Episode detection based on personalized intensity of care thresholds

Konings, Stefan R A; Bruggeman, Richard; Visser, Ellen; Schoevers, Robert A; Mierau, Jochen O; Feenstra, Talitha L

Published in:
Social Science & Medicine

DOI:
10.1016/j.socscimed.2020.113507

Social Science & Medicine 270 (2021) 113507

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: http://www.elsevier.com/locate/socscimed

Episode detection based on personalized intensity of care thresholds: a schizophrenia case study

Stefan R.A. Konings [a,*], Richard Bruggeman [b], Ellen Visser [b], Robert A. Schoevers [a], Jochen O. Mierau [c,d], Talitha L. Feenstra [e,f,g]

[a] University of Groningen, University Medical Center Groningen, Department of Psychiatry, Interdisciplinary Center Psychopathology and Emotion Regulation (ICPE), Groningen, the Netherlands
[b] University of Groningen, University Medical Center Groningen, University Center for Psychiatry, Rob Giel Research Center, Groningen, the Netherlands
[c] University of Groningen, Faculty of Economics and Business, Groningen, the Netherlands
[d] Aletta Jacobs School of Public Health, Groningen, the Netherlands
[e] University of Groningen, Faculty of Science and Engineering, Groningen Research Institute of Pharmacy, Groningen, the Netherlands
[f] Center for Nutrition, Prevention and Health Services Research, National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands
[g] University of Groningen, University Medical Center Groningen, Department of Epidemiology, Groningen, the Netherlands

A R T I C L E I N F O

Keywords:
Control charts
Schizophrenia
Mental illness
Episodes
Time series
Healthcare use
Observational data
Statistical methods

A B S T R A C T

Background: Schizophrenia Spectrum Disorder (SSD) is characterized by its chronic, episodic nature. The clear definition of such episodes is essential for various clinical and research purposes. Most current definitions of episodes in SSD are based on either hospitalizations or on symptom scales. Both have drawbacks; symptom scales are measured infrequently, while hospitalization rates are often affected by policy. This study presents an approach for defining episodes in healthcare data that does not suffer such drawbacks.
Methods: Healthcare use of 13,155 SSD patients in the Northern Netherlands with up to 12 years of follow-up was available. Patient-level structural changes in the trend of healthcare use costs were determined using Exponentially Weighted Moving Average (EWMA) control charts. Control charts restart with updated parameters after a detected structural change. Episodes were defined using these structural changes. The resulting episodes were validated by investigating their association with the Global Assessment of Functioning (GAF) scale.
Results: The mean number of episodes was 0.61 (sd: 0.60) per patient per year. For the sub-group without hospitalizations this was 0.51 (sd: 0.71). Average episode duration of the sub-group (147 days, sd: 309.4) was similar to that of the full sample (150 days, sd: 305.5). A significant inverse association was identified between GAF scores and the episode-state indicator.
Conclusions: The repeated application of EWMA control charts based on healthcare-intensity is a feasible and promising tool for quantifying patient-level healthcare episodes. The validation using GAF scores indicates that our episode indicator is associated with lower levels of global functioning. Results for individuals without hospitalizations indicate that the method is robust with regard to changes in healthcare policy.

1. Introduction

Various illnesses are characterized by a pattern of recurring episodes (Haut et al., 2006; Vajravelu et al., 2016; Warnock and Clayton, 2003), during which symptoms either become manifest or intensify. When the symptomatic burden of an illness undergoes structural changes, caregivers usually adapt treatment. Proper understanding of recurrent disease patterns could assist policymakers in resource allocation, possibly even preventing the condition of individual patients from deteriorating.

The main aim of the current study is to present a method for capturing and describing patient-level episodic patterns in healthcare use data over the full course of an illness. To demonstrate the method, the study focuses on Schizophrenia Spectrum Disorder (SSD), a burdensome illness from both a symptomatic and a societal perspective, due in part to its chronic, recurrent nature and large effect on functioning. According to a review by Emsley and colleagues (Emsley et al., 2013), the majority of patients with schizophrenia have a course of illness characterized by multiple episodes, with patients experiencing

* Corresponding author.
E-mail address: s.r.konings@umcg.nl (S.R.A. Konings).

https://doi.org/10.1016/j.socscimed.2020.113507
Received in revised form 7 October 2020; Accepted 5 November 2020
Available online 7 November 2020
0277-9536/

episodes of severe disorder followed by remission and maintenance care periods of stable disease with fewer problems.

