Reliability of the italian version of the 16-item Prodromal Questionnaire (iPQ-16) for psychosis risk screening in a young help-seeking community sample

S. Azzali [1], L. Pelizza [1], F. Paterlini [1], S. Garlassi [1], I. Scazza [1], L.R. Chiri [2], M. Poletti [1], F. Fontana [1], R. Favazzo [1], S.V. Barbanti [1], M. Fabiani [1], L. Pensieri [1], L. Cioncolini [1], E. Semrov [1], A. Raballo [3]

[1] Department of Mental Health and Pathological Addiction, Reggio Emilia Public Health-Care Centre, Reggio Emilia, Italy; [2] Department of Mental Health and Pathological Addiction, Bologna Public Health-Care Centre, Bologna, Italy; [3] Department of Psychology, Norwegian University of Science and Technology, Trondheim, Norway

Summary

Objective
Among current early screeners for psychosis-risk states, the Prodromal Questionnaire-16 items (PQ-16) is used. We aimed to assess reliability of the Italian version of the PQ-16 in a young help-seeking sample.

Methods
We included 151 individuals, aged 13-35 years, seeking help at the Reggio Emilia outpatient mental health services in a large semirural catchment area (550.000 inhabitants). Participants completed the Italian version of the PQ-16 (iPQ-16) and were subsequently evaluated with the Comprehensive Assessment of At-Risk Mental States (CAARMS). We examined test-retest reliability, internal consistency and diagnostic accuracy (i.e. sensitivity, specificity, positive and negative predictive values, and positive and negative likelihood ratios) between PQ-16 and CAARMS UHR-defined criteria using coefficient of stability (k), Cronbach's alpha and Cohen's kappa, respectively.

Results
The iPQ-16 showed excellent short term test-retest reliability (k = 0.898), high internal consistency (α = 0.810) and acceptable diagnostic accuracy (sensitivity = 73.5% and specificity = 75.9% at the proposed cut-off of $\geq$ 6 on symptom total score).

Conclusions
Psychometric properties of the iPQ-16 were satisfactory. The iPQ-16 is a suitable screening tool for routine use in mental health care services. Indeed, it is short and therefore easy to implement in routine assessment.

Key words
Ultra-High Risk • Prodrome • Early detection • Screening • Psychosis • Schizophrenia • Assessment

Introduction

Psychoses are disabling disorders and their life-changing impact is more prominent in adolescents and young adults [1]. In the last 25 years, several studies suggested that early intervention in psychosis might improve outcome and reduce psychosis treatment-related costs [2]. In this context, McGorry et al. [3] proposed the notion of Ultra-High Risk (UHR) mental states to identify subjects with prospectively high (but not inevitable) imminent risk of developing psychosis. Focusing mainly on attenuated positive symptoms, the UHR criteria are: (a) Attenuated Psychotic Symptoms (APS), which represent subthreshold positive symptoms; (b) Brief Limited Intermittent Psychotic Symptoms (BLIPS), which are transient positive symptoms that spontaneously disappear within 1 week; and (c) Genetic Risk

Correspondence
Lorenzo Pelizza
c/o CSM Petrella, via Errico Petrella 1/A, 42100 Reggio Emilia, Italy •
Tel.: +39 0522 339501 •
Fax: +39 0522 339523; •
E-mail: lorenzo.pelizza@ausl.re.it

and Functioning Deterioration syndrome (GRFD), a trait/state risk condition characterized by a history of psychosis in first-degree family members or a schizotypal personality disorder in the subject together with a low functioning for at least 1 month [4]. Translating the early detection/intervention research framework into clinical care pathways relies, in part, on the recognition of these young people at the earliest point in their help-seeking trajectory [5].

In the *Diagnostic and Statistical Manual of Mental Disorders*, fifth edition (DSM-5) [6], the Attenuated Psychosis Syndrome has been included as a new nosographical category that requires further study. With a lifetime prevalence rate of almost 13%, the incidence of Attenuated Psychosis Syndrome in the general help-seeking population (aged 13-40 years) seems to be more frequent than 7% (with a peak of approximately 22% in 11-13 year-olds) [7]. Its diagnostic criteria depend on the UHR criteria using clinical interviews (such as CAARMS) [4], which generally require extensive training to be administered and can take hours to be completed [8]. Therefore, an array of self-report screening tools has been developed to preselect potential UHR individuals for subsequent in-depth clinical assessment [9]. Accumulating empirical evidence suggests that these self-report instruments are sufficiently sensitive and specific to detect the majority of those subjects that merit a more comprehensive evaluation for UHR or First-Episode Psychosis (FEP) [10].

