inaccurate EWS for manic transitions was *'content'* (autocorrelation). For depressive transitions, three momentary states had a PPV of 0% because EWS were never detected prior to transitions: *'distracted'* (autocorrelation and standard deviation) and *'cheerful'* (standard deviation).

Fig. 9.2. Positive and negative predictive values for each early warning signal. The y-axis represents each momentary state, the x-axis the positive (PPV) and negative (NPV) predictive value, separated for manic and depressive transitions and for the two early warning signals (EWS) indicators: the autocorrelation (AR) and standard deviation (SD). The predictive values can be compared against the prevalence of the transition: the proportion of manic (32%), depressive (36%), or no transitions (68% for mania and 64% for depression). White tiles indicate that this EWS did not improve the detection of a transition above the prevalence of that transition. The color indicates the magnitude of the predictive value for that EWS: the more intense the color, the higher the predictive value. To facilitate interpretation, the EMA momentary states were assigned to summary categories based on hypothesized underlying constructs.

Individual differences

Large individual differences were found in the presence, type, and strength of EWS momentary states (see Fig. 9.3). For example, ID4 showed EWS in the autocorrelation for ten momentary states prior to a depressive transition, whereas ID2 reported none prior to their second depressive transition. Whereas the autocorrelation of ‘down’ was a particularly strong depressive EWS for ID9, this EWS was not found in any of the seven other transitions. For the four patients that reported two transitions, EWS often did not replicate. Exceptions include the autocorrelations of ‘extremely well’ (ID6) and ‘physically active’ (ID6), and the standard deviations of *physically active* (ID1 and ID7) and ‘extremely well’, ‘full of energy’, ‘physically active’, and ‘racing thoughts’ (ID7).

Fig. 9.3. *Individual differences in the type and strength of the early warning signal.* The x-axis represents each EMA momentary state, the y-axis each transition (note that four individuals had two transitions). To facilitate interpretation, the EMA momentary states were assigned to summary categories based on hypothesized underlying constructs. A colored block indicates that the EWS was significant for that transition. The color intensity indicates the strength of the EWS: the more intense the color, the stronger the EWS. Strength of the EWS was operationalized as the value of Kendall's tau. AR = autocorrelation at lag-1. EMA = ecological momentary assessment. EWS = early warning signal. sd = standard deviation.

Discussion

The present exploratory study investigated whether EWS in momentary affective and symptomatic states anticipate abrupt mood transitions in BD patients, and whether EWS might have clinical utility. Results provide preliminary support that EWS can indeed be detected in momentary states that were collected by longitudinal smartphone monitoring, and that EWS could improve the detection of mood transitions. Notably, the presence of EWS increased the probability of detecting impending transitions, but their absence could not be taken as a sign that no transition would occur in the near future. That is, although several momentary states maximized the PPV, indicating that no false positives were found, no momentary state maximized the NPV, indicating the presence of false negatives (*i.e.*, EWS were not always detected prior to transitions). Furthermore, momentary states differed in their predictive utility, and we found large inter-individual differences in the predictive capacity of EWS. Finally, the autocorrelation outperformed the standard deviation as an EWS.

The predictive values point towards several momentary states as promising indicators of nearby depressive and manic transitions. Momentary states that signaled both manic and depressive transitions were *cheerfulness*, *focusing ability*, *full of ideas*, and *worrying*. Manic transitions were most often anticipated by EWS in *racing thoughts*, *agitation*, and *full of energy*, whereas depressive transitions were most often preceded by EWS in *tiredness*. This supports the hypothesis that transitions are best anticipated by EWS in momentary states matching the underlying psychopathology⁴³. Importantly, not all momentary states constituted as accurate EWS, highlighting the importance of exploring individual items.

Contrasting to group-level risk factors that offer little guidance on the timing of relapses for individual patients, our study provides preliminary support that EWS might be used to establish *when* individual patients will relapse⁴⁴. If our results are replicated, personalized EWS as generated by smartphone gathered data could be used to alert patients and clinicians to nearby mood shifts. Timely identification of mood shifts is paramount to BD treatment, but is complicated because patients usually only recognize them when episodes have already started¹⁸. Early detection using EWS may enable early intervention that could mitigate the severity and impact of ensuing episodes⁴⁵. However, the large heterogeneity we found highlights the necessity of a personalized approach. Furthermore, the finding that EWS had limited replicability within the four individuals with two transitions tentatively suggests that continuous monitoring of various momentary states may be necessary for accurate

mood shift detection, as opposed to a less burdensome approach where personalized EWS have to be identified only once prior to a transition. As such, before EWS can find their way to clinical practice, more confirmatory research is needed.

On a critical note, even for the most accurate momentary states, EWS were not always present prior to mood shifts, and were sometimes also found in individuals without actual shifts. False positives may occur when the data show more noise or variability⁴⁶. Patients in our sample indeed showed a high frequency of mood shifts, similar to (ultra) rapid-cycling patients that experience stark fluctuations of mood most of the time⁴⁷. However, failing to anticipate actual mood shifts (false negatives) might be more problematic for clinical applications than the occasional false alarm. The absence of EWS may be explained by the fact that most transitions are often governed by only a few variables⁴⁸. Indeed, the momentary states that contained EWS differed between the individuals in our sample. False negatives may also have occurred because of the lack of stable mood episodes in the patients in our sample. That is, critical slowing down assumes transitions take place from one stable state to another, whereas the ‘stable’ state of patients in our sample may well be characterized by large mood instability⁴⁹. Finally, false negatives could occur because the timescale of the EMA assessments (five/day) may not match the timescale at which critical slowing down takes place⁵⁰. Future research may investigate how the bipolar mood system should be characterized and at what timescale it unfolds: should we attempt to anticipate distinct mood episodes, or rather mood instability in general⁵¹? The latter approach may be more promising given that many BD patients have mixed episodes, as also demonstrated by the patients in our sample. Thus, an interesting follow-up study would be to investigate EWS in more stable BD patients, who transition from stable euthymic states to manic or depressed states.

The present study is the first to investigate whether EWS anticipate transitions in BD patients in empirical data. Notably, previous studies have only examined autocorrelation and standard deviation as static indicators of future shifts^{21–23}, whereas we have prospectively examined increases in these indicators to anticipate nearby mood shifts. Further, we used a personalized and idiographic approach that has been advocated to study within-person processes²⁵. Other strengths include the relatively large sample (N=20) for idiographic studies, in which each patient could be viewed as a replication of results in other patients, and the large number of observations per person. Furthermore, the diversity in momentary states under investigation might provide new suggestions for confirmatory research.

The present results should be viewed in light of several limitations. First, our sample was diverse, consisting of BD patients with different treatment regimens and comorbid (personality) diagnoses characterized by high mood variability, which might have obscured the relation between EWS and transitions. Second, EWS studies in other fields suggest that results may be dependent on the analytical decisions regarding window size, data detrending, and the period over which the rise is calculated (two weeks)^{32,37}. Third, the question remains whether the above-threshold scores on weekly self-report symptom questionnaires adequately reflect abrupt and clinically meaningful transitions. Fourth, our estimates of the prevalence of transitions should be interpreted tentatively given our study design. However, the prevalence of transitions was quite high in this sample, rendering our approach to compare the predictive values against the prevalence rather conservative. Finally, given that our study was exploratory in nature, results may not generalize to other samples or data sets in which different methods to study EWS are employed.

To conclude, EWS show promise in signaling impending transitions to manic and depressive mood shifts in BD. Future confirmatory research may focus on examining EWS in cheerfulness, focusing ability, full of ideas, worry, energy, agitation, racing thoughts, and tiredness, as these items had the best predictive value. Further investigation into the individual characteristics that determine the clinical utility of EWS is warranted.

Supplement 9

9.1 All transitions per individual

Fig. S9.1. *Transitions in manic and depressive symptoms per individual.* Purple bars reflect the stable period that was used for computing EWS.

Table S9.1. *Ecological momentary assessment (EMA) diary items*

Dutch	English translation	Is this the first assessment you complete today?	How long did you sleep?	Anchors (far left, middle, far right)	Which prompt(s)	Reason for exclusion
Is dit de eerste meting die u invult vandaag?	Is this the first assessment you complete today?	- Yes - No	0 – 12 hours		All five prompts	
1. Hoe lang heeft u geslapen?	How long did you sleep?				If answered 'yes' on item 1	Assessed on different timescale, yielding too few data points for analysis
2. De kwaliteit van mijn slaap was...	The quality of my sleep was...			Very bad – reasonably – very good	If answered 'yes' on item 1	Assessed on different timescale, yielding too few data points for analysis
Ik voel me opgewekt	I feel cheerful			Not at all – reasonably – very much	All five prompts	
Ik voel me neerslachtig	I feel down			Not at all – reasonably – very much	All five prompts	
Ik voel me kalm	I feel calm			Not at all – reasonably – very much	All five prompts	
Ik voel me gejaagd	I feel agitated			Not at all – reasonably – very much	All five prompts	
Ik voel me bijzonder goed	I feel extremely well			Not at all – reasonably – very much	All five prompts	
Ik voel me moe	I feel tired			Not at all – reasonably – very much	All five prompts	
Ik voel me tevreden	I feel content			Not at all – reasonably – very much	All five prompts	
Ik voel me geïrriteerd	I feel irritated			Not at all – reasonably – very much	All five prompts	
Ik zit vol energie	I am full of energy			Not at all – reasonably – very much	All five prompts	
Ik zie op tegen de rest van de dag	I dread the rest of the day			Not at all – reasonably – very much	All five prompts	Responses invalid for one participant who accidentally misinterpreted the scale
Ik zit vol goede ideeën	I am full of good ideas			Not at all – reasonably – very much	All five prompts	
Ik heb het gevoel te kort te schieten	I feel inadequate			Not at all – reasonably – very much	All five prompts	
Mijn gedachten gaan snel	My thoughts are racing			Not at all – reasonably – very much	All five prompts	
Ik kan snel schakelen	I am able to focus and switch easily			Not at all – reasonably – very much	All five prompts	
Ik ben snel afgeleid	I'm distracted easily			Not at all – reasonably – very much	All five prompts	
Ik heb zin om met anderen af te spreken	I feel like socializing			Not at all – reasonably – very much	All five prompts	

Dutch	English translation	Anchors (far left, middle, far right)	Which prompt(s)	Reason for exclusion
Eigen vraag	Personal question	Not at all – reasonably – very much	All five prompts	Item differs across participants
Ik heb meer gegeten dan gewoonlijk	I have eaten more than usual	Not at all – reasonably – a lot more than usual	All five prompts	Too dependent on the occurrence of events
Sinds het vorige meetmoment heb ik gepiekerd	Since the previous prompt, I have worried	Not at all – reasonably – very much	All five prompts	
Sinds het vorige meetmoment heb ik veel gecommuniceerd	Since the previous prompt, I have communicated a lot	Not at all – reasonably – very much	All five prompts	Too dependent on the occurrence of events
Sinds het vorige meetmoment heb ik me lichamelijk ingespannen	Since the previous prompt, I have been physically active	Not at all – reasonably – very much	All five prompts	
Denk aan de meest opvallende gebeurtenis sinds het vorige meetmoment. Hoe heftig was deze gebeurtenis?	Think back on the most notable event since the previous prompt. How intense was this event?	Not at all – reasonably – very much	All five prompts	Too dependent on the occurrence of events
Is dit de laatste meting die u invult vandaag?	Is this the last assessment you complete today?	- Yes - No	All five prompts	
3. Ik heb vandaag veel afspraken gemaakt	I have made a lot of appointments today	Not at all – reasonably – a lot	If answered 'yes' on item 26	Assessed on different timescale, yielding too few data points for analysis
Sinds het vorige meetmoment, wat deed ik?	Since the previous prompt, what have I been doing? (multiple options possible)	<ul style="list-style-type: none"> - Sleeping - Household chores/groceries - Working/studying - Doing sports/walking/cycling - Something relaxed (e.g., reading, TV) - Hobby (e.g., gardening, music) - A trip (e.g., into town, concert) - Something together with others - Something intimate (e.g., cuddling, sex) - Engaging in self-care - Resting/nothing - On the way - Something else 	All five prompts	Not a continuous variable
Noteer hier eventuele opmerkingen. Noteer het ook als er iets gebeurd is dat invloed heeft op uw stemming.	Note observations here if any. Also note anything that may have influenced your mood.	Open entry	All five prompts (optional)	Not a continuous variable

9.2 Experience sampling method diary items

The item list was constructed in several steps. First, we identified relevant concepts for symptoms of bipolar disorder and searched the literature for EMA studies in patients with bipolar disorder. This yielded a first item list of 71 potentially relevant items. This list was then discussed in individual interviews with three patients and one psychiatrist. The items were finally selected on three criteria: (i) the patients and clinician recognized the item as signaling symptoms of either mania or depression; (ii) the patients and the clinician felt comfortable with the formulation of the item (*e.g.*, they could see themselves saying the sentence in daily life); (iii) both the patients, the clinician, and the research team believed the item would vary meaningfully within participants. Items that were deemed relevant but too person-specific were put on a list for participants so they might select them for their personal question.

All items were obligatory, with the exception of the comment field at the end of the questionnaire (item 29). All items were assessed five times per day, with the exceptions of the items regarding sleep (items 2-3), and the item regarding appointments (item 27). These items were only shown if participants answered ‘yes’ on item 1 or item 26. This way, we ensured that participants could still answer questions about their sleep, even though they might have skipped the first (few) assessments.

9.3 Results of the sensitivity analyses for the window size

All early warning signal (EWS) analyses were run for window sizes of 1, 2, and 3 weeks. Table S9.2 shows the predictive values, sensitivity, and specificity of the autocorrelation and standard deviation, averaged across all EMA diary items. The differences between the window sizes are quite small. One notable finding is that a three-week window is preferable when examining the autocorrelation prior to transitions to depression and mania. For the standard deviation, the two-week window was preferable.

Table S9.2. Predictive values, sensitivity, and specificity of early warning signals, for three different window sizes, averaged across all EMA diary items

	Depression			Mania			
	PPV	NPV	Sensitivity	PPV	NPV	Sensitivity	Specificity
<i>Indicator: AR</i>							
1 week window	43	65	25	45	70	25	78
2 week window	46	63	27	48	69	33	75
3 week window	59	69	36	58	73	30	84
<i>Indicator: SD</i>							
1 week window	32	61	18	42	70	30	72
2 week window	27	58	23	41	71	38	67
3 week window	18	38	15	38	68	30	67

Note. AR = autocorrelation. PPV = positive predictive value. NPV = negative predictive value. SD = standard deviation.

9.4 Calculation of positive and negative predictive values, sensitivity, and specificity

The sensitivity of EWS reflects the probability of EWS, given that a transition is present. Sensitivity was calculated per indicator (autocorrelation, standard deviation), momentary state (i) and transition type (t ; depression, mania) by dividing the number of EWS in momentary state i that anticipated transition t by the total number of transitions of type t . As an example, the autocorrelation in feeling tired anticipated four out of eight transitions towards depression, resulting in a sensitivity of .50 (*i.e.*, 50%).

$$sensitivity_{i,t} = \frac{EWS_{i,t}}{N_t}$$

The specificity of EWS reflects the probability of *no* EWS, given that there was no transition. This statistic was calculated by inspecting EWS in those individuals who did not experience a transition ($N = 7$). A specificity of 1 (*i.e.*, 100%) indicates that EWS in this particular item were never found in individuals without transitions.

$$specificity_i = 1 - \frac{EWS_i}{N}$$

Positive predictive values (PPVs) indicate the probability of a transition towards either mania or depression, given that EWS are detected. These values were calculated as follows:

$$PPV_{i,t} = \frac{sensitivity_{i,t} * prevalence_t}{sensitivity_{i,t} * prevalence_t + (1 - specificity_i) * (1 - prevalence_t)}$$

In this formula, the prevalence of a transition of type t reflects the number of transitions of type t divided by the total number of transitions. In total, we identified 22 transitions, of which 8 (36%) involved a sudden increase in symptoms of depression, 7 (32%) involved an increase in manic symptoms, and 7 (32%) were simulated in individuals without transitions. The latter ‘non-transitions’ were used to estimate the specificity of EWS.

Negative predictive values (NPVs) indicate the probability that *no* transition will occur, given that EWS are *not* detected. These values were calculated as follows:

$$NPV_{i,t} = \frac{specificity_i * (1 - prevalence_t)}{(1 - sensitivity_{i,t}) * prevalence_t + specificity_i * (1 - prevalence_t)}$$

Note that the probability of *no* transition is defined by 1 minus the prevalence of a particular type of transition, resulting in 64% (depression) and 68% (mania), respectively. The prevalence of a transition can be thought of as a set-point: in absence of any insight in EWS, the probability that a specific individual will (not) experience a transition equals the average probability of (no) transitions. EWS can be considered informative if (i) their presence considerably heightens the risk of a future transition and (ii) their absence lowers the probability that a transition will occur. Therefore, in Table S9.3, we printed PPV and NPV values that exceeded the expected probabilities ($prevalence_t$ or $1 - prevalence_t$).

Table S9.3. Predictive values, sensitivity, and specificity of early warning signals for transitions towards depression and mania

	Depression			Mania			
	PPV	NPV	Sensitivity	PPV	NPV	Sensitivity	Specificity
<i>Indicator: AR</i>							
calm	20	59	12	19	64	14	71
cheerful	100	74	38	100	71	14	100
content	23	44	38	9	42	14	29
irritated	20	59	12	41	73	43	71
thoughts racing	50	67	25	70	87	71	86
distracted	0	50	0	32	68	43	57
physically active	43	67	38	32	68	29	71
socializing	20	50	25	19	56	29	43
down	14	53	12	24	63	29	57
inadequate	43	67	38	19	64	14	71
worry	33	63	12	58	76	43	86
tired	67	75	50	48	72	29	86
agitated	50	67	25	48	72	29	86
extremely well	43	67	38	32	68	29	71
full of ideas	100	70	25	100	83	57	100
focus switch	100	74	38	100	75	29	100
full of energy	50	67	25	58	76	43	86
<i>Indicator: SD</i>							
calm	33	63	12	58	76	43	86
cheerful	0	50	0	24	63	29	57
content	20	59	12	41	73	43	71
irritated	33	63	12	48	72	29	86
thoughts racing	14	53	12	44	81	71	57
distracted	0	56	0	41	73	43	71
physically active	27	55	38	26	62	43	43
socializing	14	53	12	24	63	29	57
down	20	59	12	19	64	14	71
inadequate	20	59	12	19	64	14	71
worry	100	70	25	100	75	29	100
tired	14	53	12	24	63	29	57
agitated	33	63	12	65	81	57	86
extremely well	20	59	12	48	78	57	71
full of ideas	25	57	25	32	68	43	57
focus switch	20	50	25	26	62	43	43
full of energy	60	71	38	65	81	57	86

Note. Bold values indicate that the warning signal improved the prediction of transitions towards manic or depressive episodes when compared to the average prevalence of these transitions. AR = autocorrelation. NPV = negative predictive value. PPV = positive predictive value. SD = standard deviation.

References

1. van Meter AR, Burke C, Youngstrom EA, Faedda GL, Correll CU. The bipolar prodrome: meta-analysis of symptom prevalence prior to initial or recurrent mood episodes. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2016;55(7):543-555.
2. Dunster GP, Swendsen J, Merikangas KR. Real-time mobile monitoring of bipolar disorder: a review of evidence and future directions. *Neuropsychopharmacology*. 2021;46(1):197-208. doi:10.1038/s41386-020-00830-5
3. Michalak EE, Yatham LN, Kolesar S, Lam RW. Bipolar disorder and quality of life: a patient-centered perspective. *Quality of Life Research*. 2006;15(1):25-37. doi:10.1007/s11136-005-0376-7
4. Vachon H, Viechtbauer W, Rintala A, Myin-Germeys I. Compliance and retention with the experience sampling method over the continuum of severe mental disorders: meta-analysis and recommendations. *Journal of Medical Internet Research*. 2019;21(12):1-20. doi:10.2196/14475
5. Scheffer M, Bascompte J, Brock WA, et al. Early-warning signals for critical transitions. *Nature*. 2009;461(7260):53. doi:10.1038/nature08227
6. Scheffer M, Carpenter SR, Lenton TM, et al. Anticipating critical transitions. *Science*. 2012;338(6105):344-348. doi:10.1126/science.1225244
7. Dakos V, Scheffer M, van Nes EH, Brovkin V, Petoukhov V, Held H. Slowing down as an early warning signal for abrupt climate change. *PNAS*. 2008;105(38):14308-14312. doi:10.1073/pnas.0802430105
8. George S v., Kachhara S, Misra R, Ambika G. Early warning signals indicate a critical transition in Betelgeuse. *Astronomy and Astrophysics*. 2020;640. doi:10.1051/0004-6361/202038785
9. Drake JM, Griffen BD. Early warning signals of extinction in deteriorating environments. *Nature*. 2010;467(7314):456-459. doi:10.1038/nature09389
10. Cramer AOJ, van Borkulo CD, Giltay EJ, et al. Major depression as a complex dynamic system. *PLoS one*. 2016;11(12):1-20. doi:10.1371/journal.pone.0167490
11. Hayes AM, Andrews LA. A complex systems approach to the study of change in psychotherapy. *BMC Medicine*. 2020;18(1):1-13. doi:10.1186/s12916-020-01662-2
12. Nelson B, McGorry PD, Wichers M, Wigman JT, Hartmann JA. Moving from static to dynamic models of the onset of mental disorder. *JAMA Psychiatry*. 2017;74(5):528-534. doi:10.1001/jamapsychiatry.2017.0001
13. Olthof M, Hasselman F, Strunk G, et al. Critical fluctuations as an early-warning signal for sudden gains and losses in patients receiving psychotherapy for mood disorders. *Clinical Psychological Science*. 2019;8(1):1-11. doi:10.1177/2167702619865969
14. Kuppens P, Allen NB, Sheeber LB. Emotional inertia and psychological maladjustment. *Psychological Science*. 2010;21(7):984-991. doi:10.1177/0956797610372634
15. Wichers M, Groot PC, Psychosystems ESM group, EWS group. Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*. 2016;85:114-116. doi:10.1159/000441458
16. Wichers M, Smit AC, Snippe E. Early warning signals based on momentary affect dynamics can expose nearby transitions in depression: a confirmatory single-subject time-series study. *Journal for Person-Oriented Research*. 2020;6(1):1-15. doi:10.17505/jpor.2020.22042
17. Cabrieto J, Adolf J, Tuerlinckx F, Kuppens P, Ceulemans E. Detecting long-lived autodependency changes

- in a multivariate system via change point detection and regime switching models. *Scientific Reports*. 2018;8(1):1-15. doi:10.1038/s41598-018-33819-8
18. Murnane EL, Cosley D, Chang P, et al. Self-monitoring practices, attitudes, and needs of individuals with bipolar disorder: implications for the design of technologies to manage mental health. *Journal of the American Medical Informatics Association*. 2016;23(3):477-484. doi:10.1093/jamia/ocv165
 19. Saunders KEA, Bilderbeck AC, Panchal P, Atkinson LZ, Geddes JR, Goodwin GM. Experiences of remote mood and activity monitoring in bipolar disorder: a qualitative study. *European Psychiatry*. 2017;41:115-121. doi:10.1016/j.eurpsy.2016.11.005
 20. Bos FM, Snippe E, Bruggeman R, Doornbos B, Wichers M, van der Krieke L. Recommendations for the use of long-term experience sampling in bipolar disorder care: a qualitative study of patient and clinician experiences. *International Journal of Bipolar Disorders*. 2020;8(1):8-38. doi:10.1186/s40345-020-00201-5
 21. Bonsall MB, Geddes JR, Goodwin GM, Holmes EA. Bipolar disorder dynamics: affective instabilities, relaxation oscillations and noise. *Journal of the Royal Society Interface*. 2015;12(112):1-11. doi:10.1098/rsif.2015.0670
 22. Ortiz A, Alda M. The perils of being too stable: mood regulation in bipolar disorder. *Journal of Psychiatry and Neuroscience*. 2018;43(6):363-365. doi:10.1503/jpn.180183
 23. Curtiss J, Fulford D, Hofmann SG, Gershon A. Network dynamics of positive and negative affect in bipolar disorder. *Journal of Affective Disorders*. 2019;249:270-277. doi:10.1016/j.jad.2019.02.017
 24. Bayani A, Hadaeghi F, Jafari S, Murray G. Critical slowing down as an early warning of transitions in episodes of bipolar disorder: a simulation study based on a computational model of circadian activity rhythms. *Chronobiology International*. 2017;34(2):235-245. doi:10.1080/07420528.2016.1272608
 25. Zuidersma M, Riese H, Snippe E, Booij SH, Wichers M, Bos EH. Single-subject research in psychiatry: facts and fictions. *Frontiers in Psychiatry*. 2020;11:1-11. doi:10.3389/fpsy.2020.539777
 26. Altman EG, Hedeker D, Peterson JL, Davis JM. The Altman self-rating mania scale. *Biological Psychiatry*. 1997;42(10):948-955.
 27. Rush AJ, Trivedi MH, Ibrahim HM, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), Clinician Rating (QIDS-C), and Self-Report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biological Psychiatry*. 2003;54:573-583. doi:10.1016/S0006-3223(03)01866-8
 28. Miller CJ, Johnson SL, Eisner L. Assessment tools for adult bipolar disorder. *Clinical Psychology: Science and Practice*. 2009;16(2):188-201.
 29. Knapen S. Rhythm & blues: chronobiology in the pathophysiology and treatment of mood disorders. Published online 2019.
 30. Krieke L van der, Jeronimus BF, Blaauw FJ, et al. HowNutsAreTheDutch (HoeGekisNL): a crowdsourcing study of mental symptoms and strengths. *International Journal of Methods in Psychiatric Research*. 2016;25(2):123-144. doi:10.1002/mpr
 31. Tyler E, Jones S, Black N, Carter LA, Barrowclough C. The relationship between bipolar disorder and cannabis use in daily life: an experience sampling study. *PLoS one*. 2015;10(3):1-15. doi:10.1371/journal.pone.0118916
 32. Dakos V, Carpenter SR, Brock WA, et al. Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLoS one*. 2012;7(7):1-20. doi:10.1371/journal.pone.0041010
 33. R Core Team. R: A language and environment for statistical computing. Published online 2020.

34. Wickham H. ggplot2: elegant graphics for data analysis. Published online 2016.
35. Auguie B. gridExtra: miscellaneous functions for “grid” graphics.
36. Jäger G, Füllsack M. Systematically false positives in early warning signal analysis. *PLoS one*. 2019;14(2):1-14. doi:10.1371/journal.pone.0211072
37. Lenton TM, Livina VN, Dakos V, van Nes EH, Scheffer M. Early warning of climate tipping points from critical slowing down: comparing methods to improve robustness. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 2012;370(1962):1185-1204. doi:10.1098/rsta.2011.0304
38. Bonsall MB, Wallace-Hadrill SMA, Geddes JR, Goodwin GM, Holmes EA. Nonlinear time-series approaches in characterizing mood stability and mood instability in bipolar disorder. *Proceedings of the Royal Society B: Biological Sciences*. 2012;279(1730):916-924. doi:10.1098/rspb.2011.1246
39. McLeod A. Kendall: Kendall rank correlation and Mann-Kendall trend test.
40. Hamed KH, Rao AR. A modified Mann-Kendall trend test for autocorrelated data. *Journal of Hydrology*. 1998;204:182-196.
41. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*. 1995;57(1):289-300.
42. Altman D, Martin Bland J. Diagnostic tests 2: predictive values. *British Medical Journal*. 1994;309:102-102.
43. Wichers M, Schreuder MJ, Goekoop R, Groen RN. Can we predict the direction of sudden shifts in symptoms? Transdiagnostic implications from a complex systems perspective on psychopathology. *Psychological Medicine*. 2019;49(3):380-387. doi:10.1017/S0033291718002064
44. Fisher AJ, Medaglia JD, Jeronimus BF. Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*. 2018;115(27):1-10. doi:10.1073/pnas.1711978115
45. Morriss R. Clinical importance of inter-episode symptoms in patients with bipolar affective disorder. *Journal of Affective Disorders*. 2002;72:1-11.
46. Boettiger C, Ross N, Hastings A. Early warning signals: the charted and uncharted territories. *Theoretical Ecology*. 2013;6(3):255-264. doi:10.1007/s12080-013-0192-6
47. Kramlinger K, Post RM. Ultra-rapid and ultradian cycling in bipolar affective illness. *British Journal of Psychiatry*. 2021;168:314-323.
48. van der Maas HLJ, Molenaar PCM. Stagewise cognitive development: an application of catastrophe theory. *Psychological Review*. 1992;99(3):395-417. doi:10.1037/0033-295X.99.3.395
49. Helmich MA, Olthof M, Oldehinkel AJ, Wichers M, Bringmann LF, Smit AC. Early warning signals and critical transitions in psychopathology: challenges and recommendations. *Current Opinion in Psychology*. 2021;41:51-58. doi:10.1016/j.copsyc.2021.02.008
50. Haslbeck JMB, Ryan O. Recovering within-person dynamics from psychological time series. *Multivariate Behavioural Research*. 2021;56(2):1-32. doi:10.1080/00273171.2021.1896353
51. Hadaeghi F, Hashemi Golpayegani MR, Moradi K. Does “crisis-induced intermittency” explain bipolar disorder dynamics? *Frontiers in Computational Neuroscience*. 2013;7(116):1-2. doi:10.3389/fncom.2013.00116

10

Anticipating the direction of symptom progression using critical slowing down: a proof-of-concept study

Abstract

Background. As complex dynamic systems approach a transition, their dynamics change. This process, called critical slowing down (CSD), may precede transitions in psychopathology as well. This study investigated whether CSD may also indicate the *direction* of future symptom transitions, *i.e.*, whether they involve an increase or decrease in symptoms.

Methods. In study 1, a patient with a history of major depression monitored their mental states ten times a day for almost eight months. Study 2 used data from the TRAILS TRANS-ID study, where 122 young adults at increased risk for psychopathology (mean age 23.64 years, $SD = 0.67$, 56.6% males) monitored their mental states daily for six consecutive months. Symptom transitions were inferred from semi-structured diagnostic interviews. In both studies, CSD direction was estimated using moving-window principal component analyses.

Results. In study 1, CSD was directed towards an increase in negative mental states. In study 2, the CSD direction matched the direction of symptom shifts in 34 individuals. The accuracy of the indicator was higher in subsets of individuals with larger absolute symptom transitions. The indicator's accuracy exceeded chance levels in sensitivity analyses (accuracy 22.92% vs. 11.76%, $z = -2.04$, $p = .02$) but not in main analyses (accuracy 27.87% vs. 20.63%, $z = -1.32$, $p = .09$).

Conclusions. The CSD direction may predict whether upcoming symptom transitions involve remission or worsening. However, this may only hold for specific individuals, namely those with large symptom transitions. Future research is needed to replicate these findings and to delineate for whom CSD reliably forecasts the direction of impending symptom transitions.

