university of groningen

University Medical Center Groningen

University of Groningen

Effect of disease related biases on the subjective assessment of social functioning in Alzheimer's disease and schizophrenia patients

Jongs, Niels; Penninx, Brenda; Arango, Celso; Ayuso-Mateos, Jose Luis; van der Wee, Nic; Rossum, Inge Winter-van; Saris, Ilja M J; van Echteld, Amber; Koops, Sanne; Bilderbeck, Amy C

Published in:
Journal of Psychiatric Research

DOI:
10.1016/j.jpsychires.2020.11.013

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):
Jongs, N., Penninx, B., Arango, C., Ayuso-Mateos, J. L., van der Wee, N., Rossum, I. W., Saris, I. M. J., van Echteld, A., Koops, S., Bilderbeck, A. C., Raslescu, A., Dawson, G. R., Sommer, B., Marston, H., Vorstman, J. A., Eijkemans, M. JC., & Kas, M. J. (2022). Effect of disease related biases on the subjective assessment of social functioning in Alzheimer's disease and schizophrenia patients. *Journal of Psychiatric Research*, *145*, 302-308. https://doi.org/10.1016/j.jpsychires.2020.11.013

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.

Journal of Psychiatric Research 145 (2022) 302–308

Contents lists available at ScienceDirect

Journal of Psychiatric Research

journal homepage: www.elsevier.com/locate/jpsychires

Effect of disease related biases on the subjective assessment of social functioning in Alzheimer's disease and schizophrenia patients

Niels Jongs [a], Brenda Penninx [d], Celso Arango [e], Jose Luis Ayuso-Mateos [f,g,h], Nic van der Wee [i,j], Inge Winter-van Rossum [k], Ilja M.J. Saris [d], Amber van Echteld [k], Sanne Koops [k], Amy C. Bilderbeck [l], Andreea Raslescu [l], Gerard R. Dawson [l], Bernd Sommer [m], Hugh Marston [n], Jacob A. Vorstman [c], Marinus JC. Eijkemans [b], Martien J. Kas [a,*]

[a] Groningen Institute for Evolutionary Life Sciences, University of Groningen, the Netherlands
[b] Julius Center for Health Sciences and Primary Care, Department of Biostatistics and Research Support, University Medical Center Utrecht, the Netherlands
[c] The Hospital for Sick Children, University of Toronto, Toronto, Canada
[d] Department of Psychiatry and Amsterdam Neuroscience, VU University Medical Center, Amsterdam, the Netherlands
[e] 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
[f] Department of Psychiatry, Universidad Autónoma de Madrid, Madrid, Spain
[g] Centro de Investigación Biomédica en Red de Salud Mental, CIBERSAM, Instituto de Salud Carlos III, Madrid, Spain
[h] Hospital Universitario de La Princesa, Instituto de Investigación Sanitaria Princesa (IIS-Princesa), Madrid, Spain
[i] Department of Psychiatry, Leiden University Medical Center, the Netherlands
[j] Leiden Institute for Brain and Cognition/Psychiatric Neuroimaging, Leiden University Medical Center, the Netherlands
[k] University Medical Centre Utrecht, Department of Psychiatry, Brain Center Rudolf Magnus, Utrecht, the Netherlands
[l] P1vital Ltd., Wallingford, Oxfordshire, United Kingdom
[m] Boehringer Ingelheim Pharma GmbH & Co KG, CNS Diseases Research, Biberach an der Riss, Germany
[n] External Neurodegenerative Research, Eli Lilly and Company, Windlesham, United Kingdom

A B S T R A C T

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 (WHODAS) 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.

1. 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

* Corresponding author.
E-mail address: m.j.h.kas@rug.nl (M.J. Kas).

https://doi.org/10.1016/j.jpsychires.2020.11.013
Received 31 August 2020; Received in revised form 29 October 2020; Accepted 4 November 2020
Available online 9 November 2020
0022-3956/© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

withdrawal in neuropsychiatric patients (van der Wee et al., 2019). 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 (Podsakoff et al., 2003). 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 (Bugarski-Kirola et al., 2017; Fraguas et al., 2019; Galderisi et al., 2018). Subsequently, these distorted or miss-conceptualized 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 (Cramer et al., 1993). 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. . In the present study, we are interested in systematic errors that arise when specific symptoms, such as cognitive impairment or lack of disease insight which are well-known phenomena in participants diagnosed with Alzheimer's disease (AD), schizophrenia (SZ) or depression (Amador et al., 1993; Reddy, 2016), and cognitive impairment (Millan et al., 2012), and can affect the assessment of their own behaviour (Doyle et al., 1999; Hayhurst et al., 2014; Pennington et al., 2019; Petersen et al., 2019).

