

Reliability of the Italian version of the Brief (21-item) Prodromal Questionnaire (IPQ-B) for psychosis risk screening in a young help-seeking population

I. Scazza¹, L. Pelizza¹, S. Azzali¹,
F. Paterlini¹, S. Garlassi¹, L.R. Chiri²,
M. Poletti¹, F. Fontana¹, R. Favazzo¹,
S.V. Barbanti¹, M. Fabiani¹,
L. Pensieri¹, L. Cioncolini¹,
E. Semrov¹, A. Raballo^{3,4}

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; ⁴ Department of Medicine, Division of Psychiatry, Clinical Psychology and Rehabilitation, University of Perugia, Italy

Summary

Objective

Among current screeners for psychosis-risk mental states, the Prodromal Questionnaire-Brief (21 items) (PQ-B) is used. We aimed to assess reliability of the Italian version of the PQ-B 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-B (iPQ-B) 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-B and CAARMS UHR-defined criteria using coefficient of stability (k), Cronbach's alpha and Cohen's kappa, respectively.

Results

The iPQ-B showed excellent short term test-retest reliability ($k = 0.891$), high internal consistency ($\alpha = 0.876$) and acceptable diagnostic accuracy (sensitivity = 91.4% at the proposed cut-off of ≥ 6 on total distress score).

Conclusions

Psychometric properties of the iPQ-B were satisfactory. The iPQ-B 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 of early psychosis.

Key words

Ultra-High Risk • Prodrome • Early Detection • Screening • Psychosis • Schizophrenia • Assessment

Introduction

Specialist treatment for Ultra-High Risk (UHR) mental states of psychotic disorders can effectively reduce psychosis conversion rate ¹. However, identifying individuals with UHR remains a significant challenge ². Focusing mainly on attenuated positive symptoms, McGorry et al. (2003) ³ proposed the following UHR criteria: (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 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 ^{5,6}.

© Copyright by Pacini Editore Srl



OPEN ACCESS

Received: April 4, 2018

Accepted: June 15, 2018

Correspondence

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

Although structured interviews, such as the Comprehensive Assessment of At-Risk Mental States (CAARMS)⁴ or the Structured Interview for Prodromal States (SIPS)⁷, can reliably diagnose UHR states⁸, they generally require extensive training to be administered and can take hours to be completed⁹. Therefore, an array of self-report screening tools has been developed to preselect potential UHR individuals for subsequent in-depth clinical assessment¹⁰. 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)¹¹. The 92-item Prodromal Questionnaire (PQ-92)¹² is the most commonly used screener for psychosis risk in the literature¹⁰. However, this instrument remains rather time-consuming for routine screening⁹. Thus, Loewy et al. (2011)¹³ developed a Brief 21-item version (PQ-B), focusing on the positive symptom items of the PQ-92, since they are the essential ones for interview-based diagnoses of symptomatic prodromal syndromes (i.e. APS and BLIPS). A cut-off of ≥ 6 on the PQ-B total distress score predicted SIPS-UHR/psychosis diagnosis with high sensitivity (88%) and good specificity (68%)¹³. Overall, early intervention in young people at UHR for developing psychosis are less widespread in Italy than in other European countries¹⁴. 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., 2008: "Programma 2000")¹⁵. Therefore, translating an easy and suitable self-report screening instrument (such as the PQ-B) 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-B 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-B (iPQ-B) in identifying young people at UHR of psychosis in a help-seeking community population.

Materials and methods

Setting

As detailed in Raballo et al. (2014)¹⁶, the "Reggio Emilia At-Risk Mental States" (ReARMS) project is an early detection/intervention 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 criteria⁴ among help-seeking adolescents and young adults (13-35 years) through a mul-

ti-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)¹⁷. 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-B, 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 pathways¹⁶. 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-B 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 Loewy et al. (2011)¹³ in the validation study of the original version of the PQ-B: (a) history of past frank psychotic episodes, either affective or schizophrenic (as described in the DSM-5)¹⁸; (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 Loewy et al. (2011)¹³.

All help-seekers entering the ReARMS project agreed to participate to the research and gave their informed consent to the psychopathological evaluation, composed – among others¹⁶ – by the CAARMS (approved Italian translation by Raballo et al., 2013 [CAARMS-ITA])¹⁹ and the PQ-B (authorized Italian version by Preti and Rabal-

lo, 2011 [iPQ-B])²⁰ (Appendix I). While in chronological terms the iPQ-B 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-B scores.