Introduction

About 86% of individuals will have a mental disorder at some point in their lives¹. Given the considerable burden associated with mental disorders, there has been great interest in prevention and early intervention². Successful prevention requires a solid understanding of what it means to be “at risk” for developing psychopathological symptoms. Despite a large number of characteristics well known to predispose individuals to psychopathology, predicting at the individual level who will, or will not, develop a disorder remains largely an open question. Most individuals who are considered at risk do *not* develop a disorder^{3,4}. In order to better identify those at-risk individuals who could benefit from preventive interventions, an improved prediction of the prognosis of at-risk individuals is necessary.

One route towards improved prediction is to gain more detailed knowledge on within-individual changes occurring on the verge of disorder onset. It has been suggested that the structure of symptoms, *i.e.*, the extent to which symptoms reflect a single construct and how they covary, changes as individuals improve or worsen in terms of psychopathology^{5–7}. This could mean that alterations in the structure of symptoms may predict future symptom progression. Within-individual support for this idea is currently limited to a single case study, which showed that a relapse in depression was preceded by rising covariance between symptoms⁸. Between-individual or group-level support, provided by studies that compared symptom covariance of one group (*e.g.*, individuals prior to treatment) to another group (*e.g.*, the same individuals after treatment), is abundant but warrants cautious interpretation. This has two reasons: first, these comparisons may be subject to Berkson’s bias⁹, and second, between-individual findings do not necessarily translate to the within-individual level¹⁰. With these considerations in mind, it is noteworthy that both worsening and remitting psychopathology have been related to increased symptom covariances. Specifically, symptom covariance may be higher in individuals with persisting^{11,12} or worsening¹³ symptoms compared to individuals with remitting symptoms over time, although not all studies confirmed this^{14–16}. If this also holds within individuals, it could mean that high symptom covariances predispose individuals to psychopathology. However, symptom covariance has also been found to be *lower* in individuals before compared to after treatment^{17–24}. This could mean that high symptom covariances are linked to mental health. For instance, symptom remission may coincide with an altered appraisal of symptoms, meaning that individuals may increasingly perceive their symptoms as belonging to a unified latent construct (*i.e.*, a disorder)^{18,24,25}.

Alternatively, the association between symptom covariance and remission could be due to the simultaneous absence of symptoms (*i.e.*, floor effects). Regardless of the inferences drawn, there seems to be an apparent paradox: increased symptom covariances might relate to both symptom worsening and remission. It is at present unclear how to reconcile this. Further, it remains largely unknown whether changing symptom covariances over time can *prospectively predict* symptom remission or worsening within individuals.

A complex dynamic systems approach to psychopathology provides a framework that unites earlier findings and can be used to address the unresolved questions described above. In complex dynamic systems, transitions are often preceded by a period during which the stability within the system gradually declines, a phenomenon known as *critical slowing down*. Otherwise unpredictable transitions – such as the extinction of a species, sudden climate changes, or a sudden transition in mental health – might thus be anticipated by monitoring a system's instability^{26–28}. Critical slowing down has been shown to precede not only ecosystem and climate transitions²⁹, but also transitions between depressed and manic episodes in bipolar disorder³⁰ as well as relapse and remission of depression^{8,31,32}. This means that critical slowing down, which can be assessed in repeated assessments of mental states or symptoms of psychopathology, may foresee upcoming mental health problems. Recently, the potential of critical slowing down as a warning sign for impending transitions has been extended by noting that critical slowing down has a *direction*, meaning that it involves only a specific part of the system^{33–36}. This, in turn, means that critical slowing down could expose whether a transition is directed towards, for instance, extinction of one species or the other³³. In the context of psychopathology, exposing the direction of critical slowing down may allow for inferring whether an upcoming symptom transition is directed towards worsening or remitting symptoms. The direction of critical slowing down can be monitored using metrics similar to those described in earlier studies, namely symptom covariances (or, more specifically: the eigenvalues of the covariance matrix^{33,34,36,37}). Hence, the hypothesis that follows from a complex dynamic systems approach can be considered an extension of what was reported earlier, namely: a gradual alteration in the structure of psychopathological symptoms prospectively predicts whether a specific individual will experience a symptom transition towards remission (decrease of symptom severity) or worsening (increase of symptom severity).

The current study aimed to test the hypothesis that symptom changes within individuals – involving either an increase or decrease in symptoms over time – can be

predicted based on the direction of critical slowing down³³. Given that the application of complex dynamic systems principles to psychopathology is still in its infancy, we will approach our aim in two steps. First, we will provide a proof of concept by testing our hypothesis in a dataset in which principles from complex dynamic systems have already been confirmed^{8,38,39}. These data are time series of a single individual with a history of depression who experienced a relapse (*i.e.*, a sudden increase in symptoms), and contain multiple momentary ratings of mental states per day over a period of almost eight months. We will extend earlier findings⁸, which showed that critical slowing down preceded the relapse in depression, by investigating whether critical slowing down is indeed directed towards symptom worsening (as opposed to remission). Second, we will investigate whether the directionality of critical slowing down generalizes to symptom transitions with varying directions and magnitudes. To this end, we will repeat the analyses in a larger dataset, which consists of 134 young adults at increased risk for mental health problems who provided daily ratings on their mental states over a period of six months⁴⁰. We hypothesize that in the first dataset, the direction of critical slowing down^{33,34} points towards a relapse of symptoms. Similarly, in the second dataset, we hypothesize that the direction of critical slowing down corresponds to the change in symptoms reported by individuals (*e.g.*, strong critical slowing down towards improvement in individuals who experienced a large reduction of symptoms, and vice versa).

Methods

Study 1

The data analyzed in study 1 were extensively described elsewhere and are publicly available³⁹. Briefly, this study concerned a participant diagnosed with major depressive disorder who monitored his mood 10 times a day for almost eight consecutive months. During this period, the participant experienced a relapse in depressive symptoms⁸.

Experience sampling procedure

Experience sampling involved completing 10 questionnaires per day for a period of 239 days (almost 8 months), resulting in 1478 completed observations. Each questionnaire consisted of 50 items, of which 12 pertained to mood states. Mood-related items with negative valence (*e.g.*, feeling stressed) were rated on a 7-point Likert scale ranging from -3 to 3, whereas items with positive valence (*e.g.*, feeling content) were rated on a scale that ranged from 0

to 7. We rescaled items with negative valence to maintain a consistent interpretation. The daily assessments were complemented with weekly assessments of the depression subscale of the Symptom Checklist Revised⁴¹. The latter were used to monitor changes in symptom severity, as an indicator of a relapse in depression.

Study 2

The data analyzed in study 2 were retrieved from the TRAILS TRANS-ID study, which has been described in detail elsewhere⁴⁰. TRAILS TRANS-ID included 134 participants from an ongoing prospective cohort study, named Tracking Adolescents' Individual Lives Survey (TRAILS). TRAILS was designed to monitor mental health from childhood to adulthood through bi- or tri-annual assessments, and includes a general population and a clinical cohort⁴². Participants from the clinical cohort (TRAILS CC) had been referred to a child psychiatric outpatient clinic in the Northern Netherlands any time before the age of 11, and were therefore considered at increased risk for psychopathology. At the age of 22 years old, TRAILS CC participants were invited to a six-month daily diary study (TRAILS TRANS-ID). For analyses, we included those individuals who completed a diagnostic interview prior to and immediately after the diary study (N=122). TRAILS TRANS-ID was approved by the local Medical Ethical Committee (reference no. 2017/203). All participants provided written informed consent.

Diary procedure

For a period of six consecutive months, participants completed a questionnaire (diary) every evening concerning the past day, resulting in a maximum of 183 observations per participant. Participants received these questionnaires through a link sent in a text message to their mobile phones. Each questionnaire consisted of 58 items pertaining to positive mental states (*e.g.*, "How happy did you feel today?"), negative mental states (*e.g.*, "How anxious did you feel today?"), event appraisal (*e.g.*, "How stressful was the most stressful event that happened today?"), and substance use (*e.g.*, "How much soft drugs did you use today?"). These items were rated on a visual analogue scale ranging from 0 (not at all) to 100 (very much). A list of all diary items has been reported elsewhere⁴⁰.

Diagnostic interview

Immediately before and after the diary procedure, the short version of the Schedules for Clinical Assessment in Neuropsychiatry (mini-SCAN) was administered⁴³. The mini-SCAN is

a semi-structured diagnostic interview that assesses whether individuals meet the criteria for mental disorders, including mood, anxiety, psychotic, and substance use disorders, and attention-deficit hyperactivity disorder. Further, the mini-SCAN includes a screener for autism spectrum disorder. The mini-SCAN was complemented by the aggressive behavior subscale of the Adult Self Report⁴⁴ (ASR) to also include oppositional or antisocial behavior. The mini-SCAN and ASR were orally administered, and symptoms were scored as either absent (coded 0), subthreshold (coded 1), or clinical (coded 2)⁴⁰. The sum score of all items was considered reflective of global symptom severity and could range from 0 to 450.

Analysis

First, we selected diary items that reflected mood and were balanced in terms of valence (*i.e.*, an equal number items reflecting positive and negative mental states were chosen). This yielded 10 items in the first (single case) dataset (5 positive, 5 negative valence) and 28 items in the second dataset (14 positive, 14 negative valence). Examples of such items are “feeling relaxed” and “feeling down”. A list of all items and our motivation for selecting these items is provided in the supplement. Then, for each individual separately, we iteratively performed principal component analyses within sliding windows (*i.e.*, segments of the time series). For the first dataset ($N=1$, length = 1478 completed observations), these windows contained 150 observations, amounting to 1326 windows in total. This window size corresponds to the size adopted by Lever and colleagues³³, who used windows containing 10% of the time series length. For the second dataset ($N=122$, length = ± 183 observations per person), windows contained a maximum of 60 observations, resulting in on average 123.91 windows per person ($SD = 4.96$). Here, window size was chosen to strike a balance between overfitting (*i.e.*, using small windows, for instance containing 10% of the observations, which would complicate reliable principal component analyses) and underfitting (*i.e.*, using large windows, and potentially smoothing over potentially meaningful trends), while taking into account that on average 11.45% of the observations was missing. We inspected the influence of this methodological choice in sensitivity analyses, where we used windows of 40 and 80 observations.

Within each window, we retrieved (i) the amount of variance explained by the first principal component and (ii) the skewness of the scores projected on this first principal component. Together, these parameters formed a vector in a two-dimensional space (Fig. 10.1). The length of this vector corresponds to the variance explained by the first principal

component (*i.e.*, the largest eigenvalue of the covariance matrix). We considered this length reflective of the structure of mental states: it captures to what extent mental states are interrelated and unidimensional (Fig. 10.1). In line with earlier studies, we expected the vector to lengthen prior to symptom transitions, meaning that each window should have a larger vector than the previous windows^{33,34,45,46}. Provided that there is no change in overall variance, this corresponds to an increased symptom covariance (as previously observed in group-level studies^{17–24,47–49}). From the length and direction of the vector, we inferred the expected change in positive and negative mental states (Fig. 10.1). Repeating this procedure for each window resulted in a time series that described the expected change in mental states over time for a particular person. The trend in these time series was computed using Kendall's tau correlation coefficient. Kendall's tau is a nonparametric correlation coefficient that assesses the similarity in the rank order of two variables, without making assumptions concerning the linearity of the trend⁵⁰. A similar approach has been described before by Lever and colleagues³³.

Fig. 10.1. Example of a principal component analysis performed in a single window from the first dataset³⁹. The left figure shows the mean positive and negative mental states of the assessment occasions in this window (indicated by black dots), as well as the first principal component (dotted line). The blue arrow reflects the vector that was inferred from the principal component analysis: the length of the arrow corresponds to the proportion of variance explained by the first principal component, whereas the direction of the arrow corresponds to the skewness of the scores projected on this component. The skewness, illustrated in the right plot, is directed towards the left. Together, the length and direction of the vector can be used to infer the predicted change in mental states. Here, we would predict an increase in negative mental states (0.7) and a larger decrease in positive mental states (-0.9). A similar, more detailed explanation of this method was described by Lever and colleagues³³.

The first dataset was analyzed to provide a proof of concept. Here, we examined whether the known relapse in depression (which occurred around day 127) was preceded by a corresponding predicted increase in negative mental states. We expected that the predicted increase in negative mental states would be small long before the transition, and would rise as the transition approached. This was quantified using Kendall's tau, which is a non-parametric measure of correlation and is referred to here as the indicator. A significant, positive tau ($p < .05$) was considered indicative of a meaningful trend in the predicted change in mental states.

The second dataset was analyzed to investigate whether the aforementioned approach also holds for smaller symptom changes as experienced by at-risk individuals. Here, we analyzed whether the difference in symptom severity before and after the diary period corresponded to the predicted change in daily mental states. We focused on predicted changes in *negative* (and not positive) mental states because these changes conceptually matched our outcome (in- or decreases in symptom severity). We expected that large reductions in symptoms would coincide with a large predicted decrease in negative mental states, and vice versa. Again, trends in the predicted changes in mental states were inferred from Kendall's tau. The accuracy of the indicator was computed as the percentage of individuals for whom a change in symptoms (increase or decrease) was preceded by a corresponding trend in the predicted change in negative mental states. Compared to the symptom change in the first dataset, symptom changes in the second study were smaller, and not always of similar clinical significance. Therefore, we examined to what extent the accuracy of the indicator was dependent on the magnitude of symptom shifts or the clinical status of participants (*i.e.*, with vs. without diagnosis).

Analyses of the second dataset differed from those of the first dataset in three ways. First, the iterative principal component analyses were done on sum scores of positive and negative mental states rather than individual items. This was done because in some of the windows, items loaded ambiguously on the first principal component, which complicated deriving the orientation of this component in a two-dimensional space (Fig. 10.1). Second, in the second study, the skewness of projected scores was sometimes close to 0. This caused 180-degree shifts in the predicted direction across multiple consecutive windows, as also reported by Lever and colleagues³³. We corrected this by reversing “deviant” directions: when the direction of the indicator in windows 1 to 5 was positive, positive, negative, positive, positive – meaning that the first and last two windows suggested an increase in positive

mental states, while the third window suggested an increase in negative mental states – we corrected the third window by reversing the predicted change scores. This is equivalent to “flipping” the vector in Fig. 10.1. In the supplement, we also report the results obtained with a different skewness correction, namely removing the directions that were based on non-significant skews. A third and final difference concerns the significance testing of the indicator. Specifically, we evaluated the overall accuracy of the indicator in the second study through permutation testing, which involved shuffling the time order of each individual’s data and then computing the indicator. In this shuffled dataset, the temporal ordering of the data was lost, and therefore, we would expect the indicator to perform worse compared to the original data. Permutations were repeated 200 times in order to obtain a stable accuracy estimate. This estimate was then compared to the accuracy based on the original (non-shuffled) data⁵¹. Because accuracy was computed across individuals, we could not take a similar approach in the first study. Analyses were performed using R (version 4.0.2)⁵².

Results

Study 1

On average, the participant completed 6.2 ($SD = 1.9$) assessments per day, amounting to 1,478 diary entries in total³⁹. Prior to the relapse in depression, around day 127, there was a rising trend in the predicted change in negative mental states ($\tau = 0.68$, $p < .01$; Fig. 10.2). First, from day 54 to 88 (observations 384-614) the predicted *reduction* in negative mental states became *smaller*. From day 89 (observation 615) onwards, the indicator predicted an accumulating increase in negative mental states together with a decrease in positive mental states. Hence, starting more than 2 months prior to the relapse, a rising trend in the predicted change in negative mental states appeared to “warn” for the transition. This trend did not end immediately after the transition, supporting the idea that there might be a continuous relation between the trend in predicted changes in mental states and impending symptom change.

Fig. 10.2. *The predicted change in positive (blue) and negative (orange) mental states. The dotted line depicts the weekly assessed severity of depressive symptoms based on the SCL-90. The black line marks the relapse in depressive symptoms, which occurred around day 127 (observation 823).*

Study 2

Individuals who completed the diary period as well as both diagnostic interviews ($N=122$, 56.6% males) were on average 23.64 years old ($SD = 0.67$). They completed on average 162 diary entries (88.55%, $SD = 17.51$, range = 114-189). Prior to and immediately after the diary period, 40 (32.79% of the total sample) and 34 (27.87%) individuals met the criteria for at least one psychiatric disorder. Depression was most common ($N=25$ at baseline; $N=21$ at post), closely followed by anxiety disorder ($N=15$; $N=11$) and ADHD ($N=8$; $N=8$). The sum of the items that were rated in the diagnostic interviews ranged from 3-241 (baseline) and 2-230 (post), with average sum scores of 66.25 (baseline; $SD = 42.71$) and 69.78 (post; $SD = 46.89$). Based on the diagnostic interview data, approximately half of the sample improved in terms of their symptoms ($N=60$, 49.18%), whereas the other half reported worse symptoms at post compared to baseline ($N=58$, 47.54%) or had an equal symptom severity at baseline and post ($N=4$, 3.28%). The absolute magnitude of symptom transitions varied between 1, which has no clinical significance, and 108, which signifies an in- or decrease in the severity of half of the items assessed in the diagnostic interview (mean change = 18.46, $SD = 18.73$, median = 12.00). For the majority of individuals ($N=81$, 66.39%), symptom transitions did not

lead to a change in diagnosis. For others ($N=41$, 33.61%), symptom transitions coincided with a change in diagnostic status: 16 individuals no longer met criteria for (one of) their diagnosis and 25 met criteria for a new diagnosis.

For 65 individuals (53.28%), we found a significant change in their predicted change in negative mental states (mean absolute tau = 0.18, $SD = 0.13$, range = 0-0.56). Specifically, in 35 individuals (28.69%) we found a rising trend in the predicted change in negative mental states, which would suggest future symptom worsening (mean tau = 0.28, $SD = 0.13$, range = 0.14-0.56). The other 30 individuals (24.59%) showed a declining trend in their predicted change in negative mental states, implying symptom improvement (mean tau = -0.27, $SD = 0.10$, range = -0.45 to -0.14). Examples of both trends are depicted in Fig. 10.3.

Fig. 10.3. *Illustration of the predicted change in positive and negative mental states for two individuals.* Because windows spanned 60 observations, the earliest prediction was made for the 60th observation. A) This individual reported an increase in symptom severity based on the sum score of the diagnostic interview (baseline = 73; post = 96). This symptom increase is preceded by a gradual rise in the predicted change of negative mental states (tau = 0.25, $p < .05$). B) This individual reported a reduction in symptom severity (baseline = 63, post = 41), which was preceded by a gradual decline in the predicted change in negative mental states (tau = -0.29, $p < .01$).

The trend in the predicted change of negative mental states correctly predicted symptom changes in 34 individuals (27.87% of the entire sample), yielding a true positive rate of 52.31% (34/65). Of these individuals, 16 reported more severe symptoms at post, 17 reported less severe symptoms at post, and 1 did not change in terms of symptom severity. For this latter individual, the correct prediction reflected a non-significant trend in tau. For the others, correct predictions were reflected by significant up- or downward trends in tau, which became more pronounced as the absolute change in symptom severity increased.

This was not the case for the entire sample (*i.e.*, including those individuals for whom the indicator was not predictive of symptom change; Fig. 10.4). Individuals for whom the indicator worked did not differ from others in terms of their absolute change in symptoms (mean absolute symptom change 22.79 vs. 15.94, Cohen's $d = 1.33$, Welch's $t(40.69) = 1.43$, $p = .16$). Further, the accuracy of the indicator was not related to the likelihood to meet diagnostic criteria at baseline ($\chi^2(1) = 0.01$, $p = .93$) or at post ($\chi^2(1) = 0.21$, $p = .64$). Nevertheless, the performance of the indicator improved when evaluated in subsets of individuals with large symptom changes in either direction. Specifically, in the 0.50 quantile (*i.e.*, individuals whose absolute symptom change exceeded the median absolute symptom change), the accuracy of the indicator was 32.76%. In the 0.25 and 0.10 quantiles, accuracy equaled 26.67% and 41.67%. By definition, these latter samples are relatively small – comprising 25 and 10% of individuals (*i.e.*, $N=30$ and 12) – and hence, these results warrant cautious interpretation.

In shuffled data, where the temporal structure of the data was lost, the indicator reached an accuracy of 20.63%. This accuracy was not significantly different from the accuracy obtained in the original data (27.87%, $z = -1.32$, $p = .09$), and hence, we could not rule out the possibility that the indicator's accuracy arose from chance instead of critical slowing down. However, in sensitivity analyses where we only retained predictions that were based on significantly skewed projected scores, the indicator performed significantly better (accuracy 22.92%) compared to permutation tests (accuracy 11.76%, $z = -2.04$, $p = .02$; supplement). The accuracy of the indicator did not significantly change when considering alternative window sizes (window of 40 observations: accuracy 19.67%, $z = 1.50$, $p = .07$; window of 60 observations: accuracy 24.59%, $z = 0.58$, $p = .28$; see supplement for further details).

Fig. 10.4. *The association between the trend in the predicted change in negative mental states (tau) and change in symptom severity from baseline to post in individuals for whom the indicator was correct (N=34, orange) and in the entire sample (N=122, blue). Horizontal dotted lines depict the threshold for significant ($p < .05$) versus non-significant ($p > .05$) values of tau. For all individuals (dots) above and below these dotted lines (N=65), a change in symptoms was predicted. Negative symptom change implies improvement of psychopathology (i.e., a reduction of symptoms over time). For individuals for whom such improvement was predicted by the indicator, *larger* reductions in symptoms were related to more pronounced trends in the indicator (i.e., more negative values of tau). Vice versa, individuals for whom symptom worsening was predicted correctly showed a more pronounced indicator (i.e., more positive values of tau) as symptom change increased.*

Discussion

We investigated whether symptom changes within individuals – involving either an increase or decrease in symptoms over time – can be predicted based on the direction of critical slowing down in daily reports of mental states³³. First, we used data from a middle-aged man with a history of major depression who monitored his mood ten times per day for almost eight months. After four months, he experienced a relapse of depression, which was preceded by critical slowing down⁸. Building on this previous work, we found that critical slowing down not only anticipated the symptom transition, but also signified that the transition was directed towards an increase (as opposed to a decrease) in symptoms. We next

investigated whether the same holds on a larger scale, for more diverse symptom transitions (*i.e.*, increases and decreases of varying magnitude). To this end, we analyzed data from 122 young adults at increased risk for psychopathology who monitored their mental states daily for six consecutive months. For one in four (main analysis) to five (sensitivity analysis) individuals, the directionality of critical slowing down correctly predicted their reported symptom change. However, the indicator's accuracy only exceeded chance levels in sensitivity analysis, and therefore, the support for directionality of critical slowing down as a predictor of future symptom changes was less robust in study 2 than in study 1. In conclusion, results from study 1 and 2 tentatively support the idea that for *some* individuals, particularly those with large symptom changes, the direction of critical slowing down matches the direction of future symptom transitions.

Our findings tentatively suggest that anticipating the direction of symptom transitions by means of critical slowing down may be limited to specific individuals. We might get a better idea of who these individuals might be by taking a closer look at the literature that described parameters that determine whether the direction of critical slowing down can be used to infer the future state of a system. First, critical transitions in dynamical systems can be the result of either a positive feedback loop, or a negative feedback loop with delayed effects⁵³. The direction of critical slowing down most accurately predicts the future if the system's dynamics are governed by positive feedback loops³³. In the context of psychopathology, this means that mental states should amplify each other: *e.g.*, when feeling tired leads to feeling down, which leads to concentration problems, which leads to feeling tired again. From a complex dynamic systems perspective, such feedback loops give rise to self-sustaining states (*e.g.*, a mental disorder). Such self-sustaining states, and the feedback loops that underlie them, have been linked to both past and present mental disorders. That is, both individuals with a past mental disorder and individuals with a current mental disorder have been shown to have relatively densely connected symptoms, albeit in different studies. For example, network studies^{20–24,47–49} showed that symptoms are more densely connected in remitted individuals compared to treatment-seeking individuals with mental health problems. Provided that these group-level findings generalize to the level of the individual, this could mean that feedback loops strengthen as symptoms decline in severity. At the same time, however, cross-sectional network studies^{54–56} and dynamic network studies^{57,58} showed that individuals with a mental disorder have higher connectivity than non-affected

individuals – suggesting that those with a disorder, too, have strong feedback loops^k. These seemingly discrepant findings could be reconciled if strengthened feedback loops between mental states reflect a scar imposed by mental disorders⁵⁹. This would mean that feedback loops strengthen with illness duration – which received tentative support^{60,61}. This has consequences for the utility of the indicator presently studied, which depends on the strength of feedback loops. Specifically, it would mean that the indicator might be more suitable for individuals with a longer illness duration. In agreement with this, we found that the indicator clearly matched the symptom course in study 1, which concerned a remitted individual with a history of major depression that dated back 30 years⁶², whereas it was less robust in study 2, which concerned individuals with a shorter illness duration. Besides illness duration – and by analogy, the strength of feedback loops – the participant in study 1 differed from the participants in study 2 in the magnitude of symptom transitions. Specifically, participants in the latter study generally reported smaller, perhaps more gradual, symptom transitions compared to the participant in study 1. This touches upon a second factor that determines the accuracy of the indicator in exposing the direction of critical slowing down. That is, the indicator is more accurate when transitions reflect full collapses, as opposed to (sequential) partial collapses³³. This could mean that a transition in a small set of specific symptoms (*i.e.*, a partial collapse) is more difficult to detect than a change in almost all symptoms (*i.e.*, a full collapse). In line with this, present findings tentatively suggest that the direction of larger symptom transitions (as observed in study 1, and subsets of individuals from study 2) is more predictable than that of smaller symptom transitions. In conclusion, it is possible that critical slowing down and its direction are only detectable in individuals with a long illness duration (implying strong feedback loops) who experience relatively large transitions.

Critical slowing down: sudden versus gradual transitions

Critical slowing down is often considered prior to sudden transitions, such as the collapse of an ecosystem^{63,64}. A relapse in depression, as examined in study 1, might be of comparable impact and suddenness⁸. Critical slowing down anticipates such transitions if they occur through a cusp (or saddle node) bifurcation. This assumption can indirectly be verified, for

^k Note that both longitudinal and cross-sectional network studies should be cautiously interpreted, since comparing network characteristics between individuals with differing levels of symptom severity might be subject to Berkson's bias or collider bias⁹.

instance by testing for bimodality and hysteresis²⁹. A recent study confirmed that the data we analyzed in study 1 indeed shows such “signs of complexity”⁶⁵. Hence, the data analyzed in study 1 likely meet the requirements for observing critical slowing down. For the data analyzed in study 2, we could not verify whether symptom shifts indeed occurred through the specific types of bifurcations related to critical slowing down. Yet, this applies to the majority of applied studies into critical slowing down^{8,31,66}. For these studies, it is uncertain whether critical slowing down should be expected at all. This makes it difficult to assess whether (not) observing critical slowing down reflects a true or false positive (or negative). We addressed this ambiguity by shuffling the temporal order of the data analyzed in study 2. If critical slowing down and its direction would still be detected in such shuffled data, it likely reflects a false positive. We found that this was not the case, provided that the direction of critical slowing down is inferred using conservative criteria (sensitivity analyses). Still, it remains uncertain whether the shifts that occurred in study 2 resembled the type of shifts for which critical slowing down, and its direction, are informative. It thus requires further research to translate the mathematical assumptions of critical slowing down to empirical settings.

Strengths and limitations

A first strength of the current study is its ability to inspect critical slowing down *within individuals*. In contrast to group-level studies, we could therefore directly test the hypothesis that the direction of critical slowing down is informative of impending symptom transitions. Second, we used a variance-based method for detecting the direction of critical slowing down, which is less sensitive to the timescale of assessments compared to other metrics of critical slowing down such as the autocorrelation⁶⁷. Additionally, variance-based methods are less dependent on the amount of available data for each individual³³. It remains possible, however, that differences in the sampling frequency and duration (study 1: 1478 observations for one participant; study 2: max. 183 observations for 122 participants) contributed to differences in power, perhaps explaining why the indicator was more robust in study 1 compared to study 2. At the same time, critical slowing down only has practical relevance if it can be detected in data that are feasible to collect. Hence, although a relatively short sampling duration might have compromised statistical power in study 2, the data quantity in this study likely comes closer to what would be seen in applied settings (*e.g.*, clinical practice) compared to the data quantity in study 1.

Several limitations should be taken into account while considering present findings. First, there were no measures of symptom severity *during* the diary period in study 2, and therefore, we cannot rule out the possibility of unnoticed sudden symptom transitions during the diary period. For instance, it is possible that the symptom changes experienced by the at-risk individuals in study 2 did not evolve gradually, but rather, occurred suddenly during the diary period (*e.g.*, after 4 months). This, in turn, should affect the timing of critical slowing down: gradual development implies critical slowing down towards the end of the diary period^{68–71}, while sudden development during the diary period implies an earlier manifestation of critical slowing down (*e.g.*, between 2–4 months). Since we inspected critical slowing down across the diary period (towards the end), overlooking some symptom transitions may have led to an underestimated accuracy of the indicator. A second limitation is that we inspected only two directions of critical slowing down (towards an increase vs. decrease in symptoms). This might be an oversimplification, as more nuanced directions could also be considered (*e.g.*, towards an increase in depression vs. anxiety vs. aggression)⁷². At the same time, inspecting such fine-grained directions seems only warranted if the indicator would correctly distinguish between more global directions. Further, critical slowing down is less pronounced in high-dimensional systems compared to lower dimensional systems^{35,70}, meaning that considering many potential directions might lower the detectability of our indicator and would require a larger amount of data. Hence, reducing the complexity of our data seemed desirable.

Conclusions and future directions

In ecology, critical slowing down not only anticipates symptom transitions but also informs on the direction of these transitions³³. If the same holds for psychopathology, we could detect whether a specific individual, at a specific moment in time, is increasingly likely to experience a specific type of symptom transition (in this case, towards either improvement or worsening). Based on two studies, we found tentative support for the idea that the direction of critical slowing down might anticipate the direction of impending transitions, provided that these transitions are sufficiently large. Inconclusive findings preclude firm conclusions, and more research into critical slowing down in the context of symptom transitions is necessary. A promising step forward could be to empirically investigate to what extent critical slowing down depends on the strength of feedback loops. One way of addressing this could be to compare critical slowing down between individuals with strongly versus weakly connected

symptom networks. Second, future studies are necessary in order to explicitly address the assumptions underlying critical slowing down (*e.g.*, with respect to the bifurcation that describes the transition). This requires a translation from mathematical theory⁷³ to empirical reality, which could perhaps be facilitated by formal modeling⁷⁴. Ultimately, such a translation will improve our ability to identify those individuals whose system of mental states matches the behavior of other complex dynamic systems. These are the individuals who may benefit from the anticipatory capacity of critical slowing down. A clearer picture of when and for whom complex dynamic systems principles apply will crucially determine their clinical utility.