As demonstrated in a review by Olivares and colleagues (Olivares et al., 2013), researchers have used different definitions of relapse (i.e., the onset of a subsequent episode), but most have included hospitalizations or symptoms as part of this definition. According to the same review, only 60% of all publications in the selected set of studies using the concept of relapse also provide a definition of relapse. Of these publications, 42% use hospitalization as a component of relapse, while the most common symptom scale — the Positive and Negative Symptom Scale (PANSS) (Kay et al., 1987) — is used in 20% of the selected publications. For one third of the studies using hospitalization as a component of relapse, this was the only factor in the definition of relapse.

Each of these definitions of relapse is subject to drawbacks. The use of symptom severity requires frequent measurements of patient-reported information, although such information is substantially lacking in most large administrative datasets and registries. The frequency of hospitalizations depends strongly on the regulatory and organizational aspects of specific healthcare systems (Mansell, 2006). Definitions based on hospitalizations hence are not robust to changes in these factors, and they do not cover relapses that are treated by outpatient care exclusively.

Another approach to defining episodes has been developed within the context of studies focusing on the analysis of claims data. Healthcare data are grouped to form "episodes of care". "Episode groupers" are proprietary software applications intended for use on claims data, with the purpose of identifying episodes of care (Rosen et al., 2012). Applications of this software are benchmarking of physicians (Houchens et al., 2009; Sandy et al., 2008; Thomas et al., 2010) and studies concerning payment systems. Existing episode groupers usually split episodes according to periods of complete absence of care, using gaps ranging from 30 days to 9 months (Cerrito, 2009). The definition of episodes based on gaps in healthcare use is severely limited when applied to individuals with chronic disorders (e.g., schizophrenia). The majority of these individuals require maintenance treatment and they do not have any periods of complete absence of care.

Relapse can be considered a structural change in a relatively steady *status quo*. The method presented in this paper draws on this concept, based on control charts applied to the intensity of healthcare use. The control chart is a known tool for determining structural changes. Control charts have been discussed extensively in the statistical, quality, and process-control literature (Stoumbos and Sullivan, 2002), although they have rarely been considered within medical contexts. Neuburger and colleagues (Neuburger et al., 2017) discuss the use of Exponential Weighted Moving Average (EWMA) control charts within the context of binary clinical performance data, while Lindquist and colleagues (Lindquist et al., 2007) apply this method within the context of neuro-imaging to investigate the stability of the brain-oxygen-level-dependent signal. The method performs well for detecting small but structural changes in time-series data (Montgomery, 2007). At the time of writing, there was no known literature on the use of EWMA control charts to detect treatment episodes in chronic disorders.

The control chart method detects structural deviations from a stable health state, without *a priori* knowledge of the timing of the break. Other methods (e.g., the structural break test) require *a priori* information on the number of breaks, and they are computationally intensive. Given the chronic but varying nature of many non-communicable diseases, including schizophrenia, multiple episodes may be present in an individual's healthcare use trajectory, and there is no *a priori* knowledge of the number and timing of these episodes. For this reason, we adjusted the EWMA control chart method to detect multiple structural breaks, while allowing the definition of a stable situation to shift over time. We did this by restarting the control chart after a change was detected, shifting the hypothesized stable situation along with the direction of the structural change. We use the term "Repeated EWMA control charts" (REWMA CC) to refer to this approach.

The advantage of our proposed method over existing methods for defining episodes of care based on hospitalization or periods of absence of care is that it is more accurate in distinguishing episodes for diseases that require maintenance treatment. Furthermore, this method is customized to the individual, as it identifies variable levels of healthcare intensity within and between individuals. The method is also robust to changes in healthcare regulations, culture, and financing.

Therefore, the current study is intended to introduce a repeating EWMA control chart method for the chronological clustering and monitoring of healthcare data, to apply this method to observational data on healthcare use for schizophrenia patients, to validate the method using a symptom scale for global functioning, and to compare the method to simpler alternatives.

2. Data & methods

2.1. Data

Administrative data on diagnosis and specialized mental healthcare use were available in the Psychiatric Case Registry Northern Netherlands (PCR-NN). The catchment area consisted of three northern provinces in the Netherlands: Friesland, Drenthe, and Groningen (1.7 million inhabitants). Data were available for the period from 2000 to 2012. The sub-group of patients in the PCR-NN with a diagnosis of schizophrenia (SSD) contained 13,155 individuals. Patients were observed for up to 13 years. Healthcare use data were available on a daily basis in five different categories. Diagnoses were determined according to the DSM-IV manual, and scores of the Global Assessment of Functioning (GAF) on the fifth axis were included during follow-up. Descriptive statistics are provided in Table 1.

The GAF Scale (American Psychiatric Association, 1987) was scored by physicians, with values ranging from 1 to 100, with lower scores indicating poorer functioning. The scale is anchored by 10 broad levels of functioning. Scores were assigned at each diagnostic contact. A total of 59,529 completed GAF measurements were available for this study. The mean (SD) GAF score across all measurements was 51.2 (16.0).

