Passive digital phenotyping
Jongs, Niels

DOI:
10.33612/diss.171368248

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:
2021

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

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.
Chapter 1

Effect of disease related biases on the subjective assessment of social functioning in neuropsychiatric patients

Niels Jongs\textsuperscript{1}, Brenda Penninx\textsuperscript{4}, Celso Arango\textsuperscript{5}, Jose Luis Ayuso-Mateos\textsuperscript{6}, Nic van der Wee\textsuperscript{7,8}, Inge Winter-van Rossum\textsuperscript{9}, Ilja M. J. Saris\textsuperscript{4}, Amber van Echteld\textsuperscript{9}, Sanne Koops\textsuperscript{9}, Amy C. Bilderbeck\textsuperscript{10}, Andreea Rasleascu\textsuperscript{10}, Gerard R. Dawson\textsuperscript{10}, Bernd Sommer\textsuperscript{11}, Hugh Marston\textsuperscript{12}, Jacob A. Vorstman\textsuperscript{3}, Marinus JC Eijkemans\textsuperscript{2}, and, Martien J. Kas\textsuperscript{1}

\textsuperscript{1}Groningen Institute for Evolutionary Life Sciences, University of Groningen, the Netherlands
\textsuperscript{2}Julius Center for Health Sciences and Primary Care, Department of Biostatistics and Research Support, University Medical Center Utrecht, the Netherlands
\textsuperscript{3}The Hospital for Sick Children, University of Toronto, Toronto, Canada
\textsuperscript{4}Department of Psychiatry and Amsterdam Neuroscience, VU University Medical Center, Amsterdam, the Netherlands
\textsuperscript{5}Institute of Psychiatry and Mental Health, Department of Child and Adolescent Psychiatry, Hospital General Universitario Gregorio Marañón, CIBERSAM, IISGM, Universidad Complutense, School of Medicine, Madrid, Spain
\textsuperscript{6}Department of Psychiatry, Universidad Autónoma de Madrid, Madrid, Spain; Centro de Investigación Biomédica en Red de Salud Mental, CIBERSAM, Instituto de Salud Carlos III, Madrid, Spain; Hospital Universitario de La Princesa, Instituto de Investigación Sanitaria Princesa (IIS-Princesa), Madrid, Spain.
\textsuperscript{7}Department of Psychiatry, Leiden University Medical Center, the Netherlands
\textsuperscript{8}Leiden Institute for Brain and Cognition/Psychiatric Neuroimaging, Leiden University Medical Center, the Netherlands
\textsuperscript{9}University Medical Centre Utrecht, Department of Psychiatry, Brain Center Rudolf Magnus, Utrecht, the Netherlands
\textsuperscript{10}P1vital Ltd., Wallingford, Oxfordshire, United Kingdom
\textsuperscript{11}Boehringer Ingelheim Pharma GmbH & Co KG, CNS Diseases Research, Biberach an der Riss, Germany
\textsuperscript{12}External Neurodegenerative Research, Eli Lilly and Company, Windlesham, United Kingdom

Published in Journal of psychiatric research (2020)
ABSTRACT:

Background:
Questionnaires are the current hallmark for quantifying social functioning in human clinical research. In this study, we compared self- and proxy-rated (caregiver and researcher) assessments of social functioning in Schizophrenia (SZ) and Alzheimer’s disease (AD) patients and evaluated if the discrepancy between the two assessments is mediated by disease-related factors such as symptom severity.

Methods:
We selected five items from the WHO Disability Assessment Schedule 2.0 (WHO-DAS) to assess social functioning in 53 AD and 61 SZ patients. Caregiver- and researcher-rated assessments of social functioning were used to calculate the discrepancies between self-rated and proxy-rated assessments. Furthermore, we used the number of communication events via smartphones to compare the questionnaire outcomes with an objective measure of social behaviour.

Results:
WHODAS results revealed that both AD (p<0.001) and SZ (p<0.004) patients significantly overestimate their social functioning relative to the assessment of their caregivers and/or researchers. This overestimation is mediated by the severity of cognitive impairments (MMSE; p=0.019) in AD, and negative symptoms (PANSS; p=0.028) in SZ. Subsequently, we showed that the proxy scores correlated more strongly with the smartphone communication events of the patient when compared to the patient-rated questionnaire scores (self; p=0.076, caregiver; p<0.001, researcher-rated; p=0.046).