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 (Reddy, 2016; Siu et al., 2015), 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.

2. Methods

2.1. 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 (Kas et al., 2019) (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 (Association and others, 2000) criteria for schizophrenia (SZ) as confirmed by diagnostic interview (the Mini-International Neuropsychiatric Interview (Sheehan et al., 1998) (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 and 80 and, (2) a Mini mental state examination (Folstein et al., 1983) (MMSE) score between 20 and 26. For participants diagnosed with SZ, inclusion criteria were (1) aged between 18 and 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. For a detailed overview of the inclusion and exclusion criteria we refer to the manuscript of Bilderdeck (Bilderbeck et al., 2019).

2.2. 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)(Saris et al., 2017; van der Wee et al., 2019) from the WHO Disability Assessment Schedule 2.0 (Üstün et al., 2010). 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.

2.3. 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 (Carpenter et al., 1988; Galderisi et al., 2018). 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 (Folstein et al., 1983) 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.

2.4. 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.

2.5. 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 (Tukey, 1949) 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:

$$WHO_P = \frac{(WHO_r + WHO_c)}{2} \quad (1)$$

$$discrepancy = WHO_P - WHO_s \quad (2)$$

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).

3. 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 ($t(38) = -2.41$, $p = 0.021$).

3.1. Social functioning

The results of the group-wise comparison of the WHODAS scores are presented in Fig. 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 ($b = 4.17$, $t(132) = 6.03$, $p < 0.001$). Relative to the assessment of the caregiver (Fig. 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 ($F(2,140) = 1.49$, $p = 0.227$)(Fig. 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).

3.2. 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 Fig. 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.

Table 1
Demographic information and symptom severity per disease label. WHODAS for each assessor is also presented.

Alzheimer's disease		WHODAS						
Age (*sd*)	Sex (f/m)	Self	Caregiver	Researcher	MMSE			BEHAPP
68.94 (7.29)	24/29	6.66 (2.53)	10.83 (3.97)	9.93 (3.58)	26.77 (7.26)			$n = 23$
Schizophrenia		WHODAS			PANSS			
Age (*sd*)	Sex (f/m)	Self	Caregiver	Researcher	NT[a]	PT[a]	GT[a]	BEHAPP
30.13 (6.55)	20/41	11.43 (5.23)	13.02 (4.83)	12.32 (4.52)	14.12 (5.85)	11.51 (3.46)	25.52 (5.75)	$n = 22$

[a] NT: Negative symptoms, PT: Positive symptom, GT: General psychopathology.

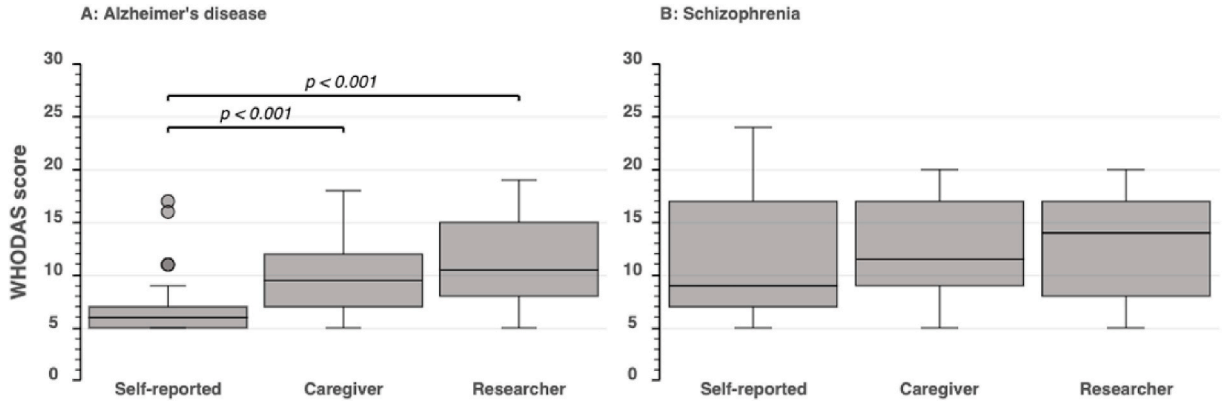

Fig. 1. (A) Self-, caregiver and researcher-rated WHODAS scores for AD patients. (B) Self-, caregiver and researcher-rated WHODAS scores for SZ patients. The results suggest that AD patients tended to significantly overestimate their social functioning relative to the assessment of the caregiver (b = 3.27, t(132) = 4.65, p < 0.001, n = 40) and researcher (b = 4.17, t(132) = 6.03, p < 0.001, n = 42). For the SZ patient the data suggest a non-significant difference between the over-estimation of social functioning (caregiver; b = 0.89, t(140) = 0.79, p = 0.708, n = 28) (researcher; b = 1.59, t(140) = 1.72, p = 0.200, n = 54).