For the purposes of this study, we defined SSD as the collection of the schizophrenic disorders (DSM-IV 295), substance-induced psychotic disorders (DSM-IV 291.3, 291.5, 292.1), psychotic disoders due to general medical conditions (DSM-IV 293.81, 293.82), delusional disorder (DSM-IV 297.1), shared psychotic disorder (DSM-IV 297.3), brief psychotic disorder (DSM-IV 298.8), and psychotic disorder not otherwise specified (DSM-IV 298.9). A diagnosis of SSD during follow-up was required as a criterion for inclusion in the study sample.

2.2. Episodes

We distinguished between two states of healthcare use: a high-intensity "in-episode" state, representing relapse, and a low-intensity "out-of-episode" state, representing stable disease. For any patient, a

Table 1
Descriptive statistics for the PCR-NN SSD population.

Variable	Mean (SD)	N
% Male	57.2%	13,155
Age at entry	41.2 (18.1)	13,155
Years of observed follow-up	6.6 (4.6)	13,155
Number of GAF measurements per patient	6.3 (4.2)	9492
Number of inpatient-care contacts per year per patient	72.3 (153.6)	8346
Number of day-treatment contacts per year per patient	6.9 (28.3)	2926
Number of home-care service contacts per year per patient	4.1 (12.5)	5028
Number of community-based daycare contacts per year per patient	8.5 (30.3)	2866
Number of outpatient contacts per year per patient	15.9 (22.9)	12,389

change in state occurred when the intensity of healthcare underwent a significant, structural shift. Taking into consideration the individual variability over time in the definition of a patient's state implies the presence of differences within and between individuals, when distinguishing between high-intensity and low-intensity care. In sensitivity analyses, we investigated the effect of different values of model parameters and of alternative methods to define the in-episode and out-of-episode states.

2.3. Cost

Specialized mental healthcare use was multiplied by the corresponding unit costs, as defined in the Dutch costing manual (Hakkaart-van Roijen et al., 2015; Tan et al., 2012). The unit costs used and their sources are provided in Table 2. All unit costs were normalized to the 2015 price level, using the national price index. Costs per day were calculated for each individual by adding up all healthcare use costs indicated for each day.

2.4. Moving average

The moving average (MA) filter is a common statistical tool used for transforming time-series data to capture a trend while filtering noise (Hooker, 1901) (Axsäter, 2015). The current study used the Exponentially Weighted Moving Average (EWMA) filter (Hunter, 1986). A parameter (α) represents the slope parameter of the geometric distribution. Our implementation of the MA filter and a table describing the properties of the filter for given α are provided in Appendix A. The baseline choice of α was 0.08, while different values were evaluated as a sensitivity analysis.

2.5. Repeated EWMA control chart

At the start of a control chart (Montgomery, 2007), there is a hypothesized stable situation, which is designated as the baseline level of healthcare intensity. The hypothesis of a stable situation is rejected once the observed signal represented by healthcare intensity crosses one of two thresholds. The thresholds are symmetric around the hypothesized stable mean, and they depend on the standard deviation (λ times σ) of the healthcare signal.

In contrast, our new approach uses an iterative procedure of consecutive control charts, to which we refer with the term "Repeated EWMA Control Chart" (REWMA CC). The REWMA CC method assumes that individuals start in the "in-control" (out-of-episode) state, except for individuals who started with a non-zero level of healthcare intensity and whose first structural change was in a downward direction. In this case, the initial state is retrospectively defined as "in-episode." Following common practice (Mohammed et al., 2008) and after evaluating various options through sensitivity analysis, a baseline choice of 3 times σ was made for the control limits of the current study.

To stabilize the results, consecutive control charts used the continued healthcare-intensity signal at the starting point, thereby eliminating the "warm-up" period for any consecutive chart. The baseline intensity of a consecutive control chart was shifted to the final intensity of the previous control chart, after which the control limits were reset and the standard deviation was re-estimated. The updated initial variance was based on the past τ days of information for the unfiltered cost data at the most recent break. The parameter τ therefore functions as a "warm-up" period for the variance in a restarted control chart. The baseline choice of τ in this study was six weeks (42 days), and varied in sensitivity analyses.

A pseudocode description of our R (version 4.0.0) implementation of the REWMA CC is given in Appendix B.

Table 2
Unit costs per type of treatment.