Appendix I

The Italian version of the Brief (21-item) Prodromal Questionnaire (iPQ-B)

(Source: 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).

(Authorized Italian version by Preti A, Raballo A. Studio CAPIRE. *Cagliari Psychosis: Investigation on Risk Emergence*, 2011).

Per cortesia, indica se hai avuto i seguenti pensieri, sentimenti ed esperienze nel corso dell'ultimo mese segnando "Sì" o "No" per ciascuna domanda. **Non tenere conto di esperienze che si verificano sotto influenza di alcol, droghe o farmaci che non ti erano stati prescritti.** Se rispondi "Sì" a una domanda indica anche quanto disagio ti ha causato quell'esperienza [quanto spiacevole è stata per te quell'esperienza].

1. Capita talvolta che gli ambienti abituali ti sembrano strani, confusi, minacciosi o irreali?

SI **NO** **Se Sì:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

2. Hai mai sentito suoni insoliti come esplosioni, schiocchi, sibili, schianti o squilli nelle tue orecchie?

SI **NO** **Se Sì:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

3. Le cose che vedi ti appaiono differenti dal modo in cui sono abitualmente (più luminose o più scure, più larghe o più piccole, comunque cambiate in qualche modo)?

SI **NO** **Se Sì:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

4. Hai avuto esperienze con la telepatia, le forze psichiche o la predizione del futuro?

SI **NO** **Se Sì:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

5. Ti sei sentito come se non avessi controllo sulle tue idee o pensieri?

SI **NO** **Se Sì:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

6. Hai difficoltà a spiegarti, perché fai troppe digressioni o devi dal filo del discorso quando parli?

SI **NO** **Se Sì:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

7. Hai l'impressione o la convinzione di essere dotato in modo particolare o di possedere un talento speciale?

SI **NO** **Se Sì:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

8. Hai l'impressione che altre persone ti stiano tenendo d'occhio o parlino di te?

SI **NO** **Se Sì:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

9. Hai talvolta sensazioni strane sulla pelle o appena al di sotto, come insetti che camminano?

SI **NO** **Se Sì:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo

10. Ti capita talvolta di essere distratto all'improvviso da suoni distanti dei quali generalmente non sei consapevole? [ai quali normalmente non presti attenzione]

SI **NO** **Se Si:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

11. Hai la sensazione che qualche persona o forza ti stia accanto anche se tu non puoi vederla?

SI **NO** **Se Si:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

12. Ti preoccupi talvolta del fatto che qualcosa nella tua mente non funzioni correttamente?

SI **NO** **Se Si:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

13. Hai mai avuto la sensazione di non esistere, o che il mondo non esiste, o di essere morto?

SI **NO** **Se Si:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

14. Qualche volta ti sei sentito confuso sulla natura reale o immaginaria di un'esperienza?

SI **NO** **Se Si:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

15. Hai delle idee o delle convinzioni che altre persone troverebbero insolite o bizzarre?

SI **NO** **Se Si:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

16. Senti che parti del tuo corpo sono cambiate in qualche modo, o che funzionano in modo diverso?

SI **NO** **Se Si:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

17. I tuoi pensieri sono talvolta così forti che puoi quasi udirli?

SI **NO** **Se Si:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

18. Ti capita di provare sfiducia o essere sospettoso riguardo alle altre persone?

SI **NO** **Se Si:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

19. Hai visto oggetti insoliti come bagliori, fiamme, lampi accecanti o figure geometriche?

SI **NO** **Se Si:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

20. Hai visto cose che altri non riescono a vedere o non sembrano notare?

SI **NO** **Se Si:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

21. Capita talvolta che le persone abbiano difficoltà a capire quello che stai dicendo?

SI **NO** **Se Si:** Quando capita, mi sento spaventato, preoccupato, o comunque la cosa mi crea problemi: Fortemente in disaccordo In disaccordo
 Indifferente D'accordo Fortemente d'accordo