Supplement 10

10.1 Items

For study 1, the experience sampling questionnaire contained 12 mood-related items. In contrast to study 2, where items assessed individuals' mood during the past day retrospectively, the items in study 1 assessed present mood (*i.e.*, at the moment of filling in the questionnaire; Table S10.1). To balance the number of positive and negative mental states, we left out two randomly selected negative mood items from analyses (I feel indecisive; I feel suspicious). This resulted in five positive and five negative mental states that were further analyzed (Table S10.1).

For study 2, the daily diary assessments consisted of 58 items, of which 52 pertained to mood (17 positive mental states, 35 negative mental states). Since items do not necessarily have the same meaning between individuals, aggregating all items would likely result in noisy variables. We therefore decided to select those items that were most likely to reflect the same two constructs (mental health vs. ill-health) across individuals. To this end, we performed a principal component analysis, accounting for the variance explained by individuals and the temporal ordering of the data. Briefly, we fitted multilevel models where each of the 52 mood-related items was predicted by an intercept and time. Both were allowed to vary between individuals. Further, we specified that observations from the same individual were autocorrelated. These autocorrelations were again allowed to vary between individuals. The residuals of these models reflected the variability in observations that was not due to either individual differences, time trends or autocorrelation. Using these residuals, we fitted a principal component analysis with two components. These two components would eventually reflect the directions in which individuals could change (*i.e.*, towards mental health/positive mental states vs. towards mental ill-health/negative mental states). We inspected the loadings of items on the two components, and selected those positive mental states with a loading $>.80$ for analysis. This yielded 14 positive mental states, which were then matched to the 14 negative mental states with the highest loadings (Table S10.1).

Table S10.1. *Items included in analyses*

	Positive mental states	Negative mental states
Study 1 (case study)	<ol style="list-style-type: none"> 1. I feel relaxed 2. I feel satisfied 3. I feel enthusiastic 4. I feel cheerful 5. I feel strong 	<ol style="list-style-type: none"> 1. I feel down 2. I feel anxious 3. I feel irritated 4. I feel lonely 5. I feel guilty
Study 2 (TRAILS TRANS-ID study)	<ol style="list-style-type: none"> 1. I could enjoy things 2. I felt energetic 3. I felt at ease with others 4. I was happy 5. My day was worth living 6. I felt good 7. I felt like others liked me 8. I felt relaxed 9. I felt appreciated 10. I am looking forward to tomorrow 11. I felt like undertaking things 12. Everything came easy 13. I was social 14. Last night I slept well 	<ol style="list-style-type: none"> 1. I felt irritated 2. I was easily upset 3. I had difficulty making decisions 4. I felt rebellious 5. I felt strange 6. I felt impatient 7. I was afraid of making mistakes 8. I was easily distracted 9. Others felt annoyed by me 10. I felt guilty 11. I had moodswings 12. I felt suspicious 13. I could not bring myself to do anything 14. I felt restless

10.2 Alternative skewness correction

As described in the analysis section of the main text, the skewness of scores projected on the first principal component was sometimes close to zero. This caused 180-degree flips in the direction of the vector we calculated within consecutive windows (see Fig. 10.1, main text). In our main results, we corrected such flips by iteratively searching for and correcting flipping patterns (*e.g.*, correcting positive – positive – negative – positive – positive to positive – positive – positive – positive). Here, we report the results obtained when deleting predictions based on vectors of which the skew was not statistically significant.

We could calculate trends in the predicted change of mental states for 96 individuals (78.69%). For others, there were too many missing predictions to calculate the tau correlation coefficient across predictions. For 37 individuals (38.54%), a significant change in their predicted change in negative mental states was found (absolute tau range = 0-0.88, mean = 0.19, *SD* = 0.18). The predicted change in negative mental states showed an increasing trend for 21 individuals (21.88%), and a decreasing trend for 17 individuals (17.71%). The trend in

the predicted change of negative mental states correctly predicted symptom changes in 22 individuals (22.92%), of whom 12 reported more severe symptoms at post, 9 reported less severe symptoms at post, and 1 did not change in terms of symptom severity. Individuals for whom the indicator worked did not differ from others in terms of their absolute change in symptoms (mean absolute symptom change 20.82 vs. 14.43, Cohen's $d = 1.31$, $t(94) = 1.57$, $p = .12$). Further, the accuracy of the indicator was not related to the likelihood to meet diagnostic criteria at baseline ($\chi^2(1) = 0.24$, $p = .63$) or at post ($\chi^2(1) < 0.01$, $p > .99$). Similar to the results reported in the main text, indicator performance improved when evaluated in subsets of individuals with large symptom changes in either direction. Specifically, in the 0.50 quantile (*i.e.*, individuals whose absolute symptom change exceeded the median absolute symptom change across individuals), the accuracy of the indicator was 29.27%. In the 0.25 and 0.10 quantiles, accuracy equaled 25.00% and 50.00%. In contrast to our main analyses, the accuracy of the indicator significantly exceeded the accuracy obtained in shuffled data, which was 11.76% ($z = -2.04$, $p = .02$). This should however be cautiously interpreted, given that we could only inspect accuracy in the subset of individuals for whom predictions could be made ($N=96$).

10.3 Alternative window sizes

For the second study, we used window sizes of 60 observations. To determine the influence of this choice on our results, we re-ran analyses with windows containing a maximum of 40 and 80 observations. The results were largely in line with what we reported in the main text. Below, the results we obtained with window sizes 40 and 80 are denoted consecutively, with subscripts where possible. A significant change in the predicted change in negative mental states was found for 56 (45.90%, absolute tau range₄₀ = 0-0.46, mean₄₀ = 0.14, SD_{40} = 0.11) and 66 individuals (54.10%, absolute tau range₈₀ = 0-0.64, mean₈₀ = 0.18, SD_{80} = 0.14). The predicted change in negative mental states showed an increasing trend for 26 (21.31%) and 31 (25.41%) individuals. For the remaining 30 (24.59%) and 35 (28.67%) individuals, the trend was decreasing. The trend in the predicted change of negative mental states correctly predicted symptom changes in 24 and 30 individuals (accuracy₄₀ = 19.67%, accuracy₈₀ = 24.59%; Table S10.2). In conclusion, the accuracy of the indicator did not seem to depend on window size.

Table S10.2. *Results of study 2 for different window sizes*

	Window size 40	Window size 60	Window size 80
Mean number of observations per window (<i>SD</i>)	36.00 (4.43)	53.44 (6.79)	70.85 (9.20)
Range of observations per window	21.67-40.00	31.36-60.00	40.78-80.00
Accuracy	19.67%	27.87%	24.59%

Note. The mean number of observations reflects the number of observations per window, averaged across windows (within individuals), averaged across individuals.

References

1. Caspi A, Houts RM, Ambler A, et al. Longitudinal assessment of mental health disorders and comorbidities across 4 decades among participants in the Dunedin birth cohort study. 2020;3(4). doi:10.1001/jamanetworkopen.2020.3221
2. Kessler RC, Aguilar-Gaxiola S, Alonso J, et al. The global burden of mental disorders: an update from the WHO World Mental Health (WMH) surveys. *Epidemiologia e Psichiatria Sociale*. 2009;18(1):23-33. doi:10.1017/S1121189X00001421
3. Cuijpers P. Examining the effects of prevention programs on the incidence of new cases of mental disorders: the lack of statistical power. *American Journal of Psychiatry*. 2003;160(8):1385-1391. <http://ajp.psychiatryonline.org>
4. Raballo A, Mechelli A, Menculini G, Tortorella A. Risk syndromes in psychiatry: a state-of-the-art overview. *Archives of Psychiatry and Psychotherapy*. 2019;21(2):7-14. doi:10.12740/APP/109506
5. Marsman M, Borsboom D, Kruis J, et al. An introduction to network psychometrics: relating Ising network models to item response theory models. *Multivariate Behavioral Research*. 2018;53(1):15-35. doi:10.1080/00273171.2017.1379379
6. van Bork R, Rhemtulla M, Waldorp LJ, Kruis J, Rezvanifar S, Borsboom D. Latent variable models and networks: statistical equivalence and testability. *Multivariate Behavioral Research*. 2019;56(2):175-198. doi:10.1080/00273171.2019.1672515
7. Cramer AOJ, van Borkulo CD, Giltay EJ, et al. Major depression as a complex dynamic system. *PLoS one*. 2016;11(12):1-20. doi:10.1371/journal.pone.0167490
8. Wichers M, Groot PC, Psychosystems ESM group, EWS group. Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*. 2016;85:114-116. doi:10.1159/000441458
9. de Ron J, Fried EI, Epskamp S. Psychological networks in clinical populations: investigating the consequences of Berkson's bias. *Psychological Medicine*. 2021;51(1):168-176. doi:10.1017/S0033291719003209
10. Hamaker EL. Why researchers should think “within-person”. A paradigmatic rationale. In: Mehl MR, Conner TS, eds. *Handbook of Research Methods for Studying Daily Life*. The Guilford Press; 2012:43-61.
11. van Borkulo C, Boschloo L, Borsboom D, Penninx BWJH, Lourens JW, Schoevers RA. Association of symptom network structure with the course of longitudinal depression. *JAMA Psychiatry*. 2015;72(12):1219-1226. doi:10.1001/jamapsychiatry.2015.2079
12. Kuranova A, Wigman JTW, Menne-Lothmann C, et al. Network dynamics of momentary affect states and future course of psychopathology in adolescents. *PLoS one*. 2021;16:1-18. doi:10.1371/journal.pone.0247458
13. van de Leemput IA, Wichers M, Cramer AOJ, et al. Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences*. 2014;111(1):87-92. doi:10.1073/pnas.1312114110
14. Hebbrecht K, Stuivenga M, Birkenhäger T, et al. Understanding personalized dynamics to inform precision medicine: a dynamic time warp analysis of 255 depressed inpatients. *BMC Medicine*. 2020;18(1). doi:10.1186/s12916-020-01867-5
15. Schwaren L, van Borkulo CD, Fried E, Goodyer IM. Assessment of symptom network density as a

- prognostic marker of treatment response in adolescent depression. *JAMA Psychiatry*. 2018;75(1):98-100. doi:10.1001/jamapsychiatry.2017.3459
16. Groen RN, Snippe E, Bringmann LF, et al. Capturing the risk of persisting depressive symptoms: A dynamic network investigation of patients' daily symptom experiences. *Psychiatry Research*. 2019;271:640-648. doi:10.1016/j.psychres.2018.12.054
17. Harvey PD, Endicott JM, Loebel AD. The factor structure of clinical symptoms in mixed and manic episodes prior to and after antipsychotic treatment. *Bipolar Disorders*. 2008;10(8):900-906. doi:10.1111/j.1399-5618.2008.00634.x
18. Fried EI, Borkulo CD van, Epskamp S, Schoevers RA, Tuerlinckx F, Borsboom D. Measuring depression over time... or not? Lack of unidimensionality and longitudinal measurement invariance in four common rating scales of depression. *Psychological Assessment*. 2016;28(11):1-14. doi:10.1037/pas0000275
19. Isa A, Bernstein I, Trivedi M, Mayes T, Kennard B, Emslie G. Childhood depression subscales using repeated sessions on children's depression rating scale- Revised (CDRS-R) scores. *Journal of Child and Adolescent Psychopharmacology*. 2014;24(6):318-324. doi:10.1089/cap.2013.0127
20. Bos FM, Fried EI, Hollon SD, et al. Cross-sectional networks of depressive symptoms before and after antidepressant medication treatment. *Social Psychiatry and Psychiatric Epidemiology*. 2018;53(6):617-627. doi:10.1007/s00127-018-1506-1
21. Bryant RA, Creamer M, O'Donnell M, et al. Acute and chronic posttraumatic stress symptoms in the emergence of posttraumatic stress disorder a network analysis. *JAMA Psychiatry*. 2017;74(2):135-142. doi:10.1001/jamapsychiatry.2016.3470
22. Beard C, Millner AJ, Forgeard MJC, et al. Network analysis of depression and anxiety symptom relationships in a psychiatric sample. *Psychological Medicine*. 2016;46(16):3359-3369. doi:10.1017/S0033291716002300
23. Blanco I, Contreras A, Chaves C, Lopez-Gomez I, Hervas G, Vazquez C. Positive interventions in depression change the structure of well-being and psychological symptoms: A network analysis. *Journal of Positive Psychology*. 2020;15(5):1-6. doi:10.1080/17439760.2020.1789696
24. Hilbert A, Herpertz S, Zipfel S, et al. Psychopathological networks in cognitive-behavioral treatments for binge-eating disorder. *Psychotherapy and Psychosomatics*. 2020;89(6):379-385. doi:10.1159/000509458
25. Fokkema M, Smits N, Kelderman H, Cuijpers P. Response shifts in mental health interventions: An illustration of longitudinal measurement invariance. *Psychological Assessment*. 2013;25(2):520-531. doi:10.1037/a0031669
26. Scheffer M. *Critical Transitions in Nature and Society*. Princeton University Press; 2009.
27. Scheffer M, Carpenter SR, Dakos V, van Nes EH. Generic indicators of ecological resilience: inferring the chance of a critical transition. *Annual Review of Ecology, Evolution, and Systematics*. 2015;46:145-167. doi:10.1146/annurev-ecolsys-112414-054242
28. Olde Rikkert MGM, Dakos V, Buchman TG, et al. Slowing down of recovery as generic risk marker for acute severity transitions in chronic diseases. *Critical Care Medicine*. 2016;44(3):601-606. doi:10.1097/CCM.0000000000001564
29. Litzow MA, Hunsicker ME. Early warning signals, nonlinearity, and signs of hysteresis in real ecosystems. *Ecosphere*. 2016;7(12). doi:10.1002/ecs2.1614
30. Bayani A, Hadaeghi F, Jafari S, Murray G. Critical slowing down as an early warning of transitions in episodes of bipolar disorder: a simulation study based on a computational model of circadian activity

- rhythms. *Chronobiology International*. 2017;34(2):235-245. doi:10.1080/07420528.2016.1272608
31. Wichers M, Smit AC, Snippe E. Early warning signals based on momentary affect dynamics can expose nearby transitions in depression: a confirmatory single-subject time-series study. *Journal for Person-Oriented Research*. 2020;6(1):1-15. doi:10.17505/jpor.2020.22042
 32. Olthof M, Hasselman F, Strunk G, et al. Critical fluctuations as an early-warning signal for sudden gains and losses in patients receiving psychotherapy for mood disorders. *Clinical Psychological Science*. 2019;8(1):1-11. doi:10.1177/2167702619865969
 33. Lever JJ, van de Leemput IA, Weinans E, et al. Foreseeing the future of mutualistic communities beyond collapse. *Ecology Letters*. 2020;23(1):2-15. doi:10.1111/ele.13401
 34. Weinans E, Lever JJ, Bathiany S, et al. Finding the direction of lowest resilience in multivariate complex systems. *Journal of the Royal Society Interface*. 2019;16(159):1-9. doi:10.1098/rsif.2019.0629
 35. Patterson AC, Strang AG, Abbott KC. When and where we can expect to see early warning signals in multispecies systems approaching tipping points: insights from theory. *The American Naturalist*. 2021;198(1). doi:10.1086/714275
 36. Dakos V. Identifying best-indicator species for abrupt transitions in multispecies communities. *Ecological Indicators*. 2018;94:494-502. doi:10.1016/j.ecolind.2017.10.024
 37. Chen S, O'Dea EB, Drake JM, Epureanu BI. Eigenvalues of the covariance matrix as early warning signals for critical transitions in ecological systems. *Scientific Reports*. 2019;9(1):1-14. doi:10.1038/s41598-019-38961-5
 38. Kossakowski JJ, Gordijn MCM, Riese H, Waldorp LJ. Applying a dynamical systems model and network theory to major depressive disorder. *Frontiers in Psychology*. 2019;10(1762):1-18. doi:10.3389/fpsyg.2019.01762
 39. Kossakowski JJ, Groot PC, Haslbeck JMB, Borsboom D, Wichers M. Data from 'Critical slowing down as a personalized early warning signal for depression.' *Journal of Open Psychology Data*. 2017;5. doi:10.5334/jopd.29
 40. Schreuder MJ, Groen RN, Wigman JTW, Hartman CA, Wichers M. Measuring psychopathology as it unfolds in daily life: addressing key assumptions of intensive longitudinal methods in the TRAILS TRANS-ID study. *BMC Psychiatry*. 2020;20(1). doi:10.1186/s12888-020-02674-1
 41. Derogatis LR. *SCL-90-R Symptom Checklist-90-R Administration, Scoring, and Procedures Manual*. PsychCorp; 1977.
 42. Huisman M, Oldehinkel AJ, de Winter A, et al. Cohort Profile: The Dutch "Tracking Adolescents" Individual Lives' Survey'; TRAILS. *International Journal of Epidemiology*. 2008;37(6):1227-1235. doi:10.1093/ije/dym273
 43. Nienhuis FJ, van de Willige G, Rijnders CAT, de Jonge P, Wiersma D. Validity of a short clinical interview for psychiatric diagnosis: the mini-SCAN. *British Journal of Psychiatry*. 2010;196(1):64-68. doi:10.1192/bjp.bp.109.066563
 44. Achenbach TM. *Manual for the Young Adult Self-Report and Young Adult Behavior Checklist.*; 1997.
 45. Dablander F, Pichler A, Cika A, Bacilieri A. Anticipating critical transitions in psychological systems using early warning signals: theoretical and practical considerations theory of critical slowing down. *Submitted*. Published online 2020:1-30.
 46. Held H, Kleinen T. Detection of climate system bifurcations by degenerate fingerprinting. *Geophysical Research Letters*. 2004;31(23):1-4. doi:10.1029/2004GL020972

47. McElroy E, Napoleone E, Wolpert M, Patalay P. Structure and connectivity of depressive symptom networks corresponding to early treatment response. *EClinicalMedicine*. 2019;8:29-36. doi:10.1016/j.eclinm.2019.02.009
48. Esfahani FZ; SH; VKF; SGP. Sensitivity of the positive and negative syndrome scale (PANSS) in detecting treatment effects via network analysis. *Innovations in Clinical Neuroscience*. 2017;14:59-67.
49. Ge F, Yuan M, Li Y, Zhang J, Zhang W. Changes in the network structure of posttraumatic stress disorder symptoms at different time points among youth survivors: a network analysis. *Journal of Affective Disorders*. 2019;259:288-295. doi:10.1016/j.jad.2019.08.065
50. Kendall MG. *A New Measure of Rank Correlation*. Vol 30.; 1938. <https://about.jstor.org/terms>
51. Isaac ERHP. *Test of Hypothesis - Concise Formula Summary*.; 2015.
52. R Core Team. R: A language and environment for statistical computing. Published online 2020.
53. Levins R. Discussion paper: the qualitative analysis of partially specified systems. *Annals New York Academy of Sciences*. 1974;231(1):123-138.
54. Santos H, Fried EI, Asafu-Adjei J, Jeanne Ruiz R. Network structure of perinatal depressive symptoms in Latinas: relationship to stress and reproductive biomarkers. *Research in Nursing and Health*. 2017;40(3):218-228. doi:10.1002/nur.21784
55. Heeren A, McNally RJ. Social anxiety disorder as a densely interconnected network of fear and avoidance for social situations. *Cognitive Therapy and Research*. 2018;42(1):103-113. doi:10.1007/s10608-017-9876-3
56. van Rooijen G, Isvoranu AM, Kruijt OH, et al. A state-independent network of depressive, negative and positive symptoms in male patients with schizophrenia spectrum disorders. *Schizophrenia Research*. 2018;193:232-239. doi:10.1016/j.schres.2017.07.035
57. Wigman JTW, van Os J, Borsboom D, et al. Exploring the underlying structure of mental disorders: cross-diagnostic differences and similarities from a network perspective using both a top-down and a bottom-up approach. *Psychological Medicine*. 2015;45:2375-2387. doi:10.1017/S0033291715000331
58. Pe ML, Kircanski K, Thompson RJ, et al. Emotion-network density in major depressive disorder. *Clinical Psychological Science*. 2015;3(2):292-300. doi:10.1177/2167702614540645
59. Rohde P, Lewinsohn PM, Seeley JR. Are people changed by the experience of having an episode of depression? A further test of the scar hypothesis. *Journal of Abnormal Psychology*. 1990;99(3):264-271.
60. Christian C, Perko VL, Vanzhula IA, Tregarthen JP, Forbush KT, Levinson CA. Eating disorder core symptoms and symptom pathways across developmental stages: a network analysis. *Journal of Abnormal Psychology*. 2019;129(2):177-191. doi:10.1037/abn0000477
61. Baez LM, Heller AS. Impact of age at onset on the phenomenology of depression in treatment-seeking adults in the STAR*D trial. *Journal of Affective Disorders*. 2020;262:381-388. doi:10.1016/j.jad.2019.10.036
62. Groot PC. Patients can diagnose too: how continuous self-assessment aids diagnosis of, and recovery from, depression. *Journal of Mental Health*. 2010;19(4):352-362. doi:10.3109/09638237.2010.494188
63. Dai L, Vorselen D, Korolev KS, Gore J. Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science*. 2012;336(6085):1175-1177.
64. Drake JM, Griffen BD. Early warning signals of extinction in deteriorating environments. *Nature*. 2010;467(7314):456-459. doi:10.1038/nature09389
65. Olthof M, Hasselman F, Lichtwarck-Aschoff A. Complexity in psychological self-ratings: implications for research and practice. *BMC Medicine*. 2020;18(317):1-16. doi:10.1186/s12916-020-01727-2

66. Maturana MI, Meisel C, Dell K, et al. Critical slowing down as a biomarker for seizure susceptibility. *Nature Communications*. 2020;11(1):1-12. doi:10.1038/s41467-020-15908-3
67. Weinans E, Quax R, van Nes EH, Leemput IA van de. Evaluating the performance of multivariate indicators of resilience loss. *Scientific Reports*. 2021;11(1):9148. doi:10.1038/s41598-021-87839-y
68. Boettiger C, Ross N, Hastings A. Early warning signals: the charted and uncharted territories. *Theoretical Ecology*. 2013;6(3):255-264. doi:10.1007/s12080-013-0192-6
69. Dakos V, Carpenter SR, van Nes EH, Scheffer M. Resilience indicators: prospects and limitations for early warnings of regime shifts. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2015;370(1659):1-10. doi:10.1098/rstb.2013.0263
70. Boerlijst MC, Oudman T, de Roos AM. Catastrophic collapse can occur without early warning: examples of silent catastrophes in structured ecological models. *PLoS one*. 2013;8(4):1-6. doi:10.1371/journal.pone.0062033
71. Kéfi S, Dakos V, Scheffer M, van Nes EH, Rietkerk M. Early warning signals also precede non-catastrophic transitions. *Oikos*. 2013;122(5):641-648. doi:10.1111/j.1600-0706.2012.20838.x
72. Wichers M, Schreuder MJ, Goekoop R, Groen RN. Can we predict the direction of sudden shifts in symptoms? Transdiagnostic implications from a complex systems perspective on psychopathology. *Psychological Medicine*. 2019;49(3):380-387. doi:10.1017/S0033291718002064
73. Bury T. Detecting and distinguishing transitions in ecological systems: model and data-driven approaches. Published online 2020.
74. Burger J, Veen D van der, Robinaugh DJ, et al. Bridging the gap between complexity science and clinical practice by formalizing idiographic theories: a computational model of functional analysis. *BMC Medicine*. 2020;18(99):1-18. doi:10.1186/s12916-020-01558-1

11

General summary and discussion

Summary of main findings

This thesis has aimed to shed light on emerging psychopathology from a complex dynamic systems perspective. The focus was on young adults, who are in a vulnerable developmental period in terms of mental health. We saw that youth differ in the complexity and dynamics of their emotional experiences (**chapters 2 and 3**), hypothesized that such dynamics may also change *within* individuals (**chapter 4**), conducted an intensive longitudinal study to investigate this and verified whether this study yielded meaningful and representative data (**chapters 5 and 6**), and finally investigated the dynamics of mental states within individuals and their relation to mental health (**chapters 7-10**). Below, I will reflect on our findings and place them in a broader context.

Between-individual studies into complex dynamic systems principles

The first two chapters describe between-individual differences in the dynamics of emotions and link these differences to subsequent psychopathology. Both used data from adolescent twins from the general population who monitored their emotions ten times a day for six consecutive days. In **chapter 2**, we showed that emotional differentiation¹ – but not emotional diversity² – predicts symptoms of psychopathology reported one year later. Specifically, individuals who report their emotions in an undifferentiated manner, meaning that they tend to lump distinct emotions (*e.g.*, feeling down, irritated, or disappointed) into a single construct (*e.g.*, feeling bad), are more likely to report worsening or persisting psychopathological symptoms compared to individuals with higher emotional differentiation. This can be understood from three perspectives. First, theories on emotion regulation propose that low emotional differentiation hampers adequate emotion regulation: if you do not know what you're feeling, you cannot know what to do about it¹. Second, a network perspective suggests that low emotion differentiation may be reflective of tightly connected emotions, meaning that emotions easily trigger each other³. This creates a vulnerable network, where an increase in one emotion (*e.g.*, feeling down) can set off a cascade of other emotions (*e.g.*, feeling anxious, irritated, etc.)⁴. Emotional experiences then become all-or-nothing phenomena, which makes sudden transitions in mental health (*e.g.*, from low to high symptom severity) more likely⁵. Finally, a complex dynamic systems perspective suggests that low emotion differentiation may reflect heightened vulnerability for a collapse of the system⁶. That is, if all emotions are experienced as a single construct (*i.e.*, feeling bad), the

system can no longer rely on alternative constructs, which in turn increases the probability of a sudden transition in mental health. Again, this implies that low emotion differentiation may predispose individuals to future psychopathology.

Despite the various theoretical explanations for the relation between emotion differentiation and psychopathology, the predictive value of emotion differentiation may be limited. Specifically, we found that the relation between emotion differentiation and future prognosis disappeared after taking into account mean negative emotionality, which has been reported in other studies on emotion dynamics as well⁷. Thus, in this study, we could not disentangle the potentially beneficial effects of being able to experience fine-grained emotions from the beneficial effect of “not feeling bad” on average. This is in line with a recent meta-analysis, which found that relatively complex dynamics – including emotion differentiation – may have limited predictive utility above and beyond simpler measures – such as the mean emotionality⁷. The reason for this is that complex dynamics conceptually and statistically overlap with the mean^{7–9}. The conceptual overlap between mean emotionality and emotion differentiation results from the fact that individuals with high levels of negative emotions, such as those with depression or neuroticism, also tend to overgeneralize their emotions (*e.g.*, feeling terrible in general)¹⁰. Hence, mean emotionality and emotion differentiation might be part of the same construct, which could resemble neuroticism – a well-established personality trait^{11,12}. The statistical overlap between mean emotionality and emotion differentiation is due to range restriction: in case of bounded measurements (*e.g.*, a questionnaire rated on a Likert scale ranging from one to seven), a low mean constrains the variance of observations¹³. As a result, mean scores and variance-dependent metrics share variance. A meta-analysis showed that this shared variance can range from 2% (autocorrelations) to 42% (Gini coefficient)⁷. Emotion differentiation falls in between, with 16% of its variance shared with mean emotionality⁷. We reported a similar overlap in **chapter 2** (correlation: -0.32, shared variance: 10%). The consequence of this overlap is that the predictor that is measured with less error (*i.e.*, mean) takes precedence over the other predictor (*i.e.*, emotion differentiation), which is less robust and therefore has a weaker relation with the outcome⁷. Hence, the effect of emotion differentiation on psychopathology (and related outcomes⁷) disappears after accounting for the more robust measures (*i.e.*, mean). This means that for predicting the course of psychopathological symptoms, emotion differentiation is redundant.

Chapter 3 also describes the relation between emotion dynamics and future

psychopathology. Here, we tested whether the autocorrelation in negative emotions – also called inertia¹⁴ – is predictive of increases in symptoms. Earlier literature on emotion dynamics described inertia as emotional rigidity, reasoning that higher autocorrelations suggest that individuals can “get stuck” in a negative emotional state^{15,16}. According to a complex dynamic systems approach to psychopathology, inertia rises as individuals become more vulnerable for future transitions. This is why rising autocorrelations (or rising inertia) were termed early warning signals (EWS). Contrary to our hypothesis, and also in apparent contrast to the study we aimed to replicate¹⁷, we did not consistently find higher autocorrelations in those individuals who reported larger symptom increases. This led us to closely inspect the models that were used in our own and this previous study. From this, we learned that in both our data and previously analyzed data¹⁷, inertia is linked to concurrent psychopathology, but not necessarily to the future course of psychopathology. A similar finding was recently reported by Houben and Kuppens¹⁸, who reported that inertia relates to concurrent depression, but not to symptoms of depression one year later reported by youth. Like in **chapter 2**, the lack of a prospective association between inertia and future psychopathology might be attributed to another, more potent, predictor in the model, namely baseline psychopathology. In **chapter 3**, this is illustrated by high correlations between baseline and follow-up ratings of symptom severity ($r = 0.79-0.86$). This leads to change scores with little variance, which in turn limits the extent to which autocorrelations can explain change scores. In conclusion, relatively straightforward and parsimonious measures such as mean negative emotions (**chapter 2**) or earlier psychopathology (**chapter 3**) suffice for the sole purpose of prediction of future psychopathology⁷. Put simply, based on between-individual studies, the best predictor of future psychopathology seems to be earlier psychopathology or overall negative emotionality. This conclusion is well known from longitudinal cohort studies^{19,20}.

Although the predictive capacity of emotion differentiation and inertia may be limited, these measures are not useless. Rather, they may advance our understanding in the emotional processes that characterize individuals with mental health problems, and may generate hypotheses about what causes and sustains such problems. For instance, both lower emotion differentiation and higher inertia have been considered reflective of emotion context insensitivity²¹. The relations between these measures and concurrent psychopathology could mean that psychopathology involves a flattened emotional landscape, where different experiences elicit the same emotion. This means that emotions become decoupled from environmental contexts and are less responsive to emotion regulation efforts:

seeking social support, for example, may not adequately resolve negative feelings. Such insight in the emotional underpinnings of psychopathology has raised several hypotheses, for instance concerning the evolutionary purpose of mental disorders²², the relation between emotional and cognitive inertia²³, and the relation between coping strategies or emotional resources and emotion dynamics^{24,25}. In sum, the utility of complex emotion dynamics, including emotion differentiation and inertia, may lie in their theoretical implications rather than their predictive capacity.

