

ARTICLE OPEN



Accuracy of self-assessment of real-life functioning in schizophrenia

Paola Rocca¹✉, Claudio Brasso¹, Cristiana Montemagni¹, Silvio Bellino¹, Alessandro Rossi², Alessandro Bertolino³, Dino Gibertoni⁴, Eugenio Aguglia⁵, Mario Amore⁶, Ileana Andriola³, Antonello Bellomo⁷, Paola Bucci⁸, Antonino Buzzanca⁹, Bernardo Carpiniello¹⁰, Alessandro Cuomo¹¹, Liliana Dell'Osso¹², Angela Favaro¹³, Giulia Maria Giordano⁸, Carlo Marchesi¹⁴, Palmiero Monteleone¹⁵, Lucio Oldani¹⁶, Maurizio Pompili¹⁷, Rita Roncone¹⁸, Rodolfo Rossi², Alberto Siracusano¹⁹, Antonio Vita^{20,21}, Patrizia Zeppegno²², Silvana Galderisi⁸, Mario Maj⁸, the Italian Network for Research on Psychoses*

A consensus has not yet been reached regarding the accuracy of people with schizophrenia in self-reporting their real-life functioning. In a large ($n = 618$) cohort of stable, community-dwelling schizophrenia patients we sought to: (1) examine the concordance of patients' reports of their real-life functioning with the reports of their key caregiver; (2) identify which patient characteristics are associated to the differences between patients and informants. Patient-caregiver concordance of the ratings in three Specific Level of Functioning Scale (SLOF) domains (interpersonal relationships, everyday life skills, work skills) was evaluated with matched-pair t tests, the Lin's concordance correlation, Somers' D , and Bland-Altman plots with limits of agreement (LOA). Predictors of the patient-caregiver differences in SLOF ratings were assessed with a linear regression with multivariable fractional polynomials. Patients' self-evaluation of functioning was higher than caregivers' in all the evaluated domains of the SLOF and 17.6% of the patients exceeded the LOA, thus providing a self-evaluation discordant from their key caregivers. The strongest predictors of patient-caregiver discrepancies were caregivers' ratings in each SLOF domain. In clinically stable outpatients with a moderate degree of functional impairment, self-evaluation with the SLOF scale can become a useful, informative and reliable clinical tool to design a tailored rehabilitation program.

npj Schizophrenia (2021)7:11; <https://doi.org/10.1038/s41537-021-00140-9>

INTRODUCTION

Patients with schizophrenia show notable impairments in everyday functioning, including deficits in social, vocational, and residential domains, even during periods of remission from active psychosis¹. Many different instruments are available for the assessment of real-life functioning, including rating scales that employ informant and self-reports², direct observations by trained clinicians³, and performance-based measures⁴.

Studies have indicated that informant reports regarding the specific behaviors reflective of community functioning may be the most reliable assessment of functioning⁵. However, many people with schizophrenia do not have informants readily available to report on their functioning⁶ or they may have limited contact with them⁷. Also, in outpatient samples, there are many behaviors to which the clinician has no access and the use of self-reports may be important to get a clearer picture of the subjective level of functioning of patients. Nonetheless, self-reports of everyday functioning on the part of people with schizophrenia have been

found to be poorly correlated with the reports of other informants and with their own performance of tests of cognition and functional abilities⁸.

Different studies have investigated the accuracy of self-appraisal in both clinical populations and healthy individuals. Healthy individuals tend to overestimate their abilities. In particular, poor performers showed a particularly positive bias, i.e., a tendency to overestimate their performance^{9,10}. On the contrary, people with mild depressive symptoms tend to be more accurate in their self-evaluation¹¹, with more severe depression symptoms associated with underestimation of functioning⁷. In studies of people with neurological conditions including multiple sclerosis¹², traumatic brain injury¹³, mild cognitive impairment, and very mild Alzheimer disease¹⁴ similar results have been found: patients with poorer neuropsychological test performance tend to underestimate their impairment.

Similarly, people with schizophrenia have substantial problems in self-reporting everyday functioning⁷, as only one-third of

¹Department of Neuroscience, Section of Psychiatry, University of Turin, Turin, Italy. ²Section of Psychiatry, Department of Biotechnological and Applied Clinical Sciences, University of L'Aquila, L'Aquila, Italy. ³Department of Neurological and Psychiatric Sciences, University of Bari, Bari, Italy. ⁴Department of Biomedical and Neuromotor Sciences, University of Bologna, Bologna, Italy. ⁵Department of Clinical and Molecular Biomedicine, Psychiatry Unit, University of Catania, Catania, Italy. ⁶Section of Psychiatry, Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics and Maternal and Child Health, University of Genoa, Genoa, Italy. ⁷Psychiatry Unit, Department of Medical Sciences, University of Foggia, Foggia, Italy. ⁸Department of Psychiatry, University of Campania "Luigi Vanvitelli", Naples, Italy. ⁹Department of Neurology and Psychiatry, Sapienza University of Rome, Rome, Italy. ¹⁰Section of Psychiatry, Department of Public Health, Clinical and Molecular Medicine, University of Cagliari, Cagliari, Italy. ¹¹Department of Molecular Medicine and Clinical Department of Mental Health, University of Siena, Siena, Italy. ¹²Section of Psychiatry, Department of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy. ¹³Psychiatric Clinic, Department of Neurosciences, University of Padua, Padua, Italy. ¹⁴Department of Neuroscience, Psychiatry Unit, University of Parma, Parma, Italy. ¹⁵Department of Medicine, Surgery and Dentistry "Scuola Medica Salernitana" Section of Neuroscience, University of Salerno, Salerno, Italy. ¹⁶Department of Psychiatry, University of Milan, Milan, Italy. ¹⁷Department of Neurosciences, Mental Health and Sensory Organs, S. Andrea Hospital, Sapienza University of Rome, Rome, Italy. ¹⁸Unit of Psychiatry, Department of Life, Health and Environmental Sciences, University of L'Aquila, L'Aquila, Italy. ¹⁹Department of Systems Medicine, Psychiatry and Clinical Psychology Unit, Tor Vergata University of Rome, Rome, Italy. ²⁰Psychiatric Unit, School of Medicine, University of Brescia, Brescia, Italy. ²¹Department of Mental Health, Spedali Civili Hospital, Brescia, Italy. ²²Department of Translational Medicine, Psychiatric Unit, University of Eastern Piedmont, Novara, Italy. *A list of authors and their affiliations appears at the end of the paper. ✉email: paola.rocca@unito.it

chronic schizophrenia patients may be able to accurately report their functional abilities⁵. This is not surprising, as lack of insight is a prevalent feature of schizophrenia and is found across cultures^{15,16}, in early and late^{17,18}, acute and non-acute¹⁹ phases of the disorder. Poor insight in schizophrenia includes unawareness of symptoms, treatment need, psychosocial consequences of illness²⁰, and alterations in cognitive processes, which involve the capacity for self-reflectiveness and resistance to excessive certainty^{21,22}. This deficit is not just the consequence of a failure to notice a problem or accept a label but a failure to make consensually valid sense of complex and potentially traumatic experience, which limits patients' abilities to form integrated sense of self²³. This lack of insight likely has multiple roots, which include symptom severity, deficits in neurocognition, social cognition and metacognition, and sociopolitical factors²³, and influences patients' self-appraisal of their performance on objective tests^{24,25} and of their own levels of real-life functioning^{6,7,26}. In particular, people with schizophrenia tend, on average, to underestimate the severity of their symptoms and to overestimate their psychosocial functioning⁵.