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, Atten-

tion 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 criteria⁴. 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 ≤ 1 month or chronic low functioning (the decline in functioning is estimated by subtracting the current SOFAS score from the highest SOFAS score in the past year); (b) APS group: sub-threshold 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 translation¹⁹, 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-B13 is a self-report questionnaire used to screen individuals for the risk of psychosis. It only takes approximately 4 minutes to be completed and comprises of 21 items recording positive symptoms experienced over the past month. For each endorsed symptom, responders rate whether they found it distressing or impairing, ranging from 1 (“strongly disagree”) to 5 (“strongly agree”), with a 4 or 5 indicating distress¹³. The PQ-B has been adopted as a screening tool using the total number of items endorsed (“symptom total score”), the number of items that are identified as distressing (“distressing item total score”) (both range 0-21), and the total distress score (range 0-105), with the latter method recommended by Loewy et al. (2011)¹³. In a recent systematic review on psychosis risk screening using the PQ in its different iterations, Savill et al. (2017)²¹ examined eight diagnostic accuracy studies using the PQ-B. Of these, one evaluated the number of distressing symptom endorsed in an UHR/psychosis-enriched sample and found a threshold of ≥ 4 distressing items as optimal cut-off²². Six studies examined the total distress score for screening: in samples with a very high prevalence ($\sim 80\%$) of UHR/psychosis individuals, a total distress score of ≥ 6 was supported¹³, whereas in similar settings with a much lower prevalence ($< 40\%$), a total distress score ≥ 18 was recommended^{22,23}. Finally, four studies adopted the total number of symptoms endorsed as cut-off: in a sample with a very high proportion of UHR/psychotic participants, a cut-off score

of ≥ 3 symptoms endorsed was supported¹³, whereas in a lower prevalence sample from a similar setting, a higher threshold of ≥ 9 was identified (albeit below 75% sensitivity)²².

Statistical analysis

Data were analyzed using the “Statistical Package for Social Science” (SPSS) 18.0 for Windows²⁴. 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. (2017)²⁵ 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-B over two weeks calculating the coefficient of stability²⁶ on a subsample of 15 participants who had scored ≥ 6 on the iPQ-B total distress score (i.e. the best recommended original cut-off proposed by Loewy et al., 2011)¹³ at the baseline assessment. This rather short-time interval was chosen to limit the possible impact of both symptomatic changes and memory effects²⁷. According to Heise (1969)²⁶, 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-B calculating the coefficient of stability within all the participants who had scored ≥ 6 on the iPQ-B total distress score at the initial assessment ($n = 123$). As additional measure of reliability, the internal consistency of the iPQ-B was assessed using the Cronbach’s α statistics within the total sample. A score above 0.65 represented a sufficient internal consistency⁶. We also examined how each PQ-B item correlated with the recommended total score (i.e. the total distress score). Correlations less than $r = 0.30$ indicated that the item might need to be removed from the questionnaire to make it more reliable²⁸. Finally, we were interested in Cronbach’s alpha value if each iPQ-B item was deleted. If this score went up after item deletion, removal should be considered to ameliorate screening tool reliability of the instrument²⁸.

Furthermore, we investigated the concurrent validity of the iPQ-B 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-B (i.e. a total distress score ≥ 6 or, as alternatives, the recommended symptom total score ≥ 3 or item distressing total score ≥ 4) and on the CAARMS (i.e. a score ≥ 3 on at least one positive symptom item) was also examined by Cohen's kappa statistics.

Finally, to explore which iPQ-B items were likely to be more predictive of CAARMS UHR diagnosis, we employed a forward stepwise logistic regression analysis, with iPQ-B item scores as independent variables and dichotomized CAARMS diagnoses (i.e. UHR- vs UHR+) as dependent variable.

Results

Over the course of the study, 151 individuals (79 females and 72 males; mean age \pm Standard Deviation [SD] = 20.00 \pm 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) ⁶.

In comparison with UHR-, UHR+ individuals showed significantly higher iPQ-B scores (Tab. I). To calculate short-term test-retest reliability, the iPQ-B was re-administered to 15 participants who had scored ≥ 6 on total distress 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 0.891 for iPQ-B total distress score, indicating good to excellent short-term test-retest reliability ²⁶.