Type of treatment	Unit costs (2015 price level)
1. Inpatient daycare	€257 (Tan et al., 2012)
2. Day-treatment daycare	€170 (Tan et al., 2012)
3. Psychiatric home-care services	€170 (Hakkaart-van Roijen et al., 2015)
4. Community-based daycare-center visits	€40 (Hakkaart-van Roijen et al., 2015)
5. Outpatient visits	€113 (Hakkaart-van Roijen et al., 2015)

2.6. Defining episodes based on REWMA CC structured data

The final step in the analysis involves defining episodes based on the structural changes identified in the time series according to the REWMA CC. If a change is in the same direction as the previous change, it is considered as an extension of the same change. Sequentially detected changes in the same direction are therefore accumulated as a single change. An episode is the period between the initial structural upward change and the next initial structural downward change.

2.7. Validation

To validate the resulting episode definitions, a regression model was estimated for the association between GAF scores for level of functioning and the episode state, controlling for several covariates. A basic model (1) was estimated.

$$GAF_i = \beta_0 + \beta_1 \cdot I_i \tag{1}$$

The extended model (2) was specified as follows:

$$GAF_i = \beta_0 + \beta_1 \cdot I_i + \beta_2 \cdot Fem_i + \beta_3 \cdot Age_i + \beta_4 \cdot Sig_i + \beta_5 \cdot Time_i + \beta_6 \cdot Time_i \cdot I_i \tag{2}$$

In this model, the GAF score (GAF_i) is a nominal scale that takes the 10 levels of global functioning. The model represents episodes with an indicator (I_i), where 0 and 1 denote "out-of-episode" and "in-episode" states, respectively. Covariates included were gender (Fem_i), age in years (Age_i), the EWMA filtered healthcare-intensity signal (Sig_i) in euros, and the number of days ($Time_i$) since the onset of the current indicator state, for each observation. We also determined the correlations between covariates.

2.8. Sensitivity analyses and comparison to other methods

To evaluate the resulting episodes, several outcome measures were defined (Appendix C). Sensitivity analyses were performed to investigate parameter and structural uncertainty. Parameter uncertainty concerns the three parameters relating to REWMA CC: α, λ times σ, and τ. The selected threshold distance was varied from 2 times σ to 5 times σ. Parameter α was evaluated over a range from 0.04 to 0.12 in steps of 0.02. Parameter τ was varied from 0 to 56 days, in steps of 14.

Results for the full sample using the baseline choice of parameters were compared to the results for the sub-group of patients without hospitalizations during follow-up.

Finally, to investigate structural uncertainty and analyze the added value of episode detection based on REWMA CC, two alternative methods for defining episodes were applied to the data. First, the onset of an episode was defined as the first healthcare contact after a pre-specified window without healthcare contacts. Second, hospitalizations were used to define episodes. The main window length for both methods was 90 days. As a sensitivity analysis for these two alternative methods, window lengths were varied from 30 to 150 days in steps of 30.

3. Results

3.1. Example: stepwise application of REWMA CC

Figure 1 shows in six steps how REWMA CC was applied to identify healthcare use episodes for a single arbitrarily selected individual. Healthcare use over time for this patient is presented in Panel A. We observed only a single period with hospitalizations. In Panel B, the healthcare use from Panel A is transformed into daily cost. The MA filter applied to the cost in Panel B is displayed in Panel C. In Panel D, the REWMA CC is shown as described, applied to the healthcare-intensity signal of a single patient, using the parameter values $\alpha = 0.08$, $\lambda = 3$, and $\tau = 42$. Integrating consecutive steps results in episodes. Panel E shows the start and end points for the structural changes corresponding to one episode. An episode starts at the first structural upward change (A) and ends at the first structural downward change (C). The resulting episodes for this example are underlined in Panel F.

Although Panel C contains high cost days in the initial years, they are only incidental, and they blend in with the noise of the filtered signal. In contrast, the high cost days after the fifth year appear to be more structural, thus resulting in additional short episodes. Our method clearly identifies episodes, even in the absence of hospitalizations. It also distinguishes an episode in periods of maintenance care, illustrating that the method enables episode detection for patients with nearly continuous health care use.

3.2. Episodes identified

Table 3 presents results for the entire sample using the main set of parameters: $\alpha = 0.08$, $\lambda = 3$ and $\tau = 42$. On average, patients had 0.61 episodes per year, with a duration of 150 days. Only 0.24% of the episodes were extremely short (less than seven days). The number of patients without a full episode cycle, consisting of an upward structural change followed by a structural downward change, was 2579 (19.6%). The mean and median cost gaps leading to a change in episode state were €18 and €25, respectively. The largest observed cost gap was €124. Outcomes determined in the sensitivity analysis using different

Table 3
Descriptive outcomes for all patients, compared to a sub-group without hospitalization.