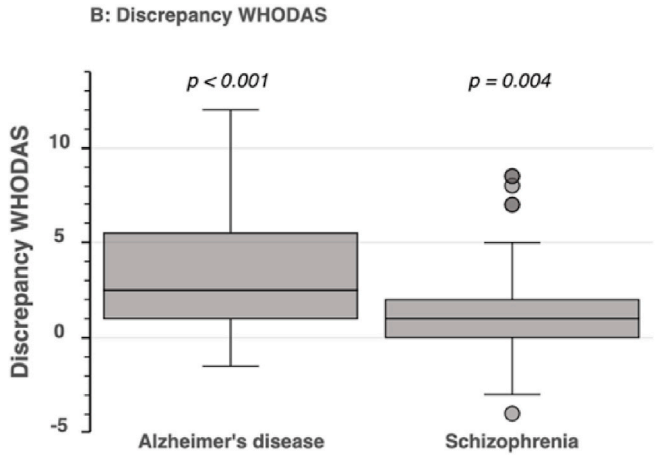

Fig. 2. Discrepancy in WHODAS scores for AD and the Schizophrenia group. Mann-Whitney *U* Test reveals that both are significantly higher than zero (AD: Mann-Whitney U = 686, n = 42, p < 0.001, SZ: Mann-Whitney U = 680.5, n = 54, p = 0.004) which suggest that both AD and SZ participants tended to overestimate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

3.3. Association between symptom severity and WHODAS discrepancies

Fig. 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, $p = 0.028$, Fig. 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, $p = 0.019$, Fig. 3B).

3.4. 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 Fig. 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)($b = -29.60$, $t(26) = -2.331$, $p = 0.014$).

4. 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-

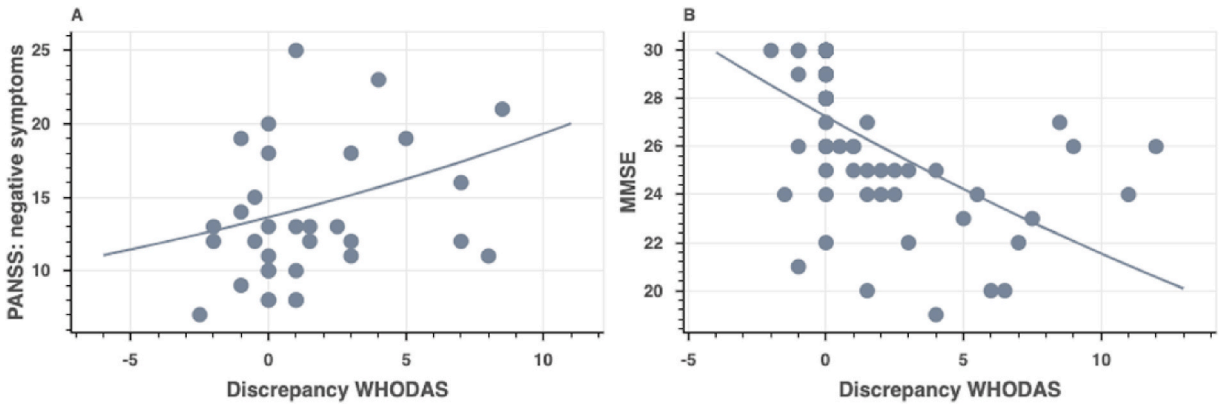

Fig. 3. (A) Negative symptoms as measured by the PANSS vs the discrepancy in WHODAS scores. Results reveal a significant positive relationship (b = 0.034, z = 2.193, p = 0.028, n = 43) which suggest that the overestimation of social functioning is associated with increased negative symptoms in SZ (B) MSSE vs the discrepancy in WHODAS scores. Results reveal a significant negative relationship (b = −0.019, z = −2.346, p = 0.019, n = 42) which suggest that increased cognitive impairment is associated with overestimation of social function in AD.