Moreover, analyses of variables influencing misestimation of self-reported real-life functioning showed that a higher level of positive symptoms and poorer cognitive and functional capacity associated with a tendency to overestimate real-life functioning⁵. Conversely, depression showed a unique relationship with real-life functioning, as it showed both adverse impacts on functioning and positive correlations with self-assessment abilities. Indeed, more severe depressive symptoms were associated with less overestimation in self-reports⁵ and with a higher degree of underestimation⁷. Besides depression, this pattern of personal appraisal is also potentially linked to insight and stigma. About this complex interplay, several studies found that self-stigma mediated the relationship between insight and depression^{27–29}. Others showed that, beyond stigma, a generally negative appraisal of one's future influences the effects of insight on mood³⁰.

The purpose of this study was to examine the concordance of schizophrenia patients' reports of their everyday life functional status with the reports of their informant and to identify which patient characteristics were associated with disagreement in these ratings between patients and caregivers.

We hypothesized that, based on previous studies, self-reported real-word functioning by people with schizophrenia would be poorly convergent with informants' reports and that higher levels of cognitive performance and real-life functioning, rated by caregivers, would predict more concordant functioning ratings between patients and informants.

RESULTS

Study population characteristics

The study population included 618 patients followed up in the 24 centers that participated in the second wave of the Italian Network for Research on Psychoses (NIRP) study between March 2016 and January 2018. Patients were all diagnosed with schizophrenia according to DSM-IV, mostly males (69.1%), aged on average 45.1±10.5 years and received 11.7±3.4 years of education. Only a minority of them lived in residential facilities (10.1%), and 34.4% were working. Antipsychotic treatment, mainly second-generation antipsychotics (69.3%), was administered to the vast majority of patients (97.4%); 54.4% were subject to polypharmacy, 34.3% to psychosocial intervention, 14.9% to psychotherapy, and 43.7% reported a relapse during the past 4 years. Psychiatric follow-up visits were scheduled monthly on about half of the patients (48.5%), whereas 17.4% of them needed a tighter control (Table 1). Illness-related factors, functional capacity, and real-life functioning mean scores are reported in detail in a previous paper³¹.

Table 1. Socio-demographic and clinical variables of the study population.

Gender (% males)	69.1
Age (years, mean ± SD)	45.1 ± 10.5
Education (years, mean ± SD)	11.7 ± 3.4
Working (%)	34.4
Currently in a residential facility (%)	10.1
Stable affective relationships (%)	18.9
<i>Current drug treatment</i>	
Antipsychotics (%)	97.4
First-generation (%)	13.1
Second-generation (%)	69.3
Both first- and second-generation (%)	15.0
Antidepressants (%)	17.6
Mood stabilizers (%)	26.0
Anxiolytics (%)	32.7
Anticholinergics (%)	9.4
Polypharmacy (%)	54.4
Any psychosocial interventions (%)	34.3
Psychotherapy (%)	14.9
Home care (%)	8.3
Relapse during past 4 years (%)	43.7
<i>Frequency of follow-up visits (%)</i>	
Monthly	48.5
Less than monthly	34.1
More than monthly	17.4
SD standard deviation	

Comparison between patient and caregiver evaluation of real-life functioning

Patients' self-evaluation of functioning was higher than caregivers' in all the evaluated items of the SLOF scale. Their mean scores were significantly different (Table 2) for all the items of the interpersonal relationship and work skills domains, and for most items of the everyday life skills. However, even if statistical significance was achieved, the mean difference between the paired scores was usually quite small in magnitude. Percent agreement and Gwet's agreement's coefficient were always high or very high: the smallest Gwet's AC was 0.782 for the "Participates in groups" item of the interpersonal relationships domain, and it ranged 0.782–0.827 in interpersonal relationships, 0.813–0.958 in everyday life skills, and 0.793–0.866 in work skills. All probabilistic benchmark intervals were 0.600–0.800 or 0.800–1.000, thus indicating an at least substantial agreement between patient and caregiver. Also, when evaluating concordance on the three domains' scores evaluated, we found that patients overestimated their functioning with respect to caregivers in all domains (Table 3 and 4). All paired differences were significant; however, the largest ones were barely over 1 point. Lin's concordance ranged 0.766–0.874 and Somers' *D* was between 0.25 and 0.27 (indicating a 25–27% probability that functioning scores self-assessed by patients exceeded those assessed by caregivers).

The Bland–Altman plots (Fig. 1) confirmed the high concordance of patient and caregiver assessments, as most subjects were placed within the limits of agreement (LOA), for all three domains. Those who lied outside the LOA were distributed quite randomly, only slightly more prevalent in the upper part of the graph (patients' scores higher than caregivers' scores) and in the central part of the mean total range (average values of functioning). These patients overall account for 17.6% of the study population

Table 2. Mean scores and concordance between patient and caregiver assessments of SLOF items.

Items interpersonal relationships	Patient mean score	Caregiver mean score	Patient-caregiver score difference	Wilcoxon matched-pair test on caregiver mean score difference (z; p value)	Percent agreement	Gwet's AC	Gwet's AC probabilistic benchmark interval
13. Accept contact with others	3.55	3.39	0.15	5.07; <0.001	0.941	0.827	0.800–1.000
14. Initiate contact with others	3.17	3.07	0.11	3.75; <0.001	0.940	0.814	0.600–0.800
15. Communicate effectively	3.80	3.62	0.18	5.70; <0.001	0.943	0.833	0.800–1.000
16. Engages in activities without prompting	3.21	3.05	0.16	4.85; <0.001	0.938	0.799	0.600–0.800
17. Participates in groups	2.91	2.78	0.13	3.83; <0.001	0.936	0.782	0.600–0.800
18. Forms and maintains friendships	2.87	2.72	0.16	4.63; <0.001	0.938	0.802	0.600–0.800
19. Asks for help when needed	3.70	3.56	0.14	3.62; <0.001	0.935	0.805	0.600–0.800
Items everyday life skills	Patient mean score	Caregiver mean score	Patient-caregiver score difference	Wilcoxon matched-pair test on patient-caregiver mean score difference (z; p value)	Percent agreement	Gwet's AC	Gwet's AC probabilistic benchmark interval
27. Household responsibilities	3.68	3.53	0.15	5.63; <0.001	0.948	0.837	0.800–1.000
28. Shopping	3.93	3.84	0.09	3.84; <0.001	0.951	0.862	0.800–1.000
29. Handling personal finances	3.60	3.45	0.15	5.55; <0.001	0.943	0.813	0.600–0.800
30. Use of telephone	4.55	4.51	0.04	2.35; 0.019	0.971	0.952	0.800–1.000
31. Traveling from residence without getting lost	4.55	4.48	0.07	2.76; 0.006	0.975	0.958	0.800–1.000
32. Use of public transportation	4.30	4.25	0.05	2.89; 0.004	0.972	0.944	0.800–1.000
33. Use of leisure time	4.17	4.02	0.15	5.27; <0.001	0.952	0.878	0.800–1.000
34. Recognizing and avoiding common dangers	4.52	4.46	0.06	2.97; 0.003	0.972	0.949	0.800–1.000
35. Self-medication	4.17	4.00	0.17	5.64; <0.001	0.955	0.889	0.800–1.000
36. Use of medical and other community services	4.28	4.15	0.13	5.33; <0.001	0.964	0.916	0.800–1.000
37. Basic reading; writing and arithmetic	4.52	4.50	0.02	0.98; 0.330	0.967	0.940	0.800–1.000
Items work skills	Patient mean score	Caregiver mean score	Patient-caregiver score difference	Wilcoxon matched-pair test on patient-caregiver mean score difference (z; p value)	Percent agreement	Gwet's AC	Gwet's AC probabilistic benchmark interval
38. Has employable skills	3.11	2.97	0.14	4.65; <0.001	0.946	0.814	0.600–0.800
39. Works with minimal supervision	3.10	2.95	0.15	5.06; <0.001	0.941	0.793	0.600–0.800
40. Is able to sustain work effort	2.92	2.79	0.13	4.98; <0.001	0.945	0.812	0.600–0.800
	4.04	3.92	0.13	4.42; <0.001	0.950	0.866	0.800–1.000

Table 2 continued

Items work skills	Patient mean score	Caregiver mean score	Patient-caregiver mean score difference	Wilcoxon matched-pair test on patient-caregiver mean score difference (<i>z</i> ; <i>p</i> value)	Percent agreement	Gwet's AC	Gwet's AC probabilistic benchmark interval
41. Appears at appointments on time	3.90	3.78	0.12	4.56; <0.001	0.947	0.846	0.800–1.000
42. Follows verbal instructions accurately	3.82	3.66	0.17	5.87; <0.001	0.945	0.834	0.800–1.000
43. Completes assigned tasks							

Table 3. Mean scores and concordance between patient and caregiver assessments of SLOF domains.