Outcome	All patients (13,155)	No hospitalizations (4809)
N with at least 1 full episode cycle	10,576	3388
Episode duration (days)	150 (305.5)	147 (309.4)
Time between episodes (days)	303 (403.9)	337 (436.6)
Number of episodes (per year)[a]	0.61 (0.60)	0.51 (0.71)
Percentage of short episodes (<7 days)	0.24%	0.02%

[a] Average number of episodes in relation to average follow-up time. Sample standard deviation reflects variance in number of episodes only.

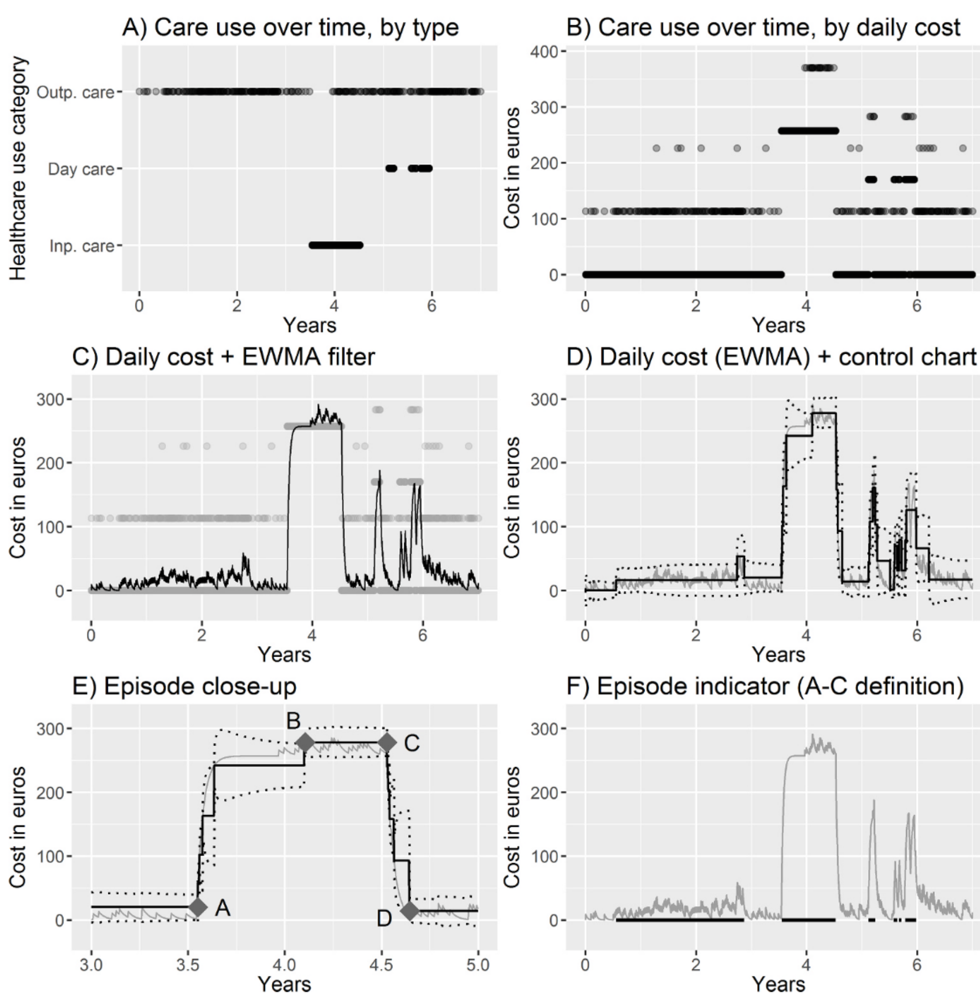

Figure 1. A collection of six panels chronologically showing the application of the REWMA CC method to obtain episodes from observational healthcare use data. Panel A: Healthcare use over time. Panel B: Cost of healthcare use over time. Panel C: Filtered cost of healthcare use over time. Panel D: Control chart around filtered cost of healthcare use. Panel E: Episode definition: episode starts at A and ends at C. Panel F: Episodes underlined in filtered healthcare use data.

parameters are presented in Appendix C. Additional information about the difference in cost levels between the in-episode and out-of-episode states is provided in Appendix D.

3.3. Validation estimates

Table 4 presents the results for a fixed effects estimation of Models 1 and 2. Results for the regular estimation are presented in Appendix E. All coefficients in the model were statistically significant. The results indicate that people in the “in-episode” state had lower GAF scores ($\beta = -0.21$), and thus poorer global functioning than did patients in the “out-of-episode” state. Moreover, patients with high mean cost at the time of receiving a GAF score generally had lower GAF scores as well ($\beta = -0.06$ per €100).

3.4. Sensitivity analyses

An overview of how the results varied according to different choices for parameters is presented in Appendix C.