Conclusion:
Here we show that the observed overestimation of WHODAS social functioning scores in AD and SZ patients is partly driven by disease-related biases such as cognitive impairments and negative symptoms, respectively. Therefore, we postulate the development and implementation of objective measures of social functioning that may be less susceptible to such biases.
INTRODUCTION

To date, the quantification of human behavioural constructs in biomedical studies predominantly rely on subjective research methods such as in-person interviews, questionnaires and self- or proxy-rated measures. The use of these methods over the past century evidently led to numerous important insights in disciplines such as psychiatry, sociology, economy, and even other disciplines in medicine. In psychiatry, for example, these methods are recently utilized to study the biological underpinnings of behavioural symptoms, such as social withdrawal in neuropsychiatric patients. Despite their wide use in research, these behavioural assessment methods have limitations that impede their objectivity. Most notably, these methods may depend on the participant’s (or the participant’s proxy) account of behaviour, and are consistently obtained post-hoc, i.e. questionnaire measures of behaviour are virtually never real-time. Observational assessments are real-time but they occur nearly always in a non-natural (e.g., clinical) setting and are relatively costly and time consuming.

Due to these limitations, current behavioural assessment methods are susceptible to various method and response biases. These biases or so-called measurement errors preclude the accurate collection of behavioural phenotypic data from humans and thereby limit our ability to understand variations in human behaviour. For example, these measurement errors interfere with the dependencies between behavioural measures and biological parameters, such as genotypes, brain activity patterns or structural brain data used to study the biological underpinnings of the observed behaviour. In addition, the search for effective treatments for social functioning (e.g. clinical trials for negative symptoms in schizophrenia) has always encountered the difficult question to determine outcomes as defined by patients or by proxies such as relatives. Subsequently, these distorted or misconceptualized dependencies between readouts might lead to wrong conclusions and explanations of behaviour and clinical trials and thus limits our understanding of the biological underpinnings of behaviour and interpretation of randomized clinical trials (RCTs).

Measurement errors in the assessment of behaviour are defined as the difference between the reported behaviour and the true behaviour which is usually unknown. These measurement errors consist of two components, a random and a systematic error component. The random component arises by factors of randomness during the collection of behavioural data. Examples are unclear questions or distortions in attention and/or motivation on the part of the participant. This random error cancels out in large sample sizes and that the trend in the observed behavioural data is not affected. Nevertheless, this component does increase the
variation or noise in the data and potentially results in poorer sensitivity of the measure, and weakened strength of association with other readouts of interest.

On the contrary, systematic errors in the assessment of behaviour do cause systematic deviations in a given measure, and for this reason can lead to potentially more misleading results. Here, we are interested in systematic errors that arise when specific symptoms, such as cognitive impairment or lack of disease insight, affect the subjective report of social behaviour in neuropsychiatric patients. Lack of insight\textsuperscript{7,8} and cognitive impairment\textsuperscript{9} are well-known phenomena in participants diagnosed with Alzheimer’s disease (AD), schizophrenia (SZ) or depression which in turn have shown to affect the assessment of their own behaviour\textsuperscript{10–13}. For instance, factors such as symptom severity, lack of insight and cognitive impairment increase the discrepancy between subjective and objective measures of behaviour in participants diagnosed with AD, SZ or depression\textsuperscript{10–13}. These objective measures of behaviour are, by definition, collected by assessment methods that do not depend on the participant’s own account of behaviour. Alternatively, behaviour is assessed by others such as caregivers, doctors or researchers. These proxy-rated accounts of behaviour are considered to be less biased and more importantly, and are likely unaffected by patient symptom severity. Therefore, the discrepancy between these self-rated and proxy-rated accounts of behaviour can be indicative of the degree of systematic error that is present in subjective assessments of behaviour.