	Patient mean score	Caregiver mean score	Patient-caregiver mean score difference (95% CI)	Paired <i>t</i> test on patient-caregiver mean score difference (<i>t</i> ; <i>p</i> value)	Lin's concordance correlation	Somers' <i>D</i> (95% CI)
Interpersonal relationships	23.23	22.19	1.04 (0.72–1.36)	6.36; <0.001	0.766	0.250 (0.185–0.312)
Everyday life skills	46.22	45.19	1.03 (0.66–1.39)	5.50; <0.001	0.874	0.270 (0.209–0.328)
Work skills	20.90	20.07	0.83 (0.56–1.11)	6.00; <0.001	0.834	0.265 (0.204–0.323)

(*n* = 83); 13.4% exceeded LOA in only one SLOF domain, 2.1% exceeded LOAs in two SLOF domains, and 2.1% in all the three SLOF domains.

Predictors of patient-caregiver discrepancies in the evaluation of real-life functioning

From the regression analyses performed using multiple imputation and multivariable fractional polynomials (MFPs, Table 4 and Fig. 2), we found that caregiver scores in each of the three domains of the SLOF analyzed were the strongest predictors of patient-caregiver discrepancies with negative coefficients. This indicates that patients' overestimation was related to low caregiver's scores; patient's estimation was less discrepant when caregiver scores were higher and for the Interpersonal Relationships and Work Skills domains at higher caregiver scores corresponded underestimation of patients. Avolition was associated with the patient-caregiver discrepancy in the evaluation of interpersonal relationships and, to a lesser extent, of work skills: more severe avolition was associated with more precise patient's self-evaluations. Expressive deficits were associated with patient-caregiver discrepancy in work skills, positive symptoms with patient-caregiver discrepancy in interpersonal relationships, and disorganization symptoms with patient-caregiver discrepancy in everyday life skills domain with a similar negative relation, i.e., more severe symptoms were associated with patient's self-evaluations closer to caregivers' ones. Finally, higher speed of processing in the MCCB was poorly associated with a higher level of patient's overestimation in the work skills domain.

DISCUSSION

This study aimed at two main goals: (a) to assess the concordance between real-world functioning self-reported by people with schizophrenia and reports generated by informants; (b) to evaluate predictors of the agreement between these two evaluations.

To our knowledge, this is the largest study carried out so far examining awareness of functional deficits in people with schizophrenia. Subjects participating in this study were living in the community and stabilized on antipsychotic treatment and were evaluated for their everyday life skills, interpersonal abilities, and work skills.

Concerning the first aim, our results demonstrated good concordance between self-reported functioning and informant reports, contrary to our research hypothesis. About 5/6 of patients self-evaluated their real-life functioning in concordance with their caregivers' estimation, whereas only 17.6% of patients were poorly concordant, mostly on one single domain of functioning. These data suggest that patients had few problems in reporting their real-life functioning and that self-report could be informative and should be considered to design rehabilitation projects.

This finding is in contrast with previous studies showing that more than half of patients were inaccurate raters of their abilities^{5,7,8,32}. These contrasting results may be owing to differences in the samples and in the types of informants. First, our patients were community-dwelling clinically stable patients, with a good service engagement and a moderate degree of functional impairment³¹. Moreover, about 2/3 of them underwent scheduled follow-up visits at least once every month for 4 years, which suggests a stable course of the illness and confirms a high service engagement that might have contributed to our patients' good awareness about their own real-life functioning and to the high concordance with informants' ratings. In other terms, as shown by a recent study presented by Vohs et al.³³, we can assume that this meaningful engagement with mental health services could improve patients' ability to appraise the consequences of the disorder they suffer from. Moreover, this

Table 4. Multivariable regressions of patient-caregiver discrepancies in SLOF domains.

	Patient-caregiver discrepancies in SLOF domains		
	Interpersonal relationships (n = 613)	Everyday life skills (n = 608)	Work skills (n = 612)
Caregiver score	−0.295 (−0.351; −0.239) <0.001	−0.244 (−0.288; −0.200) <0.001	−0.182 (−0.219; −0.144) <0.001*
PANSS positive	−0.116 (−0.190; −0.042) 0.002		
PANSS disorganization		−0.798 (−1.171; −0.425) <0.001*	
BNSS avolition	−0.264 (−0.373; −0.156) <0.001		−0.185 (−0.276; −0.094) <0.001
BNSS expressive deficits		−0.529 (−0.809; −0.249) <0.001*	
Processing speed			0.019 (0.006; 0.031) 0.005
Constant	1.042 (0.746; 1.339) <0.001	1.602 (1.223; 1.981) <0.001	1.233 (0.965; 1.502) <0.001

Note: multivariable regressions performed using multiple imputations and multivariable fractional polynomials (MFP). Linear regression coefficients are displayed with 95% confidence intervals in brackets and *p* values. * denotes cubic regression coefficients.

enhanced level of cognitive and clinical insight could have led a considerable portion of the patients in our sample to recovery, intended as a meaningful interpretation of the challenges the disorder poses and a possibility to meaningfully live with schizophrenia³⁴. As regards the informant, in our study and in the Italian context it was mostly a key caregiver, the individual most frequently and closely in contact with the patient, which shares many time and activities with patients. This may have led to a greater concordance of their viewpoints about patients' real-life functioning, not in line with previous studies that chose the high contact clinical as informant^{5,7,8,32}. We hypothesize that this type of social context influences both the patient, who is constantly made aware of her/his functional limits by the caregiver and the caregiver, who is able to accurately evaluate the patient's functioning in the wake of her/his continuous relationship with the patient. Moreover, as they share similar roles in patients' lives and therefore similar point of view, the choice of key caregivers as informants entails a high level of homogeneity among different caregivers' evaluations. Accordingly, unlike other studies where informants' evaluations were performed by non-clinicians who have different roles in the patients' lives (e.g., key caregivers, close relatives, and friends simultaneously in the same sample), we did not expect a high degree of variability depending on the informants' role in the patients' lives.

About the second aim, different types of errors may occur when asking people with schizophrenia to self-rate their functional skills, including misunderstanding the items on a rating scale, inaccurately conceptualizing normal functioning, and making inaccurate comparisons to external standards. These types of errors could derive from cognitive deficits, from the severity of specific symptomatic dimensions, and from the socio-demographic characteristics of the patients. Identifying which variables were associated with self-rated real-life functioning misestimation can be useful to determine which patients are more likely to give a biased self-evaluation.

The strongest predictors of patient-caregiver discrepancies in all the three SLOF domains were caregivers' scores. This is an expected result, because the magnitude of patient-caregiver disagreement is at least partly constrained by the reference value represented by the caregiver score; as such, in our regression models, the caregiver score was essential to adjust the predictors' coefficient estimates. Patients who received higher ratings from caregivers showed better concordance with the caregiver's evaluation. This finding is consistent with our research hypothesis and with previous studies showing that patients who have poorer functioning assessed by the informant tend to overestimate their own functioning^{5,7,8}. However, in our study inaccuracy, measured by Bland–Altman LOA, affected only 1/6 of the patients, which is a much smaller proportion compared with previous studies.