The REWMA CC became less flexible with a larger threshold. Episodes were registered less frequently, duration increased for both states, and the cost gap during a transition increased on average. Similarly, for smaller values of α, we observed fewer and longer episodes. For low τ values, the control limits started out unstable, thus leading to many short, unstable episodes. Such outcomes are undesirable for purposes of detecting structural changes. Extremely high τ values could lead to the inclusion of outdated information in the initial estimation of the control limits, assuming that the underlying variance of healthcare use changes over time as well (Appendix C).

The episode state is negatively associated with GAF scores at a high level of statistical significance (the absolute T-values for the episode indicator in all sensitivity analyses were greater than 14) for a wide range of parameter values.

Results for the sub-group without hospitalizations are also presented in Table 3. The episode durations of this sub-group were similar to those of the total sample. Episodes occurred less frequently in the group without hospitalizations, and a smaller cost gap was required to cause a change of state among these patients. Results for additional sub-groups stratified by total number of episodes, age categories, main diagnosis, and sex are presented in Appendix F.

3.5. Comparison to other methods

Results for the analyses where the REWMA CC method is compared with alternatives are provided in Appendix G. The alternative methods indicated that patients were more often in the “in-episode” state than they were in the “out-of-episode” state. This resulted in longer episodes at a similar frequency, or higher frequencies with similar episode durations. Moreover, the hospitalization method is not applicable to a considerable number of patients. Of our population of 13,155 patients with SSD, 4809 had never been hospitalized.

Table 4
Estimates for regression models (1) and (2) with fixed effects on GAF scores. REWMA CC parameters $\lambda = 3$, $\alpha = 0.08$, and $\tau = 42$; N = 59,529; GAF scores from 9492 patients.

Model (1)				
Coefficient	Estimate β_i	Standard Error	T-statistic	P-value
Episode indicator	−0.18	0.01	−15.6	<0.001
Model (2)				
Coefficient	Estimate β_i	Standard Error	T-statistic	P-value
Episode indicator	−0.21	0.01	−15.4	<0.001
Age	0.03	0.00	11.0	<0.001
Cost signal (per €100)	−0.06	0.01	−6.6	<0.001
Time in state (per year)	0.02	0.01	3.9	<0.001
Episode indicator & Time in state interaction	0.03	0.01	2.8	0.006

4. Discussion

Application of the REWMA CC method to time-series data on healthcare use allowed the definition of episodes based on within-individual variations in the intensity of healthcare use. In our example using observational data on the healthcare use of patients with SSD, this method resulted in the identification of episodes of plausible length (150 days) and frequency (0.61 per year), although the outcomes were dependent on the choice of parameters. The validation using GAF scores indicated that being in a healthcare use episode implied lower (worse) GAF scores, even after correcting for overall cost level, gender, time in treatment, and other covariates.

Sensitivity analyses demonstrated that the method was robust in sub-groups with no use of hospital care, and it outperformed alternative methods of episode definition in various aspects. The method based on hospitalization would not function properly in the group without hospital-care use, which comprised a considerable number of patients (36.6%). In addition, both of the alternative methods resulted in long episodes (more than 300 days, on average) and a large number of very short episodes (>15%). The REWMA CC approach discriminated structural levels of increased healthcare use more accurately than did the simpler alternative approaches. Moreover, the flexibility of the alternative methods was limited, as they included only a single parameter, the time gap used to define a new episode, for fine-tuning the method.

The REWMA CC has three parameters: α, λ times σ, and τ and hence allows some flexibility in use. Although the choice of parameters depends on the needs of the user, we show that episode state is negatively associated with GAF scores at a high level of statistical significance (the absolute T-values for the episode indicator in all sensitivity analyses were greater than 14) for a wide range of parameter values. Ideally, optimal parameter choices for each application could be obtained by defining a constrained objective function (e.g., in terms of a constrained number of short episodes, cost gap, test statistic, or other outcome parameter). This must be left to further research.

Our results reveal a statistically significant negative association between episode state and GAF scores. It is important to note, however, that GAF scores are subjectively measured by clinicians, and they therefore do not necessarily represent functioning from the patient’s perspective. Another important caveat is that poorly functioning patients have the choice to reject healthcare altogether. It is thus possible for patients to have severe episodes that are not registered as intensive healthcare use. However, patients with the most severe episodes may be admitted against their will if they pose a risk for the safety of themselves or others. These forced admissions are included in our data as a part of inpatient care.

Extending the results somewhat further, when applied to the intensity of healthcare use, our method assumes that care has been allocated to patients appropriately. In other words, it assumes that patients receive more intensive care when they need it, and only when they need it. This problem could be avoided by applying the method to symptom scores, provided sufficiently frequent measurements were available. Such rich datasets of daily symptom scores are currently becoming available (Bos et al., 2015).