To examine long-term test-retest reliability, the iPQ-B was administered over 1 year to 123 individuals who had scored ≥ 6 on total distress score at the baseline. Their demographic features were comparable to those of the entire sample, with a mean age of 20.10 years and a SD of 5.01 years. Sixty-three (51.2%) subjects were females. The coefficient of stability was 0.395, indicating unacceptable long-term test-retest reliability ²⁶. Across the total sample, the iPQ-B total distress score showed a Cronbach's alpha of 0.876. All item-total correlations were higher than 0.30, with the exception of item ⁹ ("Do you sometimes get strange feelings on or just beneath your skin, like bugs crawling?") (r = 0.213) (Tab. II). Therefore, most item appeared to be worthy of retention, resulting in a decrease in the alpha if deleted. Exception to this was item 9, whose deletion increased Cronbach's alpha up to the value of 0.879. Thus, removal of this item can be considered.

TABLE I. CAARMS criteria, demographic and clinical data.

	Total sample (n = 151)	UHR- (n = 81)	UHR+ (n = 70)	$\chi^2/t/Z$
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)	(47.34)	-0.572
iPQ-B symptom total score (range 0-21)	7.31 (4.90)	5.68 (4.42)	9.24 (4.76)	-4.467*
iPQ-B total distress score (range 0-105)	24.80 (18.98)	18.14 (15.95)	32.65 (19.37)	-4.717*
iPQ-B distressing item total score (range 0-21)	3.85 (3.83)	2.66 (3.07)	5.25 (4.17)	-3.966*

* $p < 0.001$.

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.

TABLE II. Internal consistency of iPQ-B.

PQ-B item	Item-total correlation	Cronbach's alpha if item deleted
1. Do familiar surroundings sometimes seem strange, confusing, threatening, or unreal to you? (PQ-B1)	.628	.857
2. Have you heard unusual sounds like banging, clicking, hissing, clapping, or ringing in your ears? (PQ-B2)	.477	.862
3. Do things that you see appear different from the way they usually do (brighter or duller, larger or smaller, or changed in some other way)? (PQ-B3)	.548	.860
4. Have you had experiences with telepathy, psychic forces, or fortune telling? (PQ-B4)	.379	.865
5. Have you felt that you are not in control of your own ideas or thoughts? (PQ-B5)	.521	.860
6. Do you have difficulty getting your point across because you ramble or go off the track a lot when you talk? (PQ-B6)	.454	.863
7. Do you have strong feelings or beliefs about being unusually gifted or talented in some way? (PQ-B7)	.340	.866
8. Do you feel that other people are watching you or talking about you? (PQ-B8)	.514	.861
9. Do you sometimes get strange feelings on or just beneath your skin, like bugs crawling? (PQ-B9)	.213	.879
10. Ti capita talvolta di essere distratto all'improvviso da suoni distanti dei quali generalmente non sei consapevole? (PQ-B10)	.509	.861
11. Hai la sensazione che qualche persona o forza ti stia accanto anche se tu non puoi vederla? (PQ-B11)	.424	.864
12. Ti preoccupi talvolta del fatto che qualcosa nella tua mente non funzioni correttamente? (PQ-B12)	.429	.864
13. Have you ever felt that you don't exist, the world does not exist, or that you are dead? (PQ-B13)	.377	.865
14. Have you been confused at times whether something you experienced was real or imaginary? (PQ-B14)	.551	.859
15. Do you hold beliefs that other people would find unusual or bizarre? (PQ-B15)	.504	.861
16. Do you feel that parts of your body have changed in some way, or that parts of your body are working differently? (PQ-B16)	.412	.864
17. Are your thoughts sometimes so strong that you can almost hear them? (PQ-B17)	.472	.862
18. Do you find yourself feeling mistrustful or suspicious of other people? (PQ-B18)	.367	.866
19. Have you seen unusual things like flashes, flames, blinding lights, or geometric figures? (PQ-B19)	.450	.863
20. Have you seen things that other people can't see or don't seem to see? (PQ-B20)	.555	.859
21. Do people sometimes find it hard to understand what you are saying? (PQ-B21)	.400	.865

iPQ-B: Italian Prodromal Questionnaire-Brief version.

Correlation *r* coefficients and Cronbach's alpha values are reported.