As to symptoms, at odds with previous studies^{5,7}, we did not find any association between self-rating of real-world functioning and depressive symptoms; instead we found that the avolition dimension of negative symptoms was the main predictor after the caregivers' scores, with more severe avolition associated with a better concordance or even with an underestimation of the SLOF Interpersonal domain functioning in patient's self-ratings. We hypothesize that these results may depend on the different instruments utilized to assess depression and negative symptoms: while we utilized an objective (based on clinicians' observations during the interview) schizophrenia specific depression rating scale, the Calgary Depression Scale for Schizophrenia³⁵ (CDSS), previous studies employed a self-reported scale, namely the Beck Depression Inventory version I or II^{5,7,32,36,37}. Moreover, the BDI-II includes items concerning anhedonia, asociality and avolition, which overlap with the experiential domain of negative symptoms that we evaluated with the Brief Negative Symptoms Scale (BNSS)³⁸. Similarly, according to the current conceptualization of negative symptoms³⁹, we assessed negative symptoms with a specific second-generation assessment scales, the BNSS, considering two factors: avolition and expressive deficits^{40,41}, whereas previous studies did not employ specific scales and considered negative symptoms as a single variable^{7,32,37,42–44}. These differences in clinical assessment could at least in part account for our contrasting results compared with previous works. We suppose that the avolition dimension of the BNSS, as experiential domain of negative symptoms including avolition, anhedonia, and asociality, similarly to depression in other studies^{45,46}, can support a self-evaluation of real-life functioning similar to the one provided by the caregiver. In other words, a higher degree of avolition could enhance a more "objective" perception of patients' limits of functioning.

The other significant predictors of discrepancies in the SLOF dimensions evaluation were illness-or cognitive-related factors that generally had weak associations with the outcome (cf. Table 3). As we found high concordance in all domains, a weak impact of predictors on disagreement might be expected, because there is little variance of disagreement to be explained, especially after adjusting for the caregivers' score. Moreover, the large size of the study population might have enhanced statistical significance even for relatively small effects, which may also be influenced by the presence of few subjects with extreme values of discrepancy. For all these reasons, the substantive significance of these associations seems quite poor and should not be emphasized. Thus, the absent or weak association of cognitive performance with disagreement on functioning assessment was an unexpected result that disconfirmed our research hypothesis.

The patients included in the present study were outpatients with stable symptoms, moderate degree of functional impairment, and a strong and stable relationship with mental services and their

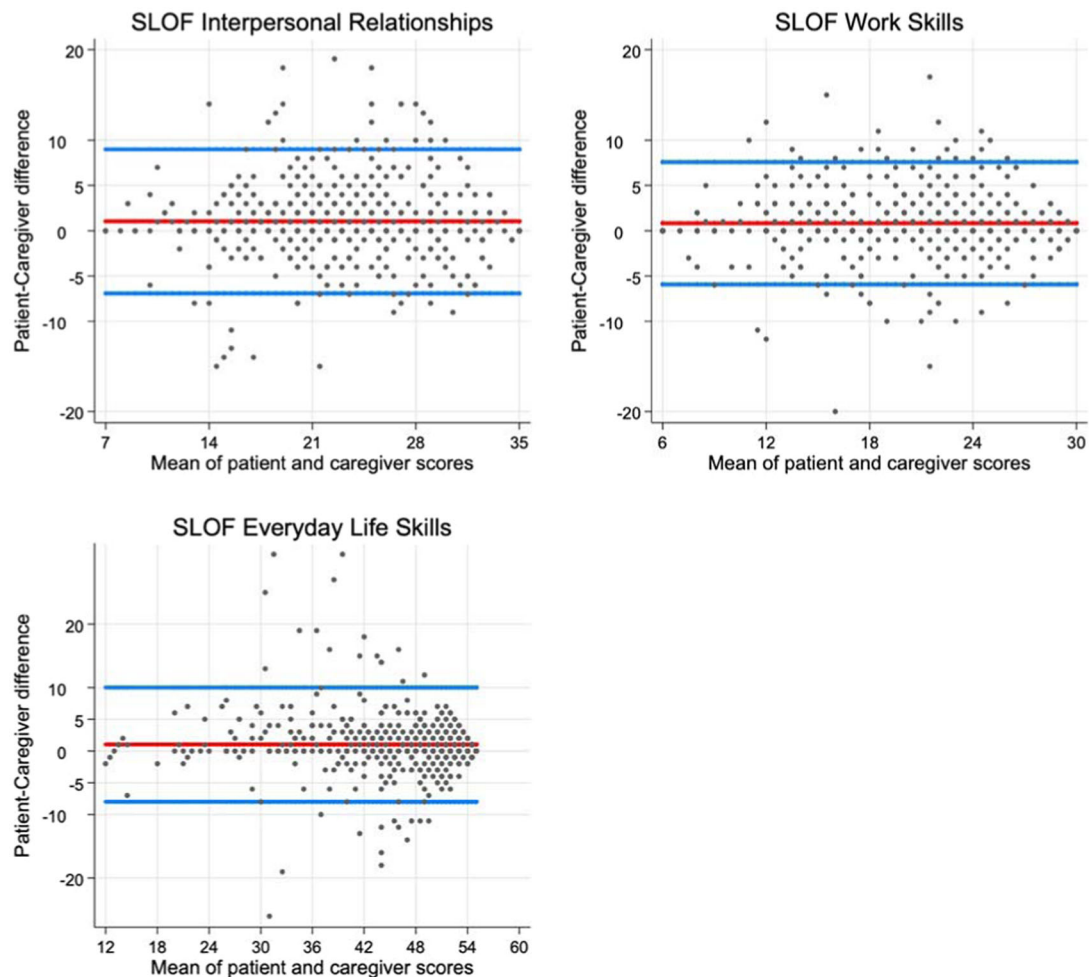


Fig. 1 Bland–Altman plots of patient-caregiver agreement on SLOF domains. Note: in these Bland–Altman plots circles represent patients; the upper and lower blue lines represent the 95% limits of agreement; the red line represents the average discrepancy between patients and caregivers. Cases lying between the blue lines are those whose patient and caregiver scores were with 95% probability concordant; those lying over the upper blue line are those whose self-reported functioning score was significantly higher than the corresponding score attributed by the caregiver; cases lying below the upper blue line are those whose self-reported functioning score was significantly lower than the corresponding score attributed by the caregiver. In the x axes, the mean of patient and caregiver scores is assumed as the more likely value of patient's functioning.

caregiver. These characteristics could have helped in achieving a high level of agreement among patients' and caregivers' evaluation. Therefore, the main limitation of the present study is that our results may not be reproducible in patients in acute phases, clinically unstable, or assessed in other clinical settings or in social context where a key caregiver is absent. In addition, the high degree of agreement made it difficult to identify predictors of patient-caregiver discrepancy, beside the functioning level itself.

Despite this limitation, this study has some important strengths: the large sample size, the naturalistic design without selection bias related to randomized controlled designs, the use of state-of-the-art statistical analysis and instruments to assess psychopathology, cognition, functional capacity, and real-world functioning.

In our sample, people with schizophrenia showed a good agreement with the caregivers' ratings of the SLOF scale. It can be assumed that this good level of concordance is partly owing to sample characteristics: clinically stable, non-hospitalized patients, with moderate degree of functional impairment, who often share background and life context with their caregivers. Moreover, a good service engagement with continuative community-based psychiatric care might enhance patients' metacognitive reflection, insight, and ability to self-appraise their real-life functioning. In this

view, in contexts in which community-based mental health care is provided, and in outpatients with clinically stable schizophrenia and a good engagement with mental health services, self-evaluation with the SLOF scale can become a useful, informative and reliable clinical tool to design a tailored rehabilitation program.