The regression analysis was performed using a large sample of 59,529 measurements across 9492 patients, and the findings are consistently in line with our expectation that the in-episode state would be negatively associated with global functioning. According to the results of the regression analysis of the second model, both the cost level of healthcare use intensity and the episode state had a strong statistically significant association with GAF scores, while controlling for several covariates. This suggests that, in general, the costs of healthcare use reflect variations between patients, while health state (such as in-episode or out-of-episode) represents variations within a given individual.

One strength of the REWMA CC method is that it is robust to system changes. We identified similar episode durations when comparing the sub-group with inpatient care to the group without inpatient care. The method is thus also applicable to individuals without hospital care, and it will continue to work even in light of policy changes affecting the supply of hospital beds. The group without inpatient care appears to have been more stable, as it had relatively fewer episodes and relatively less overall individual variation. Both of these results make sense. To further ensure the robustness of the REWMA CC method, an application of the method to time-series data from a different cohort, preferably managed in a different health care system, could be a topic for future research.

A further advantage of the REWMA CC is that it allows new data to be appended to existing data without re-estimating the entire model. This is useful within the context of monitoring, as it allows for quick feedback after data input, while requiring less computational power, as compared to other methods (e.g., structural break tests). Moreover, the EWMA control chart is robust to non-normality (Stoumbos and Sullivan, 2002).

The method could be used in various clinical and policy applications, for instance, the clustering of patients based on patterns in healthcare time-series data. This classification of patients could then be used to assign tailored treatment through updated guidelines. Feedback on treatment patterns, as well as comparing healthcare providers, could be highly valuable for clinicians and healthcare providers to improve their practice. Additionally, prediction of healthcare patterns could support the efficient allocation of physicians and other resources such as inpatient beds.

Furthermore, the method could be used to investigate outcomes with respect to further deinstitutionalization. First, it could serve for the evaluation of the reduction of specialized mental healthcare in favor of treatment in primary care. Second, predicted patterns could assist the decision making of which patients should be referred back to primary care, and for which patients treatment at a specialized mental health care institution is cost-effective.

In summary, we present a novel method for defining episodes in time-series data on healthcare use. The method defines episodes regardless of hospitalization or other type of healthcare use, as episodes are based entirely on structural changes in the input data. In addition, the proposed method distinguishes well between high and low intensity of care, as opposed to distinguishing between care and absence of care. This characteristic is a major advantage within the context of chronic disorders requiring nearly continuous maintenance care. As a result, the method allows for a new and more detailed perspective in the analysis of healthcare trajectories for chronic and recurring illnesses, such as SSD. Moreover, it is robust to changes in healthcare regulations, culture, and financing, thereby allowing for the consistent identification of episodes in datasets with real-world observations of healthcare use.

Funding

The authors declare that they received no specific funding for this work. The data infrastructure used to obtain the research data and the time spent on this research by TL Feenstra were partly funded by an unrestricted grant from Stichting De Friesland, Leeuwarden, the Netherlands (grant number DS29). The grant conditions guarantee full freedom of research and have no influence on publications.

Credit author statement

Stefan Konings: Conceptualization, Methodology, Software, Formal Analysis, Writing – Original Draft. **Richard Bruggeman**: Conceptualization, Methodology, Writing – Original Draft. **Ellen Visser**: Resources, Data Curation, Writing – Original Draft. **Robert Schoevers**: Writing – Review & Editing, Supervision. **Jochen Mierau**: Conceptualization, Methodology, Writing – Original Draft. **Talitha Feenstra**: Conceptualization, Methodology, Writing – Original Draft, Supervision.

Conflict of interest statement

The authors declare that there is no conflict of interest.

Acknowledgments

The authors would like to thank the RoQua members who were involved in providing the PCR-NN dataset and facilitated the corresponding data infrastructure.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2020.113507.

References

American Psychiatric Association, 1987. Revised (DSM-III-R). In: Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Association, Washington, DC.

Axsäter, S., 2015. Inventory Control. Springer. https://doi.org/10.1007/978-3-319-15729-0.

Bos, F.M., Schoevers, R.A., aan het Rot, M., 2015. Experience sampling and ecological momentary assessment studies in psychopharmacology: a systematic review. Eur. Neuropsychopharmacol 25, 1853–1864. https://doi.org/10.1016/j.euroneuro.2015.08.008.

Cerrito, P., 2009. Text Mining Techniques for Healthcare Provider Quality Determination: Methods for Rank Comparisons. IGI Global, Hershey, PA.

Emsley, R., Chiliza, B., Asmal, L., Harvey, B.H., 2013. The nature of relapse in schizophrenia. BMC Psychiatr. 13, 50. https://doi.org/10.1186/1471-244X-13-50.