At the proposed PQ-B total distress score cut-off of ≥ 6 (13), 125 participants (82.7%) scored positive; of these, 64 (50.4%) also scored ≥ 3 on any positive CAARMS item (i.e. meeting the UHR threshold). Altogether, 6 participants (8.6%) with any CAARMS positive score ≥ 3 were missed by this PQ-B cut-off, and 61 (49.6%) were falsely identified. Cohen's kappa was 0.141, consistent with a slight agreement²⁹. With regard to the diagnostic accuracy at the proposed PQ-B cut-off of ≥ 6 on total distress score, sensitivity was 91.4%, specificity 24.7%, PPV 51.2%, NPV 76.9%, LR+ 1.21, and LR- 0.35. Thus, at this threshold, the iPQ-B total distress score was slightly better in ruling out than in ruling in possible UHR status, changing post-test probability

to a small (but sometimes important) degree³⁰. Considering the proposed PQ-B cut-off of ≥ 3 on symptom total score¹³, sensitivity was 91.4%, specificity 28.4%, PPV 52.4%, NPV 79.3%, LR+ 1.28, LR- 0.30, and Cohen's kappa 0.177. Finally, examining the diagnostic accuracy measures of the recommended PQ-B cut-off threshold of ≥ 4 on distressing item total score²¹, sensitivity was 61.4%, specificity 75.3%, PPV 68.2%, NPV 69.3%, LR+ 1.89, LR- 0.51, and Cohen's kappa 0.356 (consistent with a fair agreement)²⁹.

Using a conditional forward stepwise method, five items (iPQ-B5, iPQ-B6, iPQ-B8, iPQ-B9, and iPQ-B13) were entered into the regression model with a statistically significant power (Tab. III). Although iPQ-B 5, 6, 8, and 13 had

TABLE III. Logistic regression of dichotomized CAARMS UHR diagnoses by iPQ-B items.

iPQ-B item	B	SE	Wald	df	p	OR
PQ-B5	0.195	0.089	4.849	1	0.028	1.216
PQ-B6	0.189	0.097	3.795	1	0.048	1.208
PQ-B8	0.401	0.089	20.115	1	0.000	1.493
PQ-B9	-0.443	0.116	14.522	1	0.000	0.642
PQ-B13	0.237	0.135	3.113	1	0.045	1.268
Constant	-0.226	0.229	1.240	1	0.285	0.775

Overall model fit test → $\chi^2 = 62.653$, $p = 0.000$

Associated strength → Cox-Snell R² = 0.228, Nagelkerke R² = 0.319

iPQ-B: Italian Prodromal Questionnaire – Brief version, CAARMS: Comprehensive Assessment of At-Risk Mental States; UHR: Ultra-High Risk mental states; B: regression coefficient; SE: standard error; Wald: Wald statistic value; df: degree of freedom; p: statistical significance; and OR: odd ratio

positive regression coefficients, iPQ-B 9 showed a negative one. The percentage of correct diagnosis using this model for predicting CAARMS UHR diagnosis was 76.7%.

Discussion

Aim of the current was to evaluate the reliability of PQ-B 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-B) 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-B 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-B total scores. On a dimensional level – as expected on the basis of the PQ-B item composition - these findings suggest that increasing PQ-B 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-B with regard to internal consistency of the total distress score ($\alpha = 0.876$). Removal of item 9 (“Do you sometimes get strange feelings on or just beneath your skin, like bugs crawling?”), which resulted in a slight increase in Cronbach's alpha value up to 0.879, can be considered. Similarly, in a comparable adolescent/young adult

help-seeking sample (age between 12 and 35 years), Loewy et al. (2011)¹³ found an excellent PQ-B internal consistency with a Cronbach's alpha of 0.853. Moreover, Xu et al. (2016)³¹ showed an overlapping internal consistency ($\alpha = 0.897$) in Chinese help-seeking individuals (aged 15-45 years) visiting a general mental health setting. Therefore, PQ-B appears to be reliably good in different samples and cultures. Moreover, in our sample, iPQ-B demonstrated a Cronbach's alpha value that we consider as satisfactory internal consistency for a screener that has to come before a clinical interview⁶. In a recent validation study of the Italian version of the PQ-92, Kotzalidis et al. (2017)²⁵ 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 0.891, indicating a good to excellent short-term (two-week) test-retest reliability of the iPQ-B.

Based on what Kotzalidis et al. (2017)²⁵ suggested, we also addressed longer term test-retest reliability administering the iPQ-B over 1 year to all the participants who had scored ≥ 6 on total distress score at first assessment. We found a coefficient of stability = 0.395, indicating unacceptable long-term test-retest reliability. According to Michel et al. (2014)²⁷, 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 participants⁶.