The 92-item Prodromal Questionnaire (PQ-92) [11] is the most commonly used screener in the literature [9]. However, this instrument remains rather time-consuming for routine screening [12]. Thus, Ising et al. [10] developed a 16-item version of the PQ (PQ-16), which showed good psychometric properties for screening large help-seeking sample in general mental health care services for nonpsychotic disorders [13]. In young adults, a cut-off score of ≥ 6 on symptom total score predicts CAARMS UHR/psychosis diagnosis with high sensitivity (87%) and specificity (87%) [10].

Overall, early intervention in young people at UHR for developing psychosis are less widespread in Italy than in other european countries [1]. In particular, some pilot programmes have focused specifically on early detection and intervention in UHR young adults, aged 18-30 years (see Cocchi et al. [14]). Therefore, translating an easy and suitable self-report screening instrument (such as the PQ-16) into Italian language could lead to the implementation of specific services for UHR individuals within the framework of Italy's National Health Service. To the best of our knowledge, no psychometric evaluation study on the PQ-16 in an Italian clinical sample has been reported in the literature to date. Thus, the current study was designed to test the reliability of the Italian version of the PQ-16 (iPQ-16) in identifying young people at UHR of psychosis in a help-seeking community population.

Materials and methods

Setting
As detailed in Raballo et al. [15], the "Reggio Emilia At-Risk Mental States" (ReARMS) project is an early detection infrastructure implemented under the aegis of the "Regional Project on Early Detection and Intervention in Psychosis" in the Reggio Emilia Department of Mental Health. This project aims: (a) to identify people with FEP and individuals at high clinical risk according to UHR criteria [4] among help-seeking adolescents and young adults (13-35 years) through a multi-step procedure; and (b) to provide evidence-based interventions that are supposed to be effective in UHR/FEP subjects (i.e. intensive case management, family psycho-education, individual cognitive-behavioral therapy, pharmacological treatment, as appropriate). The first filtering step included a pre-clinical triage service, conducted by trained non-medical personnel, using the "Screening Schedule" for Psychosis (SS) [16]. Such triage was mainly meant to maximise appropriate referrals to the ReARMS project and avoid over-inclusion of subjects clearly outside the severity threshold for presumed psychosis risk spectrum. The second step included a comprehensive multidimensional battery including the iPQ-16, followed by the administration of the CAARMS to define the clinical status (i.e. psychosis risk, psychosis, or neither) and the consequent access to the ReARMS clinical-therapeutic pathways [15]. Complying with the declaration of Helsinki, relevant ethical approvals were locally sought for the study.

Participants
For the purpose of the study (i.e. field-testing the reliability of the iPQ-16 in identifying UHR mental states), we focused on adolescent and young adult help-seekers, aged 13-35 years, who were consecutively referred to all of child/adolescent and adult mental health services of the Reggio Emilia Department of Mental Health between September 2012 and September 2017. In the present research, inclusion criteria were: (a) specialist help-seeking; (b) age between 13 and 35 years; and (c) presence of UHR criteria defined by the CAARMS (i.e. APS, BLIPS, and/or GRFD) [4] at the initial assessment. Individuals who were below the CAARMS UHR threshold were considered as CAARMS UHR negative cases.

The exclusion criteria were modeled on the psychometric approach adopted by Ising et al. [10] in the validation study of the original Dutch version of the PQ-16: (a) his-

PQ-16

Nome ....
Data di nascita ....
Data ....

Se hai un´ etá compresa tra 12 e 35, per favore compila il questionario

Il questionario esplora aspetti dei tuoi pensieri, sentimenti ed esperienze. Per cortesia, leggi attentamente ogni affermazione e indica se sei d´accordo o in disaccordo cercando "vero" o "falso" sulla destra. Cerca di rispondere ad ogni domanda. Nel caso tu risponda VERO, indica nell´ultima colonna il livello di disagio associato. Per favore, rispondi a tutte le domande.

Se hai risposto "Vero", quanto disagio hai provato?