A second nuancing notion pertaining to **chapters 2 and 3** is that the lack of group-level associations between complex emotion dynamics and future psychopathology does not rule out the possibility of *within-individual* associations between both characteristics. That is, it is still possible that emotion dynamics change within individuals as they approach a shift in psychopathology. This discrepancy of between- versus within-individual findings is called Simpson's paradox²⁶: what holds between individuals does not necessarily hold within individuals (supplementary Fig. S11.1). Therefore, **chapters 2 and 3** could only provide an indirect exploration of the within-individual processes that follow from a complex systems approach to psychopathology. A more direct investigation of these processes requires an intensive longitudinal design, where individuals repeatedly monitor their mental state¹ for several consecutive months (**chapters 5 and 6**). Before introducing this design, however, I will first elaborate on the type of questions that an intensive longitudinal design can address.

Within-individual hypotheses and a matching study design

Chapter 4 notes the key hypotheses that follow from a complex dynamic approach to psychopathology and describes the type of design that can address these hypotheses. We hypothesized that transitions in symptoms, such as the onset or sudden worsening of psychopathology, are preceded by EWS. Further, we hypothesized that the type of mental states wherein EWS occur is predictive of the type of symptoms that arise later on. A similar idea received support in ecological and simulation studies^{27–29}. For instance, in a simulated ecosystem consisting of a predators and prey species, a sudden transition towards highly abundant prey was preceded by EWS in the prey population but not by EWS in the predator

¹ Throughout this discussion, I will distinguish between emotions (which were assessed in chapters 2 and 3) and mental states (chapters 5–10). The latter cover emotions, but also cognitions (e.g., worrying, having trouble concentrating), and behaviours (e.g., engaging in social interactions, drinking alcohol).

population²⁸. The reason that EWS may predict the type of transition that is about to develop is that critical slowing down – the process that underlies EWS – does not equally affect all parts of the system. Rather, it has a certain direction, meaning only those variables that align with the direction of critical slowing down will show EWS prior to a transition. If this would hold in the context of psychopathology, the mental state showing EWS could match the future state of the system. For instance, EWS in “feeling down” could signal an upcoming depression whereas EWS in “feeling irritated” could warn for behavioral problems.

We explored the directionality of critical slowing down in **chapters 3, 9 and 10**. In **chapter 3**, we indeed found that autocorrelations in “feeling suspicious” were positively related to subsequent symptom increases in the domain interpersonal sensitivity, but not to increases in other domains. This tentatively (and indirectly) confirmed our idea that mental states congruent with the psychopathological symptoms about to develop would be particularly suitable for detecting EWS. However, as we did not find a similar pattern of results for symptom increases in anxiety, depression, and somatic complaints, the support for our hypothesis was modest at best. In **chapters 9 and 10**, we investigated the directionality of critical slowing down *within* individuals using three different intensive longitudinal datasets. Before discussing the findings from these chapters, I will first introduce the data that we collected in order to investigate the hypotheses raised in **chapter 4**.

In **chapter 4** we posited that an intensive longitudinal design where individuals monitor their mental states for several consecutive months allows for examining whether EWS foresee symptom transitions within individuals, as well as the directions of such transitions. Ideally, individuals should be at-risk for various types of symptoms – meaning that transitions can have multiple directions. This design is described in **chapter 5**, which introduces the TRAILS TRANS-ID study that was set-up by our team. Briefly, we recruited 134 participants from the clinical cohort of TRAILS³⁰, who were considered at increased risk for psychopathological symptoms across the entire diagnostic spectrum. This, because they had been in contact with an outpatient secondary psychiatric facility already early in life. Participants provided daily ratings of 58 mental states for a period of six consecutive months – resulting in over one million observations in total. Participants’ compliance, reflected in the proportion of completed ratings (88.4%), was higher than the compliance in less intensive studies³¹, supporting the feasibility of this type of design.

It is possible that the high compliance obtained in the TRAILS TRANS-ID study was due an exceptionally high motivation to contribute to the study, which would limit the

representativeness of our sample. For instance, it could be that only those individuals who were relatively happy and healthy were willing to participate. Alternatively, it is also possible that those with relatively severe mental health problems were more likely to participate, for example, because they were interested in monitoring their mental health over time or because they hoped that their participation in our study could indirectly help others with similar problems. **Chapter 6** shows that neither of these selection biases were present. Further, it shows that TRAILS TRANS-ID participants were at a disadvantage relative to the general population on nearly every characteristic. Hence, despite considerable adversities in most aspects of life (including physical, psychological, and social characteristics), individuals are willing and able to participate in intensive longitudinal designs. Compliance in the diary study was not related to most participant characteristics, although individuals with more internalizing problems and less externalizing problems and antisocial behavior who did not smoke completed more assessments, and thus, might be overrepresented. Despite these small biases within this at-risk group, we concluded that intensive longitudinal designs are not reserved for only the (un)happy few.

Within-individual studies into complex dynamic systems principles

With the data collected in the TRAILS TRANS-ID study, I was able to study symptom transitions, and the direction thereof, within individuals. These symptom transitions have been proposed to happen between relatively stable states, which can be labelled mental health and mental disorder³². In **chapter 7**, I investigated these stable states^m. Specifically, I examined whether stable states are indeed restricted to the extremes of the symptom severity spectrum, or instead, can also occupy the grey area in between these extremes. Part of this grey area reflects subthreshold symptoms of psychopathology, which are characterized by insufficient severity, duration, or impairment to qualify as symptoms of a mental disorder, but nevertheless cause substantial suffering^{33,34}. Subthreshold psychopathology is often implicitly considered as an unstable, transient condition^{35–45} that marks the transition from one stable state (mental health) to another (mental disorder). Yet, we found that the stability of psychopathological symptoms does not differ across severity levels. This means that, just like mental health

^m Throughout this thesis, *stable states* reflect global equilibria that describe individual's overall mental health. *Mental states*, in contrast, reflect momentary assessments of emotions, cognitions, and behaviours (e.g., "at this moment, I am feeling happy").

and mental disorder, subthreshold symptoms may reflect a persisting condition in which individuals can “get stuck”. This finding matched the observations we made while collecting data, namely, that many individuals seemed to consistently experience problems in daily life despite not meeting the formal diagnostic threshold for any clinical diagnosis as assessed with the mini-SCAN⁴⁶. In complex dynamic systems terms, these individuals might be trapped into a maladaptive stable state, which should be destabilized in order to facilitate a transition towards a healthier state⁴⁷.

The shift from one stable state to another may resemble a critical transition. In **chapter 8**, we examined such transitions and investigated whether they are preceded by EWS in negative mental states. We found that 20 out of 122 individuals (16.4%) experienced a sudden drop in mental health over the course of six months. Most negative mental states did not show EWS prior to these drops, resulting in an overall sensitivity of 5.0%. In a matched subset of individuals without a drop in mental health, we found fewer EWS. Hence, EWS in negative mental states seldomly reflected a true positive – meaning that they seldomly occurred when we would theoretically expect them – but were less likely to reflect a false positive (*i.e.*, a false alarm) – meaning that EWS hardly occurred when we would *not* expect them (supplementary Fig. S11.2). This could either mean that EWS for transitions in mental health in at-risk youth do not exist – implying that the true and false positives described in **chapter 8** were chance findings – or that the detection of EWS is subject to a strict set of conditions. For instance, it is possible that the idea of critical slowing down prior to transitions in mental health only holds for a very specific set of individuals or psychopathological domains (*e.g.*, remitted individuals who experience a relapse in depression^{48,49}). Other possibilities include EWS only being detectable on certain timescales (*e.g.*, hour-to-hour assessments of mental states, as opposed to daily assessments)^{50,51}, in a specific set of variables²⁸, or for specific types of symptom transitions (*e.g.*, large shifts towards an alternate stable state that follow from accumulated instability)^{52–54}. Given the potential impact of these parameters on our findings, we next investigated EWS in a different dataset, which consisted of 20 individuals with bipolar disorder who experienced transitions between depressed and manic episodes (**chapter 9**). This study examined EWS on a different timescale (five assessments per day, as opposed to one assessment per day), in different variables (mental states reflecting symptoms of bipolar disorder, such as “feeling agitated” and “having racing thoughts”, as opposed to more generic mental states), for different types of transitions (shifts between depressed and manic episodes as opposed to drops in

mental health), using a different statistical approach (moving window analyses, as opposed to generalized additive models). Despite these methodological differences, the results of the study were largely in line with what we reported in the previous chapter: EWS were present, but were not exclusive to individuals with transitions and differed substantially between individuals. To enable a direct comparison between the findings from **chapters 8 and 9**, I calculated overall sensitivity and specificity for **chapter 9** (similar to **chapter 8**, but different from the average values that are reported in **chapter 9**ⁿ; supplementary Table S11.1). Compared to **chapter 8**, the overall likelihood of detecting EWS was six times larger (28.1% vs. 4.5%) whereas the percentage of EWS that occurred in individuals who experienced a transition relative to the total number of EWS that were detected was similar (54.6-56.6% vs. 55.6%). As a result, we found a higher sensitivity but a lower specificity of EWS in **chapter 9** (overall sensitivity for transitions towards depression and mania: 26.5% and 32.8%, overall specificity: 74.8%) compared to **chapter 8** (sensitivity: 5.0%, specificity: 96.0%). This means that transitions in bipolar disorders were easier to detect, but also more likely to yield a false positive, compared to the drops in mental health experienced by at-risk youth. What is preferable in the trade-off between sensitivity and specificity depends on the costs of missing a transition. These costs may be higher in patients with a bipolar disorder, for whom transitions were defined based on their clinical significance, compared to at-risk youth, for whom transitions were defined based on change point analyses and likely less impairing. Thus, the results from **chapters 8 and 9** are in line with what would be preferred from a clinical point of view: at-risk youth rarely receive a false alarm (at the cost of the sensitivity of EWS), whereas bipolar patients more often receive a false alarm (in the interest of a higher sensitivity of EWS). That said, the sensitivity and specificity of EWS are nowhere near what is recommended in the literature – namely, around 80%^{55,56} – and therefore, far from clinical usefulness.

The limited clinical utility of EWS is further emphasized by their poor predictive

ⁿ Average sensitivity, specificity, and predictive values reflect the mean values across mental states. These values are generally higher than the overall values because they are more sensitive to high outliers (e.g., predictive values of 100%, which happen when EWS in a mental state never reflect a false positive/negative).

^o Note that chapter 9 describes both the autocorrelation and the standard deviation as EWS. To facilitate the comparison with findings from chapter 8, which solely examined autocorrelations, my focus here is on the results we obtained with autocorrelations as EWS. The overall inferences, however, also apply to the standard deviation.

values. These predictive values may be particularly interesting from an applied perspective: if EWS are detected, what is the probability of an ensuing transition towards reduced mental health? I compared these predictive values to the probability of a transition if we had not monitored EWS (*i.e.*, the prevalence of transitions). Both for at-risk youth and for patients with bipolar disorder, the presence of EWS only marginally increased the probability of a transition (in at-risk youth: from 16.4% to 19.7%, in bipolar patients with a transition towards depression: from 36.0% to 37.1%, in bipolar patients with a transition towards mania: from 32.0% to 38.0%). Together with the low sensitivity of EWS, this suggests the clinical utility of EWS is questionable.

The limited clinical utility of EWS in at-risk youth and bipolar patients becomes even more apparent when compared to alternative approaches to identifying individuals at increased risk for mental health problems. For instance, in adolescent inpatients, suicidal crises could be predicted two weeks in advance with good accuracy based on daily diary data that spanned a period of two weeks (sensitivity: 80%, specificity: 96%)⁵⁷. An eight-week diary study in individuals diagnosed with bipolar disorders reported similar findings, showing that suicidal ideation can be predicted one week in advance (sensitivity: 88%, specificity: 95%)⁵⁸. Future risk for other psychiatric problems, including depression, psychosis, and substance use, has also been identified with reasonable accuracy^{20,59–62}. It should be noted, however, that the design and inferences from the latter studies differ from those in the present thesis, and therefore, we cannot directly compare sensitivity and specificity across studies. For instance, most earlier studies employed machine learning techniques and included baseline psychopathology as a predictor, which puts them at an advantage when compared to EWS, which do not incorporate baseline psychopathology. Further, most earlier studies aimed to distinguish between individuals with versus without the outcome of interest (*e.g.*, future suicidal ideation), thereby operating at the between-individual level. In contrast, EWS could theoretically monitor risk within individuals, which may yield insight in the proximity of sudden increases in mental health problems for specific individuals. However, findings from **chapters 8 and 9** do not support this promise of EWS. In conclusion, the question whether someone is at risk for mental health problems may at present be more resolvable through the risk calculators that have been developed by between-individual studies than through EWS^{61,63,64}.

So far, I have only discussed the overall predictive accuracy of EWS across *all* individuals with a transition in mental health relative to those without. However, it might well

be that the utility of EWS is specific to certain individuals, for whom the sensitivity, specificity, and predictive value of EWS might be more favorable. Such individuals could, for instance, be characterized by a particular diagnosis or data with a specific type (*e.g.*, emotions, thoughts, or physical experiences), length, or resolution (*e.g.*, many assessments per day for several consecutive months⁴⁹). Whether such characteristics play a role in the detection of EWS could not be investigated in the studies described in **chapters 8 and 9**, and hence remains a highly important topic for further research. In conclusion, based on current findings, it is likely that *if* EWS have clinical utility, this would be limited to specific individuals. Thus, if EWS would ever find their way into clinical practice, this would first require a clear delineation of the type of individuals who may benefit from monitoring EWS.

A final part the hypotheses laid down in **chapter 4** concerned the anticipation of specific symptom clusters (*e.g.*, depression vs. mania). Specifically, we hypothesized that the mental states wherein EWS manifest would foretell the direction of symptom transitions. This hypothesis was based on the notion that critical slowing down – the phenomenon that underlies EWS – has a certain direction. In **chapter 9**, we found partial support for the idea that EWS may foretell which symptom cluster will arise. Specifically, we found that transitions towards a manic episode experienced by patients with bipolar disorder were most often preceded by EWS in mental states conceptually related to mania (*i.e.*, having racing thoughts, feeling agitated, and feeling full of energy). Transitions towards depression, conversely, were most often preceded by EWS in feeling tired. Again, findings were not conclusive, primarily because EWS in mental states indicative of depression were not exclusive to transitions towards depression, nor were mental states indicative of mania exclusive to transitions towards mania. In **chapter 10**, we adopted a different approach to monitoring the direction of critical slowing down. Instead of inspecting the dynamics of individual mental states (*i.e.*, within-individual changes in the variance and autocorrelation), as we did in **chapter 9**, we examined changes in the within-individual covariance structure of multiple mental states. Specifically, we computed the proportion of variance explained by the first principal component and the skewness of the scores projected on this component. Together, these parameters reflect the direction of critical slowing down, which in turn informs the direction of impending transitions⁶⁵. In a dataset where critical slowing down was already established^{49,66}, we confirmed the hypothesis that the direction of critical slowing down matches the direction of a future transition (in this case, towards a relapse in depression). Next, we examined whether this finding generalizes to more subtle symptom changes

occurring over the course of six months. These symptom changes were studied using the TRAILS TRANS-ID data – which were also used in **chapters 5-8**. We found marginal support for a link between the direction of critical slowing down and the progression of symptoms. Specifically, the indicated direction matched the actual direction of symptom changes only in 20-25% of the individuals, and this performance only exceeded chance levels in sensitivity analyses. Like in **chapters 8 and 9**, it follows that critical slowing down (and its derivatives, *i.e.*, EWS) does not anticipate every shift in psychopathological symptoms.

Although our approach in **chapter 10** was developed in an entirely different system – namely, an ecosystem consisting of plants and pollinators⁶⁵ – its underpinnings are not new to research in psychiatry. For instance, the amount of variance explained by the first principal component has also been considered indicative of the unidimensionality of a measure⁶⁷. Such unidimensionality can mean that different mental states are experienced as a single, unified construct. In other words, high unidimensionality may signal that distinct emotions (*e.g.*, feeling down, irritated, or disappointed) all lie on the same continuum (*e.g.*, feeling bad). This may sound familiar. Indeed, we investigated a similar phenomenon in **chapter 2** (there labeled emotional complexity). We have thus come full circle: we started out with exploring whether between-individual differences in the diversity of emotions predict prognosis in terms of mental health, and ended with examining whether rising unidimensionality (or declining diversity)^{3,5} within persons predicts symptom improvement or worsening over the course of six months. This is not to say we are back at the start, but rather, that seemingly different concepts such as emotional complexity and the direction of critical slowing down sometimes have a similar operationalization. Likewise, a single operationalization can be understood from different perspectives. This may provide a new angle for understanding a well-known concept, which could be considered the core of this thesis: like studies on the directionality of critical slowing down may inform studies on emotion dynamics (and vice versa), complex dynamic systems theory may inform psychiatry.

Wrapping-up

This thesis explored a complex dynamic systems approach to psychopathology, described the data that are required to test the hypotheses that follow from this approach, and finally also examined several of these hypotheses. The first two empirical chapters, which both described between-individual studies, strongly link with studies into emotion dynamics. Through the lens of complex dynamic systems, these chapters indirectly addressed

whether diversity of the system indeed protects against sudden shifts in psychopathology (**chapter 2**) and whether critical slowing down may precede such shifts (**chapter 3**). Both studies were on a between-individual level, and therefore, they could only indirectly test complex dynamic systems principles. **Chapters 4-6** described that studying within-individual processes requires intensive longitudinal designs, and further showed that such designs are feasible, yield potentially meaningful data, and are not subject to selection bias. Finally, **chapters 7-10** provided an illustration of the type of inferences that can be drawn from intensive longitudinal data. Specifically, in **chapter 7** we reported that stable symptom states might manifest not just as mental health or a mental disorder, but also as subthreshold psychopathology. **Chapters 8 and 9** showed that sudden shifts between stable symptom states are sometimes preceded by EWS, but sensitivity, specificity, and predictive values of these EWS are currently too low to support their widespread clinical implementation. In **chapter 10**, we showed that critical slowing down might be informative for the direction of future symptom shifts, but this only held for a subset of individuals (namely, those with larger absolute shifts) and clinical utility therefore remains speculative. Taken together, this thesis aimed to understand the progression of psychopathology from a complex dynamic perspective. This endeavor brought to light several lessons as well as follow-up questions that will help moving our field forward, which I will address below.

Lessons learned

Lessons learned from within-individual research

This thesis started off with a description of between-individual studies, followed by studies that highlighted the potential of within-individual designs. The final chapters illustrated the results and inferences that within-individual designs may yield. An often-proposed advantage of these designs is that they can accommodate the person-specific nature of psychopathology⁶⁸⁻⁷⁰. This is because within-individual designs compare individuals to themselves rather than to others, thus allowing each individual to have a unique expression of psychopathological symptoms⁶⁸. In support of this personalized approach, Wright and Woods⁷¹ wrote: *“There may well be as many manifestations of psychopathology as there are persons to experience them”*. Evidently, there exists a tension between this person-specific nature of psychopathology on the one hand, and our wish to establish generic principles that may explain the development of symptoms on the other hand. The latter wish led me to

summarize over one million observations into a two-by-two table (**chapter 8**). Based on this table, I concluded that the complex dynamic systems principles we investigated do not hold for everybody – if they hold at all. This might mean that we should abandon the quest for generic, one-size-fits-all complex dynamic systems principles in the context of psychiatry and instead allow for more personalization. I see two ways of doing so. First, we could investigate whether complex dynamic systems principles hold for a specific subset of individuals, for instance defined by their demographic or clinical characteristics (*e.g.*, having a history of mental disorder) or by their data characteristics (*e.g.*, reporting daily mental states with sufficient autocorrelation)^{72,73}. Second, we could incorporate the person-specific nature of psychopathology into our designs and analyses. This could mean, for instance, allowing individuals to monitor mental states that matter to *them* (as opposed to the mental states that we as researchers think that matter) or investigating the dynamics of mental states at a timescale tailored to the individual. Such adaptations may challenge drawing inferences that apply to more than one individual, and therefore may primarily serve settings that favor person-specificity above generalizability (*e.g.*, clinical settings).

Lessons learned from EWS in other fields

The majority of earlier studies evaluated EWS based on their performance in a single system (*e.g.*, a single ecosystem, a financial market, or the climate). In contrast, we studied 40 different “systems” (individuals) in **chapter 8** (20 with and 20 without a transition in mental health) and 20 different systems in **chapter 9**. The scope of our studies thus went well beyond what is typical in the literature. A notable exception is a study by Burthe and colleagues⁷⁴, who studied 126 different ecosystems. In each system, the authors identified transitions (*i.e.*, turning points in the abundance of species) and EWS (*i.e.*, rising autocorrelations and variances). They found that EWS had a relatively low sensitivity (autocorrelation: 23.8%, variance: 20.0%, supplementary Table S11.1). Of note is that EWS more often reflected a false positive (*i.e.*, EWS in absence of a transition) than a false negative (*i.e.*, no EWS in presence of a transition). This is different from our findings in **chapters 8 and 9**, which showed that false positives are less common than false negatives. To illustrate, the probability of a true positive in our data (autocorrelation: 55%, variance: 37%) was much higher compared to the probability of true positives in ecosystems (autocorrelation: 21%, variance: 15%)⁷⁴. Hence, a first lesson that can be drawn from EWS in ecology – where they have a relatively long history⁷⁵ – is that EWS are not perfectly predictive^{74,76}. The tendency to evaluate single

systems may have obscured this imperfection: if EWS precede the single transition that is evaluated, they may appear to be flawless.

Based on the limited sensitivity of EWS and the considerable frequency of false positives, Burthe and colleagues⁷⁴ concluded that *“using variance and autocorrelation as early warnings may be wishful thinking”*. However, they also acknowledged that the type of transitions that were evaluated may not have matched the requirements for detecting EWS. This stresses a second lesson that we can draw from other fields: the definition and operationalization of transitions play a key role in the investigation of EWS. If the transition is ill-defined, the sensitivity of EWS will be underestimated. Why are well-defined transitions so important for evaluating EWS? In short: because critical slowing down (the foundation of EWS) is restricted to transitions that have certain mathematical properties. Specifically, critical slowing down only happens when gradient systems undergo so-called zero-eigenvalue bifurcations. It is beyond the scope of this thesis to elaborate on the technical aspects of these bifurcations, but it is worth noting that they can involve not only the “classical” critical transition (*i.e.*, fold or saddle node bifurcations) but also more gradual changes in systems (*e.g.*, Hopf bifurcations, transcritical bifurcations)^{53,54,77}. At the same time, abrupt transitions in systems are not necessarily classical critical transitions, but can also follow from, for instance, large external shocks⁵⁴. It follows that the empirical investigation of critical slowing down prior to sudden transitions is far from straightforward: not all sudden transitions fulfill the mathematical requirements for detecting critical slowing down, and simultaneously, some gradual transitions *do* fulfill those requirements (see Fig. 11.1 for a schematic illustration)^{53,78}. This may contribute to both false positives (*i.e.*, detecting EWS in absence of a transition) as well as false negatives (*i.e.*, not detecting EWS in presence of a transition). Perhaps some of our findings in **chapters 8 and 9** can be explained by this, as we found that EWS did not occur for every transition (false negatives), and also occurred in individuals who did not experience a symptom transition (false positives).

Fig. 11.1. *Schematic illustration of when EWS are (not) to be expected.* This illustration was inspired by a figure from Boettiger *et al.*⁵³ and conveys that (i) not every sudden transition is preceded by EWS and (ii) some gradual transitions are preceded by EWS. Both situations are often ignored.

Despite the crucial role of transitions in the theory on EWS, most empirical studies did not formally define transitions. Rather, it seems that any large change in a system has been considered as a potential critical transition that might be anticipated by EWS⁷⁹. This is problematic because it misleads the evaluation of EWS. That is, judging EWS by their ability to foresee events they are not (theoretically) expected to foresee is much like judging fish by their ability to climb trees. To arrive at a reliable conclusion about the utility of EWS, we should therefore aim to better characterize the transitions that systems exhibit. I see two approaches to achieving this. First, we could attempt to evaluate whether our time series are sufficiently similar to simulated time series that contain a classical critical transition (*i.e.*, a zero-eigenvalue bifurcation)⁸⁰. This could be achieved, for instance, using deep learning algorithms⁸¹. Second, we could check whether our time series show so-called hallmarks of non-linearity, namely *alternative stable states* and *hysteresis*^{79,82}. Such hallmarks considerably increase the likelihood that EWS anticipate transitions in the time series⁷⁹. The idea of *alternative stable states* implies bistability⁸³: individuals can reside one stable state or the other, but not in the region in between (Fig. 11.2). Since systems cannot occupy the unstable region in between two stable states, shifts between stable states may occur in a stepwise version (*e.g.*, sudden onset, progression, or relapse of symptoms)^p. Hence, bistability is a prerequisite for critical transitions. *Hysteresis* is strongly linked to the concept of alternative stable states⁸², and refers to the phenomenon that a system's response to an environmental driver depends on its current state⁷⁹. For instance, in the context of psychiatry, this could

^p Note that this assumes a fast-responding system and a relatively slow driver⁹³.

mean that the effect of chronic stress (a hypothetical environmental driver) on an individual's mood depends on the individual's current state (*e.g.*, mental health vs. mental disorder). It is noteworthy that chronic stress (*e.g.*, poverty or a somatic disease) differs from the stress evoked by daily events that cause temporary dips or uplifts in mood (*i.e.*, perturbations). Whereas chronic stress may affect the overall stability of a system's state (*e.g.*, mental health), day-to-day events cause fluctuations around this state (*e.g.*, feeling happy or sad, depending on what just happened). As noted previously, these fluctuations may linger as the system's state loses stability. For instance, if chronic stress accumulates (*e.g.*, from one to five imaginary units), the healthy state may become increasingly unstable, and the effect of daily events on an individual's mood may prolong. At the individual's tipping point, the healthy state becomes so unstable that only a minor perturbation can evoke a transition towards a disordered state. Hysteresis refers to the phenomenon that this transition can only be reverted if the environmental driver is reduced *beyond* the level at which the system reached its previous tipping point⁴. This means that chronic stress levels should be reduced beyond five (*e.g.*, to three) in order to facilitate a transition back to mental health (Fig. 11.2). In other words, across a certain range of chronic stress, the system can be in either one state (*e.g.*, mental health) or the other (*e.g.*, mental disorder). Hence, the relation between the system's state and the driver is nonlinear.

There are several approaches for assessing whether alternative stable states and hysteresis are indeed present. For instance, the existence of alternative stable states implies that mental states should be distributed around two modes (*i.e.*, bimodality)⁶. Such bimodality can be inferred from a histogram of the observations or from comparing the fit of models that assume one versus two states⁵. For instance, based on longitudinal assessments of depressive symptoms, Hosenfeld and colleagues⁵ found that the majority (66%) of individuals diagnosed with depression showed bimodal distributions in their symptom ratings. Hysteresis can be examined by checking whether the relation between a hypothetical driver (*e.g.*, chronic stress) and outcome (*e.g.*, severity of psychopathological symptoms) is better captured by a regression model with or without an interaction term^{6,79}. There are more refined methods for checking hysteresis, but these are also data demanding (*e.g.*, require long time series with numerous shifts between stable states)⁸². In any case, it should be noted that the outlined methods for inspecting alternative stable states and hysteresis are at best indirect indications for the presence of a fold bifurcation⁶. This means that it is not possible to *guarantee* the presence of a classical critical transition. Nevertheless,

indirectly inspecting these hallmarks of non-linearity may still be better (more informative) than not inspecting this at all.

Fig. 11.2. *Alternative stable states and hysteresis in the context of psychiatry.* If transitions in mental health indeed resemble the critical transitions observed in other complex dynamic systems, they may occur as follows: a gradual change in a driver (e.g., chronic stress) causes an individual (purple dot) to reach their tipping point (red dot). The dotted purple line illustrates uplifts and dips in mood (i.e., fluctuations around the individual's stable state). Despite these perturbations away from the stable state, the stable state is eventually restored. Close to the tipping point, only a small perturbation (short grey arrows) is needed to “push” the individual beyond the tipping point (i.e., right from or above the red dot). This triggers a transition towards the alternate stable state. The black dotted line reflects an unstable region, meaning that the individual cannot remain in a state in between mental health and mental disorder. For a certain range of chronic stress levels, the system can reside either in the healthy or the disordered state. This reflects hysteresis: in order to reverse a transition, the environmental driver should be lowered *beyond* the level at which the other tipping point is reached. Note that the labels in this illustration are arbitrary: alternative drivers could be conceived, and the stable states could also reflect, for instance, mental health and subthreshold psychopathology.

Besides alternative stable states and hysteresis, there are several other assumptions that should be met in order to apply complex dynamic system principles such as EWS. First, the system under consideration should show signs of feedback mechanisms^{72,84}. This means that there should be autocorrelation: the system's current condition should be dependent on its previous condition. This in turn implies that the system behaves non-randomly, which can be checked using Bartels rank test. As an example, Olthof and colleagues⁸⁵ showed that most mental states in the dataset that we also analyzed in **chapter 10** fluctuate non-randomly, suggesting that the assumption of feedback mechanisms was met. Second, EWS are only detectable if they are measured in the right variables^{28,29,86}. Whether a variable is “right” depends on the system under consideration, and more specifically on (i) the direction of perturbations and (ii) the direction of critical slowing down^{28,29}. For instance, EWS in “feeling

anxious” will only appear if (i) the events that an individual experiences cause fluctuations in feeling anxious and (ii) critical slowing down manifests in feeling anxious, meaning that feeling anxious should be one of the variables that show an increasingly slow return to equilibrium as the tipping point approaches. Further, EWS may be particularly pronounced in variables that are directly (and not indirectly) affected by an environmental driver⁸⁷. Speculatively, this could mean that EWS in a mental state that precipitates anxiety (*e.g.*, worrying) might be stronger and earlier than EWS in feeling anxious. It follows that selecting the variables in which EWS should be monitored requires a solid understanding of the system, *e.g.*, with respect to the processes that take place as a system approaches its tipping point. A final requirement for detecting EWS pertains to the length and resolution of time series data^{53,88,89}. Recommendations regarding the amount of data that is required for detecting EWS substantially vary across studies and simulation settings: the sensitivity of EWS in time series of 100 observations may range anywhere between 50%⁹⁰ and 90%⁹¹. In practice, EWS have been applied in time series of 40 observations⁸⁹, but also in much longer time series (*e.g.*, 2,600-27,000 observations^{89,92}). Hence, although the length and resolution of time series matter⁹¹, it is difficult to establish an optimal number and frequency of observations for detecting EWS. Again, what is optimal may depend on the system under consideration: the extinction of coral reef may take several decades to unfold and therefore requires long-term time series, whereas transitions in small lakes may take place over several months and can be examined in shorter, fine-grained data⁹³. This illustrates that knowing the timescale at which processes take place is crucial for determining the optimal length and resolution of data (this topic is further discussed in section *Timescales*). In conclusion, a final lesson that can be drawn from other fields is that the detectability of EWS depends on the methodological and analytical choices that are made – namely, the type of variables that are monitored as well as the frequency and duration of monitoring. These choices should be informed by the system’s characteristics, but also by what is practically feasible. Importantly, tailoring methodological choices to the system may be easier when a single system is studied (*e.g.*, an ecosystem) compared to when multiple systems are studied (*e.g.*, in the context of mental health, where each individual represents a unique system). A universal optimal methodological approach to detecting EWS in the context of mental health thus seems unlikely. Even a less generic approach is challenging, since the knowledge that could guide methodological choices is still largely lacking. Hence, we are only in the beginning of exploring complex dynamic systems principles, and the modest results reported in this thesis should not be considered definite.