Here, we first compared ratings of social behaviour among neuropsychiatric patients diagnosed with schizophrenia (SZ) or probable Alzheimer’s Disease (AD) as provided by three sources: (1) self-report, (2) study researchers, and (3) caregivers. Given the lack of disease insight and cognitive impairments in these psychiatric populations, we expected that patients would overestimate their social functioning relative to the assessment of others (e.g., caregivers). Subsequently, we assessed to what extent the discrepancy between the different assessors is mediated by patient symptom severity. Based on previous observations\textsuperscript{7,14}, we expected that increases in disease severity would be associated with increased overestimation of function on the part of the patient relative to informants. Finally, we examined if the patient, researcher- or caregiver-rated assessment correlated more strongly with an objective indicator of social behaviour that was derived from real-world passive remote monitoring of smartphone communication behaviour of the patient.
METHODS

Sample

The participants in this study were recruited through several clinical sites in Spain and the Netherlands as part of the Psychiatric Ratings using Intermediate stratified Markers\textsuperscript{15} (PRISM) consortium. The data analysed here were collected from 53 patients with a diagnoses of Alzheimer’s disease (AD) according to the criteria as outlined by the National Institute on Aging (NIA) and the Alzheimer’s Association (AA) (NIA-AA), and 61 patients meeting DSM-IV\textsuperscript{16} criteria for schizophrenia (SZ) as confirmed by diagnostic interview (the Mini-International Neuropsychiatric Interview\textsuperscript{17} (MINI)). Data in this study was collected with ethical approval and written informed consent was provided by all subjects.

Inclusion criteria for participants diagnosed with AD were (1) men and women aged between 50 to 80 and, (2) a Mini mental state examination\textsuperscript{18} (MMSE) score between 20 to 26. For participants diagnosed with SZ, inclusion criteria were (1) aged between 18 to 45 years, (2) stable medication dosage at least 8 weeks prior to recruitment and, (3) a diagnosis of SZ with a disease duration of no longer than 15 years. AD and SZ patients were excluded if they presented very severe disease symptoms (e.g. a score of $\geq 22$ on the 7-item PANSS positive symptom factor for schizophrenia, a score $< 20$ on the MMSE for Alzheimer’s Disease), had a current DSM-IV diagnosis of Major Depressive Disorder as assessed by the MINI or scored $\geq 16$ on the QIDS-SR16, suffered from drug or alcohol dependence within the three years prior to screening or had any contraindications for MRI studies. Additional exclusion criteria were (1) comorbid mental disorders that required intervention or treatment, (2) neurological diseases affecting the central nervous system and, (3) clinically important systematic illness that affects the ability to complete the study assessments. AD patients with a cerebrovascular accident based on patient history or imaging (where available) were excluded from recruitment.

Measure of social functioning

To assess the level of social functioning in AD and SZ patients we used a subset of five items (items 1 - 4 from the getting along domain; item 6 from the participation domain\textsuperscript{19}) from the WHO Disability Assessment Schedule 2.0\textsuperscript{20}. This five-item subset of the WHODAS assessed to what extend participants were able to engage in interpersonal relations and community related activities. Responses on these five items were on a Likert scale and ranged between 1 (no problems) to 5 (extreme or cannot do). We summed the responses on these five items to calculate a total score; higher scores here represent decreased social functioning. To acquire assessments from different parties, we asked caregivers and researchers
to assess the patient’s overall social functioning by using this five-item version of the WHODAS. For 40 AD and 28 SZ participants we collected caregiver-rated WHODAS score and for 42 AD and 54 SZ we collected researcher-rated WHODAS scores. To assess to what extend AD and SZ participants tended to overestimate their social functioning we calculated the discrepancies between self-rated and a combined researcher/caregiver-rated WHODAS score.

Clinical measures of symptom severity

To assess the severity of symptoms in AD and SZ participants we utilized two different questionnaires. For SZ, we utilized the PANSS to assess the severity of negative, positive and general psychopathology symptoms. Previous studies indicate that between negative and positive symptoms, the former is more strongly associated with decreased social functioning in SZ. Hence, in our analysis we focused on the severity of negative symptoms and their association with the discrepancies between self-rated and researcher/caregiver-rated WHODAS subset cores. The sum score on the negative symptom domain of the PANSS ranges from 7 to 49 points. Higher scores on this domain are indicative of more negative symptoms. PANSS data was available for a total of 43 SZ participants. Despite the focus on the severity of negative symptoms in SZ, we will also report the statistics for the association between the positive symptoms and general psychopathology scores.

For AD, we used the MMSE to assess the severity of cognitive impairment. The MSSE assesses 8 different domains of cognitive functioning and is scored on a 0- to 30-point scale. Higher scores here are indicative of better cognitive functioning. MMSE data was available for a subset of 45 AD participants.