		Vero	Falso	Nessuno	Lieve	Medio	Grave
1.	Ho perso interesse per cose che prima mi piacevano.	☐	☐	☐ 0	☐ 1	☐ 2	☐ 3
2.	Spesso mi sembra di vivere degli eventi esattamente come mi fossero già capitati prima (déjà vu)	☐	☐	☐ 0	☐ 1	☐ 2	☐ 3
3.	A volte sento l'odore o il sapore di cose che gli altri non riescono a sentire.	☐	☐	☐ 0	☐ 1	☐ 2	☐ 3
4.	Spesso sento suoni insoliti come esplosioni, schiocchi, sibili, schianti o squilli nelle orecchie	☐	☐	☐ 0	☐ 1	☐ 2	☐ 3
5.	A volte mi sono sentito confuso sulla natura reale o immaginaria di una esperienza	☐	☐	☐ 0	☐ 1	☐ 2	☐ 3
6.	Guardandomi allo specchio o guardando un'altra persona, ho visto il volto cambiare davanti ai miei occhi.	☐	☐	☐ 0	☐ 1	☐ 2	☐ 3
7.	Divento estremamente ansioso quando incontro qualcuno per la prima volta.	☐	☐	☐ 0	☐ 1	☐ 2	☐ 3
8.	Ho visto cose che altri apparentemente non riescono a vedere	☐	☐	☐ 0	☐ 1	☐ 2	☐ 3
9.	I miei pensieri sono talvolta così forti che posso quasi udirli	☐	☐	☐ 0	☐ 1	☐ 2	☐ 3
10.	A volte colgo un significato speciale nelle pubblicità, nelle vetrine dei negozi o nel modo in cui sono disposti gli oggetti intorno a me.	☐	☐	☐ 0	☐ 1	☐ 2	☐ 3
11.	A volte mi sono sentito come se non avessi controllo sulle mie idee o pensieri	☐	☐	☐ 0	☐ 1	☐ 2	☐ 3
12.	Talvolta mi capita di essere distratto all'improvviso da suoni distanti ai quali normalmente non presto attenzione	☐	☐	☐ 0	☐ 1	☐ 2	☐ 3
13.	Ho sentito cose che gli altri non riescono a sentire, come voci di gente che bisbiglia o che parla.	☐	☐	☐ 0	☐ 1	☐ 2	☐ 3
14.	Ho spesso la sensazione che gli altri ce l'abbiano con me.	☐	☐	☐ 0	☐ 1	☐ 2	☐ 3
15.	Ho avuto la sensazione che qualche persona o forza mi stia accanto anche se non posso vederla	☐	☐	☐ 0	☐ 1	☐ 2	☐ 3
16.	Ho l'impressione che parti del mio corpo siano cambiate in qualche modo, o che funzionino in modo diverso da prima	☐	☐	☐ 0	☐ 1	☐ 2	☐ 3

Hai risposto a tutti gli item? Grazie per aver compilato il questionario

FIGURE 1. *iPQ-16.*

tory of past frank psychotic episodes, either affective or schizophrenic (as described in the DSM-5) [6]; (b) history of previous exposure to antipsychotics; (c) current substance dependence; (d) severe learning disability or known mental retardation (Intelligence Quotient < 70); (e) neurological disease or any other medical disorder associated with psychiatric symptoms; (f) poor fluency in the Italian language; and (g) residence outside the catchment area. All these exclusion criteria have been applied after the SS administration in order to select a sample comparable to one assessed by Ising et al. [10].

All help-seekers entering the ReARMS project agreed to participate to the research and gave their informed consent to the psychopathological evaluation, composed – among others15 – by the CAARMS (approved Italian translation by Raballo et al. – CAARMS-ITA – [17]) and the PQ-16 (Italian adaptation by Raballo et al. – iPQ-16 – [18]) (Fig. 1). While in chronological terms the iPQ-16 was administered after the SS for psychosis, the meaning of its administration was different (i.e. zooming in on prodromal experiences before the CAARMS-based interview) and the CAARMS assessors were blinded to the iPQ-16 scores.