The overall picture that emerges from ecological studies into EWS is that applying complex dynamic systems principles in real-world systems (such as mental health) requires detailed insight in the dynamic properties of the system. This includes, for instance, insight in the type of transitions that take place, the timescale at which such transitions unfold, and the variables that may show EWS. The necessity of such a priori understanding is somewhat ironic, given that EWS have frequently been put forward as tools that allow for anticipating system behavior despite limited knowledge of the system (*e.g.*, in terms of underlying mechanisms)⁶. This is also what seemed attractive in the context of psychopathology: EWS would enable us to anticipate symptom progression without necessitating detailed insight into the mechanisms that underlie mental ill-health (*e.g.*, genes, neurophysiological processes, social and psychological characteristics, etc.)⁹⁴. Yet, a deeper look into the requirements for detecting critical slowing down shows that EWS are not agnostic. That is, their successful application requires insight in the dynamics of a system. Perhaps this is why simulation studies^{77,90,95} – where the mathematical properties of systems are known in advance – are usually more successful in detecting EWS compared to empirical studies^{74,76,79,92}.

Lessons learned from EWS in psychiatry

The initial promise of a complex dynamic systems approach to mental health was to anticipate symptom transitions within individuals, in real-time. I will reflect on this promise from clinical, practical, ethical, and theoretical viewpoints. The first three viewpoints consider the scenario that EWS would indeed accurately anticipate transitions in mental health – which is not supported by present findings. Hence, my thinking in what follows is based on the assumption that, in time, we will be able to fulfill the promise of a complex dynamic systems approach to mental health. The theoretical perspective zooms in on the discrepancy between the inferences we would like to draw and the inferences that we are currently able to draw from EWS. Further, it highlights the gap between the assumptions related to EWS and our current understanding of psychopathology. This gap should be addressed by future research in order to definitively reject (or accept) the hypothesis that complex dynamic systems principles may translate to psychopathology.

Clinical viewpoint: early intervention

The idea that transitions in mental health may be anticipated by EWS means that if an individual would repeatedly monitor their mental states (*e.g.*, every evening), they (or their

clinician) would be able to infer their vulnerability for a sudden change in their mental health. This has clear clinical utility: if mental health problems could be recognized *before* they arise or worsen, we could try to avert them. For instance, early interventions may promote social support, psycho-education, and vocational functioning, thereby fostering resilience and preventing worsening of symptoms⁹⁶. An illustrative example of such early intervention efforts concerns headspace⁹⁷, which was developed in Australia and recently emerged in the Netherlands, where it is named @ease⁹⁸. Headspace provides a soft entry to mental health care through “one-stop shop” locations where youth receive help from both clinical and non-clinical personnel⁹⁷. Such a service may be particularly useful for individuals whose complaints are not (yet) severely impairing. For 60% of youth, engaging in headspace significantly improves psychological distress and/or psychosocial functioning⁹⁹. Given that individuals attend headspace for highly varying reasons and number of sessions, this improvement rate can be considered a good outcome⁹⁹. In more controlled settings such as randomized controlled trials, early interventions have also proven effective. For instance, meta-analyses of randomized controlled trials showed that early interventions reduce the subsequent incidence of mental disorders over one year by 19% (depression¹⁰⁰), 54% (psychosis¹⁰¹), and 43% (anxiety disorders¹⁰²). Early interventions are most profitable for youth, in whom psychiatric problems often emerge for the first time¹⁰³. We therefore investigated adolescents around 18 years old (**chapters 2 and 3**) and young adults between 21 and 24 years old (**chapters 5-8, 9-10**). If future research succeeds to anticipate changes in their mental health with sufficient sensitivity and specificity, we could potentially bend the outcome of the downward spiral that marks the progression of mental health problems.

Practical viewpoint: collecting the data

From a practical point of view, studying complex dynamic systems principles is demanding: it requires high-resolution, long-term time series of mental states to measure the dynamics, such as EWS. Such data are easily simulated⁷², but much harder to collect. Although this thesis showed that individuals are willing to keep diaries of their mental states for prolonged periods of time, it should be noted that acquiring these data requires a considerable investment of time and money on the part of participants and researchers. A practical work-around might be to rely on passively collected data, for instance by monitoring social interactions, physical activity, or smartphone data¹⁰⁴. As an example, a meta-analysis reported weak support for a relation between passively monitored social interactions (inferred from the duration and

frequency of phone calls) and symptoms of depression, but stronger support for the relation between time spent at home (inferred from GPS tracking) and depression¹⁰⁵. These passively collected data have the additional benefit that they are perhaps more “objective” (*i.e.*, less prone to measurement error) compared to self-reported data. Still, passively collected data have their own particular challenges. For instance, the use of physiological data for monitoring critical slowing down would force us to rethink the system we are evaluating. At present, I conceptualized this system as a network of interacting mental states (*e.g.*, I felt down, I worried, I used (soft)drugs). This fits well with how psychopathology is defined in diagnostic manuals, namely, as an enduring and impairing pattern of emotions, cognitions, and behaviours. Passively monitored data, such as phone usage, thus have a more indirect relation to psychopathology compared to assessments of mental states. In conclusion, although passively collected data have some practical advantages over self-report data and may be promising for the purpose of prediction (*e.g.*, using machine learning), their suitability for monitoring EWS and understanding mental health remains to be confirmed.

An alternative solution to the data demands imposed by a complex dynamic systems approach is to further improve our current assessments of mental states. It has been noted, for instance, that momentary assessments (*e.g.*, “how do you feel right now?”) may be more suitable for detecting sudden transitions compared to retrospective assessments (*e.g.*, “how did you feel during the past day?”)⁹⁰. Second, the context in which certain mental states occur should be taken into account⁹. This might also yield more insight in the perturbations that individuals are exposed to. Finally, we could consider further personalizing our designs. This would, for instance, mean that a certain construct (*e.g.*, anxiety) is measured with different items for different individuals. Such personalization might limit the generalizability or widespread application of intensive longitudinal designs, but this is likely compensated for by the reduced measurement error in the data. In conclusion, there is room for improvement in the intensive longitudinal assessment of mental states. Although the present thesis yielded modest findings with regard to the predictive utility of warning signals in intensive longitudinal data, it is also among the first works to show that collecting such data (in an at-risk population) is possible. The empirical work in this thesis therefore awaits follow-up research that may further improve the quality of mental state assessments.

Ethical viewpoint: knowledge is power versus ignorance is bliss

The reasoning throughout this thesis has been that attempting to anticipate impending

psychopathology is worthwhile. However, it can be questioned whether it is ethical to let individuals know that they are likely to experience mental health problems in the near future. And if so, how will individuals react to such knowledge? Individuals may try to avert the prognosis that they were given, for instance by seeking social support or engaging in pleasant activities. Alternatively, individuals could also submit to their prognosis, fueled by a sense of inevitability. This could turn the prognosis into a self-fulfilling prophecy. As an illustration, it has been shown that individuals who have been falsely informed that they are genetically predisposed to develop depression retrospectively report more severe depressive symptoms and have less faith in their ability to cope with stress compared to individuals who did not receive such information^{106,107}. Knowledge about the risk for mental disorders may also negatively impact how individuals are treated by others. For instance, parents of children who were declared at increased risk for metabolic disorders report higher levels of stress and anxiety compared to parents of children who screened negative¹⁰⁸. A similar reaction may hold in the context of mental disorders. Despite these concerns, both the general population and at-risk populations have a positive attitude towards knowing their risk for psychopathology in advance^{109,110}. Whether they have something to gain from this knowledge depends on the circumstances. If psycho-education (*e.g.*, an explanation of what it means to be at increased risk), preventive interventions or early treatments are both available and effective, being warned for potential onset or worsening of mental health problems in advance may be beneficial^{106,107,109}. Otherwise, knowing one's risk for psychopathology may do more harm than good.

Theoretical viewpoint: expectation versus reality

Our hopeful expectation was that EWS would enable a real-time prediction of whether and when individuals will experience a transition towards a particular cluster of psychopathological symptoms. Over the past years, however, I learned that the inferences we can draw from EWS are more modest. First, it has been shown that EWS cannot inform on the exact timing of transitions in mental health^{54,77}. Such transitions happen at a tipping point, which reflects a situation where the instability of the system's current state (*e.g.*, mental health) has accumulated to such an extent that a transition towards an alternate state (*e.g.*, mental disorder) becomes inevitable⁶. Importantly, however, complex dynamic systems never reach their tipping point. Rather, transitions occur already *before* the tipping point, due to a stochastic (unpredictable) environmental perturbation⁵⁴. In case the system's current

state is relatively stable, a large perturbation is needed to provoke a transition. Conversely, if the system is highly unstable, a small perturbation can cause the system to pass its tipping point. In the context of mental health, for instance, a minor event can cause an already vulnerable individual to experience a transition towards psychopathology. Given the role of these environmental perturbations in eliciting transitions, the predictive capacity of EWS is inherently limited. That is, EWS may monitor vulnerability, but cannot tell when the system is pushed towards an alternate state. This means that we may theoretically be able to infer that a transition towards increased mental health problems is increasingly likely, but we cannot predict *when* it will occur. This limitation is not specific to EWS in the context of mental health, but rather, applies to the broader field of complex dynamic systems. The extent to which it limits the theoretical usefulness of EWS is, however, debatable: if systems are only mildly perturbed – meaning that the daily events that people experience do not affect their mood too much – systems may come close to their tipping point, and we may theoretically be able to predict transitions with sufficient accuracy. Whether this holds in the context of mental health is still unsure, as we know little about (the impact of) the daily events that people experience. Unfortunately, monitoring these daily events is not straightforward. As an illustration, the study we conducted (TRAILS TRANS-ID, described in **chapter 5**) allowed for monitoring daily events by means of taking notes, but we noticed that participants differed in their tendency to use this option (some never took any notes, others provided detailed reports of what happened today). Hence, verifying the role of perturbations in transitions in mental health is important but complicated, and requires further research.

A second concern that I became more aware of during the past years is that EWS do not have reference values, meaning that it is impossible to interpret their absolute value¹¹¹. Hence, there are no absolute thresholds or cut-offs for EWS. EWS are therefore only informative in a relative sense – for instance, as indicators of increasing vulnerability (*i.e.*, declining resilience). In acknowledgement of this, EWS have been relabeled as “resilience indicators”^{54,112–114}. This term may better capture the capacity of EWS, because it stresses that EWS merely indicate that present resilience is lower than previous resilience. In acknowledgement of the limited predictive capacity of EWS, the literature on EWS describes these signals as anticipatory instead of predictive⁹⁰: EWS may *precede* transitions, but are not suited for *predicting* if and when these transitions will occur¹¹¹. Again, this limitation is not specific to our studies, but applies to all studies into EWS. Also, the lack of reference values can be overcome. For instance, it has been suggested that an increased autocorrelation of

two standard deviations (relative to the autocorrelation during a normative, stable situation) can be considered meaningful or significant (e.g., stable period: autocorrelation of 0.3 ± 0.2 , instable period: autocorrelation of 0.6 ± 0.2)¹¹⁵. This introduces new challenges, such as the delineation of a stable period, but also illustrates that – with further empirical research – it might be possible to derive reference values (and decision thresholds) for EWS.

A third important element of the promise of a complex dynamic systems approach is the notion of *symptom transitions*. Such transitions can involve a sudden increase or decrease in the severity of psychopathological symptoms. From repeated assessments of symptom severity, we know that these transitions indeed happen^{48,49,116–119}. However, as already explained in section *Lessons learned from EWS in other fields*, it is unknown whether these transitions indeed resemble the type of transitions that are preceded by critical slowing down, and by analogy, EWS. In other words, in most empirical settings, it is unclear whether EWS should be expected at all⁷⁹. As already touched upon, this complicates rejecting the hypothesis that EWS precede transitions: the absence of EWS can also be attributed to ill-defined transitions. It is important to add that the transitions that would theoretically be optimal for detecting EWS – namely, transitions that occur through a fold bifurcation (section *Lessons learned from EWS in other fields*) – are not necessarily clinically meaningful. This is because a fold bifurcation is defined by mathematics, whereas clinically meaningful transitions can be defined as significant deviations from a norm. This norm can be derived from comparisons between the general population and populations who meet diagnostic criteria, or from the individual's past. Although transitions that occur through fold bifurcations and clinically meaningful transitions *may* overlap, they do not do so by definition. Thus, it is difficult to derive an optimal definition of transitions in mental health: clinically meaningful transitions are not necessarily preceded by critical slowing down (and, by analogy, EWS), whereas theoretically meaningful transitions that are by definition preceded by critical slowing down may not capture all important symptom changes (and therefore lack clinical utility). In the end, how to define transitions depends on one's inferential goals. From a theoretical perspective, for instance, transitions should be defined by their statistical properties (e.g., by the type of bifurcation that is involved), the reasoning being that definitive conclusions about the utility EWS can only be drawn if certain assumptions pertaining to critical slowing down are met. A more pragmatic, clinical perspective, in contrast, would favor transitions with a clear clinical meaning, the reasoning being that EWS will only be useful if they anticipate clinically significant outcomes. This is also the perspective I took in this thesis: **chapter 8**

concerns sudden drops in mental health (*i.e.*, sudden decreases in day-to-day mental health relative to individuals' average), **chapter 9** considers changes in symptoms exceeding clinical cut-offs, and **chapter 10** involves changes in symptom severity over the course of six months of varying magnitude. I could not verify whether (or to what extent) this definition violated the assumptions of critical slowing down, and hence, future work is needed to determine whether and where fold bifurcations and clinically meaningful transitions intersect.

A fourth and final part of the promise that EWS convey concerns the anticipation of transitions *in real-time*. With this, I mean that we should theoretically be able to infer critical slowing down without already knowing the future. In practice, however, EWS are often investigated in historical data, where the future is known. The studies described in this thesis are no different in this regard. We first detected transitions (**chapters 8 and 9**) – which allowed us to know the future – and subsequently examined whether EWS preceded those transitions. This approach matches the vast majority of empirical studies into EWS, and provides an idea of the probability of EWS given that there is a transition (*i.e.*, $P(\text{EWS}|\text{transition})$, or the sensitivity). From an applied perspective, however, it is perhaps more interesting to know the probability of a transition, given that we observe EWS (*i.e.*, $P(\text{transition}|\text{EWS})$, or the positive predictive value). This probability can only be inferred if we inspect EWS in absence of transitions. For instance, we could treat every point in a time series as a potential transition-point, and subsequently inspect whether EWS precede this point. For the TRAILS TRANS-ID data, this would entail making ± 184 predictions per person (one for every time point), instead of just one prediction. Such an approach would likely yield numerous false positives, but also provides a more accurate representation of what we would face if we truly examined EWS in real-time. Unfortunately, empirical studies on EWS that investigated systems without transitions are scarce (but see Burthe *et al.*⁷⁴). This led others to warn for the “the prosecutor’s fallacy”¹²⁰ and the “conjunction fallacy”¹²¹, which both describe that exclusively focusing on systems with transitions may lead to confusing the sensitivity of EWS with their predictive value. This confusion may have contributed to an overly optimistic idea of what EWS have to offer^{52,120}: if transitions are always preceded by EWS, it can be erroneously reasoned that EWS always reflect a true positive. **Chapters 8 and 9** addressed the existing gap in the literature by also examining EWS in individuals without transitions. However, we decided to inspect their EWS only for a single non-transition time point. This created a fair comparison to those individuals with a transition (for whom we also inspected EWS prior to a single time point), but also limited the number of chance findings.

Hence, strictly speaking, we may have underestimated the frequency of false positives. Nevertheless, the approach we took reflects a first step in the right direction, namely away from only confirming that EWS happen prior to transitions and towards investigating whether EWS are indeed specific to transitions. A next step could be to investigate EWS in real-time (*i.e.*, by updating the prediction upon each new observation), which would be a more rigorous test of the idea that EWS may reveal what we do not know yet.

It can be concluded that if EWS are informative of future transitions in mental health at all, they do not predict in real-time, but rather anticipate in hindsight. This is not a limitation of EWS per se, but applies to all current efforts to predict mental health: it seems that life can only be understood backwards, but must be lived forwards (a famous quote from Kierkegaard, later linked to prediction in mental health research by Nelson and Hartman)¹²². At present, therefore, our expectations of EWS do not match reality. This partly has to do with limitations inherent to EWS, namely that a precise indication of the timing of transitions is impossible. Other concerns have more to do with how EWS have been studied so far and can be overcome with further research: EWS do not yet have reference values, it is unclear whether transitions in mental health should be expected to follow from critical slowing down, and the predictive value of EWS is still ambiguous (let alone their capacity to predict in real-time). One may wonder whether we should (or could) have been aware of these limitations in advance. In an ideal world, yes. In reality, however, it takes time (in my experience, several years) and (unexpected) results in order to fully understand how things work. Therefore, I consider the lessons laid out above a result of the empirical work in this thesis, rather than knowledge that should have been present beforehand.

Prospects of a complex dynamic systems approach to psychopathology

The above paragraphs may have given a pessimistic impression of the utility of EWS in the context of mental health. Yet, I agree with Scheffer¹¹³, who noted that *“the glass is not half empty, but rather half full, with many unexplored possibilities of filling it further”* (p.162). Below, I will discuss several of those possibilities. A first important goal for future research is to improve our understanding of mental health as a system. Of what elements does this system consist and does it indeed check the many boxes of complexity (*i.e.*, alternative stable states, sudden transitions, hysteresis, the presence of feedback loops; see also section

Lessons learned from EWS in other fields)^{72,85,123,124}? An important related issue concerns the determination of the timescale at which processes evolve. This topic resurfaced in most chapters, and therefore deserves some elaboration (section *Timescales*). Provided that we can successfully characterize mental health as complex dynamic system, we could look further into EWS (section *Alternative EWS*). A final goal for future research is to look beyond EWS, and investigate stable states in mental (ill-)health (section *Stable states in mental health*).

Timescales

Complex dynamic systems invariably concern interactions between elements that change over time (*e.g.*, interactions between species, cells, or mental states). Hence, to measure complex dynamic systems, it is essential to define the timescale at which processes take place. By analogy, measuring EWS requires a priori insight in the timescale that is relevant for determining a variable's return to equilibrium (*i.e.*, a stable state). Especially the autocorrelation, which was central to **chapters 3 and 7-9**, crucially depends on the time-interval between assessments⁵⁰. As an illustration, your mood now will likely be correlated to your mood in five minutes, but less so to your mood a week from now. This implies that analyzing autocorrelations in data with differing temporal resolutions should yield different findings. In practice, however, autocorrelations were found to have similar associations with symptoms of depression whether based on second-to-second assessments¹²⁵, minute-to-minute assessments¹²⁶, hourly assessments (*i.e.*, ten assessments per day)²³, or daily assessments (*i.e.*, one assessment per day)¹²⁷. Although all of these studies interpreted autocorrelation in the same way – namely, as reflecting emotional inertia – it is likely that the carry-over of emotions from second-to-second has a different meaning compared to the carry-over over the course of days. However, theories on emotion that explicitly mention the timescale at which certain dynamics take place are presently absent, and hence, it remains unknown what timescale best captures the phenomenon of interest. To further complicate the matter, it might well be that such timescales differ between different mental states. For instance, Verduyn¹²⁸ reported that sadness persists much longer than shame, fear, or disgust. This challenges study designs or statistical models that generalize to more than a single mental state. I see two solutions to deal with the timescale issue.

First, we could use continuous-time models to identify how parameters such as autocorrelations vary across time¹²⁹. This would allow for analyzing how autocorrelations – or other dynamics, such as cross-lagged correlations – vary across different time intervals and

across mental states. By definition, autocorrelations exponentially decline: they peak at the smallest possible time interval and then decay¹³⁰ (supplementary Fig. S11.3). The speed of this decay may vary between mental states, and may reveal how quickly a departure from the mean – for instance, feeling more anxious than usual – dies out. Cross-lagged correlations, in contrast, peak at the interval at which the effect of one mental state (*e.g.*, feeling tired) on the other (*e.g.*, feeling irritated) is highest. This peak might reflect the most informative time interval, and can be inferred algebraically using a continuous-time framework^{130,131}. Continuous-time models are often presented as an alternative to discrete-time models, which have been used throughout this thesis. The latter implicitly assume that processes only occur at fixed moments in time (*e.g.*, receiving a salary every month). This also means that dynamics (*e.g.*, autocorrelations) are considered only at a single time interval, and hence, are either present or absent. Continuous-time models, in contrast, assume that processes vary continuously over time (*e.g.*, most psychological processes, including emotions). This means that dynamics are considered to vary across time, which is in line with what would be expected from complex dynamic systems. Importantly, however, these models need some variability in terms of the intervals between assessments in order to yield reliable estimates. Continuous-time models are therefore less suitable to the TRAILS TRANS-ID data, which were collected with fixed 24-hour intervals. Additionally, continuous-time models may deal less well with unexpected, large deviations (*e.g.*, a sudden, brief increase in negative mental states, for instance due to a major life event)¹³². A final challenge related to these models is that they are more complex compared to their discrete-time counterparts. Nevertheless, continuous-time models are gradually finding their way into research on emotion dynamics^{133–136}. To my knowledge, however, continuous-time models have not yet been used to infer the optimal time interval at which mental states affect each other (cross-correlation) or themselves (autocorrelation; apart from illustrative examples)¹³¹.

A second solution to the timescale issue could be to shift our focus away from metrics like autocorrelations and towards metrics that are less sensitive to the temporal resolution of data (*e.g.*, means and variances)^{8,91}. This suits the recent call for more parsimonious metrics in between-individual studies on emotion dynamics⁷. Briefly, it was reported that between-individual differences in symptoms of depression⁷, borderline personality disorder⁷, well-being⁷, and personality⁹, are better explained by the mean and variance of emotions than by more complex measures such as autocorrelations. However, this does not necessarily imply that the mean and variance also outperform rising autocorrelations in terms of predictive capacity *within* individuals. Based on simulation studies, for instance, rising autocorrelations

have been considered a more robust EWS compared to rising variances⁹¹. Empirical studies confirmed this: the probability of true positive EWS is higher for the autocorrelation (21%) compared to the variance (15%)⁷⁴. We found a similar pattern of results in **chapter 9** (probability of true positive EWS for the autocorrelation: 55%, variance: 37%). Further, we found that rising variances have a lower overall specificity (69%) for detecting transitions in bipolar disorder compared to rising autocorrelations (75%) (**chapter 9**, supplement of this chapter). Another reason to be hesitant about rising variances as an EWS is that the amplitude of daily dips or uplifts in mood strongly depend on the events that people experience (*e.g.*, a birthday party). Complex dynamic systems theory assumes that such events – or perturbations – are stochastic. Although this assumption may hold, for example, for ecosystems that are randomly perturbed by rain, it may not hold in the context of mental health. People are not exposed to random events only, but rather, can exert some control over what happens to them (*e.g.*, when deciding whether or not to join the party). A likely consequence is that the variance of mental states changes in ways that would not be expected from complex dynamic systems theory. For instance, as an individual becomes more vulnerable for depression, they may withdraw from events. This would lead to a declining trend in the variance of their mental states, whereas critical slowing down predicts an opposite pattern. Autocorrelations are less sensitive to trends in perturbations, and therefore, more likely to still show a rising trend⁹¹. Hence, rising variances as an EWS have their own particular limitations, and are therefore unlikely to be a flawless solution to the limitations imposed by the autocorrelation. Whether a rising mean reflects a viable alternative to the autocorrelation remains an open question. Typically, means are not considered as warning signals for upcoming transitions in the context of complex dynamic systems. This is because, theoretically, means are not expected to change prior to critical transitions¹³⁷. This is the very reason why critical transitions attracted so much attention in the recent decades: they cannot be foreseen by obvious warning signs, such as rising means, but rather follow from an instability that has accumulated ‘under the surface’¹³⁸. From a clinical point of view, however, it seems likely that at least for some individuals, the onset or worsening of psychopathology is anticipated by a rising trend in, for instance, feeling down, anxious, or irritated. In support of this idea, Smit *et al.*¹³⁹ showed that rises in feeling restless may predict a relapse in depression two months in advance. Similarly, rising trends in negative mental states precede suicidal crises in patients with bipolar disorder⁵⁸. Hence, from a clinical perspective, further investigating how individuals’ average mental states change prior to transitions in mental health might be worthwhile.

Alternative EWS

Provided that we can obtain a solid understanding of mental health as a system, and that this system indeed features the (mathematical) properties of complex dynamic systems, further investigation of EWS is reasonable. Particularly, it would be interesting to explore alternative (and hopefully, more sensitive) metrics of critical slowing down. A first step in this direction could be to consider multivariate EWS. In contrast to the EWS described in **chapters 8 and 9**, which were based on individual negative mental states (*e.g.*, feeling down), multivariate EWS could comprise multiple mental states (*e.g.*, feeling down, irritated, and anxious). Such a multivariate signal is likely more reliable than univariate signals, which are based on assessments of single mental states with limited reliability^{7,86,90}. Examples of multivariate EWS include the maximum value of the covariance matrix¹⁴⁰, the autocorrelation of the first principal component scores¹⁴¹, and the eigenvalue of the covariance matrix (*i.e.*, the amount of variance explained by the first principal component, used in **chapter 10**)^{141,142}. The accuracy of such multivariate signals depends on the properties of the system – *e.g.*, the type of noise it is exposed to⁷² – and therefore, exploring alternative metrics is only warranted once we have a better understanding of the system we are dealing with. Also, multivariate EWS do not overcome the limitations inherent to EWS – namely, being unable to predict the precise timing of transitions and being solely interpretable in a relative sense (section *Lessons learned from EWS in psychiatry, Theoretical viewpoint*). Investigating alternative EWS therefore only has clinical utility if these inherent limitations are overcome.

Stable states in mental health

Although this discussion has so far mostly focused on EWS, the potential merits of a complex dynamic systems approach to understanding psychopathology reach beyond EWS. Specifically, it is interesting to further look into the idea that mental health manifests in distinct stable states (attractors). Such stable states are typically labelled mental health and mental disorder, but could also represent subthreshold symptoms of psychopathology (**chapter 7**). So far, I have only considered the scenario that both stable states reflect fixed-point attractors (*i.e.*, equilibria to which the individual returns again and again, for instance a stable state of depression or a state of well-being). However, this idea likely (over)simplifies the manifestation of psychopathology in several respects. First, it is possible that some stable states are better described as *limit cycles* as opposed to fixed-point attractors. Limit cycles reflect an attractor that consists of a cyclic pattern (*e.g.*, predator-prey dynamics)⁶.

For instance, bipolar disorder (particularly, the rapid cycling subtype that we investigated in **chapter 9**) might be represented as a limit cycle attractor, where manic and depressive symptoms alternate, rather than two different fixed-point attractors representing manic and depressive episodes. This would mean that the switches between manic and depressive episodes do not reflect critical transitions between attractors, but rather, *are* the attractor. Other disorders that might show cyclic patterns may be those that are triggered by specific environmental stimuli (*e.g.*, specific phobias, post-traumatic stress disorder). If these disorders are indeed better described as limit cycles than as fixed-point attractors, this means that EWS would be expected prior to the onset of the cyclic pattern (*e.g.*, depression, mania, depression, etc.) rather than prior to the onset of one part of this pattern (*e.g.*, depression).

A second refinement to the idea of alternative stable states is that some states in mental health may reflect *long transients* instead of stable, fixed-point attractors. Long transients reflect states that are close to an unstable state and seem stable, which is why they have been termed quasi-stable states^{6,143}. Long transients differ from fixed-point attractors in that they constitute an intermediate state rather than a stable long-term state. As an example, an ultra-high risk state for psychosis could be considered a long transient: individuals may be at ultra-high risk for multiple years before they eventually transition to psychosis¹⁴⁴. Unlike the stable states I considered so far, long transients can give rise to a sudden transition in absence of any environmental change¹⁴⁵. The distinction between fixed-point attractors and long transients is interesting because it has implications for how the state could be modified^{6,143,145}. Altering a fixed-point attractor requires that it is destabilized, which should facilitate a transition towards an alternate (more adaptive) state. In the context of psychiatry, this may involve breaking the patterns that maintain a maladaptive attractor (*e.g.*, through reappraisal and improving emotion regulation strategies) and strengthening a more adaptive attractor (*e.g.*, through behavioral activation)⁴⁷. Altering a long transient, in contrast, can be attained by rapidly intervening on the driver⁹³. In practice, this could mean removing a stressful environment. This, in turn, may avert a transition towards a lasting, unfavorable state⁹³.

Despite the fact that long transients may precipitate an unfavorable and potentially persisting state, long transients have long been considered innocent and uninteresting⁹³. This perhaps parallels how subthreshold psychopathology has sometimes been viewed, namely, as trivial in light of the impairments associated with severe psychopathology¹⁴⁶. However, whether or not subthreshold psychopathology may reflect a long transient remains

speculative until some adjacent questions are answered. For instance, if subthreshold psychopathology resembles a long transient, there should be another, more stable state that develops over a longer timescale. Further, long transients may be particularly likely in slowly responding systems – which implies that they could be specific to certain psychopathological domains or individuals⁹³. Both of these open questions bring back an issue described earlier, namely that we know little about the timescale at which dynamics related to mental health take place (section *Timescales*). Therefore, the possibility that there are long transients in the context of psychopathology can only be investigated once we have better understanding of the timescale at which mental (ill-)health develops.

A third consideration related to the idea of stable states in mental health is that reducing mental health to just two states may not do justice to the variety of symptom expressions that exist (*e.g.*, internalizing symptoms, such as depression and anxiety, versus externalizing symptoms, such as behavioral problems and substance use). This led us to propose a three-dimensional stability landscape in **chapter 4**. However, before attempting to model this complex, three-dimensional system, we should first improve our understanding of bistability in the context of psychopathology. For instance, it would be interesting to know when and/or for whom the idea of bistability holds. It could be that under some circumstances, systems are bistable (and hence, vulnerable for sudden transitions), whereas other circumstances give rise to more continuous shifts between alternate states. This suggestion is in line with the observation that psychopathological symptoms shift suddenly for some individuals, and gradually for others³⁶. It also matches our finding that although almost all at-risk young adults studied in **chapters 5-8 and 10** reported a change in their symptom severity from baseline to post, sudden drops in their daily reports of mental health were rare (prevalence 16.4%, **chapter 8**). Cusp models allow for addressing differences in individuals' susceptibility to sudden transitions (Fig. 11.3). Specifically, cusp models suggest that the degree to which a system is likely to show gradual versus more sudden transition is dependent on a third variable. A compelling candidate variable could be the connectivity of the overall network of mental states, as it has been proposed that more densely connected networks are more likely to undergo critical transitions compared to networks with weaker connectivity⁴. Cusp models have already been applied in other areas of psychology. For instance, according to the catastrophe theory of attitudes, transitions in individuals' attitudes (*e.g.*, from being pro towards against Zwarte Piet) happen gradually in case involvement is low, but resemble critical transitions at high levels of involvement^{147,148}. Whether a similar

principle holds in the context of psychopathology – meaning that the suddenness of transitions would depend on the connectivity of the network – awaits empirical investigation.