Smartphone measure of social functioning

To assess the degree of social functioning in a more objective manner smartphone data was collected from a subset of participants. The smartphone data used in this study was collected by the BEHAPP smartphone application. BEHAPP is a passive behavioural monitoring application for Android that collects data by utilizing the embedded sensors in participants’ own smartphones. BEHAPP is used for scientific research that aims to provide objective, quantitative and longitudinal measures of human (social) behaviour to classify, for example, mental health disorders based on digital behavioural profiles, develop digital biomarkers to study disease progression and treatment efficacy, and identify early indicators of disease that allow prediction of disease onset, relapse and remission.

A subset of 19 AD patients and 16 SZ patients provided written consent to install the BEHAPP application on their own smartphone and to allow passive monitoring of their activities. Smartphone data collected over a consecutive period of 14 days
was used to extract the total number of communication related events measured (i.e. usage of communication apps such as WhatsApp, Telegram or Skype) and was used as a more objective measure of social behaviour.

**Statistical analysis**

The statistical analysis for this study was conducted in a stepwise manner and aims to describe the association between self-, researcher and caregiver-rated WHODAS subset score and an objective measure of social behaviour as collected by BEHAPP. First, a two-way ANOVA was used to assess the main effect of the different assessors (patient, researcher or caregiver) on the WHODAS scores for AD and SZ. Subsequently, a post-hoc test with a Tukey correction was used to evaluate the difference between the assessors within each disease label.

In order to assess to what extent participants tended to overestimate or underestimate their social functioning, we calculated the discrepancy between self-rated and researcher/caregiver-rated WHODAS scores. This discrepancy is calculated by combining the researcher- and caregiver-rated WHODAS scores. In this manner we were able to obtain a single discrepancy score between the self-rated and proxy-rated (researcher- and caregiver-rated) WHODAS scores. This discrepancy was calculated in the following manner for participants of whom researcher and caregiver-rated WHODAS subset scores were available:

\[
WHO_p = \frac{(WHO_r + WHO_c)}{2} \tag{1}
\]

\[
\text{discrepancy} = WHO_p - WHO_s \tag{2}
\]

Where \(WHO_r\) denotes the researcher-rated WHODAS score, \(WHO_c\) the caregiver, and \(WHO_s\) the self-rated. First, we calculated the average WHODAS score for the proxy-rated assessments (1) and subsequently calculated the discrepancy (2). For participants with either researcher or caregiver-rated WHODAS scores, discrepancies were calculated by subtracting the self-rated from the researcher or caregiver-rated WHODAS scores. By applying this approach, discrepancy data was available for 54 SZ and 42 AD participants. Given the Likert scale nature of this measure we used a Mann-Whitney U Test to evaluate if the discrepancies are significantly higher than zero. Next, we evaluated if the discrepancies are associated with the severity of symptoms in AD and SZ participants. Four generalized linear models (GLMs) with a Poisson distribution were utilized to assess the main effect of the discrepancies on the severity of symptoms (PANSS negative, positive and general psychopathology and MMSE).
Finally, we studied the association between the self-, researcher- and caregiver-rated WHODAS scores and the more objective measure of behaviour as generated by BEHAPP. A GLM was utilized to assess to what extend lower WHODAS scores are associated with decreased number of communication events. A total of two GLMs were fitted with the different WHODAS scores (average proxy- and self-rated) as a main effect and an interaction with disease label (AD or SZ).

**RESULTS**

Demographic and symptomatic information about the two patient populations is presented in Table 1. The average age of the AD and SZ patients who agreed to passively monitor behaviour by using their own smartphone is 67.74 ± 7.62 and 31.59 ± 6.40, respectively, and is non-significantly different relative to participants that have not participated in BEHAPP. AD patients that did participate in BEHAPP had similar MMSE scores (24.30 ± 2.03 vs 23.60 ±2.12) than those who did not participate (t(42) = 1.13, p = 0.26). However, for the SZ patients the general psychopathology as measured by the PANSS was significantly lower (23.07 ± 4.11 vs 26.89 ±6.14) in the participants participating in BEHAPP compared to those that did not participate (t(38) = -2.41, p = 0.021).

**Table 1 | Demographic information and symptom severity per disease label.** Average (SD) WHODAS scores for each assessor is also presented.