Measures

The CAARMS is a semi-structured clinical interview designed to cover different aspects of attenuated psychopathology as well as functioning (via the integrated Social and Occupational Functioning Assessment Scale – SOFAS – module) [4]. It takes approximately 1-1.5 hours to be administered and consists of 27 items (each one scored in terms of frequency/duration – 0-6 – and intensity – 0-6 –). Those items are clustered in seven subscales: (a) "Positive Symptoms"; (b) "Cognitive Change, Attention and Concentration"; (c) "Emotional Disturbance"; (d) "Negative Symptoms"; (e) "Behavioral Change"; (f) "Motor/Physical Changes"; and (g) "General Psychopathology". The CAARMS "Positive Symptoms" subscale, which covers delusions, hallucinations and thought disorder, is used to determine the UHR criteria [4]. UHR status is defined as follows: (a) GRFD group: schizotypal personality disorder in the subject or history of psychosis in a first-degree family member associated with 30% drop in functioning for $\leq$ 1 month or chronic low functioning (the decline in functioning is estimated by subtracting the actual SOFAS score from the highest SOFAS score in the past year); (b) APS group: subthreshold positive psychotic symptoms within the past 12 months; and (c) BLIPS group: criteria for psychotic disorder met for < 7 day and remitting spontaneously (i.e. without antipsychotic medication).

CAARMS interviews are conducted by specialized personnel including clinical psychologists and psychiatrists, who underwent collective supervision by the main author of the approved Italian translation [17], who was trained at Orygen, the National Youth Research Center in Melbourne, Australia. The inter-rater reliability of these assessments was ensured by regular CAARMS scoring workshops and supervision sessions.

The PQ-16 [10] is a self-report instrument specifically designed to detect people at risk of psychosis. It is composed of nine items on perceptual aberrations/hallucinations, five items on unusual thought content/delusions, and two negative symptoms. This instrument only takes approximately 3 minutes to be completed and assesses the presence of positive and negative symptom items on a true/false Likert-scale, according to the individual subjective experience in the last month. Moreover, distress is added on a 4-point scale for each endorsed item (from 0 = "no distress" to 3 = "severe distress"). The PQ-16 can be scored by the total number of symptoms endorsed (range 0-16) or the sum of distress scores (range 0-48) [10 19]. When using the symptom total

TABLE I. *CAARMS criteria, demographic and clinical data.*

	Total sample (n = 151)	UHR- (n = 81)	UHR+ (n = 70)	χ^2/t/Z
Gender (female)	79 (52.3%)	41 (50.6%)	38 (54.3)	0.203
Ethnic group (Caucasian)	130 (86.1%)	69 (85.2%)	61 (87.1%)	0.012
Mother tongue (Italian)	138 (91.4%)	76 (93.8%)	62(88.6%)	0.735
Age	20.00 (5.78)	20.26 (6.44)	19.54 (4.53)	0.541
Years of Education	11.34 (2.39)	11.47 (2.40)	11.19 (2.38)	0.726
DUI (in weeks)	69.59 (51.00)	66.39 (54.65)	72.95 (47.34)	-0.572
iPQ-16 symptom total score (range 0-16)	5.48 (3.66)	4.19 (3.48)	7.04 (3.25)	-5.206 *
iPQ-16 total distress score (range 0-48)	14.03 (10.72)	10.41 (10.21)	18.42 (9.81)	-5.207 *

* p < 0.001. Frequencies and percentages, mean (standard deviation), chi-squared (χ^2) test (with Yates correction), Student's t test, and Mann-Whitney U test (Z) values are reported.

score, a cut-off threshold of ≥ 6 appeared appropriate in general mental health settings [10]. Using the total distress score may improve the accuracy of the instrument [19]. In general mental health help-seeking populations, a total distress threshold of ≥ 8 was supported [19 20].

Statistical analysis

Data were analyzed using the "Statistical Package for Social Science" (SPSS) 18.0 for Windows [21]. For the specific purposes of this study, the sample was dichotomized as follows: UHR+ (i.e. those who were above CAARMS UHR threshold [that is APS, BLIPS and/ or GRFD]), and UHR (i.e. those who are below such threshold) [4]. The two subgroups were compared on socio-demographic, clinical, and psychopathological parameters. Categorical data were analysed using Chi-squared test with Yates' correction. Quantitative variables were examined using the Mann-Whitney's U test or the Student's t-test – as appropriate –.