METHODS

Participants

We used the 4-year follow-up database of the NIRP study, involving at the baseline 921 community-dwelling, clinically stable patients^{47,48}. Twenty-four out of the 26 Italian university psychiatric clinics and/or mental health departments that joined the baseline study participated in the follow-up. All patients recruited by the 24 participating centers for the baseline study were invited to participate.

Patients with a diagnosis of schizophrenia according to the DSM-IV and confirmed with the Structured Clinical Interview for DSM-IV-Patient version (SCID-I-P) were included in the follow-up study. Exclusion criteria were (a) a history of head trauma with loss of consciousness in the 4-year interval between the baseline and the follow-up assessments; (b) progressive cognitive deterioration possibly owing to dementia or other neurological illness diagnosed in the last 4 years; (c) a history of alcohol and/or substance abuse in the last six months; (d) current pregnancy or lactation;

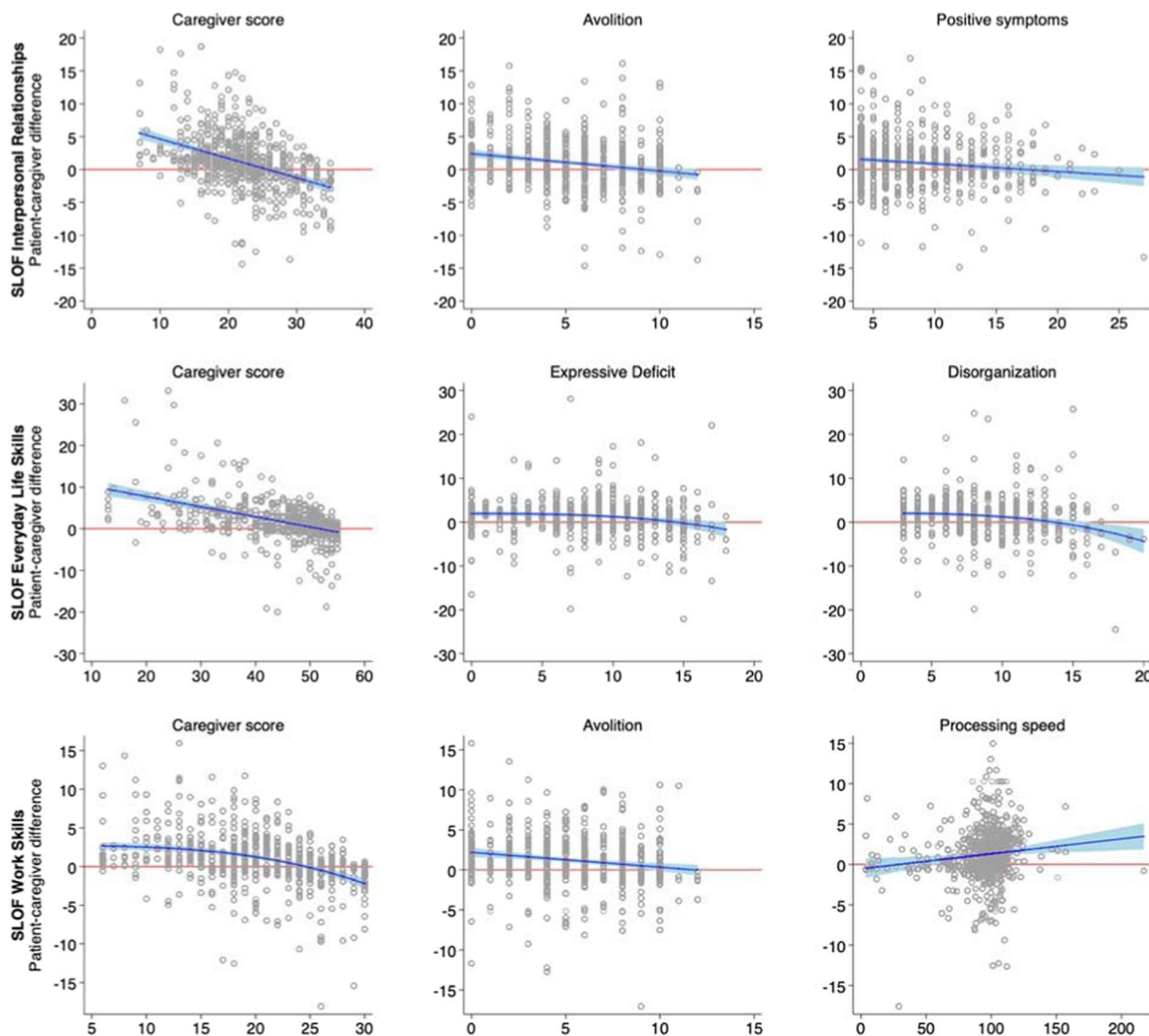


Fig. 2 Scatterplots of patient-caregiver differences in the three SLOF domains and their significant predictors, with the estimated regression curve. Note: in these plots, circles represent patients, whose coordinates are given by patient-caregiver difference (Y axis) and values of its significant predictors according to multivariable fractional polynomial regression (X axis). The blue line represents the estimated regression curve, with 95% confidence interval as shaded area.

(e) inability to provide informed consent, and (f) treatment modifications and/or hospitalization owing to symptom exacerbation in the last three months.

For the participants in the baseline study who could not be traced or were deceased, investigators had to fill in a specific form reporting clinical information available at the last contact and, if available, the cause of death in case of decease. After receiving a comprehensive explanation of the study procedures and goals, a written informed consent to participate in the follow-up procedures was asked to all patients. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. All procedures involving human patients were approved by local Ethics Committees of the participating centers, and recruitment was carried out from March 2016 to December 2017.

Procedures

The assessments of the enrolled patients were completed following the same schedule used in the baseline study⁴⁷: (1) socio-demographic information, psychopathological, and neurological assessments on the first

day; (2) neurocognitive, social cognition, and functional capacity assessments on the second day; (3) according to the patient's preference, assessment of personal resources and perceived stigma were carried out on the third day morning, or in the afternoon of any of the days. For real-life functioning assessment, the patient's key caregiver, was invited to join one of the scheduled sessions.

Evaluation of illness-related factors

A clinical form was filled in with information on disease course and treatments in the previous 4 years, using all available sources of information (patients, relatives, medical records, and mental health workers).

The Positive and Negative Syndrome Scale (PANSS)⁴⁹ was used to rate symptom severity. Positive symptoms were assessed using four items of the PANSS: P1 (delusions), P3 (hallucinatory behavior), P5 (grandiosity), G9 (unusual thought content). Disorganization was assessed using three items of the PANSS scale: P2 (conceptual disorganization), N5 (difficulty in abstract thinking), and G11 (poor attention). We used the consensus 5-factor solution proposed by Wallwork et al.⁵⁰; for both dimensions, higher scores represent greater symptom severity. Negative symptoms were

assessed using the Italian version of the Brief Negative Symptom Scale (BNSS)³⁸; the scores of the two domains Avolition (sum of anhedonia, asociality, and avolition) and expressive deficit (sum of blunted affect and alogia) were used in statistical analyses (higher scores correspond to greater severity)^{40,41}. Depressive symptoms were evaluated by the CDSS³⁵; the total score was used in data analyses (higher scores correspond to greater severity of depression).

The Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Consensus Cognitive Battery (MCCB)⁵¹ was used to assess the following neurocognitive domains: processing speed, attention/vigilance, working memory, verbal learning, visual learning, and reasoning and problem solving (for all domains, higher scores represent better cognitive functioning). The assessment of social cognition was partly included in the MCCB (Mayer-Salovey-Caruso Emotional Intelligence Test, MSCEIT, managing emotion section)⁵², which examines the regulation of emotions in oneself and in one's relationships with others by presenting vignettes of various situations, along with ways to cope with the emotions depicted in these vignettes. This test was integrated by the Facial Emotion Identification Test⁵³, measuring emotion recognition, and The Awareness of Social Inference Test (TASIT)⁵⁴, which is a theory of mind (ToM) test consisting of seven scales (positive emotions, negative emotions, sincere, simple sarcasm, paradoxical sarcasm, sarcasm enriched, lie), organized into three sections: emotion recognition; social inference (minimal); social inference (enriched).

Assessment of personal resources

The Resilience Scale for Adults⁵⁵, a self-administered scale, was used to assess perception of self, perception of the future, social competence and family cohesion (higher scores correspond to higher resilience). The Service Engagement Scale⁵⁶ was employed to measure patients' levels of difficulty to engage with mental health services (higher total score represents greater difficulty).

Evaluation of context-related factors

The Internalized Stigma of Mental Illness⁵⁷ questionnaire evaluated the experience of internalized stigma (higher total score corresponds to greater experience of internalized stigma). The number of incentives was registered as a count variable, ranging from 0 to 4, and includes the availability of a disability pension, access to family practical and financial support, and registration in the unemployment list.

Assessment of functional capacity and real-life functioning

The short version of the University of California San Diego (UCSD) Performance-based Skills Assessment Brief (UPSA-B)⁵⁸ was used to assess functional capacity. It assesses "financial skills" and "communication skills". Participants receive scaled scores for each of the subscales (range = 0–50), which are summed to create an overall score ranging from 0 to 100. Higher scores indicate better functional capacity.

Real-life functioning was evaluated using the Specific Level of Functioning Scale (SLOF, Italian version)^{59–61} a hybrid instrument that explores many aspects of functioning. It consists of a 43-item self- or informant-rated scale of a person's behavior and functioning, which assesses the following domains: physical efficiency, skills in self-care, interpersonal relationships, social acceptability, everyday life skills (e.g., shopping, using public transportation), and working skills. Each of the 43 items are rated on a five-point Likert scale, indicating the level of assistance the participant needs to perform the task, with higher score indicating better functioning. The SLOF scale differs from the other outcome measures in emphasizing patient's current functioning and observable behavior, as opposed to inferred mental or emotional states, and focuses on a person's skills, assets, and abilities rather than deficits. Moreover, the SLOF does not include items relevant to psychiatric symptomatology or cognitive dysfunctions. SLOF interpersonal relationships, everyday life skills, and work skills domains were included in statistical analyses. In this follow-up study, the patients her/himself and her/his key caregiver separately answered to the 43 SLOF items. The patient's key caregiver, preferably the same interviewed in the baseline study, was chosen as informer as usually, this is the individual most frequently and closely in contact with the patient in the Italian context.

Training of researchers and inter-rater reliability

Researchers were trained by the coordinating center two months before the beginning of the follow-up recruitment to ensure consistency with the baseline data collection procedures. The inter-rater reliability was evaluated by Cohen's kappa for categorical variables, and intraclass correlation coefficient (ICC) for continuous variables. For items showing a small degree of variation among patients, the percentage of perfect agreement was calculated. An excellent inter-rater agreement was found for the SCID-I-P (Cohen's kappa = 0.91). Good to excellent agreement was observed for BNSS (ICC = 0.74–0.97), PANSS (ICC = 0.60–0.98, percentage agreement = 64–100%), CDSS (ICC = 0.76–0.98), and MCCB (ICC = 0.98). Further details can be found in Galderisi et al.³¹.

Statistical analysis

The characteristics of the study population were summarized by reporting continuous variables as mean ± standard deviation and categorical variables as percentages. Patient-caregiver agreement was first calculated on each of the 24 items belonging to the interpersonal, activities, and work domains of the SLOF scale, and subsequently on these three domains. As the SLOF items are ordinal variables on the 1–5 range, agreement on the items' level was assessed by reporting the patients' and caregivers' mean score and their difference, the two-sided Wilcoxon matched-pair test of the null hypothesis of equality of means, the percentage agreement, the Gwet's agreement coefficient (AC) and its related probabilistic benchmark interval. Gwet's AC coefficient is recently preferred to Cohen's kappa family of coefficients because it has been demonstrated to be more robust, especially in the presence of skewed data, and to be able to avoid the paradox of negative agreement^{62–65}. Moreover, Gwet's ACs are categorized into the scheme provided by Landis and Koch⁶⁶ as slight (0.00–0.20), fair (0.21–0.40), moderate (0.41–0.60), substantial (0.61–0.80), and almost perfect (0.81–1.00) using a probabilistic assignment, that takes into account the variance of the estimate³⁸. We applied ordinal weighting to Gwet's AC calculation in order to assign an increasing penalty when disagreement between raters was larger. The three domains of the SLOF are discrete-continuous variables, therefore for these variables the patient-caregiver agreement was assessed by reporting the patients' and caregivers' mean score and their difference, the two-sided matched-pair *t* test of the null hypothesis of equality of means, the Lin's concordance correlation and Somers's *D*, which assigns the probability that patients are less or more likely to overestimate their caregiver's evaluation on the same domain. The Bland-Altman plots and LOA for each domain were also obtained, to understand whether disagreement between patient and caregiver occurred at particular values along the domains' score range.

Finally, to investigate which characteristics could predict disagreement, we performed multivariable linear regressions of the patient-caregiver score difference for each SLOF domain on a set of covariates including gender, age, education, positive symptoms, disorganization, avolition, expressive deficit, depression, the six MCCB items of neurocognition, the five items of social cognition and functional capacity, adjusted for the caregiver's score. These regression models were carried out using MFP on 10 imputed data sets obtained through multiple imputation of predictors' missing data using chained equations. MFP procedure verified which functional form (linear or nonlinear, i.e., quadratic, cubic, square root, etc.) of each variable best represented its relationship with the outcome, and simultaneously selected those significantly associated to the outcome. Stata 15.1 was used for all analyses, specifically the user-written procedures `kappaetc`⁶³, `concord`⁶⁷, `scomersd`⁶⁸, and `mfpmi`⁶⁹ allowed to estimate Gwet's AC, Lin's concordance correlation, Somers' *D* and MFPs, respectively.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available as they contain information that could compromise the privacy of research participants.

CODE AVAILABILITY

The code that supports the findings of this study is publicly available at: https://github.com/dingib/NIRP/blob/main/SLOF_concordance.do

Received: 4 October 2020; Accepted: 12 January 2021;
Published online: 15 February 2021