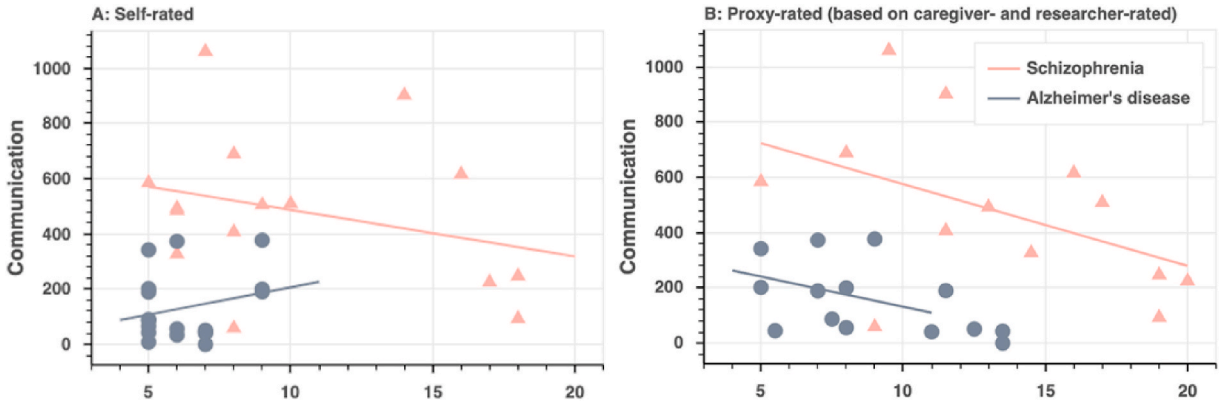

Fig. 4. Comparison of smartphone-based measure of communication and self or proxy rated WHODAS assessments. (A) Self-rated (b = −16.85, t(29) = −1.471, p = 0.076, n = 27) and (B) proxy-rated (b = −29.60, t(26) = −2.331, p = 0.014, n = 27) WHODAS plotted against the number of communication events as registered by the BEHAPP application.

related biases.

Comparisons of self-rated versus proxy 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 RCTs for negative symptoms in which social functioning is assessed in these patient populations.

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 (Doyle et al., 1999; Hayhurst et al., 2014; Pennington et al., 2019; Petersen et al., 2019) 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 (Osatuke et al., 2008). 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 (Siu et al., 2015). 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 (Ready et al., 2006). 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 (Üstün et al., 2010) 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 (Boldingh et al., 2004; Bruera et al., 2001) and psychological wellbeing (Bassett et al., 1990) 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 (Magaziner, 1997). Furthermore, it is important to note that, in the present study, the age of inclusion of AD patients reflects better the natural history of the disorder than that of SZ patients. Also, the DSM-IV criteria used in the study are based on an older version of the current manual. In addition, it is important to emphasize the lack of a control group in the present study. The reason for this is that validated measures for a fine-grained assessment of schizophrenic and/or Alzheimer's related symptoms in healthy controls is lacking within the PRISM project. With current methods such as the PANNS and the MSSE the variability on these scores is limited and therefor, a comparison between healthy controls and proxy-based assessments is limited. However, without doubt we recognize the value of such a comparison since it informs about to what extend SZ and AZ patients differ in the overestimation of their social functioning relative to healthy controls.

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.

Author statement

BWJHP, CA, NvdW, GRD, ACB, BS, HM, and MJK designed the PRISM clinical study protocol. BWJHP, CA, JLA-M, NvdW, IWvR, IMJS, AvE, SK, ACB, AR, GRD, and MJK were involved in the implementation, execution and/or patient recruitment and assessments of the clinical study. NJ performed the data analysis under the supervision of JAV, MJCE and MJK. NJ and MJK wrote the first version of the manuscript. All authors reviewed the manuscript prior to submission and their feedback was implemented to the final version of the manuscript. The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to them containing information that could compromise research participant privacy/consent. None of the experiments was preregistered.

Open practices statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to them containing information that could compromise research participant privacy/consent. None of the experiments was preregistered.

Declaration of competing interest

CA 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. BP has received (non-related) research funding from Jansen Research and Boehringer Ingelheim during the conduct of the study. MK has received (non-related) research funding from Novartis during the conduct of the study. BS was fully employed by Boehringer Ingelheim during the conduct of the study. HM was fully employed by Eli Lilly and Company during the conduct of the study. AR, ACB and GRD were fully employed by P1Vital during the conduct of the study. All other authors declare no conflict of interests.