Following the psychometric approach adopting by Kotzalidis et al. [22] in the validation study of the Italian version of the PQ-92 in order to compare their and our results, in the present research we measured short-term test-retest reliability of the iPQ-16 over two weeks calculating the coefficient of stability [23] on a subsample of 15 participants who had scored ≥ 6 on the iPQ-16 total symptom score at the baseline assessment. This rather short-time interval was chosen to limit the possible impact of both symptomatic changes and memory effects [24]. According to Heise [23], we interpreted test-retest reliability coefficients as follows: ≥ 0.90 excellent reliability, 0.81-0.90 good reliability, 0.71-0.80 acceptable reliability, 0.61-0.70 questionable reliability, 0.51-0.60 poor reliability, and ≤ 0.50 unacceptable reliability.

Moreover, we examined long-term test-retest reliability of the iPQ-16 calculating the coefficient of stability within all the participants who had scored ≥ 6 on the iPQ-16 symptom total score at initial assessment (n = 116). As additional measure of reliability, the internal consistency of the iPQ-16 was assessed using the Cronbach's α statistics within the total sample. A score above 0.65 represented a sufficient internal consistency [10].

Furthermore, we investigated the concurrent validity of the iPQ-16 by comparing its results to CAARMS outcomes. In the total sample, we examined diagnostic accuracy measures (i.e. sensitivity, specificity, positive and negative predictive values – PPV and NPV –, and positive and negative likelihood ratios – LR+ and LR- –, that balance sensitivity against specificity). As an additional measure of concurrent validity, the correspondence of positive results on the iPQ-16 (i.e. a symptom total score ≥ 6 or, as alternative, a total distress score ≥ 8) and on the CAARMS (i.e. a score ≥ 3 on at least one positive symptom item) was also examined by Cohen's kappa statistics.

Results

Over the course of the study, 151 individuals (79 females and 72 males; mean age ± Standard Deviation [SD] = 20.00 ± 5.78) consecutively participated at the intake interview within the ReARMS protocol. Table I shows screening outcomes and demographic characteristics of the total sample and the two subgroups, i.e. UHR+ (n = 70) and UHR- (n = 81). No significant differences were found in terms of gender, ethnic group, mother tongue, age, years of education, and Duration of Untreated Illness (DUI).

In comparison with UHR-, UHR+ individuals showed significantly higher iPQ-16 symptom total score and total distress score (Tab. I).

To calculate short-term test-retest reliability, the iPQ-16 was re-administered to 15 participants who had scored ≥ 6 on symptom total score at the first assessment. Their socio-demographic characteristics were comparable to those of the total sample, with a mean age of 19.94 years and a SD of 4.89 years. Eight (53%) participants were females. The coefficient of stability was 0.898 for iPQ-16 symptom total score, indicating good to excellent short-term test-retest reliability [23].

To examine long-term test-retest reliability, the iPQ-16 was administered over 1 year to 116 individuals who had scored ≥ 6 on symptom total score at the baseline. Their demographic features were comparable to those of the entire sample, with a mean age of 20.20 years and a SD of 5.59 years. Sixty (51.7%) subjects were females. The coefficient of stability was 0.486, indicating unacceptable long-term test-retest reliability [23].

Across the total sample, the iPQ-16 symptom total score showed a Cronbach's alpha of 0.810.

At the proposed PQ-16 symptom total score cut-off of ≥ 6 [10], 68 participants (45%) scored positive; of these, 50 (73.5%) also scored ≥ 3 on any positive CAARMS item (i.e. meeting the UHR threshold). Altogether, 20 participants (28.6%) with any CAARMS positive score ≥ 3 were missed by this PQ-16 cut-off, and 20 (13.2%) were falsely identified. Cohen's kappa was 0.483, consistent with a fair agreement [25]. With regard to the diagnostic accuracy at the proposed PQ-16 cut-off of ≥ 6 on symptom total score, sensitivity was 73.5%, specificity 75.9%, PPV 71.4%, NPV 77.8%, LR+ 3.05, and LR- 0.35. Thus, at this threshold, the iPQ-16 symptom total score was slightly better in ruling out than in ruling in possible UHR status, changing post-test probability to a small (but sometimes important) degree [26].

Considering the proposed PQ-16 cut-off of ≥ 8 on total distress score [19 20], sensitivity was 85.3%, specificity 51.8%, PPV 59.2%, NPV 81.1%, LR+ 1.77, LR- 0.28, and Cohen's kappa 0.344.