REFERENCES

- Leung, W. W., Bowie, C. R. & Harvey, P. D. Functional implications of neuropsychological normality and symptom remission in older outpatients diagnosed with schizophrenia: a cross-sectional study. *J. Int. Neuropsychol. Soc.* **14**, 479–488 (2008).
- Leifker, F. R., Patterson, T. L., Heaton, R. K. & Harvey, P. D. Validating measures of real-world outcome: the results of the VALERO expert survey and RAND panel. *Schizophr. Bull.* **37**, 334–343 (2011).
- Kleinman, L. et al. Development and psychometric performance of the schizophrenia objective functioning instrument: an interviewer administered measure of function. *Schizophr. Res.* **107**, 275–285 (2009).
- Harvey, P. D., Velligan, D. I. & Bellack, A. S. Performance-based measures of functional skills: usefulness in clinical treatment studies. *Schizophr. Bull.* **33**, 1138–1148 (2007).
- Sabbag, S. et al. Predictors of the accuracy of self-assessment of everyday functioning in people with schizophrenia. *Schizophr. Res.* **137**, 190–195 (2012).
- Patterson, T. L., Semple, S. J., Shaw, W. S., Grant, I. & Jeste, D. V. Researching the caregiver: family members who care for older psychotic patients. *Psychiatr. Ann.* **26**, 772–784 (1996).
- Bowie, C. R. et al. Self-assessment of functional status in schizophrenia. *J. Psychiatr. Res.* **41**, 1012–1018 (2007).
- Sabbag, S. et al. Assessing everyday functioning in schizophrenia: not all informants seem equally informative. *Schizophr. Res.* **131**, 250–255 (2011).
- Dunning, D. & Story, A. L. Depression, realism, and the overconfidence effect: are the sadder wiser when predicting future actions and events? *J. Pers. Soc. Psychol.* **61**, 521–532 (1991).
- Ehrlinger, J., Johnson, K., Banner, M., Dunning, D. & Kruger, J. Why the unskilled are unaware: further explorations of (absent) self-insight among the incompetent. *Org. Behav. Hum. Decis. Process.* **105**, 98–121 (2008).
- Alloy, L. B. & Abramson, L. Y. Judgment of contingency in depressed and non-depressed students: sadder but wiser? *J. Exp. Psychol. Gen.* **108**, 441–485 (1979).
- Carone, D. A., Benedict, R. H. B., Iii, F. M., Fishman, I. & Weinstock-Guttman, B. Interpreting patient/informant discrepancies of reported cognitive symptoms in MS. *J. Int. Neuropsychol. Soc.* **11**, 574 (2005).
- Spikman, J. M. & van der Naalt, J. Indices of impaired self-awareness in traumatic brain injury patients with focal frontal lesions and executive deficits: implications for outcome measurement. *J. Neurotrauma* **27**, 1195–1202 (2010).
- Ryu, S. Y. et al. Self-and informant-reported cognitive functioning and awareness in subjective cognitive decline, mild cognitive impairment, and very mild Alzheimer disease. *Int. J. Geriatr. Psychiatry* **35**, 91–98 (2020).
- Schennach, R. et al. Insight in schizophrenia course and predictors during the acute treatment phase of patients suffering from a schizophrenia spectrum disorder. *Eur. Psychiatry* **27**, 625–633 (2012).
- Wang, Y. et al. Insight in Chinese schizophrenia patients: a 12-month follow-up. *J. Psychiatr. Ment. Health Nurs.* **18**, 751–757 (2011).
- Koren, D. et al. The nature and evolution of insight in schizophrenia: a multi-informant longitudinal study of first-episode versus chronic patients. *Schizophr. Res.* **151**, 245–251 (2013).
- Vohs, J. L. et al. Metacognitive capacity as a predictor of insight in first-episode psychosis. *J. Nerv. Ment. Dis.* **203**, 372–378 (2015).
- Braw, Y. et al. Comparison of insight among schizophrenia and bipolar disorder patients in remission of affective and positive symptoms: analysis and critique. *Eur. Psychiatry* **27**, 612–618 (2012).
- Amador, X. F. et al. Awareness of illness in schizophrenia and schizoaffective and mood disorders. *Arch. Gen. Psychiatry* **51**, 826–836 (1994).
- Martin, J. M., Warman, D. M. & Lysaker, P. H. Cognitive insight in non-psychiatric individuals and individuals with psychosis: an examination using the Beck Cognitive Insight Scale. *Schizophr. Res.* **121**, 39–45 (2010).
- Riggs, S. E. et al. Assessment of cognitive insight: a qualitative review. *Schizophr. Bull.* **38**, 338–350 (2012).
- Lysaker, P. H., Pattison, M. L., Leonhardt, B. L., Phelps, S. & Vohs, J. L. Insight in schizophrenia spectrum disorders: relationship with behavior, mood and perceived quality of life, underlying causes and emerging treatments. *World Psychiatry* **17**, 12–23 (2018).
- Medalia, A. & Thysen, J. A comparison of insight into clinical symptoms versus insight into neuro-cognitive symptoms in schizophrenia. *Schizophr. Res.* **118**, 134–139 (2010).
- Silberstein, J. M., Pinkham, A. E., Penn, D. L. & Harvey, P. D. Self-assessment of social cognitive ability in schizophrenia: association with social cognitive test performance, informant assessments of social cognitive ability, and everyday outcomes. *Schizophr. Res.* **199**, 75–82 (2018).
- McKibbin, C., Patterson, T. L. & Jeste, D. V. Assessing disability in older patients with schizophrenia: results from the WHODAS-II. *J. Nerv. Ment. Dis.* **192**, 405–413 (2004).
- Belvederi Murri, M. et al. The “insight paradox” in schizophrenia: magnitude, moderators and mediators of the association between insight and depression. *Schizophr. Bull.* **42**, 1225–1233 (2016).
- Lien, Y. J. et al. Insight, self-stigma and psychosocial outcomes in schizophrenia: a structural equation modeling approach. *Epidemiol. Psychiatr. Sci.* **15**, 1–10 (2016).
- Valiente, C. et al. Insight in paranoia: the role of experiential avoidance and internalized stigma. *Schizophr. Res.* **164**, 214–220 (2015).
- MacDougall, A. G., Vandermeer, M. R. & Norman, R. M. Negative future self as a mediator in the relationship between insight and depression in psychotic disorders. *Schizophr. Res.* **165**, 66–69 (2015).
- Galderisi, S. et al. The interplay among psychopathology, personal resources, context-related factors and real-life functioning in schizophrenia: stability in relationships after 4 years and differences in network structure between recovered and non-recovered patients. *World Psychiatry* **19**, 81–91 (2020).
- Gould, F. et al. Self-assessment in schizophrenia: accuracy of evaluation of cognitive and everyday functioning. *Neuropsychology* **29**, 675 (2015).
- Vohs, J. L. et al. Metacognitive reflection and insight therapy for early psychosis: a preliminary study of a novel integrative psychotherapy. *Schizophr. Res.* **195**, 428–433 (2018).
- Leonhardt, B. L. et al. Recovery and serious mental illness: a review of current clinical and research paradigms and future directions. *Expert Rev. Neurother.* **17**, 1117–1130 (2017).
- Addington, D., Addington, J. & Maticka-Tyndale, E. Assessing depression in schizophrenia: the Calgary Depression Scale. *Br. J. Psychiatry* **163**, 39–44 (1993).
- Beck, A. T., Robert, A. S. & Gregory, K. B. Manual for the beck depression inventory-II. *Psychological Corporation* **1**, 82 (1996).
- Durand, D. et al. Factors influencing self-assessment of cognition and functioning in schizophrenia: implications for treatment studies. *Eur. Neuropsychopharmacol.* **25**, 185–191 (2015).
- Mucci, A. et al. The Brief Negative Symptom Scale (BNSS): independent validation in a large sample of Italian patients with schizophrenia. *Eur. Psychiatry* **30**, 641–647 (2015).
- Marder, S. R. & Galderisi, S. The current conceptualization of negative symptoms in schizophrenia. *World Psychiatry* **16**, 14–24 (2017).
- Kirkpatrick, B. et al. The Brief Negative Symptom Scale: psychometric properties. *Schizophr. Bull.* **37**, 300–305 (2011).
- Strauss, G. P. et al. Next-generation negative symptom assessment for clinical trials: validation of the Brief Negative Symptom Scale. *Schizophr. Res.* **142**, 88–92 (2012).
- Ermel, J. et al. Self versus informant reports on the specific levels of functioning scale: relationships to depression and cognition in schizophrenia and schizoaffective disorder. *Schizophr. Res. Cogn.* **9**, 1–7 (2017).
- Vidarsdottir, O. G. et al. Social and non-social measures of cognition for predicting self-reported and informant-reported functional outcomes in early psychosis. *Scand. J. Psychol.* **60**, 295–303 (2019).
- Simonsen, C. et al. Self-rated disability in first treated episode of psychosis: a 1-year follow-up study. *Compr. Psychiatry* **85**, 48–54 (2018).
- Misdrabi, D. et al. Depression in schizophrenia: the influence of the different dimensions of insight. *Psychiatry Res.* **216**, 6–12 (2014).
- Grover, S. et al. Relationship of depression with cognitive insight and socio-occupational outcome in patients with schizophrenia. *Int. J. Soc. Psychiatry* **6**, 181–194 (2017).
- Galderisi, S. et al. The influence of illness-related variables, personal resources and context-related factors on real-life functioning of people with schizophrenia. *World Psychiatry* **13**, 275–287 (2014).
- Galderisi, S. et al. Interplay among psychopathologic variables, personal resources, context-related factors, and real-life functioning in individuals with schizophrenia: a network analysis. *JAMA Psychiatry* **75**, 396–404 (2018).
- Kay, S. R., Fiszbein, A. & Opler, L. A. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr. Bull.* **13**, 261–276 (1987).
- Wallwork, R. S., Fortgang, R., Hashimoto, R., Weinberger, D. R. & Dickinson, D. Searching for a consensus five-factor model of the Positive and Negative Syndrome Scale for schizophrenia. *Schizophr. Res.* **137**, 246–250 (2012).
- Nuechterlein, K. H. et al. The MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. *Am. J. Psychiatry* **165**, 203–213 (2008).