Hakkaart-van Roijen, L., Van der Linden, N., Bouwmans, C., Kanters, T., Tan, S.S., 2015. Methodologie van kostenonderzoek en referentieprijzen voor economische evaluaties in de gezondheidszorg. Opdracht Van Zorginstituut Nederland. https://www.zorginstituutnederland.nl/publicaties/publicatie/2016/02/29/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg. (Accessed 30 September 2020).

Haut, S.R., Bigal, M.E., Lipton, R.B., 2006. Chronic disorders with episodic manifestations: focus on epilepsy and migraine. Lancet Neurol. 5, 148–157. https://doi.org/10.1016/S1474-4422(06)70348-9.

Hooker, R.H., 1901. Correlation of the marriage-rate with trade. J. Roy. Stat. Soc. 64, 485–492. https://doi.org/10.1017/CBO9781139170116.010.

Houchens, R.L., McCracken, S., Marder, W., Kelley, R., 2009. The Use of an Episode Grouper for Physician Profiling in Medicare A Preliminary Investigation, vol. 49. Thomson Healthcare for MedPAC, Washington, DC.

Hunter, J.S., 1986. The exponentially weighted moving average. J. Qual. Technol. 18, 203–210. https://doi.org/10.1080/00224065.1986.11979014.

Kay, S.R., Fiszbein, A., Opler, L.A., 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr. Bull. 13, 261–276. https://doi.org/10.1093/schbul/13.2.261.

Lindquist, M.A., Waugh, C., Wager, T.D., 2007. Modeling state-related fMRI activity using change-point theory. Neuroimage 35, 1125. https://doi.org/10.1016/j.neuroimage.2007.01.004.

Mansell, J., 2006. Deinstitutionalisation and community living: progress, problems and priorities. J. Intellect. Dev. Disabil. 31, 65–76. https://doi.org/10.1080/13668250600686726.

Mohammed, M.A., Worthington, P., Woodall, W.H., 2008. Plotting basic control charts: tutorial notes for healthcare practitioners. BMJ Qual. Saf. 17, 137–145. https://doi.org/10.1136/qshc.2004.012047.

Montgomery, D.C., 2007. Introduction to Statistical Quality Control. John Wiley & Sons, New York, NY.

Neuburger, J., Walker, K., Sherlaw-Johnson, C., van der Meulen, J., Cromwell, D.A., 2017. Comparison of control charts for monitoring clinical performance using binary data. BMJ Qual. Saf. 26, 919. https://doi.org/10.1136/bmjqs-2016-005526.

Olivares, J.M., Sermon, J., Hemels, M., Schreiner, A., 2013. Definitions and drivers of relapse in patients with schizophrenia: a systematic literature review. Ann. Gen. Psychiatr. 12, 32. https://doi.org/10.1186/1744-859X-12-32.

Rosen, A.B., Aizcorbe, E.L.A., Cutler, M., 2012. Comparing commercial systems for characterizing episodes of care. BEA Working Paper 2012-7. http://www.bea.gov/papers/pdf/meg_vs_etg_paper.pdf. (Accessed 30 September 2020).

Sandy, L.G., Rattray, M.C., Thomas, J.W., 2008. Episode-based physician profiling: a guide to the perplexing. J. Gen. Intern. Med. 23, 1521. https://doi.org/10.1007/s11606-008-0684-z.

Stoumbos, Z.G., Sullivan, J.H., 2002. Robustness to non-normality of the multivariate EWMA control chart. J. Qual. Technol. 34, 260. https://doi.org/10.1080/00224065.2002.11980157.

Tan, S.S., Bouwmans-Frijters, C.A.M., Hakkaart-van Roijen, L., 2012. Handleiding voor kostenonderzoek: methoden en referentieprijzen voor economische evaluaties in de gezondheidszorg. Tijdschr. Gezondheidswetenschappen (TSG) 90, 367.

Thomas, F., Caplan, C., Levy, J.M., Cohen, M., Leonard, J., Caldis, T., et al., 2010. Clinician feedback on using episode groupers with Medicare claims data. Health Care Financ. Rev. 31, 51.

Vajravelu, S., O'Brien, K.K., Moll, S., Solomon, P., 2016. The impact of the episodic nature of chronic illness: a comparison of fibromyalgia, multiple sclerosis and human immunodeficiency virus. Edorium J. Disabil. Rehabil. 2, 53–65. https://doi.org/10.5348/D05-2016-11-RA-7.

Warnock, J.K., Clayton, A.H., 2003. Chronic episodic disorders in women. Psychiatr. Clin. 26, 725–740. https://doi.org/10.1016/S0193-953X(03)00042-X.