Discussion

Aim of the current was to evaluate the reliability of PQ-

16 in an Italian clinical sample of young people at UHR of psychosis. Introducing and promoting the routinary use of the Italian version of a validated assessment tool to detect UHR subjects in the general help-seeking population (such as the iPQ-16) could positively impact on the implementation of specific services for early detection and intervention on UHR individuals within the framework of Italy's National health Service. In the current study, we therefore examined test-retest reliability and internal consistency of the iPQ-16 in consecutive young help-seekers attending all of child/adolescent and adult mental health services of the Reggio Emilia Department of Mental Health.

In comparison with UHR-, UHR+ individuals showed significantly higher iPQ-16 total scores. On a dimensional level – as expected on the basis of the PQ-16 item composition – these findings suggest that increasing PQ-16 scores are associated with the severity of both psychotic and general psychopathology, as well as the intensity of distress related to prodromal symptoms.

We found excellent reliability of the iPQ-16 with regard to internal consistency of the symptom total score ($\alpha = 0.810$). Similarly, in a Dutch adult help-seeking sample attending to a secondary mental health care service, the PQ-16 showed good internal consistency, with a Cronbach's alpha of 0.774 [10]. Internal consistency in our clinical sample was even better than the high reliability reported for the Chinese PQ-16 version ($\alpha = 0.750$) in 579 college students [20] and the Dutch PQ-16 version ($\alpha = 0.790$) in 176 help-seeking adolescents (aged 12-17 years) attending one of the three Child and Adolescent Mental Health Services in Rotterdam in the Netherlands [13]. In our sample, iPQ-16 demonstrated a Cronbach's alpha value that we consider as satisfactory internal consistency for a screener that has to come before a clinical interview [10].

In a recent validation study of the Italian version of the PQ-92, Kotzalidis et al. [22] re-administered the instrument to 15 individuals two weeks after first assessment and found excellent short-term test-retest reliability(coefficient of stability = 0.942 for PQ-92 total score). Similarly, we found a coefficient of stability equal to 0.898, indicating a good to excellent short-term (two-week) test-retest reliability of the iPQ-16.

Based on what Kotzalidis et al. [22] suggested, we also addressed longer term test-retest reliability administering the iPQ-16 over 1 year to all the participants who had scored ≥ 6 on symptom total score at first assessment. We found a coefficient of stability = 0.486, indicating unacceptable long-term test-retest reliability. According to Michel et al. [24], this finding suggests that the self-report screening questionnaire assessed a fluctuating condition rather than a trait characteristic, i.e. a condition itself that varied between test and retest. When examining these results, some methodological peculiarities of the current study shall be considered. Indeed, ReARMS is a clinically project providing evidence-based interventions that are supposed to be effective in UHR individuals (i.e. intensive case-management, family psycho-education, individual cognitive-behavioral therapy within the framework of assertive community treatment). Precisely because providing the optimal treatment for the help-seekers was the main ethical mandate in our clinical setting, our treatments were not controlled (e.g. against placebo group or other treatments), but evenly delivered to all UHR participants.

In the Dutch study validating the PQ-16, Ising et al. [10] observed an excellent concurrent validity with CAARMS diagnoses in a sample of adult help-seekers for non-psychotic disorders in general mental health care services. A cut-off of ≥ 6 on symptom total score had a high sensitivity (87%) and high specificity (87%) in discriminating between people with UHR/psychosis and individuals without CAARMS diagnosis, with a PPV of 44%. With regard to the diagnostic accuracy at the proposed PQ-16 cut-off of ≥ 6 on symptom total score [5], sensitivity in our sample (73.5%) was lower than previously reported for the PQ in its various versions [9]. However, this result was similar to that (62%) observed by Kotzalidis et al. [22] in the validation study of the Italian version of the 92-item PQ. Moreover, at the proposed PQ-16 cut-off of ≥ 6, our PPV (approximately 71%) was higher than previously reported, with values ranging between 29% and 44% [9 10 20 22]. In particular, PPV was only equal to 38% in the validation study of the Italian version of the 92-item PQ [22]. The difference between these findings may be the result of differences in selection procedures. In fact, first screening procedure in the ReARMS protocol included a triage service using the SS for psychosis [16], which probably excluded a certain amount of true negative cases.