52. Mayer J. D. & Salovey P., Caruso D. R. *Mayer-Salovey-Caruso emotional intelligence test (MSCEIT) item booklet*. UNH Personality Lab. 26, https://scholars.unh.edu/personality_lab/26 (2002).
53. Kerr, S. L. & Neale, J. M. Emotion perception in schizophrenia: specific deficit or further evidence of generalized poor performance? *J. Abnorm. Psychol.* **102**, 312 (1993).
54. McDonald, S. et al. Reliability and validity of The Awareness of Social Inference Test (TASIT): a clinical test of social perception. *Disabil. Rehabil.* **28**, 1529–1542 (2006).
55. Friberg, O., Hjemdal, O., Rosenvinge, J. H. & Martinussen, M. A new rating scale for adult resilience: what are the central protective resources behind healthy adjustment. *Int. J. Methods Psychiatr. Res.* **12**, 65–76 (2003).
56. Tait, L., Birchwood, M. & Trower, P. A new scale (SES) to measure engagement with community mental health services. *J. Ment. Health* **11**, 191–198 (2002).
57. Ritscher, J. B., Otilingam, P. G. & Grajales, M. Internalized stigma of mental illness: psychometric properties of a new measure. *Psychiatry Res.* **121**, 31–49 (2003).
58. Mausbach, B. T., Harvey, P. D., Goldman, S. R., Jeste, D. V. & Patterson, T. L. Development of a brief scale of everyday functioning in persons with serious mental illness. *Schizophr. Bull.* **33**, 1364–1372 (2007).
59. Montemagni, C., Rocca, P., Mucci, A., Galderisi, S. & Maj, M. Italian version of the “Specific Level of Functioning”. *J. Psychopathol.* **21**, 287–296 (2015).
60. Mucci, A. et al. The Specific Level of Functioning Scale: construct validity, internal consistency and factor structure in a large Italian sample of people with schizophrenia living in the community. *Schizophr. Res.* **159**, 144–150 (2014).
61. Schneider, L. C. & Struening, E. L. SLOF: a behavioral rating scale for assessing the mentally ill. *Soc. Work Res.* **19**, 9–2 (1983).
62. Wongpakaran, N., Wongpakaran, T., Wedding, D. & Gwet, K. L. A comparison of Cohen’s Kappa and Gwet’s AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. *BMC Med. Res. Methodol.* **13**, 61 (2013).
63. Klein, D. Implementing a general framework for assessing interrater agreement in Stata. *Stata J.* **18**, 871–01 (2018).
64. Quarfoot, D. & Levine, R. A. How robust are multirater interrater reliability indices to changes in frequency distribution? *Am. Stat.* **70**, 373–384 (2016).
65. Gwet, K. L. *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters*. Advanced Analytics, LLC (2014).
66. Landis, J. R. & Koch, G. G. The measurement of observer agreement for categorical data. *Biometrics* **33**, 159–174 (1977).
67. Steichen, T. J. & Cox, N. J. A note on the concordance correlation coefficient. *Stata J.* **2**, 183–189 (2002).
68. Newson R. *SCSOMERSD: Stata module to compute rank statistics for scenario comparisons* (2012).
69. Morris T. & Royston P. *MFPML: Stata module to build multivariable fractional polynomial models in multiply imputed data* (2016).

ACKNOWLEDGEMENTS

We thank all the patients and caregivers participating in the study. This work was supported by the Italian Ministry of Education, the Italian Society of Psychopathology, and the Italian Society of Biological Psychiatry.

THE ITALIAN NETWORK FOR RESEARCH ON PSYCHOSES

Paola Bozzatello¹, Cristina Badino¹, Benedetta Giordano¹, Piergiuseppe Di Palo³, Vitalba Calia³, Marco Papalino³, Stefano Barlati^{20,21}, Giacomo Deste^{20,21}, Anna Ceraso^{20,21}, Federica Pinna¹⁰, Benedetta Olivieri¹⁰, Daniela Manca¹⁰, Giuseppe Piegari⁸, Francesco Brandò⁸, Luigi Giuliani⁸, Carmen Aiello⁸, Laura Fusar Poli⁵, Carmen Concerto⁵, Teresa Surace⁵, Mario Altamura⁷, Stefania Malerba⁷, Flavia Padalino⁷, Pietro Calcagno⁶, Martino Belvederi Murri⁶, Andrea Amerio⁶, Francesca Pacitti¹², Valentina Socci¹⁸, Alessia Lucaselli², Laura Giusti¹⁸, Anna Salza¹⁸, Donatella Ussorio¹⁸, Felice Iasevoli²³, Carla Gramaglia²², Eleonora Gambaro²², Eleonora Gattoni²², Elena Tenconi¹³, Enrico Collantoni¹³, Paolo Meneguzzo¹³, Paolo Ossola¹⁴, Matteo Tonna¹⁴, Maria Lidia Gerra¹⁴, Claudia Carmassi¹², Barbara Carpita¹², Ivan Mirko Cremone¹², Giulio Corrivetti²⁴, Giammarco Cascino²⁴, Francesca Marciello²⁴, Roberto Brugnoli¹⁷, Anna Comparelli¹⁷, Valentina Corigliano¹⁷, Nicoletta Girardi¹⁷, Tommaso Accinni¹⁷, Luca Carlone¹⁷, Andrea Fagiolini¹¹, Arianna Goracci¹¹, Simone Bolognesi¹¹, Giorgio Di Lorenzo¹⁹, Cinzia Niolu¹⁹ and Michele Ribolsi¹⁹

²³Department of Neuroscience, Reproductive Sciences and Dentistry, University of Naples Federico II, Naples, Italy. ²⁴Department of Mental Health of Salerno, Salerno, Italy.

AUTHOR CONTRIBUTIONS

PR, CB, CM, SB, AR, AB, SG, and MM contributed to the conception and the design of the work, all authors contributed in the acquisition or in the interpretations of data for the work; DG performed the statistical analyses; PR, CB, DG, and CM drafted the work and all authors revised it critically for important intellectual content; all authors gave their final approval of the version to be published; all authors agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

COMPETING INTERESTS

The authors declare that they have no competing interests as defined by Nature Research, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41537-021-00140-9>.

Correspondence and requests for materials should be addressed to P.R.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021