Fig. 11.3. *Illustration of the cusp model to illustrate that bistability in mental health may depend on a third variable (e.g., connectivity). It can be seen that the system's response to increasing stress depends on the connectivity of the system: when connectivity is low, the response is gradual, whereas high levels of connectivity facilitate a sudden transition.*

Concluding remarks

It should be emphasized, once more, that both the exploration of alternative EWS (e.g., multivariate EWS) and the further investigation of alternative states in mental health (e.g., through cusp models) are only warranted once we have a better understanding of the system we are dealing with. Without such understanding, we will be unable to infer whether the absence of EWS (or other complex dynamic systems properties) are due to violated assumptions (i.e., a true negative) or insufficient data quality or quantity (i.e., a false negative, supplementary Fig. S11.2). Therefore, the priority of future research in this field should be to characterize the system of mental health in terms of its complex features (e.g., bimodality, hysteresis) and its dynamics (i.e., timescales of interest)^{72,85,123,124}. This quest can be approached from many different angles. For instance, advanced statistics may allow for determining system behavior through continuous-time modeling^{50,131}, formal modeling⁸⁰, or by explicitly testing for complex features^{5,6,79,85}. Such tests have already been applied in ecological and climate systems, and may accurately identify those systems that are likely to exhibit EWS⁷⁹. Further, clinical insight is required to determine the types of elements (e.g., types of mental states) of which a system of mental health consists. This insight might be attained through a collaboration between scientists, clinicians, and patients. Finally, we

should pay attention to the psychometric and practical matters that come with measuring mental health as a complex dynamic system. For instance, we might increase the reliability of our measurements by further personalizing our designs or by replacing individual mental states with a smaller set of homogeneous constructs (*e.g.*, derived by principal component analyses)¹⁴⁹.

If we fail to characterize the system of mental health, EWS might await the same fate as catastrophe theory, which flourished in the 1970s and attempted to characterize and predict sudden shifts in various systems using the same mathematical underpinnings as EWS^{147,150}. Catastrophe theory was quickly embraced as a tool that could yield “*profound insight into the universe*”¹⁵¹ and has been applied to all sorts of sudden shifts in biology, physics, social systems, health, and financial systems¹⁴⁷. However, as most of the resulting theoretical models appeared to ignore fundamental (mathematical) assumptions underlying catastrophe theory, they were quickly disregarded as “*rubbish*”^{151,152}. These strong critiques were later nuanced, with the reconciling notion that we should not throw out the baby with the bathwater¹⁵³. The lesson that can be drawn is that we should be reluctant in applying “generic” principles – including EWS, or more generally, catastrophe theory – to real-world systems in which the assumptions that underlie the generic principles are not fully understood. An improved understanding of what we study and appreciation of the assumptions that come with taking a complex dynamic systems perspective will hopefully prevent history from repeating itself.

With this thesis, I hope to have contributed some insights in how complex dynamic systems translate to psychopathology. Many parts of this translation are still unresolved, and require future research. I believe that this research is worth conducting, not only because of the potential clinical merits of complex dynamic systems principles, but also because of their theoretical merits. That is, considering mental health as a complex dynamic system may lead us to re-think established concepts and generate hypotheses about what triggers and sustains psychopathology. This thesis is therefore not the endpoint, but rather, the beginning of a scientific journey.

Supplement 11

Fig. S11.1. *Schematic illustration of Simpson’s paradox²⁶, adapted from Hamaker¹⁵⁴*. The large black circle reflects a potential group-level association between emotion differentiation and psychopathology: between-individual differences in emotion differentiation do not relate to between-individual differences in psychopathology. The colored circles reflect potential within-individual associations, each color representing a different individual. It can be seen that, as an individual’s emotion differentiation declines, their level of psychopathology may increase. Any combination of between- and within-individual relations is possible; this figure serves only to illustrate why between-individual associations need not generalize to the individual level.

	EWS present	EWS absent
Transition present	True positive	False negative
Transition absent	False positive/false alarm	True negative

Fig. S11.2. *Schematic illustration of true and false positives and negatives in the context of EWS*. Transition here refers to a transition in mental health (either a drop in mental health, as reported in chapter 8, or a shift between manic and depressed episodes, as reported in chapter 9). Ideally, true positives/negatives should be frequent, whereas false positives/negatives should be seldom.

Table S11.1. Overview of the sensitivity and specificity of EWS across studies

Data	EWS: autocorrelation		EWS: standard deviation	
	<i>Present</i>	<i>Absent</i>	<i>Present</i>	<i>Absent</i>
<i>Chapter 8 – at-risk youth</i>				
Transition	35	665	-	-
No transition	28	672	-	-
Sens./spec.	5.0%	96.0%	-	-
PPV/NPV	19.7%	83.8%	-	-
<i>Chapter 9 – bipolar patients</i>				
Transition towards depression	36	100	22	114
Transition towards mania	39	80	47	72
No transition	30	89	37	82
Sens./spec. depression	26.5%	74.8%	16.2%	68.9%
Sens./spec. mania	32.8%	74.8%	39.5%	68.9%
PPV/NPV depression	37.1%	64.4%	22.6%	59.4%
PPV/NPV mania	38.0%	70.3%	37.0%	70.5%
<i>Ecosystems (Burthe et al.⁷⁴)</i>				
Transition	19	61	16	64
No transition	72	-	90	-
Sens./spec.	23.8%	-	20.0%	-

Note. Sens. and spec. refer to the sensitivity ($P(\text{EWS}|\text{transition})$) and specificity ($P(\text{no EWS}|\text{no transition})$) of EWS; PPV and NPV refer to the positive predictive value ($P(\text{transition}|\text{EWS})$) and negative predictive value ($P(\text{no transition}|\text{no EWS})$). For readability, I did not denote the probability of true positive EWS (*i.e.*, true positives/total no. of positives). In chapter 8, we did not investigate the standard deviation as an EWS. The data pertaining to ecosystems serve to illustrate how our findings compare to those obtained in other systems. The data concerning EWS in ecosystems have been reported by Burthe and colleagues⁷⁴. As the authors did not report true negatives, I could not calculate the specificity, PPV, and NPV of EWS for these systems.

Fig. S11.3. *Autocorrelations and cross-lagged correlations may vary across time intervals.* Autocorrelations reflect the carry-over effects of mental state 1 and 2 at different time intervals. The cross-lagged correlation reflects the extent to which mental state 1 affects mental state 2 at later point in time. The studies in this thesis inspected dynamics at a fixed interval (*i.e.*, 24 hours in chapters 7 and 8-10, 3 hours in chapter 9), which potentially limits the effects we could detect: how much autocorrelation 1 and 2, for instance, differ, may be influenced by the time interval.

References

1. Barrett LF, Gross J, Christensen TC, Benvenuto M. Knowing what you're feeling and knowing what to do about it: mapping the relation between emotion differentiation and emotion regulation. *Cognition and Emotion*. 2001;15(6):713-724. doi:10.1080/02699930143000239
2. Quoidbach J, Gruber J, Mikolajczak M, Kogan A, Kotsou I, Norton MI. Emodiversity and the emotional ecosystem. *Journal of Experimental Psychology: General*. 2014;143(6):2057-2066. doi:10.1037/a0038025
3. Curtiss J, Fulford D, Hofmann SG, Gershon A. Network dynamics of positive and negative affect in bipolar disorder. *Journal of Affective Disorders*. 2019;249:270-277. doi:10.1016/j.jad.2019.02.017
4. Cramer AOJ, van Borkulo CD, Giltay EJ, et al. Major depression as a complex dynamic system. *PLoS one*. 2016;11(12):1-20. doi:10.1371/journal.pone.0167490
5. Hosenfeld B, Bos EH, Wardenaar KJ, et al. Major depressive disorder as a nonlinear dynamic system: bimodality in the frequency distribution of depressive symptoms over time. *BMC Psychiatry*. 2015;15(1):1-9. doi:10.1186/s12888-015-0596-5
6. Scheffer M. *Critical Transitions in Nature and Society*. Princeton University Press; 2009.
7. Dejonckheere E, Mestdagh M, Houben M, et al. Complex affect dynamics add limited information to the prediction of psychological well-being. *Nature Human Behaviour*. 2019;3(5):478-491. doi:10.1038/s41562-019-0555-0
8. Jahng S, Wood PK, Trull TJ. Analysis of affective instability in ecological momentary assessment: indices using successive difference and group comparison via multilevel modeling. *Psychological Methods*. 2008;13:354-375. doi:10.1037/a0014173.supp
9. Wendt LP, Wright AGC, Pilkonis PA, et al. Indicators of affect dynamics: structure, reliability, and personality correlates. *European Journal of Personality*. 2020;34(6):1060-1072. doi:10.1002/per.2277
10. van der Gucht K, Dejonckheere E, Erbas Y, et al. An experience sampling study examining the potential impact of a mindfulness-based intervention on emotion differentiation. *Emotion*. 2019;19(1):123-131. doi:10.1037/emo0000406
11. Larsen RJ, Ketelaar T. Personality and susceptibility to positive and negative emotional states. *Journal of Personality and Social Psychology*. 1991;61(1):132-140. doi:10.1037/0022-3514.61.1.132
12. Costa PT, McCrae RR. Normal personality assessment in clinical practice: the NEO personality inventory. *Psychological Assessment*. 1992;4(1):5-13. doi:10.1037/1040-3590.4.1.5
13. Mestdagh M, Pe M, Pestman W, Verdonck S, Kuppens P, Tuerlinckx F. Sidelining the mean: the relative variability index as a generic mean-corrected variability measure for bounded variables. *Psychological Methods*. 2018;23(4):690-707. doi:10.1037/met0000153
14. Suls J, Green P, Hillis S. Emotional reactivity to everyday problems, affective inertia, and neuroticism. *Personality and Social Psychology Bulletin*. 1998;24(2):127-136.
15. Koval P, Kuppens P, Allen NB, Sheeber L. Getting stuck in depression: The roles of rumination and emotional inertia. *Cognition and Emotion*. 2012;26(8):1412-1427. doi:10.1080/02699931.2012.667392
16. Kuppens P, Allen NB, Sheeber LB. Emotional inertia and psychological maladjustment. *Psychological Science*. 2010;21(7):984-991. doi:10.1177/0956797610372634
17. van de Leemput IA, Wichers M, Cramer AOJ, et al. Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences*. 2014;111(1):87-92.

- doi:10.1073/pnas.1312114110
18. Houben M, Kuppens P. Emotion dynamics and the association with depressive features and borderline personality disorder traits: unique, specific, and prospective relationships. *Clinical Psychological Science*. 2020;8(2):226-239. doi:10.1177/2167702619871962
 19. Manchia M, Pisanu C, Squassina A, Carpiello B. Challenges and future prospects of precision medicine in psychiatry. *Pharmacogenomics and Personalized Medicine*. 2020;13:127-140. doi:10.2147/PGPM.S198225
 20. Hatton CM, Paton LW, McMillan D, Cussens J, Gilbody S, Tiffin PA. Predicting persistent depressive symptoms in older adults: a machine learning approach to personalised mental healthcare. *Journal of Affective Disorders*. 2019;246(December 2018):857-860. doi:10.1016/j.jad.2018.12.095
 21. Rottenberg J, Hindash AC. Emerging evidence for emotion context insensitivity in depression. *Current Opinion in Psychology*. 2015;4:1-5. doi:10.1016/j.copsyc.2014.12.025
 22. Rottenberg J. Mood and emotion in major depression. *Current Directions in Psychological Science*. 2005;14(3):167-170.
 23. Koval P, Kuppens P, Allen NB, Sheeber L. Getting stuck in depression: the roles of rumination and emotional inertia. *Cognition and Emotion*. 2012;26(8):1412-1427. doi:10.1080/02699931.2012.667392
 24. Kashdan TB, Barrett LF, McKnight PE. Unpacking emotion differentiation. *Current Directions in Psychological Science*. 2015;24(1):10-16. doi:10.1177/0963721414550708
 25. Alessandri G, de Longis E, Eisenberg N, Hobfoll SE. A multilevel moderated mediational model of the daily relationships between hassles, exhaustion, ego-resiliency and resulting emotional inertia. *Journal of Research in Personality*. 2020;85. doi:10.1016/j.jrp.2020.103913
 26. Kievit RA, Frankenhuis WE, Waldorp LJ, Borsboom D. Simpson's paradox in psychological science: a practical guide. *Frontiers in Psychology*. 2013;4. doi:10.3389/fpsyg.2013.00513
 27. Dakos V. Identifying best-indicator species for abrupt transitions in multispecies communities. *Ecological Indicators*. 2018;94:494-502. doi:10.1016/j.ecolind.2017.10.024
 28. Boerlijst MC, Oudman T, de Roos AM. Catastrophic collapse can occur without early warning: examples of silent catastrophes in structured ecological models. *PLoS one*. 2013;8(4):1-6. doi:10.1371/journal.pone.0062033
 29. Patterson AC, Strang AG, Abbott KC. When and where we can expect to see early warning signals in multispecies systems approaching tipping points: insights from theory. *The American Naturalist*. 2021;198(1). doi:10.1086/714275
 30. Huisman M, Oldehinkel AJ, de Winter A, et al. Cohort profile: the Dutch "Tracking Adolescents' Individual Lives' Survey"; TRAILS. *International Journal of Epidemiology*. 2008;37(6):1227-1235. doi:10.1093/ije/dym273
 31. Vachon H, Viechtbauer W, Rintala A, Myin-Germeys I. Compliance and retention with the experience sampling method over the continuum of severe mental disorders: meta-analysis and recommendations. *Journal of Medical Internet Research*. 2019;21(12):1-20. doi:10.2196/14475
 32. Nelson B, McGorry PD, Wichers M, Wigman JT, Hartmann JA. Moving from static to dynamic models of the onset of mental disorder. *JAMA Psychiatry*. 2017;74(5):528-534. doi:10.1001/jamapsychiatry.2017.0001
 33. Wesselhoeft R, Sørensen MJ, Heiervang ER, Bilenberg N. Subthreshold depression in children and adolescents - a systematic review. *Journal of Affective Disorders*. 2013;151(1):7-22. doi:10.1016/j.jad.2013.06.010

34. Rai D, Skapinakis P, Wiles N, Lewis G, Araya R. Common mental disorders, subthreshold symptoms and disability: longitudinal study. *British Journal of Psychiatry*. 2010;197(5):411-412. doi:10.1192/bjp.bp.110.079244
35. Lee YY, Stockings EA, Harris MG, et al. The risk of developing major depression among individuals with subthreshold depression: a systematic review and meta-analysis of longitudinal cohort studies. *Psychological Medicine*. 2018;49:92-102. doi:10.1017/s0033291718000557
36. Shankman SA, Lewinsohn PM, Klein DN, Small JW, Seeley JR, Altman SE. Subthreshold conditions as precursors for full syndrome disorders: a 15-year longitudinal study of multiple diagnostic classes. *Journal of Child Psychology and Psychiatry*. 2009;50(12):1485-1494. doi:10.1111/j.1469-7610.2009.02117.x.Subthreshold
37. Bosman RC, ten Have M, de Graaf R, Muntingh AD, van Balkom AJ, Batelaan NM. Prevalence and course of subthreshold anxiety disorder in the general population: a three-year follow-up study. *Journal of Affective Disorders*. 2019;247:105-113. doi:10.1016/j.jad.2019.01.018
38. Norén Selinus E, Molero Y, Lichtenstein P, et al. Subthreshold and threshold attention deficit hyperactivity disorder symptoms in childhood: psychosocial outcomes in adolescence in boys and girls. *Acta Psychiatrica Scandinavica*. 2016;134:533-545. doi:10.1111/acps.12655
39. Papachristou E, Oldehinkel AJ, Ormel J, et al. The predictive value of childhood subthreshold manic symptoms for adolescent and adult psychiatric outcomes. *Journal of Affective Disorders*. 2017;212:86-92. doi:10.1016/j.jad.2017.01.038
40. Fergusson DM, Horwood LJ, Ridder EM, Beautrais AL. Subthreshold depression in adolescence and mental health outcomes in adulthood. *Archives of General Psychiatry*. 2005;62:66-72.
41. Tuithof M, ten Have M, van Dorsselaer S, Kleinjan M, Beekman A, de Graaf R. Course of subthreshold depression into a depressive disorder and its risk factors. *Journal of Affective Disorders*. 2018;241(July):206-215. doi:10.1016/j.jad.2018.08.010
42. Kaymaz N, Drukker M, Lieb R, et al. Do subthreshold psychotic experiences predict clinical outcomes in unselected non-help-seeking population-based samples? A systematic review and meta-analysis, enriched with new results. *Psychological Medicine*. 2012;42(11):2239-2253. doi:10.1017/S0033291711002911
43. Cukor J, Wyka K, Jayasinghe N, Difede J. The nature and course of subthreshold PTSD. *Journal of Anxiety Disorders*. 2010;24(8):918-923. doi:10.1016/j.janxdis.2010.06.017
44. Janssens M, Boyette L, Lou, Heering HD, et al. Developmental course of subclinical positive and negative psychotic symptoms and their associations with genetic risk status and impairment. *Schizophrenia Research*. 2016;174(1-3):177-182. doi:10.1016/j.schres.2016.03.028
45. Fusar-poli P, Bonoldi I, Yung AR, et al. Predicting psychosis. *Archives of General Psychiatry*. 2012;69(3):220-229.
46. Nienhuis FJ, van de Willige G, Rijnders CAT, de Jonge P, Wiersma D. Validity of a short clinical interview for psychiatric diagnosis: the mini-SCAN. *British Journal of Psychiatry*. 2010;196(1):64-68. doi:10.1192/bjp.bp.109.066563
47. Hayes AM, Andrews LA. A complex systems approach to the study of change in psychotherapy. *BMC Medicine*. 2020;18(1):1-13. doi:10.1186/s12916-020-01662-2
48. Wichers M, Smit AC, Snippe E. Early warning signals based on momentary affect dynamics can expose nearby transitions in depression: a confirmatory single-subject time-series study. *Journal for Person-Oriented Research*. 2020;6(1):1-15. doi:10.17505/jpor.2020.22042

49. Wichers M, Groot PC, Psychosystems ESM group, EWS group. Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*. 2016;85:114-116. doi:10.1159/000441458
50. Haslbeck JMB, Ryan O. Recovering within-person dynamics from psychological time series. *Multivariate Behavioural Research*. 2021;56(2):1-32. doi:10.1080/00273171.2021.1896353
51. Wen H, Ciamarra MP, Cheong SA. How one might miss early warning signals of critical transitions in time series data: a systematic study of two major currency pairs. *PLoS one*. 2018;13(3):1-22. doi:10.1371/journal.pone.0191439
52. Boettiger C, Hastings A. Quantifying limits to detection of early warning for critical transitions. *Journal of The Royal Society Interface*. 2012;9(75):2527-2539. doi:10.1098/rsif.2012.0125
53. Boettiger C, Ross N, Hastings A. Early warning signals: the charted and uncharted territories. *Theoretical Ecology*. 2013;6(3):255-264. doi:10.1007/s12080-013-0192-6
54. Dakos V, Carpenter SR, van Nes EH, Scheffer M. Resilience indicators: prospects and limitations for early warnings of regime shifts. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2015;370(1659):1-10. doi:10.1098/rstb.2013.0263
55. First MB, Botteron KN, Castellanos FX. *Consensus Report of the APA Work Group on Neuroimaging Markers of Psychiatric Disorders*.; 2012. <https://www.researchgate.net/publication/261507750>
56. Large MM, Ryan CJ, Singh SP, Paton MB, Nielssen OB. The predictive value of risk categorization in schizophrenia. *Harvard Review of Psychiatry*. 2011;19(1):25-33. doi:10.3109/10673229.2011.549770
57. Czyz EK, Yap JRT, King CA, Nahum-Shani I. Using intensive longitudinal data to identify early predictors of suicide-related outcomes in high-risk adolescents: practical and conceptual considerations. *Assessment*. Published online 2020. doi:10.1177/1073191120939168
58. Thompson W, Gershon A, O'Hara R, Bernert R, Depp C. The prediction of study-emergent suicidal ideation in bipolar disorder: a pilot study using ecological momentary assessment data. *Bipolar Disorders*. 2014;16(7):669-677. doi:10.1111/bdi.12218
59. Birmaher B, Merranko JA, Gill MK, et al. Predicting personalized risk of mood recurrences in youths and young adults with bipolar spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2020;(February). doi:10.1016/j.jaac.2019.12.005
60. Yuen HP, Mackinnon A, Hartmann J, et al. Dynamic prediction of transition to psychosis using joint modelling. *Schizophrenia Research*. 2018;202:333-340. doi:10.1016/j.schres.2018.07.002
61. Raket LL, Jaskolowski J, Kinon BJ, et al. Dynamic Electronic Health Record Detection (DETECT) of individuals at risk of a first episode of psychosis: a case-control development and validation study. *The Lancet Digital Health*. 2020;2(5):e229-e239. doi:10.1016/S2589-7500(20)30024-8
62. Fisher AJ, Soyster PD. Generating accurate personalized predictions of future behavior: a smoking exemplar. *Preprint*. Published online 2021.
63. Fusar-Poli P, Rutigliano G, Stahl D, et al. Development and validation of a clinically based risk calculator for the transdiagnostic prediction of psychosis. *JAMA Psychiatry*. 2017;74(5):493-500. doi:10.1001/jamapsychiatry.2017.0284
64. Arango C, Dragioty E, Solmi M, et al. Risk and protective factors for mental disorders beyond genetics: an evidence-based atlas. *World Psychiatry*. Published online 2021.
65. Lever JJ, van de Leemput IA, Weinans E, et al. Foreseeing the future of mutualistic communities beyond collapse. *Ecology Letters*. 2020;23(1):2-15. doi:10.1111/ele.13401

66. Kossakowski JJ, Groot PC, Haslbeck JMB, Borsboom D, Wichers M. Data from ‘Critical slowing down as a personalized early warning signal for depression.’ *Journal of Open Psychology Data*. 2017;5. doi:10.5334/jopd.29
67. Fried EI, Borkulo CD van, Epskamp S, Schoevers RA, Tuerlinckx F, Borsboom D. Measuring depression over time... or not? Lack of unidimensionality and longitudinal measurement invariance in four common rating scales of depression. *Psychological Assessment*. 2016;28(11):1-14. doi:10.1037/pas0000275
68. Barlow DH, Nock MK. Why can’t we be more idiographic in our research? *Perspectives on Psychological Science*. 2009;4(1):19-21.
69. Os J van, Delespaul P, Wigman JTW, Myin-Germeys I, Wichers M. Psychiatry beyond labels: introducing contextual precision diagnosis across stages of psychopathology. *Psychological Medicine*. 2013;43(07):1563-1567. doi:10.1017/s0033291713000937
70. van Os J, Delespaul P, Wigman J, Myin-Germeys I, Wichers M. Beyond DSM and ICD: introducing “precision diagnosis” for psychiatry using momentary assessment technology. *World Psychiatry*. 2013;12(2):113-117.
71. Wright AGC, Woods WC. Personalized models of psychopathology. *Annual Review of Clinical Psychology*. 2020;16:49-74. doi:10.1146/annurev-clinpsy-102419-125032
72. Weinans E, Quax R, van Nes EH, Leemput IA van de. Evaluating the performance of multivariate indicators of resilience loss. *Scientific Reports*. 2021;11(1):9148. doi:10.1038/s41598-021-87839-y
73. Johnson SL, Nowak A. Dynamical patterns in bipolar depression. *Personality and Social Psychology Review*. 2002;6(4):380-387. doi:10.1207/S15327957PSPR0604_12
74. Burthe SJ, Henrys PA, Mackay EB, et al. Do early warning indicators consistently predict nonlinear change in long-term ecological data? *Journal of Applied Ecology*. 2016;53(3):666-676. doi:10.1111/1365-2664.12519
75. Wissel C. A universal law of the characteristic return time near thresholds. *Oecologia*. 1984;65:101-107.
76. Gsell AS, Scharfenberger U, Özkundakci D, et al. Evaluating early-warning indicators of critical transitions in natural aquatic ecosystems. *Proceedings of the National Academy of Sciences of the United States of America*. 2016;113(50):1-9. doi:10.1073/pnas.1608242113
77. Kéfi S, Dakos V, Scheffer M, van Nes EH, Rietkerk M. Early warning signals also precede non-catastrophic transitions. *Oikos*. 2013;122(5):641-648. doi:10.1111/j.1600-0706.2012.20838.x
78. Scheffer M, Carpenter SR, Lenton TM, et al. Anticipating critical transitions. *Science*. 2012;338(6105):344-348. doi:10.1126/science.1225244
79. Litzow MA, Hunsicker ME. Early warning signals, nonlinearity, and signs of hysteresis in real ecosystems. *Ecosphere*. 2016;7(12). doi:10.1002/ecs2.1614
80. Burger J, Veen D van der, Robinaugh DJ, et al. Bridging the gap between complexity science and clinical practice by formalizing idiographic theories: a computational model of functional analysis. *BMC Medicine*. 2020;18(99):1-18. doi:10.1186/s12916-020-01558-1
81. Bury TM, Sujith RI, Pavithran I, et al. Deep learning for early warning signals of regime shifts. *Preprint*. Published online 2021. doi:10.1101/2021.03.28.437429
82. Schröder A, Persson L, de Roos AM. Direct experimental evidence for alternative stable states: A review. *Oikos*. 2005;110(1):3-19. doi:10.1111/j.0030-1299.2005.13962.x
83. Haslbeck J, Ryan O. Recovering bistable systems from psychological time series. *Preprint*. Published online 2021.

84. Gelfand LA, Engelhart S. Dynamical systems theory in psychology: assistance for the lay reader is required. *Frontiers in Psychology*. 2012;3(382):1-3. doi:10.3389/fpsyg.2012.00382
85. Olthof M, Hasselman F, Lichtwarck-Aschoff A. Complexity in psychological self-ratings: implications for research and practice. *BMC Medicine*. 2020;18(317):1-16. doi:10.1186/s12916-020-01727-2
86. Bury T. Detecting and distinguishing transitions in ecological systems: model and data-driven approaches. Published online 2020.
87. Clements CF, Ozgul A. Including trait-based early warning signals helps predict population collapse. *Nature Communications*. 2016;7. doi:10.1038/ncomms10984
88. Dakos V, Carpenter SR, Brock WA, et al. Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLoS one*. 2012;7(7):1-20. doi:10.1371/journal.pone.0041010
89. Dakos V, Scheffer M, van Nes EH, Brovkin V, Petoukhov V, Held H. Slowing down as an early warning signal for abrupt climate change. *PNAS*. 2008;105(38):14308-14312. doi:10.1073/pnas.0802430105
90. Dablander F, Pichler A, Cika A, Bacilieri A. Anticipating critical transitions in psychological systems using early warning signals: theoretical and practical considerations theory of critical slowing down. *Submitted*. Published online 2020:1-30.
91. Dakos V, van Nes EH, D'Odorico P, Scheffer M. Robustness of variance and autocorrelation as indicators of critical slowing down. *Ecology*. 2012;93(2):264-271. doi:10.1890/11-0889.1
92. Lenton TM, Livina VN, Dakos V, van Nes EH, Scheffer M. Early warning of climate tipping points from critical slowing down: comparing methods to improve robustness. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 2012;370(1962):1185-1204. doi:10.1098/rsta.2011.0304
93. Hughes TP, Linares C, Dakos V, van de Leemput IA, van Nes EH. Living dangerously on borrowed time during slow, unrecognized regime shifts. *Trends in Ecology and Evolution*. 2013;28(3):149-155. doi:10.1016/j.tree.2012.08.022
94. Wichers M. The dynamic nature of depression: a new micro-level perspective of mental disorder that meets current challenges. *Psychological Medicine*. 2014;44(7):1349-1360. doi:10.1017/S0033291713001979
95. Bayani A, Hadaeghi F, Jafari S, Murray G. Critical slowing down as an early warning of transitions in episodes of bipolar disorder: a simulation study based on a computational model of circadian activity rhythms. *Chronobiology International*. 2017;34(2):235-245. doi:10.1080/07420528.2016.1272608
96. Patel V, Flisher AJ, Hetrick S, McGorry P. Mental health of young people: a global public-health challenge. *The Lancet*. 2007;369. doi:10.1016/S0140
97. McGorry P, Trethowan J, Rickwood D. Creating headspace for integrated youth mental health care. *World Psychiatry*. 2019;18(2):140-141. doi:10.1002/wps.20619
98. Leijdesdorff S, Postma MR, van Kersbergen L, Marchetta N, van Amelsvoort T. No boundaries: a 2 year experience in a specialized youth mental health care program in the Netherlands. *Early Intervention in Psychiatry*. 2020;14(2):228-234. doi:10.1111/eip.12887
99. Rickwood BA DJ, Mazzer KR, Parker BA AG, Tanti BSW CJ, McGorry PD. Changes in psychological distress and psychosocial functioning in young people accessing headspace centres for mental health problems. *Medical Journal of Australia*. 2015;202(10):537-543. doi:10.5694/mja14.01
100. Cuijpers P, Pineda BS, Quero S, et al. Psychological interventions to prevent the onset of depressive disorders: a meta-analysis of randomized controlled trials. *Clinical Psychology Review*. 2021;83.