In our sample, specificity (approximately 76%) is consistent with that previously reported. Indeed, specificity values were also good to excellent in the Dutch study validating the original PQ-16 at a cut-off of 6 or more endorsed prodromal symptoms [10] and in the validation study of the Italian version of the 92-item PQ [22] (87% and 82%, respectively). Otherwise, our NPV (approximately 78%) was lower than previously reported, with values ranging between 90% and 100% [10 20]. Kotzalidis et al. [22] found a NPV of 91% in the validation study of the Italian version of the 92-item PQ. The difference between these findings may be the result of the same differences in selection procedures previously mentioned.

Compared to the PQ-16 cut-off of ≥ 6 on symptom total score, the proposed ≥ 8 threshold on total distress score [19] increased sensitivity value up to 85%, but with a significant decrease in specificity (approximately 52%).

However, according to Loewy et al. [11], for screening purposes grater weighting should be given to sensitivity over specificity as part of a two-step screening process. Using the proposed PQ-16 cut-off of ≥ 6 on symptom total score, a sensitivity value of 73% is quite low and means that a number of people who would appropriate for early intervention services are not being identified. Considering this argument, the cut-off of ≥ 8 on total distress score might be more appropriate.

Limitations

Firstly, a possible limitation of this research is that the iPQ-16 was completed in a population plausibly "enriched" for the target diagnoses, i.e. young help-seekers with clinical features of possible psychosis. Therefore, the current field-test of the iPQ-16 was not meant to identify cut-offs applicable to the general population, in which the psychometric endorsement of so-called psychotic-like experiences might occasionally occur, yet with transient temporal pattern, not necessarily accompanied by distress or treatment seeking, and not inevitably followed by a transition to psychosis [1][10]. Indeed, a certain number of false positives would be identified.

Another limitation is that since the SS for psychosis [16] was used in the eligibility triage for the ReARMS protocol (i.e. before the iPQ-16 administration), this is likely to impact the generalizability of our findings. Indeed, the PQ would ideally be used as the first step in a 2-stage screening process [11]. Therefore, by excluding a certain amount of true negative cases in the pre-PQ step, this would reduce the specificity of the screener.

Finally, although the ability of the iPQ-16 to include cases appears to be less than its ability to exclude them, it may still miss some cases worthy of further investigation [22]. In this respect, Ising et al. [10] included 4-point scale questions on distress following each PQ-16 item to examine if this enhanced the PPV of the instrument. However, in the present research the PPV was good.

Conclusions

The Italian version of the PQ-16 showed satisfying psychometric properties, comparable to 92-item homologue [22]. However, yet optimal cut-off to improve concurrent validity and, consequently, economic and clinical usefulness has still to be determined through multi-centric testing. Moreover, the iPQ-16 seems to be a suitable screening tool for routine use in mental health care services. Indeed, it is short (taking only few minutes to be completed) and therefore easy to implement in routine assessment. Finally, the iPQ-16 can be helpful in identifying potential psychotic symptoms for further exploration in an early phase, especially in young adults and adolescents with low functioning.

Acknowledgements

Conflict of interest
None.

References

1 Pelizza L, Azzali S, Garlassi S, et al. *Adolescents at Ultra-High Risk for psychosis in Italian neuropsychiatry services: prevalence, psychopathology and transition rate*. Eur Child Adolesc Psychiatry 2017 (in press).

2 Pelizza L, Raballo A, Semrov E, et al. *Validation of the early detection Primary Care Checklist in an Italian community help-seeking sample: the "Checklist per la Valutazione dell'Esordio Psicotico"*. Early Interv Psychiatry 2017 (in press).

3 McGorry PD, Yung AR, Phillips LJ. *The "close-in" or ultra-high risk model: a safe and effective strategy for research and clinical intervention in pre-psychotic mental disorder*. Schizophr Bull 2003;29:771-90.

4 Yung AR, Yuen HP, McGorry PD, et al. *Mapping the onset of psychosis: the comprehensive assessment of at-risk mental states*. Aust N Z J Psychiatry 2005;39:964-71.

5 Pelizza L, Raballo A, Semrov E, et al. *Identification of young people at Ultra-High Risk (UHR) of developing psychosis: validation of the "Checklist per la Valutazione dell'Esordio Psicotico" for use in primary care setting*. Journal of Psychopathology 2016;22:172-9.