In the original study validating the PQ-B, Loewy et al. (2011)¹³ observed a good to excellent concurrent validity with CAARMS diagnoses in a sample of adolescent and young adult help-seekers attending to an Early Intervention Psychosis (EIP) service. A cut-off of ≥ 6 on total distress score had a high sensitivity (88%) and good specificity (68%) in discriminating between people with UHR/psychosis and individuals without CAARMS diagnosis.

With regard to the diagnostic accuracy at the proposed PQ-B cut-off of ≥ 6 on total distress score¹³, sensitivity in our sample (91.4%) was substantially in line with previously reported for the PQ in its various versions¹⁰. However, this result was much higher to that (62%) observed by Kotzalidis et al. (2017)²⁵ in the validation study of the Italian version of the 92-item PQ. Moreover, at the proposed PQ-B cut-off of ≥ 6 , our PPV (51.2%) was consistent with previously reported, with values ranging between 29% and 44%²¹. In particular, PPV was equal to 38% in the validation study of the Italian version of the 92-item PQ²⁵. 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 psychosis¹⁷, which probably excluded a certain amount of true negative cases.

In our sample, specificity (approximately 25%) was lower than previously reported. Indeed, specificity values were good to excellent in the original study validating the PQ-B at a cut-off of ≥ 6 on total distress score¹³ and in the validation study of the Italian version of the 92-item PQ²⁵ (68% and 82%, respectively). Likewise, our NPV (approximately 77%) was good, but slightly lower than previously reported, with values ranging between 90% and 100%^{2,21}. In this regards, Kotzalidis et al. (2017)²⁵ 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.

However, according to Loewy et al. (2011)¹³, for screening purposes, greater weighting should be given to sensitivity over specificity as part of a two-step screening process. Indeed, low sensitivity scores mean that a certain number of people who would appropriate for early intervention are not being identified. Consequently, in most cases, having a few more false positives is less of an issue than missing appropriate individuals from a clinical perspective.

Compared to the PQ-B cut-off of ≥ 6 on total distress score, the proposed ≥ 3 threshold on symptom total score (range 0-21)¹³ slightly increased specificity value up to 28.4%, while maintaining a 91.4% sensitivity. Finally, using the recommended PQ-B cut-off of ≥ 4 on distressing item total score (range 0-21)²¹, even if specificity increased to 75.3%, a sensitivity value of 61.4% is quite low and means that a relevant number of people who would appropriate for early intervention services are not being identified.

Among all iPQ-B items, five symptoms (i.e. iPQ-B5 ["Have you felt that you are not in control of your own ideas or thoughts?"], iPQ-B6 ["Do you have difficulty getting your point across because you ramble or go off the track a lot when you talk?"], iPQ-B8 ["Do you feel that other people are watching you or talking about you?"], iPQ-B9 ["Do you sometimes get strange feelings on or just beneath your skin, like bugs or crawling?"], and iPQ-B13 ["Have you ever felt you don't exist, the world does not exist, or that you are dead?"]) correctly predicted UHR- vs UHR+ diagnoses (76.7% of correct diagnostic ascription in the logistic regression model). It should be noted that iPQ-B9 showed a significant negative correlation coefficient, thus making its absence full of meaning. This is in line with the above mentioned result emerging from Cronbach's statistics, which concluded for the removal of iPQ-B item 9. Although this item is intended to be an attenuated psychotic symptom, the results of the logistic regression suggest that on a group-level, subjects endorsing it might actually report paresthetic sensations (e.g. feelings of pins and needles is when their arms or legs "fall asleep") or somato-vegetative expressions of anxiety. In their early pilot study, Yung et al. (1996)³² reported that three kinds of attenuated positive symptoms (i.e. perceptual abnormalities, suspiciousness, and delusional mood) could account for 62%, 71%, and 62%, respectively, of symptoms that prodromal individuals experienced.

Limitations

Firstly, a possible limitation of this study is that the iPQ-B 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-B 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^{2,14}. Indeed, a certain number of false positives would be identified.

Another limitation is that since the SS for psychosis¹⁷ was used in the eligibility triage for the ReARMS protocol (i.e. before the iPQ-B 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 process¹³. 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 investigation²⁷. In this respect, Loewy et al. (2011)¹³ included 4-point scale questions on distress following each PQ-B item to examine if this enhanced the PPV of the instrument.