- doi:10.1016/j.cpr.2020.101955
101. van der Gaag M, Smit F, Bechdolf A, et al. Preventing a first episode of psychosis: meta-analysis of randomized controlled prevention trials of 12month and longer-term follow-ups. *Schizophrenia Research*. 2013;149(1-3):56-62. doi:10.1016/j.schres.2013.07.004
 102. Moreno-Peral P, Conejo-Cerón S, Rubio-Valera M, et al. Effectiveness of psychological and/or educational interventions in the prevention of anxiety: a systematic review, meta-analysis, and meta-regression. *JAMA Psychiatry*. 2017;74(10):1021-1029. doi:10.1001/jamapsychiatry.2017.2509
 103. Kessler RC, Angermeyer M, Anthony JC, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry*. 2007;6:168-176.
 104. Insel TR. Digital phenotyping: a global tool for psychiatry. *World Psychiatry*. 2018;17(3):276-277. doi:10.1002/wps.20550
 105. Rohani DA, Faurholt-Jepsen M, Kessing LV, Bardram JE. Correlations between objective behavioral features collected from mobile and wearable devices and depressive mood symptoms in patients with affective disorders: systematic review. *Journal of Medical Internet Research mHealth and uHealth*. 2018;6(8):1-21. doi:10.2196/mhealth.9691
 106. Lebowitz MS, Ahn W kyoung. Blue genes? Understanding and mitigating negative consequences of personalized information about genetic risk for depression. *Journal of Genetic Counseling*. 2018;27(1):204-216. doi:10.1007/s10897-017-0140-5
 107. Lebowitz MS, Ahn WK. Testing positive for a genetic predisposition to depression magnifies retrospective memory for depressive symptoms. *Journal of Consulting and Clinical Psychology*. 2017;85(11):1052-1063. doi:10.1037/ccp0000254
 108. Hewlett J, Waisbren SE. A review of the psychosocial effects of false-positive results on parents and current communication practices in newborn screening. *Journal of Inherited Metabolic Disease*. 2006;29(5):677-682. doi:10.1007/s10545-006-0381-1
 109. Lawrie SM, Fletcher-watson S, Whalley HC, McIntosh AM. Predicting major mental illness: ethical and practical considerations. *British Journal of Psychiatry*. 2019;5:1-5. doi:10.1192/bjo.2019.11
 110. Wang JL, Eccles H, Nannarone M, Schmitz N, Patten S, Lashewicz B. Does providing personalized depression risk information lead to increased psychological distress and functional impairment? Results from a mixed-methods randomized controlled trial. *Psychological Medicine*. Published online 2020:1-9. doi:10.1017/S0033291720003955
 111. Biggs R, Carpenter SR, Brock WA. Turning back from the brink: detecting an impending regime shift in time to avert it. *Proceedings of the National Academy of Sciences*. 2009;106(3):826-831. doi:10.1073/pnas.0811729106
 112. Rector JL, Gijzel SMW, van de Leemput IA, van Meulen FB, Olde Rikkert MGM, Melis RJF. Dynamical indicators of resilience from physiological time series in geriatric inpatients: lessons learned. *Experimental Gerontology*. 2021;149. doi:10.1016/j.exger.2021.111341
 113. Scheffer M, Carpenter SR, Dakos V, van Nes EH. Generic indicators of ecological resilience: inferring the chance of a critical transition. *Annual Review of Ecology, Evolution, and Systematics*. 2015;46:145-167. doi:10.1146/annurev-ecolsys-112414-054242
 114. Dai L, Vorselen D, Korolev KS, Gore J. Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science*. 2012;336(6085):1175-1177.

115. Drake JM, Griffen BD. Early warning signals of extinction in deteriorating environments. *Nature*. 2010;467(7314):456-459. doi:10.1038/nature09389
116. Hayes AM, Laurenceau JP, Feldman G, Strauss JL, Cardaciotto LA. Change is not always linear: the study of nonlinear and discontinuous patterns of change in psychotherapy. *Clinical Psychology Review*. 2007;27(6):715-723. doi:10.1016/j.cpr.2007.01.008
117. Tang TZ, DeRubeis RJ, Hollon SD, Amsterdam J, Shelton R. Sudden gains in cognitive therapy of depression and depression relapse/recurrence. *Journal of Consulting and Clinical Psychology*. 2007;75(3):404-408. doi:10.1037/0022-006X.75.3.404
118. Heinzl S, Tominschek I, Schiepek G. Dynamic pattern in psychotherapy- discontinuous changes and critical instabilities during the treatment of obsessive compulsive disorder. *Nonlinear Dynamics, Psychology, and Life Sciences*. 2014;18(2):155-176.
119. Helmich MA, Wichers M, Olthof M, et al. Sudden gains in day-to-day change: revealing nonlinear patterns of individual improvement in depression. *Journal of Consulting and Clinical Psychology*. 2020;88(2):119-127. doi:10.1037/ccp0000469
120. Boettiger C, Hastings A. Early warning signals and the prosecutor's fallacy. *Proceedings of the Royal Society B: Biological Sciences*. 2012;279(1748):4734-4739. doi:10.1098/rspb.2012.2085
121. Spears BM, Futter MN, Jeppesen E, et al. Ecological resilience in lakes and the conjunction fallacy. *Nature Ecology & Evolution*. 2017;1(11):1616-1624. doi:10.1038/s41559-017-0333-1
122. Nelson B, Hartmann JA. Prediction in mental health research and its limits (or why life can only be understood backwards but must be lived forwards). *Early Intervention in Psychiatry*. 2018;12(5):767-770. doi:10.1111/eip.12530
123. van der Maas HLJ, Molenaar PCM. Stagewise cognitive development: an application of catastrophe theory. *Psychological Review*. 1992;99(3):395-417. doi:10.1037/0033-295X.99.3.395
124. Ladyman J, Lambert J, Wiesner K. What is a complex system? *European Journal for Philosophy of Science*. 2013;3(1):33-67. doi:10.1007/s13194-012-0056-8
125. Kuppens P, Sheeber LB, Yap MBH, Whittle S, Simmons JG, Allen NB. Emotional inertia prospectively predicts the onset of depressive disorder in adolescence. *Emotion*. 2012;12(2):283-289. doi:10.1037/a0025046
126. Koval P, Pe ML, Meers K, Kuppens P. Affect dynamics in relation to depressive symptoms: variable, unstable or inert? *Emotion*. 2013;13(6):1132-1142. doi:10.1037/a0033579
127. Brose A, Schmiedek F, Koval P, Kuppens P. Emotional inertia contributes to depressive symptoms beyond perseverative thinking. *Cognition and Emotion*. 2015;29(3):527-538. doi:10.1080/02699931.2014.916252
128. Verduyn P, Lavrijsen S. Which emotions last longest and why: the role of event importance and rumination. *Motivation and Emotion*. 2014;39(1):119-127. doi:10.1007/s11031-014-9445-y
129. Kuiper RM, Ryan O. Drawing conclusions from cross-lagged relationships: re-considering the role of the time-interval. *Structural Equation Modeling*. 2018;25(5):809-823. doi:10.1080/10705511.2018.1431046
130. Dormann C, Griffin MA. Optimal time lags in panel studies. *Psychological Methods*. 2015;20(4):489-505. doi:10.1037/met0000041
131. Ryan O, Kuiper RM, Hamaker EL, Ryan O. A continuous time approach to intensive longitudinal data: what, why, and how? In: *Continuous Time Modeling in the Behavioral and Related Sciences*. Springer; 2018:27-54.
132. Loossens T, Tuerlinckx F, Verdonck S. A comparison of continuous and discrete time modeling of affective

- processes in terms of predictive accuracy. *Scientific Reports*. 2021;11(1). doi:10.1038/s41598-021-85320-4
133. Heller AS, Davidson RJ, Fox AS. Parsing affective dynamics to identify risk for mood and anxiety disorders. *Emotion*. 2019;19(2):283-292. doi:10.1037/emo0000440
 134. Ebner-Priemer UW, Houben M, Santangelo P, et al. Unraveling affective dysregulation in borderline personality disorder: a theoretical model and empirical evidence. *Journal of Abnormal Psychology*. 2015;124(1):186-198. doi:10.1037/abn0000021
 135. Wood J, Oravecz Z, Vogel N, et al. Modeling intraindividual dynamics using stochastic differential equations: age differences in affect regulation. *The journals of gerontology Series B, Psychological sciences and social sciences*. 2017;73(1):171-184. doi:10.1093/geronb/gbx013
 136. Westermann S, Grezellschak S, Oravecz Z, Moritz S, Lüdtke T, Jansen A. Untangling the complex relationships between symptoms of schizophrenia and emotion dynamics in daily life: findings from an experience sampling pilot study. *Psychiatry Research*. 2017;257:514-518. doi:10.1016/j.psychres.2017.08.023
 137. van Nes EH, Arani BMS, Staal A, et al. What do you mean, “tipping point”? *Trends in Ecology & Evolution*. 2016;31:622-632.
 138. Scheffer M, Bascompte J, Brock WA, et al. Early-warning signals for critical transitions. *Nature*. 2009;461(7260):53. doi:10.1038/nature08227
 139. Smit AC, Snippe E, Wichers M. Increasing restlessness signals impending increase in depressive symptoms more than 2 months before it happens in individual patients. *Psychotherapy and Psychosomatics*. 2019;88(4):249-251. doi:10.1159/000500594
 140. Suweis S, D’Odorico P. Early warning signs in social-ecological networks. *PLoS ONE*. 2014;9(7). doi:10.1371/journal.pone.0101851
 141. Held H, Kleinen T. Detection of climate system bifurcations by degenerate fingerprinting. *Geophysical Research Letters*. 2004;31(23):1-4. doi:10.1029/2004GL020972
 142. Chen S, O’Dea EB, Drake JM, Epureanu BI. Eigenvalues of the covariance matrix as early warning signals for critical transitions in ecological systems. *Scientific Reports*. 2019;9(1):1-14. doi:10.1038/s41598-019-38961-5
 143. Morozov A, Abbott K, Cuddington K, et al. Long transients in ecology: Theory and applications. *Physics of Life Reviews*. 2020;32:1-40. doi:10.1016/j.plrev.2019.09.004
 144. McGorry PD, Hartmann JA, Spooner R, Nelson B. Beyond the “at risk mental state” concept: transitioning to transdiagnostic psychiatry. *World Psychiatry*. 2018;17(June):133-142.
 145. Francis TB, Abbott KC, Cuddington K, et al. Management implications of long transients in ecological systems. *Nature Ecology and Evolution*. 2021;5(3):285-294. doi:10.1038/s41559-020-01365-0
 146. Coyne JC. Self-reported distress: analog or ersatz depression? *Psychological Bulletin*. 1994;116(1):29-45.
 147. Zeeman EC. Catastrophe theory. *Scientific American*. 1976;234(4):65-83.
 148. van der Maas HLJ, Kolstein R, van der Pligt J. Sudden transitions in attitudes. *Sociological Methods and Research*. 2003;32(2):125-152. doi:10.1177/0049124103253773
 149. Bulteel K, Tuerlinckx F, Brose A, Ceulemans E. Improved insight into and prediction of network dynamics by combining VAR and dimension reduction. *Multivariate Behavioral Research*. 2018;53(6):853-875. doi:10.1080/00273171.2018.1516540
 150. René Thom. *Stabilité Structurelle et Morphogenèse: Essai d'une Théorie Générale Des Modèles*. Benjamin; 1972.

151. Zahler RS, Sussmann HJ. Claims and accomplishments of applied catastrophe theory. *Nature*. 1977;269:759-763.
152. Stewart I. Elementary catastrophe theory. *IEEE Transactions on Circuits and Systems*. 1983;30(8):578-586. doi:10.1109/TCS.1983.1085392
153. Rosser JB. The rise and fall of catastrophe theory applications in economics: was the baby thrown out with the bathwater? *Journal of Economic Dynamics and Control*. 2007;31(10):3255-3280. doi:10.1016/j.jedc.2006.09.013
154. Hamaker EL. Why researchers should think “within-person”. A paradigmatic rationale. In: Mehl MR, Conner TS, eds. *Handbook of Research Methods for Studying Daily Life*. The Guilford Press; 2012:43-61.

Nederlandse samenvatting

Dankwoord/acknowledgements

Curriculum vitae

List of publications

SHARE publications

Nederlandse samenvatting

Achtergrond

Veel mensen krijgen ooit in hun leven te maken met een psychische stoornis, zoals depressie, verslaving, of ADHD¹. Voor 75% van hen begint die stoornis in de adolescentie en vroege volwassenheid². De mentale gezondheid van jongvolwassenen is dus kwetsbaar. Het is echter onmogelijk om op individueel niveau te voorspellen *of* en *wanneer* die kwetsbaarheid toeneemt. Dat maakt het lastig om op het juiste moment in te grijpen, en zo het beloop en de gevolgen van psychische problemen gunstig te beïnvloeden.

Hoe kunnen we er op tijd achter komen of en wanneer een specifiek persoon psychische klachten zal ontwikkelen? Dit proefschrift onderzoekt deze vraag door te kijken naar wat we kunnen leren van de kwetsbaarheid en weerbaarheid van andere systemen, zoals ecosystemen, samenlevingen, en het klimaat. Zulke systemen worden “complexe dynamische systemen” genoemd: ze bestaan uit veel verschillende elementen die door interacties een evenwicht creëren³. Een interessante eigenschap van dit soort systemen is dat ze plotselinge transities (*i.e.*, plotse veranderingen) kunnen vertonen⁴. Het systeem verandert dan van het ene evenwicht (bijv. een tropisch klimaat) naar het andere evenwicht (bijv. een ijstijd)^{4,5}. Dit soort transities komen als een donderslag bij heldere hemel: ze kunnen veroorzaakt worden door schijnbaar kleine, onschuldige gebeurtenissen, maar hebben grote, soms onomkeerbare, gevolgen. Dit heeft veel onderzoekers gemotiveerd om te zoeken naar manieren om kritische transities van tevoren te zien aankomen^{4,6–8}. Als dat zou kunnen, wordt het wellicht mogelijk om transities met nadelige gevolgen, zoals het uitsterven van soorten, het uitbreken van oorlogen, of plotse klimaatveranderingen, te voorkomen. Een doorbraak op dit gebied is het inzicht dat kritische transities vooraf worden gegaan door een toenemende kwetsbaarheid binnen het systeem^{9–11}. Dit wordt kritische vertraging genoemd: naarmate de kwetsbaarheid van het systeem toeneemt, herstelt het steeds trager van verstoringen¹¹. Het systeem heeft dan steeds meer moeite om het evenwicht te herstellen. Op een gegeven moment is de kwetsbaarheid zo hoog dat zelfs een kleine verstoring een transitie naar een ander evenwicht kan veroorzaken. Welk ander evenwicht er ontstaat hangt af van de richting van kritische vertraging, ofwel de “zwakke plek” van het systeem^{12,13}.

De aanwezigheid van kritische vertraging betekent dat de transities die kunnen voorkomen in complexe dynamische systemen beter voorspelbaar zijn dan ze in eerste instantie lijken: de donderslag vindt vaak plaats in een hemel die geleidelijk aan steeds

donkerder (kwetsbaarder) werd. Als we die toenemende kwetsbaarheid kunnen meten, kunnen we de transitie dus voorzien. Een uitdaging daarbij is dat het meten van kwetsbaarheid langdurige en frequente metingen (tijdsseries data) vereist. Op basis van dit soort data kunnen tekenen van kritische vertraging, ofwel vroege waarschuwingssignalen, worden gevonden¹⁴. Een belangrijk kenmerk van die signalen is dat ze in theorie voorafgaan aan alle kritische transities – ongeacht het systeem waarin die transities plaatsvinden⁴. Het is bijvoorbeeld al bekend dat vroege waarschuwingssignalen voorafgaan aan transities in ecosystemen^{15–17}, maar ook aan epileptische aanvallen¹⁸, het instorten van samenlevingen¹⁹, en de uitbraak van epidemieën^{20,21}. De cruciale vraag voor de psychiatrie is of vroege waarschuwingssignalen ook voorspellend zijn voor transities in mentale gezondheid. Dit soort transities kunnen bijvoorbeeld een plotse toename in psychische klachten betreffen: het systeem verandert dan van een gezonde staat, gekenmerkt door weinig psychische klachten, naar een ongezonde staat, die gekenmerkt wordt door meer psychische klachten. Als vroege waarschuwingssignalen inderdaad voorafgaan aan transities in mentale gezondheid, zou het mogelijk zijn om kwetsbaarheid voor toekomstige klachten binnen personen te monitoren.

Vroege waarschuwingssignalen in de psychiatrie: de belofte en de uitdaging

Als we in staat zouden zijn om met vroege waarschuwingssignalen kwetsbaarheid voor toekomstige psychische klachten te monitoren heeft dat veelbelovende implicaties: we zouden weten wie er wanneer kwetsbaar is voor psychische stoornissen, zodat we die stoornissen mogelijk kunnen voorkomen. Wellicht is het ook mogelijk om met vroege waarschuwingssignalen te voorspellen voor welke toekomstige klachten mensen het meest kwetsbaar zijn. Dit is met name relevant voor jongvolwassenen, bij wie het risico op psychische klachten vaak nog “pluripotent” is^{22,23}. Dat wil zeggen dat de kwetsbaarheid kan leiden tot een breed scala aan problemen, variërend van stemmings- en angstklachten tot psychoses en gedragsproblemen. Om te onderzoeken of vroege waarschuwingssignalen inderdaad nuttig zijn voor het vroegtijdig detecteren van kwetsbaarheid voor een specifiek type klachten is het nodig dat jongvolwassenen langdurig hun stemming bijhouden, in een periode waarin een transitie in mentale gezondheid waarschijnlijk is. Bij aanvang van mijn promotieonderzoek was nog onduidelijk of het verzamelen van dit soort data haalbaar was. Het zou bijvoorbeeld kunnen dat langdurige dagelijkse metingen simpelweg te belastend zijn. Daarnaast is het mogelijk dat alleen relatief gezonde, of juist relatief getroebleerde, jongvolwassenen bereid zijn om langdurig hun stemming bij te houden. Een deel van dit

proefschrift is daarom gewijd aan de vraag of en voor wie we mentale gezondheid op deze manier kunnen onderzoeken.

Overzicht van dit proefschrift

In dit proefschrift is onderzocht in hoeverre principes die gelden in complexe dynamische systemen de ontwikkeling van psychopathologie kunnen voorspellen en verklaren. In de **hoofdstukken 2 en 3** hebben we gekeken naar verschillen *tussen* personen in het beloop van klachten. Kunnen we voorspellen wie er later psychische klachten ontwikkelt, en wie niet? Vervolgens wordt in **hoofdstuk 4** beschreven hoe de ontwikkeling van psychopathologie *binnen* personen onderzocht zou kunnen worden, en waarom dat interessant en nuttig is. In dit hoofdstuk worden verschillende hypothesen beschreven die volgen uit het idee dat mentale gezondheid beschouwd kan worden als complex dynamisch systeem. In **hoofdstukken 5 en 6** wordt de studie beschreven die we hebben uitgevoerd om deze hypothesen te toetsen. In deze studie hebben 134 jongvolwassenen met een verhoogd risico voor psychische problemen gedurende zes maanden dagelijks hun stemming bijgehouden. Op basis van de resulterende data hebben we onderzocht (i) hoe stabiele evenwichten in mentale gezondheid eruitzien (**hoofdstuk 7**), (ii) of vroege waarschuwingssignalen transities in mentale gezondheid kunnen voorspellen (**hoofdstuk 8**) en (iii) of vroege waarschuwingssignalen kunnen voorspellen in welke richting symptomen van psychopathologie zich zullen ontwikkelen (**hoofdstuk 10**). Daarnaast hebben we vroege waarschuwingssignalen ook in een andere populatie onderzocht, namelijk mensen met een bipolaire stoornis die transities tussen een manische en depressieve toestand rapporteerden (**hoofdstuk 9**). De belangrijkste bevindingen van het proefschrift zijn hieronder samengevat.

Individuele verschillen in kwetsbaarheid

In **hoofdstukken 2 en 3** hebben we onderzocht of we op basis van herhaalde metingen van stemming toekomstige mentale problemen kunnen voorspellen. Beide hoofdstukken maken gebruik van een studie waarin jongeren (N=401, ±18 jaar oud) uit de algemene populatie tien keer per dag, zes dagen achter elkaar, vragen over hun stemmingen hebben beantwoord (bijv. “Hoe somber voel je je nu, op een schaal van 1 tot 7?”). Een jaar later zijn symptomen van psychopathologie gemeten. In **hoofdstuk 2** vonden we dat jongeren met een genuanceerde beleving van emoties (bijv. “Ik voel me teleurgesteld”, “Ik voel me geïrriteerd”, “Ik voel me beschaamd”, enz.) een gunstiger beloop van symptomen hadden vergeleken met jongeren die

emoties minder goed leken te onderscheiden (bijv. “Ik voel me slecht”). Vanuit psychologisch perspectief kan dit verklaard worden doordat vage emoties lastig te reguleren zijn: als je niet precies weet hoe je je voelt, is het lastiger om adequaat op je emoties te reageren²⁴. Vanuit het perspectief van complexe dynamische systemen kan onze bevinding verklaard worden op basis van het gegeven dat systemen met een hoge diversiteit beter beschermd zijn tegen instorten⁴. Een belangrijke kanttekening bij onze bevinding is dat toekomstige mentale problemen ook voorspeld konden worden met een relatief simpele maat, namelijk de gemiddelde mate waarin jongeren überhaupt negatieve emoties rapporteerden. Anders gezegd: voor het voorspellen van toekomstige mentale problemen volstaat de gemiddelde hoeveelheid negatieve emoties, en is het dus niet nodig om de complexiteit van emoties te berekenen. Dit sluit aan bij eerder onderzoek²⁵. Samenvattend kan gesteld worden dat emotionele complexiteit wellicht inzicht geeft in wat mensen kwetsbaar maakt voor psychische klachten, maar niet noodzakelijk is voor het voorspellen van die klachten.

In **hoofdstuk 3** hebben we op een andere manier gekeken naar de samenhang tussen emoties en toekomstige psychische klachten. Specifiek hebben we onderzocht of emotionele inertie de toename in psychische klachten voorspelt. Emotionele inertie beschrijft de instandhouding van emoties, ofwel de mate waarin het ervaren van een emotie *nu* voorspellend is voor het ervaren van diezelfde emotie op *een later moment*^{26,27}. Lage emotionele inertie suggereert dat emoties snel verdwijnen. Het mentale evenwicht herstelt zich dan snel, wat betekent dat het systeem stabiel is en dus niet kwetsbaar voor een plotse transitie. Hoge emotionele inertie wijst er juist op dat emoties lang blijven hangen, wat duidt op een instabiel systeem dat kwetsbaar is voor een plotse transitie. In **hoofdstuk 3** hebben we onderzocht of jongeren met een hogere inertie ook een grotere toename in psychische klachten rapporteerden, waarmee we tevens een belangrijk onderzoek naar complexe dynamische systeem-principes in de psychiatrie hebben gerepliceerd²⁸. We vonden geen verband tussen inertie en de toekomstige toename in klachten. Wél vonden we dat inertie gerelateerd is aan *huidige* psychische klachten: hoe meer emoties blijven hangen, des te meer mentale problemen iemand rapporteert. Deze bevindingen deden we ook toen we de data uit het eerdere onderzoek opnieuw analyseerden met de methode die we gebruikten in ons eigen onderzoek²⁸.

Samengenomen laten **hoofdstuk 2 en 3** zien dat psychopathologie samengaat met minder complexe emoties – in die zin dat verschillende negatieve emoties beleefd worden als slechts één dimensie (bijv. “helemaal niet slecht” – “heel slecht”) – en met meer rigide

emoties – in die zin dat negatieve emoties relatief lang blijven hangen. Deze karakteristieken zijn echter niet noodzakelijk (emotionele complexiteit) of bruikbaar (emotionele inertie) voor het voorspellen van het beloop van psychische klachten over de periode van een jaar. Het nut van emotionele complexiteit en emotionele inertie voor het begrijpen van psychopathologie ligt dus niet zozeer in hun voorspellende waarde, maar meer in hun bijdrage aan ons inzicht in emotionele processen die samengaan met psychische klachten.

Mentale gezondheid als complex dynamisch systeem: waarom en hoe

Hoofdstuk 4 beschrijft dat mentale gezondheid beschouwd kan worden als een complex dynamisch systeem. Dit perspectief op mentale gezondheid zou betekenen dat (i) mentale gezondheid gekenmerkt wordt door alternatieve evenwichten (bijv. “gezond” en “ziek”), (ii) plotse overgangen tussen deze evenwichten voorafgegaan worden door vroege waarschuwingssignalen en (iii) het type psychische klachten waarvoor iemand kwetsbaar is afgeleid kan worden uit vroege waarschuwingssignalen. Om deze hypothesen te onderzoeken is het nodig dat mensen langdurig hun stemming bijhouden in een periode waarin ze kwetsbaar zijn voor een plotse transitie in hun mentale gezondheid. Wij hebben daarom de TRAILS TRANS-ID studie opgezet, die beschreven wordt in **hoofdstuk 5**. Voor deze studie hebben we 134 jongvolwassenen (± 23 jaar oud) gevraagd om gedurende zes maanden dagelijks een vragenlijst over hun stemming in te vullen via hun smartphone (dit wordt ook wel een dagboekstudie genoemd). De deelnemers maakten deel uit van het klinische cohort van een lopende cohortstudie (TRAILS²⁹). Ze waren in hun jeugd in aanraking gekomen met de geestelijke gezondheidszorg, en daarom gemiddeld kwetsbaarder voor psychische problemen dan leeftijdsgenoten uit de algemene bevolking. Direct voor, na, en een jaar na de dagboekstudie hebben we middels een semi-gestandaardiseerd interview gecheckt of de deelnemers voldeden aan de diagnostische criteria voor een psychische stoornis. Daaruit bleek dat, bij ieder interview, ongeveer 30% van onze deelnemers klinische problemen ervoer. Stemmingsstoornissen waren daarbij het meest voorkomend (19-28%), gevolgd door angststoornissen (10-18%) en ADHD (3-7%).

Een belangrijke bevinding uit **hoofdstuk 5** is dat de data die verzameld zijn in de TRAILS TRANS-ID studie voldoen aan een aantal voorwaarden die essentieel zijn voor analyse van de data. Het bleek bijvoorbeeld dat de dagelijkse metingen een consistente betekenis hadden over de tijd, wat betekent dat het mogelijk is om patronen binnen personen te onderzoeken. Daarnaast bleek dat een intensieve dataverzameling zoals de onze – waarbij

elke deelnemer 183 vragenlijsten invult – haalbaar is. Dat volgde onder andere uit het feit dat de deelnemers gemiddeld 89% van de vragenlijsten invulden. Dat betekent niet noodzakelijk dat een dergelijke dataverzameling geschikt is voor *elke* populatie. Het is bijvoorbeeld mogelijk dat de deelnemers van TRAILS TRANS-ID bijzonder gemotiveerd waren om mee te doen aan onderzoek. Misschien waren onze deelnemers relatief gezond ten opzichte van de populatie die we onderzochten, of juist relatief ongezond. Meestal kunnen dat soort mogelijkheden niet onderzocht worden omdat er geen gegevens zijn van de mensen die *niet* meedoen aan onderzoek. Omdat onze deelnemers binnen een bestaande cohortstudie waren geworven konden wij echter wél onderzoeken of onze deelnemers representatief waren voor de bredere populatie (*i.e.*, kwetsbare jongvolwassenen). In **hoofdstuk 6** hebben we daarom gekeken of onze deelnemers verschilden van (i) jongvolwassenen uit het klinische cohort van TRAILS die zich *niet* hadden aangemeld voor onze studie en (ii) jongvolwassenen uit het algemene populatie cohort van TRAILS. We vonden nauwelijks verschillen tussen onze deelnemers en de mensen die niet hadden meegedaan, wat betekent dat onze intensieve dataverzameling niet voorbehouden was aan de “(un)happy few”. Daarnaast vonden we dat onze deelnemers in bijna alle opzichten verschilden van de algemene populatie: ze hadden meer psychologische, lichamelijke, economische, en sociale problemen. Onze deelnemers waren dus duidelijk kwetsbaar, maar dat belette hen niet om succesvol deel te nemen aan een intensieve, langdurige studie.

Stabiliteit en instabiliteit van mentale gezondheid binnen individuen

Met de data van de TRAILS TRANS-ID studie konden we enkele hypothesen onderzoeken die volgen uit het idee dat mentale gezondheid zich zou kunnen gedragen als een complex dynamisch systeem. Allereerst hebben we gekeken naar alternatieve evenwichten (stabiele staten) in mentale gezondheid. Meestal wordt er uitgegaan van twee verschillende evenwichten, namelijk “gezond” en “ziek”. Daarmee wordt geïmpliceerd dat de toestand daartussenin – waarin mensen wel problemen ervaren, maar van onvoldoende ernst of duur voor een klinische diagnose – tijdelijk en instabiel is. In **hoofdstuk 7** vonden we dat die implicatie niet klopt: de stabiliteit van psychische klachten hangt niet samen met de ernst van die klachten. Dat betekent dat de tussenfase net zo (in)stabiel is als de extremen (“gezond” en “ziek”). Mensen kunnen dus – ongewenst – blijven hangen in een evenwicht waarin ze milde psychische klachten ervaren.

De ontwikkeling (of remissie) van psychopathologie kan gezien worden als een

transitie van het ene evenwicht (bijv. “gezond”) naar het andere evenwicht (bijv. “ziek”). In **hoofdstuk 8** hebben we onderzocht of zo’n transitie vooraf wordt gegaan door vroege waarschuwingssignalen. In de 20 mensen die een plotselinge dip in hun mentale gezondheid hadden, vonden we slechts sporadisch waarschuwingssignalen. De sensitiviteit van vroege waarschuwingssignalen was dus laag: slechts 5% van de mogelijke signalen voorspelde de dip in mentale gezondheid. Bovendien vonden we ook “valse alarmen” bij mensen *zonder* dip in hun mentale gezondheid (specificiteit: 96%). Het is mogelijk dat onze bevindingen het gevolg zijn van de groep mensen die we hebben onderzocht (kwetsbare jongvolwassenen), de tijdsschaal van onze metingen (dagelijkse metingen gedurende zes maanden), of de statistische methode die we hebben gebruikt. Daarom hebben we in **hoofdstuk 9** een vergelijkbare studie gedaan, maar dan in 20 mensen met een bipolaire stoornis die transities tussen een bipolaire en manische toestand rapporteerden. Deze mensen hadden gedurende vier maanden lang vijf keer per dag hun stemming bijgehouden (± 491 metingen per persoon). Ondanks het verschil in de doelgroep, tijdsschaal, en statische methode die we in deze studie hanteerden, waren de resultaten in grote lijnen vergelijkbaar met onze bevindingen in **hoofdstuk 8**. We vonden iets vaker waarschuwingssignalen in mensen met een transitie (sensitiviteit: 27% en 33% voor transities richting depressie en manie), maar ook iets vaker fout positieve waarschuwingssignalen (specificiteit: 75%). Bipolaire transities waren dus iets makkelijker te herkennen door vroege waarschuwingssignalen, maar leverden ook vaker een vals alarm. Samengenomen tonen **hoofdstukken 8 en 9** aan dat het klinisch nut van vroege waarschuwingssignalen op dit moment nog onbewezen is. Dat kan betekenen dat vroege waarschuwingssignalen voor transities in mentale gezondheid simpelweg niet bestaan – wat impliceert dat de signalen die wij vonden toevalsbevindingen waren. Een andere mogelijkheid is dat vroege waarschuwingssignalen alleen van toepassing zijn voor een specifieke doelgroep, die bijvoorbeeld gekenmerkt wordt door specifieke (transities in) psychische klachten of door bepaalde kenmerken in de data die het makkelijker maken om waarschuwingssignalen te detecteren. Er is vervolgonderzoek nodig om dat soort vragen te kunnen beantwoorden.