<table>
<thead>
<tr>
<th>Alzheimer’s disease</th>
<th>WHODAS</th>
<th>PANSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (sd)</td>
<td>Sex (f/m)</td>
<td>Self</td>
</tr>
<tr>
<td>68.94 (7.29)</td>
<td>24/29</td>
<td>6.66 (2.53)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Schizophrenia</th>
<th>WHODAS</th>
<th>PANSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (sd)</td>
<td>Sex (f/m)</td>
<td>Self</td>
</tr>
<tr>
<td>30.13 (6.55)</td>
<td>20/41</td>
<td>11.43 (5.23)</td>
</tr>
</tbody>
</table>

*NT: Negative symptoms, PT: Positive symptom, GT: General psychopathology

**Social functioning**

The results of the group-wise comparison of the WHODAS scores are presented in Figure 1. The one-way ANOVA revealed a significant main effect of assessor on the WHODAS scores for AD patients (F(2,132) = 20.72, p < 0.001). Post-hoc analyses using the Tukey post-hoc correction for significance indicated that AD patients tended to significantly overestimate their social functioning relative to the assessment of the caregiver (b = 3.27, t(132) = 4.65, p < 0.001) and researcher
Effect of disease related biases on the subjective assessment of social functioning in neuropsychiatric patients

Relative to the assessment of the caregiver (Figure 1A), AD patients tended to overestimate their social functioning on average with 3.26 points (33%). This contrast is slightly higher between the self- and researcher-rated WHODAS scores. Relative to the researcher-rated scores, AD patients overestimated their social functioning on average with 4.17 points (39%) (self-: 6.66 ± 2.53; Caregiver-: 9.93 ± 3.58; researcher-rated: 10.83 ± 3.97).

There was no significant effect of assessor on the WHODAS scores for SZ patients ($F(2,140) = 1.49, p = 0.227$)(Figure 1B). Relative to the caregiver- ($b = 0.89$, $t(140) = 0.79$, $p = 0.708$) and researcher-rated ($b = 1.59$, $t(140) = 1.72$, $p = 0.200$)

WHODAS scores, SZ patients did not report their social function significantly different compared to caregiver or researcher(self-: 11.43 ± 5.23; Caregiver-: 12.32 (+7%) ± 4.52; researcher-rated: 13.02 (+12%) ± 4.83).

**Discrepancies in WHODAS scores**

Next, we calculated individual discrepancies between the self-rated WHODAS and the caregiver- and researcher-rated WHODAS (see methods). Positive scores on this measure indicate that patients rated their social function as better compared to ratings provided by their informant(s). These discrepancies are depicted in Figure 2. Mann-Whitney U Tests showed that for both AD and SZ patients, the discrepancy in WHODAS scores was significantly higher than zero (AD: Mann-
Whitney U = 686, n = 42, p < 0.001; SZ: Mann-Whitney U = 680.5, n = 54, p = 0.004), such that both AD (3.44 ± 3.30) and SZ (1.18 ± 2.74) patients reported their social functioning to be significantly higher than that reported by their caregivers and researchers.

**Figure 2 | Discrepancy in WHODAS scores between proxy assessors and patients.** Mann-Whitney U Test reveals that both are significantly higher than zero (AD: Mann-Whitney U = 686, n = 42, p < 0.001, SZ: Mann-Whitney U = 680.5, n = 54, p = 0.004) which suggest that both AD and SZ participants tended to overestimate their social functioning.

**Figure 3 | Symptom severity plotted against WHODAS discrepancies.** (A) Negative symptoms as measured by the PANSS vs the discrepancy in WHODAS scores. Results reveal a significant positive relationship ($b = 0.034$, $z = 2.193$, $p = 0.028$) which suggest that the overestimation of social functioning is associated with increased negative symptoms in SZ. (B) MSSE vs the discrepancy in WHODAS scores. Results reveal a significant negative relationship ($b = -0.019$, $z = -2.346$, $p = 0.019$) which suggest that increased cognitive impairment is associated with overestimation of social function in AD.
Effect of disease related biases on the subjective assessment of social functioning in neuropsychiatric patients

**Association between symptom severity and WHODAS discrepancies**

Figure 3 depicts the relation between the severity of symptoms and the discrepancies between different WHODAS scores. GLM with a Poisson distribution revealed that there was a significant and positive relationship between PANSS negative symptom scores and the size of the WHODAS discrepancy in SZ participants ($b = 0.034$, $z = 2.193$, $p = 0.028$, Figure 3A).