6 American Psychiatric Association (APA). *Diagnostic and statistical manual of mental disorders*. Arlington: American Psychiatric Publishing Inc 2013 (V edition).

7 Schultze-Lutter F, Michel C, Ruhrmann S, et al. *Prevalence and clinical significance of DSM-5-Attenuated Psychosis Syndrome in adolescents and young adults in the general population: the Bern Epidemiological At-Risk (BEAR) study*. Schizophr Bull 2014;40:1499-508.

8 Brandizzi M, Schultze-Lutter F, Masillo A, et al. *Self-report attenuated psychotic-like experiences in help-seeking adolescents and their association with age, functioning, and psychopathology*. Schizophr Res 2014;160:110-17.

9   Addington J, Stowkowy J, Weiser M. *Screening tools for clinical high risk for psychosis.* Early Interv Psychiatry 2015;9:345-56.

10 Ising HK, Veling W, Loewy RL, et al. *The validity of the 16-item version of the Prodromal Questionnaire (PQ-16) to screen for Ultra High Risk of developing psychosis in the general help-seeking population.* Schizophr Bull 2012;38:1288-96.

11 Loewy RL, Bearden CE, Johnson JK, et al. *The prodromal questionnaire (PQ): preliminary validation of a self-report screening measure for prodromal and psychotic syndromes.* Schizophr Res 2005;79:117-25.

12 Loewy RL, Pearson R, Vinogradov S, et al. *Psychosis risk screening with the Prodromal Questionnaire - Brief version (PQ-B).* Schizophr Res 2011;129:42-6.

13 de Jong Y, Mulder CL, Boon AE, et al. *Screening for psychosis risk among adolescents in child and adolescent mental health services: a description of the first step with the 16-item version of the Prodromal Questionnaire (PQ-16).* Early Interv Psychiatry 2016 (in press).

14 Cocchi A, Meneghelli A, Preti A. *Programma 2000: celebrating 10 years of activity of an Italian pilot programme on early intervention in psychosis.* Aust N Z J Psychiatry 2008;42:1003-12.

15 Raballo A, Chiri LR, Pelizza L, et al. *Field-testing the early intervention paradigm in Emilia-Romagna: the Reggio Emilia At Risk Mental State (ReARMS) project.* Early Interv Psychiatry 2014;8(Suppl 1):s.88.

16 Jablensky A, Sartorius N, Ernberg G, et al. *Schizophrenia: manifestations, incidence and course in different cultures, a World Health Organization ten-country study.* Psychol Med Monogr Suppl 1992;20:1-97.

17 Raballo A, Semrov E, Bonner Y, et al. *Traduzione e adattamento italiano della CAARMS (the Comprehensive Assessment of At Risk Mental States).* Bologna 2013.

18 Raballo A, Kotzalidis GD, Savoja V, et al. *PQ-16.* Roma: Università "La Sapienza" 2012.

19 Savill M, D'Ambrosio J, Cannon TD, et al. *Psychosis risk screening in different populations using the Prodromal Questionnaire: a systematic review.* Early Interv Psychiatry 2017 (in press).

20 Chen F, Wang L, Heeramun-Aubeeluck A, et al. *Identification and characterization of college students with attenuated psychosis syndrome in China.* Psychiatry Res 2014;216:346-50.

21 *SPSS for Windows, rel. 18.0.* Chicago, IL: SPSS Inc 2010.

22 Kotzalidis GD, Solfanelli A, Piacentino D, et al. *The Italian version of the 92-item Prodromal Questionnaire: concurrent validity with the SIPS and factor analysis in a sample of 258 outpatients aged 11-36 years.* Schizophr Res 2017 (in press).

23 Heise DR. *Separating reliability and stability in test-retest correlation.* Am Sociol Rev 1969;34:93-9.

24 Michel C, Schultze-Lutter F, Schimmelmann BG. *Screening instruments in child and adolescent psychiatry: general and methodological considerations.* Eur Child Adolesc Psychiatry 2014;23:725-7.

25 Landis JR, Koch GG. *The measurement of observer agreement for categorical data.* Biometrics 1997;33:159-74.

26 Jaeschke R, Guyatt G, Sackett DL. *Users' guides to the medical literature. III. how to use an article about a diagnostic test. A. Are the Results of the Study Valid? Evidence-based medicine working group.* JAMA 1994;271:389-91.