Conclusions

The Italian version of the PQ-B showed satisfying psychometric properties, comparable to 92-item homologue²⁵. 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-B 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-B 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

The current study received no specific grant or other funding from any funding agency, commercial or not-for-profit sector. It also had no relationship that may pose conflict of interest. The authors have no competing interests. ReARMS project is partly financed through a special regional fund: "Progetto Esordi Psicotici della Regione Emilia Romagna". Our special thanks go to all the staff and service users of all the outpatient community mental health services of the Reggio Emilia Department of Mental Health and Pathological Addiction.

Conflict of Interest

The authors are no conflict of interests.

References

- Fusar-Poli P, Bonoldi I, Yung AR, et al. *Predicting psychosis: meta-analysis of transition outcomes in individuals at high clinical risk*. Arch Gen Psychiatry 2012;69:220-9.
- Lorenzo P, Silvia A, Federica P, et al. *The Italian version of the 16-item Prodromal Questionnaire (iPQ-16): field-test and psychometric features*. Schizophr Res 2018 (in press).
- 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.
- 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.
- 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.
- Azzali S, Pelizza L, Paterlini F, et al. *Reliability of the 16-item Prodromal Questionnaire (iPQ-16) for psychosis risk screening in a young help-seeking community sample*. Journal of Psychopathology 2018;24:16-23.
- Miller TJ, McGlashan TH, Rosen JL, et al. *Prodromal assessment with the structured interview for prodromal syndromes and the scale of prodromal symptoms: predictive validity, interrater reliability, and training to reliability*. Schizophr Bull 2003;29:703-15.
- Fusar-Poli P, Cappucciati M, Rutigliano G, et al. *At risk or not at risk? Meta-analysis of the prognostic accuracy of psychometric interviews for psychosis prediction*. World Psychiatry 2015;14:322-32.
- 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;160:110-7.
- Addington J, Stowkowy J, Weiser M. *Screening tools for clinical high risk for psychosis*. Early Interv Psychiatry 2015;9:345-56.
- 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;Jul 26. [Epub ahead of print]
- 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.
- 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.
- 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 2018;27:725-37.
- 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.
- 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.
- 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.
- American Psychiatric Association (APA). *Diagnostic and statistical manual of mental disorders, V edition*. Arlington: American Psychiatric Publishing Inc. 2013.
- Raballo A, Semrov E, Bonner Y, et al. *Traduzione e adattamento italiano della CAARMS (the Comprehensive Assessment of At Risk Mental States)*. Bologna: Regione Emilia-Romagna 2013.
- Preti A, Raballo A. *PQ-B: authorized Italian version*. Cagliari: Studio CAPIRE 2011.
- 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* 2018;12:3-14.
- ²² Kline E, Thompson E, Bussell K, et al. *Psychotic-like experiences and distress among adolescents using mental health services*. *Schizophr Res* 2014;152:498-502.
- ²³ Thompson E, Kline E, Reeves G, et al. *Identifying youth at risk for psychosis using the behavior assessment system for children*. *Schizophr Res* 2013;151:238-44.
- ²⁴ SPSS Inc. *SPSS for Windows*, rel. 18.0. Chicago, IL: SPSS Inc. 2010.
- ²⁵ 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;189:50-6.
- ²⁶ Heise DR. *Separating reliability and stability in test-retest correlation*. *Am Sociol Rev* 1969;34:93-9.
- ²⁷ 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.
- ²⁸ Green SB, Yang Y, Alt M, et al. *Use of internal consistency coefficients for estimating reliability of experimental task scores*. *Psychon Bull Rev* 2016;23:750-63.
- ²⁹ Landis JR, Koch GG. *The measurement of observer agreement for categorical data*. *Biometrics* 1997;33:159-74.
- ³⁰ 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.
- ³¹ Xu L, Zhang T, Zheng L, et al. *Psychometric properties of Prodromal Questionnaire – Brief version among Chinese help-seeking individuals*. *PLoS One* 2016;11:e0148935;
- ³² Yung AR, McGorry PD, McFarlane CA, et al. *Monitoring and care of young people at incipient risk of psychosis*. *Schizophr Bull* 1996;22:283-303.