Analysis of the positive symptoms ($b = -0.026$, $z = -1.427$, $p = 0.154$) and general psychopathology ($b = -0.023$, $z = -1.921$, $p = 0.055$) revealed that both are non-significantly related with WHODAS discrepancy in SZ patients. Similarly, a GLM with a Poisson distribution revealed that there was a significant and positive relationship between MMSE scores and the size of the WHODAS discrepancy in AD participants ($b = -0.019$, $z(31) = -2.346$, $p = 0.019$, Figure 3B).

**Association between BEHAPP and WHODAS**

The association between the number of communication events as registered by the BEHAPP application and the self-rated and the proxy-rated (based on caregiver- and researcher-rated scores) WHODAS scores is depicted in Figure 4. Noteworthy is the significant difference in communication events (t-test; $t(19) = 4.51$, $p < 0.001$) between AD (133 ± 125) and SZ (480 ± 275) diagnosed participants.

GLM analysis revealed a non-significant main effect of the self-rated WHODAS on the number of smartphone communication events ($b = -16.85$, $t(29) = -1.471$, $p = 0.076$). In contrast, proxy-rated WHODAS scores were significantly and negatively related to the number of communication events (i.e., improved ratings of social function were related to more communication events) ($b = -29.60$, $t(26) = -2.331$, $p = 0.014$).
DISCUSSION

Here we provide evidence to suggest that, relative to the WHODAS subset assessment by caregivers and researchers, subjects diagnosed with AD and SZ tend to overestimate their level of engagement in interpersonal relations and community related activities. Furthermore, our results suggest that this overestimation of social functioning is mediated by the severity of patient symptoms, namely, by the severity of cognitive impairment in AD, and level of negative symptomology in SZ. Together, our findings suggest that the validity of self-rated assessments of social functioning in SZ and AD patients are affected by disease-related biases.

In the present study, the patient self-rated score for social functioning was compared with patient assessments performed by the caregiver and the researcher. Comparisons of these scorings indicate that the assessments from the caregiver and researcher are different from the patient self-rated score, and pose the question which of these values represent the ground truth? While the patient score is suggesting overestimation of their level of social functioning relative to the scores of the caregiver and researcher, the self-rated assessment may very well reflect the subjective experience of social functioning by the patient. To further investigate this difference in scoring, we also assessed social behaviour of the patient on the basis of social measures derived from their smartphone. For that purpose, we compared the WHODAS assessments (by patient, caregiver, and researcher) with an objective measure of social behaviour, namely by monitoring smartphone communication activities in a subgroup of the patients in this study. Compared to the patient scores, this analysis revealed a relatively stronger correlation between the proxy-rated WHODAS scores and the more objective measure of social functioning collected through passive smartphone monitoring. These correlations suggest that the proxy-rated (i.e. caregiver, family, practitioner) assessments have a better basis in patients to objectively quantify behavioural phenotypes, as compared with self-rated assessments made by patients themselves. These findings provide important considerations for clinical studies, such as RCTs for negative symptoms in which social functioning is assessed in these patient populations.

The difference in self- versus proxy-rated scoring may be explained by impact of symptoms in patients, such as cognitive impairments and lack of disease insight. A preceding study confirmed this argument by demonstrating that proxy data is relatively more reliable while assessing the severity of depressive symptoms in AD patients that suffered from severe cognitive impairments and lack of disease insight. In addition, researchers recently showed higher inter-rater reliability for proxy-based assessments of behaviour relative to self-reports. This inter-rater reliability was evaluated by using different assessment methods that intended to measure the same behavioural construct (i.e. physical health, cogni-
Effect of disease related biases on the subjective assessment of social functioning in neuropsychiatric patients

Results of this study revealed higher correlations between the different assessment methods for proxy-based assessments relative to self-reports. In contrast, several studies suggest that the assessment of behaviour by proxies is also driven by a variety of different biases\textsuperscript{27,28}, demographical factors\textsuperscript{29,30} and patient characteristics\textsuperscript{31} and that these have to be taken into account while interpreting behavioural data collected via proxy-based assessments\textsuperscript{32}.