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.

Dr. Arango has also received funding support by the Spanish Ministry of Science and Innovation. Instituto de Salud Carlos III (SAM16-PE07CP1, PI16/02012, PI19/024), co-financed by ERDF Funds from the European Commission, "A way of making Europe", CIBERSAM. Madrid Regional Government (B2017/BMD-3740 AGES-CM-2), European Union Structural Funds. Fundación Familia Alonso and Fundación Alicia Koplowitz.

References

Amador, X.F., Strauss, D.H., Yale, S.A., Flaum, M.M., Endicott, J., Gorman, J.M., 1993. Assessment of insight in psychosis. Am. J. Psychiatr. https://doi.org/10.1176/ajp.150.6.873.

Association, A.P., others, 2000. Diagnostic and Statistical Manual of Mental Disorders, 4th text revision ed. Am. Psychiatr. Assoc, Washington, DC.

Bassett, S.S., Magaziner, J., Hebel, J.R., 1990. Reliability of proxy response on mental health indices for aged, community-dwelling women. Psychol. Aging. https://doi.org/10.1037/0882-7974.5.1.127.

Bilderbeck, A.C., Penninx, B.W.J.H., Arango, C., van der Wee, N., Kahn, R., Winter-van Rossum, I., Hayen, A., Kas, M.J., Post, A., Dawson, G.R., 2019. Overview of the clinical implementation of a study exploring social withdrawal in patients with schizophrenia and Alzheimer's disease. Neurosci. Biobehav. Rev. https://doi.org/10.1016/j.neubiorev.2018.06.019.

Boldingh, E.J., Jacobs-Van Der Bruggen, M.A., Lankhorst, G.J., Bouter, L.M., 2004. Assessing pain in patients with severe cerebral palsy: development, reliability, and validity of a pain assessment instrument for cerebral palsy. Arch. Phys. Med. Rehabil. https://doi.org/10.1016/j.apmr.2003.06.029.

Bruera, E., Sweeney, C., Calder, K., Palmer, L., Benisch-Tolley, S., 2001. Patient preferences versus physician perceptions of treatment decisions in cancer care. J. Clin. Oncol. https://doi.org/10.1200/JCO.2001.19.11.2883.

Bugarski-Kirola, D., Blaettler, T., Arango, C., Fleischhacker, W.W., Garibaldi, G., Wang, A., Dixon, M., Bressan, R.A., Nasrallah, H., Lawrie, S., Napieralski, J., Ochi-Lohmann, T., Reid, C., Marder, S.R., 2017. Bitopertin in negative symptoms of