In **hoofdstuk 10** hebben we tot slot onderzocht of kritische vertraging – het proces dat ten grondslag ligt aan vroege waarschuwingssignalen – het beloop van psychische klachten voorspelt. Op basis van studies in ecosystemen verwachtten we dat we de *richting* van kritische vertraging zouden kunnen afleiden, ofwel de zwakke plek van een systeem^{12,13}. We hebben dit idee eerst onderzocht in een studie waarin kritische vertraging reeds was

aangetoond^{30,31}. Dit betrof een casus studie van een man die bijna acht maanden lang tien keer per dag zijn stemming bijhield, en in die periode een terugval in depressie meemaakte. We vonden dat de richting van kritische vertraging – namelijk, een voorspelde toename van negatieve emoties en een afname van positieve emoties – overeenkwam met de richting van de transitie in zijn mentale gezondheid. Vervolgens hebben we hetzelfde principe onderzocht in de TRAILS TRANS-ID studie. We vonden dat de richting van kritische vertraging voor 20-25% van de mensen overeenkwam met de richting waarin hun klachten zich ontwikkelden. In combinatie met de eerste studie wijst dat erop dat kritische vertraging niet voor iedere verandering in mentale gezondheid van toepassing is. De uitdaging voor vervolgonderzoek is het beschrijven en herkennen van die transities waarvoor kritische vertraging – en andere principes uit complexe dynamische systemen – wél nuttig zijn.

Conclusies en vervolgonderzoek

De emoties van jongeren verschillen in complexiteit en inertie, en dat soort verschillen hangen samen met hun psychische klachten. Met herhaalde, dagelijkse vragenlijsten kan onderzocht worden hoe emoties *binnen* personen veranderen. Zulk onderzoek is relatief intensief, maar kan desondanks succesvol worden uitgevoerd bij jongvolwassenen met een verhoogd risico op psychische klachten. De resulterende data kunnen mogelijk gebruikt worden om op individueel niveau te voorspellen of en wanneer iemand toenemende psychische problemen zal ervaren. Het idee dat mentale gezondheid zich gedraagt als een complex dynamisch systeem biedt daarvoor een theoretisch kader. Volgens dit idee worden plotse veranderingen in mentale gezondheid voorafgegaan door kritische vertraging – en dus door vroege waarschuwingssignalen. In dit proefschrift is onderzocht in hoeverre vroege waarschuwingssignalen nuttig zijn voor het voorzien van toekomstige veranderingen in mentale gezondheid. Het bleek dat kritische vertraging en daaruit voortvloeiende vroege waarschuwingssignalen minder generiek zijn dan vaak wordt gedacht. In tegenstelling: kritische vertraging komt alléén voor in systemen die een specifieke soort verandering tussen twee stabiele staten doormaken en die tevens op de juiste manier worden gemeten (*i.e.*, met de juiste frequentie en in de juiste variabelen)^{32,33}. Om een definitief oordeel te kunnen vellen over het nut van vroege waarschuwingssignalen in de psychiatrie, is het dus van belang om te weten in hoeverre mentale gezondheid voldoet aan de aannamen die aan dat soort signalen ten grondslag liggen. Een eerste concrete stap in die richting zou kunnen zijn om de tijdsschaal waarop psychische processen plaatsvinden te onderzoeken. Daarnaast is

het belangrijk om stabiele staten en transitie daartussen beter te begrijpen. Als we mentale gezondheid beter kunnen beschrijven als een systeem, zal het ook mogelijk worden om beter in te schatten in hoeverre principes uit andere systemen te vertalen zijn naar mentale gezondheid. Uiteindelijk zou dat kunnen bijdragen aan ons vermogen om psychopathologie op een gepersonaliseerde manier te onderzoeken en te behandelen.

Referenties

1. Caspi A, Houts RM, Ambler A, et al. Longitudinal assessment of mental health disorders and comorbidities across 4 decades among participants in the Dunedin birth cohort study. 2020;3(4). doi:10.1001/jamanetworkopen.2020.3221
2. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry*. 2005;62:593-602. doi:10.1001/archpsyc.62.6.593
3. Richardson MJ, Dale R, Marsh KL. Complex dynamical systems in social and personality psychology. In: *Handbook of Research Methods in Social and Personality Psychology*. ; 2014:253-282. doi:10.1017/cbo9780511996481.015
4. Scheffer M. *Critical Transitions in Nature and Society*. Princeton University Press; 2009.
5. Dakos V, Scheffer M, van Nes EH, Brovkin V, Petoukhov V, Held H. Slowing down as an early warning signal for abrupt climate change. *PNAS*. 2008;105(38):14308-14312. doi:10.1073/pnas.0802430105
6. Litzow MA, Hunsicker ME. Early warning signals, nonlinearity, and signs of hysteresis in real ecosystems. *Ecosphere*. 2016;7(12). doi:10.1002/ecs2.1614
7. Burthe SJ, Henrys PA, Mackay EB, et al. Do early warning indicators consistently predict nonlinear change in long-term ecological data? *Journal of Applied Ecology*. 2016;53(3):666-676. doi:10.1111/1365-2664.12519
8. Olde Rikkert MGM, Dakos V, Buchman TG, et al. Slowing down of recovery as generic risk marker for acute severity transitions in chronic diseases. *Critical Care Medicine*. 2016;44(3):601-606. doi:10.1097/CCM.0000000000001564
9. Scheffer M, Carpenter SR, Lenton TM, et al. Anticipating critical transitions. *Science*. 2012;338(6105):344-348. doi:10.1126/science.1225244
10. Scheffer M, Bascompte J, Brock WA, et al. Early-warning signals for critical transitions. *Nature*. 2009;461(7260):53. doi:10.1038/nature08227
11. Wissel C. A universal law of the characteristic return time near thresholds. *Oecologia*. 1984;65:101-107.
12. Lever JJ, van de Leemput IA, Weinans E, et al. Foreseeing the future of mutualistic communities beyond collapse. *Ecology Letters*. 2020;23(1):2-15. doi:10.1111/ele.13401
13. Weinans E, Lever JJ, Bathiany S, et al. Finding the direction of lowest resilience in multivariate complex systems. *Journal of the Royal Society Interface*. 2019;16(159):1-9. doi:10.1098/rsif.2019.0629
14. Dakos V, Carpenter SR, Brock WA, et al. Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLoS one*. 2012;7(7):1-20. doi:10.1371/journal.pone.0041010
15. Veraart AJ, Faassen EJ, Dakos V, van Nes EH, Lürling M, Scheffer M. Recovery rates reflect distance to a tipping point in a living system. *Nature*. 2012;481(7381):357-359. doi:10.1038/nature10723
16. Drake JM, Griffen BD. Early warning signals of extinction in deteriorating environments. *Nature*. 2010;467(7314):456-459. doi:10.1038/nature09389
17. Dai L, Vorselen D, Korolev KS, Gore J. Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science*. 2012;336(6085):1175-1177.
18. Maturana MI, Meisel C, Dell K, et al. Critical slowing down as a biomarker for seizure susceptibility. *Nature*

-
- Communications*. 2020;11(1):1-12. doi:10.1038/s41467-020-15908-3
19. Scheffer M, van Nes EH, Bird D, Bocinsky RK, Kohler TA. Loss of resilience preceded transformations of pre-Hispanic Pueblo societies. 2021;118(18):2021. doi:10.1073/pnas.2024397118/-/DCSupplemental
 20. Harris MJ, Hay SI, Drake JM. Early warning signals of malaria resurgence in Kericho, Kenya. *Biology Letters*. 2020;16(3). doi:10.1098/rsbl.2019.0713
 21. Kaur T, Sarkar S, Chowdhury S, Sinha SK, Jolly MK, Dutta PS. Anticipating the novel coronavirus disease (COVID-19) pandemic. *Frontiers in Public Health*. 2020;8(521):1-13. doi:10.1101/2020.04.08.20057430
 22. McGorry PD, Hartmann JA, Spooner R, Nelson B. Beyond the “at risk mental state” concept: transitioning to transdiagnostic psychiatry. *World Psychiatry*. 2018;17:133-142.
 23. van Os J. The dynamics of subthreshold psychopathology: implications for diagnosis and treatment. *American Journal of Psychiatry*. 2013;170:695-698.
 24. Barrett LF, Gross J, Christensen TC, Benvenuto M. Knowing what you’re feeling and knowing what to do about it: mapping the relation between emotion differentiation and emotion regulation. *Cognition and Emotion*. 2001;15(6):713-724. doi:10.1080/02699930143000239
 25. Dejonckheere E, Mestdagh M, Houben M, et al. Complex affect dynamics add limited information to the prediction of psychological well-being. *Nature Human Behaviour*. 2019;3(5):478-491. doi:10.1038/s41562-019-0555-0
 26. Suls J, Green P, Hillis S. Emotional reactivity to everyday problems, affective inertia, and neuroticism. *Personality and Social Psychology Bulletin*. 1998;24(2):127-136.
 27. Kuppens P, Allen NB, Sheeber LB. Emotional inertia and psychological maladjustment. *Psychological Science*. 2010;21(7):984-991. doi:10.1177/0956797610372634
 28. van de Leemput IA, Wichers M, Cramer AOJ, et al. Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences*. 2014;111(1):87-92. doi:10.1073/pnas.1312114110
 29. Huisman M, Oldehinkel AJ, de Winter A, et al. Cohort profile: the Dutch “TRacking Adolescents” Individual Lives’ Survey”; TRAILS. *International Journal of Epidemiology*. 2008;37(6):1227-1235. doi:10.1093/ije/dym273
 30. Kossakowski JJ, Groot PC, Haslbeck JMB, Borsboom D, Wichers M. Data from ‘Critical slowing down as a personalized early warning signal for depression.’ *Journal of Open Psychology Data*. 2017;5. doi:10.5334/jopd.29
 31. Wichers M, Groot PC, Psychosystems ESM group, EWS group. Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*. 2016;85:114-116. doi:10.1159/000441458
 32. Boettiger C, Hastings A. Quantifying limits to detection of early warning for critical transitions. *Journal of The Royal Society Interface*. 2012;9(75):2527-2539. doi:10.1098/rsif.2012.0125
 33. Boettiger C, Ross N, Hastings A. Early warning signals: the charted and uncharted territories. *Theoretical Ecology*. 2013;6(3):255-264. doi:10.1007/s12080-013-0192-6

Dankwoord/acknowledgements

Van alle dingen die leuk zijn aan een promotieonderzoek, is de kans om samen te werken met allerlei ontzettend slimme en gedreven mensen het allerleukst. Een aantal van hen wil ik hier graag bedanken: zonder hen was dit proefschrift er niet geweest. Daarnaast zijn er veel mensen die indirect hebben bijgedragen aan de totstandkoming van dit proefschrift, vanwege het simpele feit dat ze me gelukkig maken. Ook hen wil ik graag bedanken.

Allereerst **Marieke**, dankjewel voor je vertrouwen in mij en voor je optimisme wanneer ik dacht dat mijn resultaten niet zoveel voorstelden. Wat is het fijn om zo'n scherpe, enthousiaste en nuchtere promotor te hebben. Ik zal niet vergeten dat ik bij jou altijd kon binnenlopen met de meest gedetailleerde vragen (die ene P-waarde, in dat ene paper waar we het twee weken geleden over gehad hadden), en dat jij dan altijd direct wist waar het over ging én me een nuttig advies kon geven. Dankjewel voor kansen die je me hebt gegeven. Ik kijk ernaar uit om met je terug te kijken en je te laten weten wat er geworden is van de student die je ooit onder je hoede nam.

Catharina, jij wist ondanks een volle agenda letterlijk op ieder moment van de week tijd voor me te maken: ontzettend bedankt voor zoveel toewijding. Jouw zeer scherp afgestelde radar voor vaag, abstract, of pompeus taalgebruik heeft mijn werk ontzettend verbeterd. Ik heb daarnaast veel geleerd van jouw klinische, methodologische, en statistische expertise – bijvoorbeeld met betrekking tot de problematiek die we zagen bij de deelnemers aan ons onderzoek. Tot slot heb ik geprofiteerd van wat ik níet van jou kon leren, en wat ik dus zelf mocht uitzoeken en uitleggen. Dankjewel dus, voor het feit dat je me hebt vrijgelaten en voor het vertrouwen dat daarvoor nodig is. Daarnaast ben ik je ontzettend dankbaar voor je hulp en advies rond het zetten van de volgende stap in mijn werkzame leven. Ik kijk ernaar uit om in de toekomst te blijven samenwerken!

Hanneke, jij laat zien dat je zowel bescheiden als getalenteerd kunt zijn, en in die zin ben je een voorbeeld. Op kantoor, met koffie in de stad, of tijdens een wandeling door de Drentse natuur: jij was altijd bereid mijn (soms onzekere) getwijfel aan te horen en me bemoedigend advies te geven. Van jou heb ik ook geleerd dat de wetenschap niet ten koste hoeft te gaan aan andere belangrijke dingen in het leven. Samen met Marieke en Catharina heb je altijd benadrukt dat nieuwsgierigheid en enthousiasme de beste drijfveren zijn voor onderzoek. Dat is de beste les die je me kon meegeven: dankjewel.

De leden van de beoordelingscommissie, **prof. Peter Kuppens**, **prof. Anna Lichtwarck-Aschoff**, en **prof. Ulrich Ebner-Priemer**, wil ik graag bedanken voor het lezen en beoordelen van dit proefschrift.

De **onderzoekers van het ICPE** creëerden een ambitieuze en leerzame werkomgeving, waar tegelijkertijd ruimte was voor borrels, online koffiemomenten, en een schrijfweek in de Franse Alpen (heel veel dank, **Hans**). **Lino**, **Anne Margit**, and **Yugyun**: it has been a pleasurable and unburdening experience to share my thoughts with you. Bij het **TRANS-ID team** heb ik me altijd op mijn plek gevoeld. **Harriëtte**, bedankt voor je daadkrachtige en oplossingsgerichte optreden toen dat nodig was. **Evelien**, jij bood altijd een luisterend oor en een open blik. Tijdens meetings herinnerde jij ons er vaak aan dat de beste aanpak afhangt van de onderzoeksvraag – ik hoop dat ik dat advies altijd onthoud. **Laura**, dankzij jou ben ik VAR modellen en GAMs veel beter gaan begrijpen. Het is een zegen om iemand in het team te hebben die ingewikkelde dingen zo helder uit kan leggen. **Sandip**, thank you so much for your intelligent and cheerful contribution to TRANS-ID. It has been a great pleasure to collaborate with you. **Larisa**, it's exciting to know that the data we collected have reached scholars outside of Groningen. I admire your dedication and optimism, despite the pandemic times in which you arrived.

Het **TRAILS consortium** was onmisbaar voor de TRAILS TRANS-ID studie. **Aukelien** en **Marjan**, dankzij jullie ervaring met dataverzameling en projectmanagement hadden Robin en ik een vliegende start bij het opzetten van de TRAILS TRANS-ID studie. Ook **Margo** en **Esther** hebben achter de schermen enorm geholpen, waarvoor veel dank. Het **Roqua team** ben ik dankbaar voor hun essentiële technische bijdrage aan onze studie. **Bianca**, bedankt voor je hulp bij de dataverzameling. Uiteraard wil ik ook de **deelnemers aan de TRAILS TRANS-ID studie** bedanken. Zes maanden lang dagelijks je stemming bijhouden is niet niks: dank voor die enorme inspanning. Weet dat ik – en vele onderzoekers met mij – versteld sta van de data die jullie hebben geproduceerd.

Mijn proefschrift is het resultaat van vele samenwerkingen. Ten eerste de **co-auteurs van TwinssCan**, veel dank voor het beschikbaar maken van de TwinssCan data en voor jullie feedback op hoofdstukken 2 en 3. **Rutger**, bedankt voor je input bij hoofdstuk 4. **Arnout**, dankzij jouw statistische talent kon ik transities zoeken die met het blote oog moeilijk te detecteren waren. **Robin**, jij hebt ontzettend veel overstijgende input gegeven – in die zin dat je vaak nadacht over het grotere verhaal – heel veel dank daarvoor. **Fionneke**,

Richard, Lian, Evelien, Bennard, Benno, en **Sandip**, jullie gaven me de kans om vroege waarschuwingssignalen bij mensen met een bipolaire stoornis te onderzoeken in hoofdstuk 9. Dat betekende tevens een fijne samenwerking met **Fionneke** en **Sandip**. Thank you for your enthusiasm and inspiring input: it was a pleasure to work in a team with such complementary skills. **Els**, jij sprak over complexe dynamische systemen in een voor mij nieuwe taal – daar heb ik veel van geleerd. Bedankt voor je bijdrage aan hoofdstuk 10.

YAMFAR(M), (M)AMAFRY, of welk ander anagram dan ook: wat geweldig om samen met jullie in hetzelfde schuitje te zitten. Jullie hebben mij niet alleen professioneel, maar zeker ook persoonlijk veel gebracht: dankzij jullie weet ik wat fermenteren is (**YAM**), waarom en hoe je over je salaris moet onderhandelen (**FAR**), en hoe hoog mijn ape/sloth/T-rex index precies is. **Yoram**, jij blijft me verrassen met verborgen talenten zoals vuurtjes stoken, cocktails maken, en DJ-vaardigheden. **Arnout**, jij lijkt altijd bezig met data verzamelen, of het nu gaat om de lekkerste espresso, de winnende strategie in een bordspelletje, of om je eigen stemming. Het is ontzettend fijn dat jij dingen zo goed vanuit een ander perspectief kan bekijken: dat heeft vaak waardevolle adviezen of stof tot nadenken opgeleverd. **Marmar**, wat lief dat jij je naam hebt opgeofferd zodat ik de mijne kon behouden. Wat ben jij een creatieve geest en wat kun je goed tegenwicht bieden aan de bravoure van je oude kamergenoten. Ik kijk er nu al naar uit om je in Noorwegen op te zoeken. **Fionneke**, bij jou kan ik altijd advies inwinnen, of het nu gaat om een abstract of een kledingkeuze. En wat kun je goed mensen met elkaar verbinden: de wetenschap verdient meer mensen zoals jij! **Anouk**, mijn eerste kennismaking met jou bleek veelzeggend, want je vroeg me of ik het leuk vind om hard te lopen. Jazeker. Wat was het gezellig om met jou, Robin, Peter, Bertus, en Elske op donderdagmiddag door het Noorderplantsoen te rennen. Jouw aanstekelijke enthousiasme en tomeloze energie komen vast goed van pas bij de neurochirurgie. **Robin**, ik had me geen betere partner in crime kunnen wensen. Van jouw talent om uitdagingen of problemen te voorzien, en ze bovendien niet uit de weg te gaan, heb ik enorm geprofiteerd en hopelijk ook geleerd. Je durft dingen te bevragen die de meeste mensen omzeilen – zowel professioneel als persoonlijk – en dat bewonder ik zeer. Dat jij als paranymf aan mijn zijde staat geeft vertrouwen in de verdediging. **Maurits**, wat is het ontzettend leuk om met jou door Groningen en Drenthe te racen en intussen meer te leren over het échte wielrennen (en hoe het werkt met de bonificatiepunten) en over open science (en waarom we artikelen niet meer bij bepaalde uitgevers in moeten dienen). Ik ben er blij mee dat je ondanks mijn niet-geïsoïneerde vertoning toch naast me durft te rijden! Dankjewel voor al je geduld, je reepjes onderweg, en voor het vangen van de wind.

Christa, Karin, Lisanne, Imke, Stephanie, en Mara, wat is het waardevol om vriendinnen te hebben die ik al zo lang ken. Jullie gezelschap voelt ontzettend vertrouwd, mede dankzij alle gebeurtenissen die we samen hebben meegemaakt en de willekeurige dingen die door de jaren heen heel waardevol zijn geworden en onlosmakelijk met jullie verbonden zijn (ik denk aan: Speuld, chocoswing, Castle on the Hill, Avalon, dansjes, en nog heel veel meer). **Charlotte**, ik bewonder jouw vermogen om een abstract verhaal te vertalen naar een ontwerp: dank voor de prachtige omslag. **Lotte**, jou ken ik al mijn hele bewuste leven. Er is niemand die zo goed kan luisteren als jij. Ik voel me bij jou altijd gesteund en begrepen en kan alleen maar hopen dat we als we straks écht oud zijn nog steeds eindeloos kletsen met thee en koekjes. **José**, jij hebt niet voor niets altijd mensen om je heen. Het is heerlijk om met jou door de polder of een bos te wandelen (en dan om de haverklap te pauzeren omdat jij een vogeltje hebt gespot).

Mijn broers, **Wouter** en **David**. **Wouter**, of moet ik zeggen: Van Binsbergen, wat kan ik ontzettend om jou lachen. Ik ben regelmatig jaloers op je scherpe geest en je vermogen om direct een inconsistentie in mijn redeneringen te spotten. Het is heel leuk om jouw zus te zijn. **David**, als iedereen jouw morele kompas had zag de wereld er vast een stuk mooier uit! Ik ben ontzettend trots op jou en op de moed die het vergt om in Dresden promotieonderzoek te doen en kijk nu al uit naar jouw verdediging. **Pap en mam**, zonder jullie onvoorwaardelijke vertrouwen in mijn kunnen was ik vast niet aan dit avontuur begonnen. Jullie hebben me geleerd dat ik niets moet (inmiddels een gevleugelde uitspraak in Groningen). Even belangrijk bleek de overtuiging dat inzet en toewijding altijd beloond worden. Dank voor zoveel steun en wijze raad: jullie zijn fantastisch. Tot slot, lieve **Bart**, in een proefschrift over gevoeletjes ben jij het allerbelangrijkste gevoeletje. Dankjewel voor al je nuchtere woorden, je optimisme op momenten dat ik het kan gebruiken, je idealisme, en bovenal je vertrouwen in onze liefde.

Curriculum vitae

Marieke Schreuder (Rotterdam, 16 September 1992) completed the bachelor Psychobiology at the University of Amsterdam and an extracurricular honors program cum laude in 2013. This study sparked her interest for what makes people feel, think, and behave the way they do, and more specifically, what happens if these processes are burdening. She therefore continued with a research master in Brain and Cognitive Sciences at the University of Amsterdam and studied, among other things, habitual and goal-directed behavior in mice. To also gain understanding in mental disorders from a clinical perspective, she pursued a master in Clinical Psychology at Leiden University. Here, she had the opportunity to do a clinical internship at the department of mood disorders at the Amsterdam University Medical Center. After graduating cum laude in 2015 and 2017, Marieke continued as a PhD student at the Interdisciplinary Center Psychopathology and Emotion Regulation (ICPE, University Medical Center Groningen) under the supervision of prof. dr. Marieke Wichers, dr. Catharina Hartman, and dr. Hanneke Wigman. Here, Marieke studied the development of mental disorders in young adults at increased risk for psychopathology. This project provided an excellent opportunity for combining scientific and clinical interests, and resulted in the current thesis. After having obtained her PhD degree, Marieke will continue as a postdoctoral researcher at the Katholieke Universiteit Leuven, to which she is looking forward with much enthusiasm.

List of publications

Peer-reviewed publications

Groen, R.N., Arizmendi, C., Wichers, M., **Schreuder, M.J.**, Gates, K.T., Hartman, C.A., & Wigman, J.T.W. (2021). Shared and individual-specific daily stress-reactivity in a cross-diagnostic at-risk sample. Conditionally accepted at Journal of Abnormal Psychology.

Schreuder, M.J., Wigman, J.T.W., Groen, R.N., Wichers, M., & Hartman, C.A. (2021). On the transience or stability of subthreshold psychopathology. *Scientific Reports*, 11, 1-7. doi: 10.1038/s41598-021-02711-3

Schreuder, M.J., Hartman, C.A., Menne-Lothmann, C., Decoster, J., van Winkel, J., Delespaul, P., De Hert, M., Derom, C., Thiery, E., Rutten, B.P.F., Jacobs, N., van Os, J., Wigman, J.T.W., & Wichers, M. (2020). Early warning signals in psychopathology: what do they tell? *BMC Medicine*, 18, 1-11. doi: 10.1186/s12916-020-01742-3

Schreuder, M.J., Wichers, M., Hartman, C.A., Menne-Lothmann, C., Decoster, J., van Winkel, J., Delespaul, P., De Hert, M., Derom, C., Thiery, E., Rutten, B.P.F., Jacobs, N., van Os, J., & Wigman, J.T.W. (2020). Lower emotional complexity as a prospective predictor of psychopathology in adolescents from the general population. *Emotion*, 1-8. doi: 10.1037/emo0000778

Schreuder, M.J.*, Groen, R.N.*, Wigman, J.T.W., Hartman, C.A., & Wichers, M. (2020). Measuring psychopathology as it unfolds in daily life: addressing key assumptions of intensive longitudinal methods in the TRAILS TRANS-ID study. *BMC Psychiatry*, 20, 1-14. doi: 10.1186/s12888-020-02674-1

Wichers, M., **Schreuder, M.J.**, Goekoop, R., & Groen, R.N. (2019). Can we predict the direction of sudden shifts in symptoms? Transdiagnostic implications from a complex systems perspective on psychopathology. *Psychological Medicine*, 1-8. doi: 10.1017/S0033291718002064

Schreuder, M.J., Schirmbeck, F., Meijer, C., & de Haan, L. (2017). Childhood trauma, neuroticism and obsessive-compulsive symptoms in patients with psychotic disorder. *Psychiatry Research*, 254, 48-53. doi: 10.1016/j.psychres.2017.04.030

*Contributed equally

Submitted for publication

Schreuder, M.J., Hartman, C.A., Groen, R.N., Smit, A.C., Wichers, M., & Wigman, J.T.W. (2021). Anticipating transitions in mental health in at-risk youth: a large-scale diary study into early warning signals.

Schreuder, M.J., Wigman, J.T.W., Groen, R.N., Weinans, E., Wichers, M., & Hartman, C.A. (2021). Anticipating the direction of symptom progression using critical slowing down: a proof-of concept study.

Groen, R.N., Wigman, J.T.W., Vos, M., **Schreuder, M.J.**, Wichers, M., & Hartman, C.A. (2021). How a general vulnerability for psychopathology during adolescence manifests in young adults' daily lives.

Bos, F.M.* , **Schreuder, M.J.***, George, S.V., Doornbos, B., Bruggeman, R., van der Krieke, L., Haarman, B.C.M., Wichers, M., & Snippe, E. (2021). Anticipating manic and depressive shifts in patients with bipolar disorder using early warning signals. Preprint: doi: 10.31234/osf.io/v5b7t

Helmich, M.A., Smit, A.C., Bringmann, L.F., **Schreuder, M.J.**, Oldehinkel, A.J., Wichers, M., & Snippe, E. (2021). Detecting impending symptom transitions using early warning signals in individuals receiving treatment for depression. Preprint: doi: 10.31234/osf.io/vf86s

Groen, R.N.* , **Schreuder, M.J.***, Wigman, J.T.W., Wichers, M., & Hartman, C.A. (2021). For whom do intensive longitudinal studies work? Participation and compliance in a six-month daily diary study among individuals at risk for mental health problems.

Morosan, L., Wigman, J.T.W., Groen, R.N., **Schreuder, M.J.**, Wichers, M., & Hartman, C.A. (2021). The impact of social relationships during adolescence on interpersonal functioning in young adulthood: a macro- and micro-level investigation using the TRAILS TRANS-ID study.

*Contributed equally

SHARE publications

This thesis is published within the **Research Institute SHARE** (Science in Healthy Ageing and healthcaRE) of the University Medical Center Groningen/University of Groningen.

Further information regarding the institute and its research can be obtained from our website: <https://umcgresearch.org/w/share>

More recent theses can be found in the list below.

(supervisors are between brackets)

2021

Hulshof TA

Quantitative appraisal of trial methods in reviews: an investigation of antipsychotic drug effects in dementia

(prof SU Zuidema, dr HJ Luijendijk)

Broekstra R

Values of trust and participation in scientific data repositories

(prof RP Stolk, prof S Otten, dr ELM Maeckelberghe, dr JL Aris-Meijer)

Du Y

CT lung cancer screening in China; the necessity to optimize target population

(prof GH de Bock, prof M Oudkerk, dr GA Sidorenkov)

Gol JM

Betwixt and between; medical care for functional somatic symptoms

(prof JGM Rosmalen, prof ROB Gans, prof RC Oude Voshaar)

Teegne BS

The heart of the matter; discovery of new genetic loci for heart rate variability and its relationship with blood pressure and mortality

(prof H Snieder, dr H Riese, dr IM Nolte)

Vries AM de

Regaining control during dialysis

(prof AV Ranchor, prof GJ Navis, dr MJ Schroevers, dr R Westerhuis)

Berg KS van der

Determinants of adverse health outcomes in late-life depression; the role of vitamine D and frailty

(prof RC Oude Voshaar, dr RM Marijnissen, dr JM Hegeman)

Asefa NG

Genetics and epidemiology of glaucoma and myopia

(prof H Snieder, prof NM Jansonius)

Kalteren WS

Neonatal anemia and red blood cell transfusions: finding the optimal balance

(prof AF Bos, dr EMW Kooi)

Elfrink AKE

Evaluating quality of care in liver surgery; the Dutch hepato biliary audit project

(prof Jm Klaase, dr DJ Grünhagen, dr NFM Kok)

Spahrkäs SS

Beating cancer-related fatigue with the Untire app

(prof M Hagedoorn, prof R Sanderman, dr A Looijmans)

Slagers AJ

Beyond the tendon; psychological factors during rehabilitation of lower limb tendon injuries

(prof J Zwerver, dr I van den Akker-Scheek, prof JHB Geertzen, dr IHF Reininga)

Kan K

Going beyond cost-effectiveness: analyzing routine mental healthcare data and stakeholders' perspectives to improve depression care

(prof RA Schoevers, prof TL Feenstra, prof E Buskens, dr F Jörg)

Ee IB van

The art of living with cancer at an advanced age

(prof JPJ Slaets, prof M Hagedoorn, dr CHM Smits)

Loohuis AMM

App-based treatment for female urinary incontinence; evidence based eHealth as an alternative to care-as-usual

(prof MY Berger, dr MH Blanker, de H van der Worp)

Wijnen A

Rehabilitation policies following total hip arthroplasty; across borders

(prof SK Bulstra, dr M Stevens, dr D Lazovic)

Narmandakh A

The night before the day after; connections between sleep problems and anxiety across the lifespan

(prof AJ Oldehinkel, dr AM Roest)

Birkenhäger-Gillesse EG

‘More at Home with Dementia’; effects of psychosocial interventions in the community and in nursing homes

(prof SU Zuidema, prof WP Achterberg, dr SIM Janus)

Sheferaw ED

Improving respectful maternity care in Ethiopia

(prof J Stekelenburg, prof TH van den Akker, dr YM Kim)

For earlier theses visit our website