The present findings suggest that the overestimation based on the WHODAS social functioning scores in patients may be affected by disease-related biases. Indeed, correlation analysis revealed that higher negative symptom scores on the PANSS are associated with stronger deviations in WHODAS scores between SZ patients and their proxy-rated scores. Comparable findings were observed for AD patients and their cognitive performance on the MMSE; the lower the patient’s MMSE score (lower cognitive performance), the stronger the deviations in WHODAS scores between AD patients and their proxy-rated scores. In light of these findings, our observations are consistent with those from previous studies\textsuperscript{10–13} showing the effects of disease-related symptoms on behavioural data collected through self-rated assessment methods.

A variety of disease-related factors should be considered in the context of our findings. First, lack of disease insight, unawareness, and the denial of symptoms are well-established phenomena in subjects diagnosed with AD and SZ\textsuperscript{33}. Not only are these phenomena highly distressing for caregivers, they are linked to severity and cognitive impairments for both disorders. Second, it has been shown that the lack of insight in patients is associated with the overestimation of social functioning relative to patients with milder symptoms\textsuperscript{14}. Comparable results are found for AD. Relative to assessment of caregivers, AD patients tend to overestimate their independent functioning in everyday life and this overestimation is partly explained by the severity of cognitive impairments and lack of insight\textsuperscript{34}. Altogether, these results support previous findings which suggest that increased lack of insight in SZ and AD patients is related to the relative overestimation of one’s own social functioning. These findings are in line with the relative overestimation of social functioning with increasing disease severity in the present study, and suggest that quality of life and social functioning assessments by SZ and AD patients are likely to be affected by their cognitive impairments and negative symptoms.

In summary, here we show that relative to the assessment of caregiver and researchers participants diagnosed with AD and SZ tend to overestimate their social functioning. Furthermore, we showed that this overestimation is mediated by the severity of cognitive impairments and negative symptoms. Important to note is that social functioning in the present study is defined as the ability to...
engage in interpersonal relations and community related activities as measured by the 5 WHODAS\textsuperscript{20} items. Further research is needed to evaluate whether our results are generalizable to other constructs of social functioning. For example, it has been shown that for more subjective constructs such as pain\textsuperscript{35,36} and psychological wellbeing\textsuperscript{29} proxy-based assessments are less valid. Increased agreement on proxy-based assessments is observed between different assessors if the construct is less dependent on the judgement or perception of the patients and more on objective behaviours such as the count of specific events\textsuperscript{37}. Although our results indicate that caregiver- and/or researcher-based WHODAS assessments of social functioning are consistent and relate to objective smartphone measures, care should be taken in assuming that these proxy measures adequately capture the ground truth of individuals’ social function.

To gain a better understanding of the variation in human social behaviour and in underlying biological mechanisms, we propose to further investigate the use of objective methods to quantify social functioning. Recently, researchers started to utilize the smartphone as an objective tool to quantify behaviour, including social behaviour. This so-called digital phenotyping utilises the wide variety of sensors in a smartphone to passively monitor behaviour in a longitudinal manner. Key features of this approach are that 1) data is collected in real-time, 2) in the participant’s natural environment, and 3) without the need for any self- or proxy reporting, thereby addressing some of the most important challenges inherent to current behavioural research. Here we showed that the quantification of a relatively simple event such as communication events as registered by a smartphone may provide potential indicators of social functioning. However, further studies are needed to develop and optimize novel objective, longitudinal, quantitative and real-world measures of social functioning.

ACKNOWLEDGEMENTS

The PRISM project (www.prism-project.eu) leading to this application has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 115916. This Joint Undertaking receives support from the European Union’s Horizon 2020 research and innovation programme and EFPIA. This publication reflects only the authors’ views neither IMI JU nor EFPIA nor the European Commission are liable for any use that may be made of the information contained therein.
Effect of disease related biases on the subjective assessment of social functioning in neuropsychiatric patients

Conflict of Interests:
Dr. Arango. has been a consultant to or has received honoraria or grants from Acadia, Angelini, Gedeon Richter, Janssen Cilag, Lundbeck, Minerva, Otsuka, Roche, Sage, Servier, Shire, Schering Plough, Sumitomo Dainippon Pharma, Sunovion and Takeda.
REFERENCES


Chapter 1

Effect of disease-related biases on the subjective assessment of social functioning in neuropsychiatric patients