schizophrenia—results from the phase III FlashLyte and DayLyte studies. Biol. Psychiatr. https://doi.org/10.1016/j.biopsych.2016.11.014.
Carpenter, W.T., Heinrichs, D.W., Wagman, A.M.I., 1988. Deficit and nondeficit forms of schizophrenia: the concept. Am. J. Psychiatr. https://doi.org/10.1176/ajp.145.5.578.
Cramer, D., Biemer, P.P., Groves, R.M., Lyberg, L.E., Mathiowetz, N.A., Sudman, S., 1993. Measurement errors in surveys. Br. J. Sociol. https://doi.org/10.2307/591421.
Doyle, M., Flanagan, S., Browne, S., Clarke, M., Lydon, D., Larkin, C., O'Callaghan, E., 1999. Subjective and external assessments of quality of life in schizophrenia: relationship to insight. Acta Psychiatr. Scand. 99, 466–472. https://doi.org/10.1111/j.1600-0447.1999.tb00994.x.
Folstein, M.F., Robins, L.N., Helzer, J.E., 1983. The mini-mental state examination. Arch. Gen. Psychiatr. https://doi.org/10.1001/archpsyc.1983.01790060110016.
Fraguas, D., Díaz-Caneja, C.M., Pina-Camacho, L., Umbricht, D., Arango, C., 2019. Predictors of placebo response in pharmacological clinical trials of negative symptoms in schizophrenia: a meta-regression analysis. Schizophr. Bull. https://doi.org/10.1093/schbul/sbx192.
Galderisi, S., Mucci, A., Buchanan, R.W., Arango, C., 2018. Negative symptoms of schizophrenia: new developments and unanswered research questions. The Lancet Psychiatry. https://doi.org/10.1016/S2215-0366(18)30050-6.
Hayhurst, K.P., Massie, J.A., Dunn, G., Lewis, S.W., Drake, R.J., 2014. Validity of subjective versus objective quality of life assessment in people with schizophrenia. BMC Psychiatr. 14, 1–8. https://doi.org/10.1186/s12888-014-0365-x.
Kas, M.J., Penninx, B., Sommer, B., Serretti, A., Arango, C., Marston, H., 2019. A quantitative approach to neuropsychiatry: the why and the how. Neurosci. Biobehav. Rev. 97, 3–9. https://doi.org/10.1016/j.neubiorev.2017.12.008.
Magaziner, J., 1997. Use of proxies to measure health and functional outcomes in effectiveness research in persons with Alzheimer disease and related disorders. Alzheimer Dis. Assoc. Disord.
Millan, M.J., Agid, Y., Brüne, M., Bullmore, E.T., Carter, C.S., Clayton, N.S., Connor, R., Davis, S., Deakin, B., Derubeis, R.J., Dubois, B., Geyer, M.A., Goodwin, G.M., Gorwood, P., Jay, T.M., Joëls, M., Mansuy, I.M., Meyer-Lindenberg, A., Murphy, D., Rolls, E., Saletu, B., Spedding, M., Sweeney, J., Whittington, M., Young, L.J., 2012. Cognitive dysfunction in psychiatric disorders: characteristics, causes and the quest for improved therapy. Nat. Rev. Drug Discov. 11, 141–168. https://doi.org/10.1038/nrd3628.
Osatuke, K., Ciesla, J., Kasckow, J.W., Zisook, S., Mohamed, S., 2008. Insight in schizophrenia: a review of etiological models and supporting research. Compr. Psychiatr. 49, 70–77. https://doi.org/10.1016/j.comppsych.2007.08.001.
Pennington, C., Ball, H., Swirski, M., 2019. Functional cognitive disorder: diagnostic challenges and future directions. Diagnostics 9, 131. https://doi.org/10.3390/diagnostics9040131.
Petersen, J.Z., Porter, R.J., Miskowiak, K.W., 2019. Clinical characteristics associated with the discrepancy between subjective and objective cognitive impairment in depression. J. Affect. Disord. https://doi.org/10.1016/j.jad.2018.12.105.
Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y., Podsakoff, N.P., 2003. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J. Appl. Psychol. 88, 879–903. https://doi.org/10.1037/0021-9010.88.5.879.
Ready, R.E., Ott, B.R., Grace, J., 2006. Insight and cognitive impairment impairment and alzheimer ' s disease patients. Am. J. Alzheimer's Dis. Other Dementias 21, 242–248.
Reddy, M.S., 2016. Lack of insight in psychiatric illness: a critical appraisal. Indian J. Psychol. Med. https://doi.org/10.4103/0253-7176.183080.
Saris, I.M.J., Aghajani, M., van der Werff, S.J.A., van der Wee, N.J.A., Penninx, B.W.J.H., 2017. Social functioning in patients with depressive and anxiety disorders. Acta Psychiatr. Scand. 352–361. https://doi.org/10.1111/acps.12774.
Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., Dunbar, G.C., 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J. Clin. Psychiatr.
Siu, C.O., Harvey, P.D., Agid, O., Waye, M., Brambilla, C., Choi, W.K., Remington, G., 2015. Insight and subjective measures of quality of life in chronic schizophrenia. Schizophr. Res. Cogn. 2, 127–132. https://doi.org/10.1016/j.scog.2015.05.002.
Tukey, J.W., 1949. Comparing individual means in the analysis of variance. Biometrics. https://doi.org/10.2307/3001913.
Üstün, T.B., Chatterji, S., Kostanjsek, N., Rehm, J., Kennedy, C., Epping-Jordan, J., Saxena, S., von Korff, M., Pull, C., 2010. Developing the world health organization disability assessment schedule 2.0. Bull. World Health Organ. https://doi.org/10.2471/BLT.09.067231.
van der Wee, N.J.A., Bilderbeck, A.C., Cabello, M., Ayuso-Mateos, J.L., Saris, I.M.J., Giltay, E.J., Penninx, B.W.J.H., Arango, C., Post, A., Porcelli, S., 2019. Working definitions, subjective and objective assessments and experimental paradigms in a study exploring social withdrawal in schizophrenia and Alzheimer's disease. Neurosci. Biobehav. Rev. 97, 38–46. https://doi.org/10.1016/j.neubiorev.2018.06.020.