An evaluation of psychometric properties of the client's questionnaire of the Wisconsin Quality of Life Index-Canadian version (CaW-QLI)

Pablo Diaz[1], Céline Mercier[2], Raymonde Hachey[3], Jean Caron[4] & Geneviève Boyer[5]

[1]*Schizophrenia & Continuing Care Program, Centre for Addiction & Mental Health, Toronto, ON, Canada;* [2]*Department of Psychiatry, McGill University, Douglas Hospital Research Centre, Montreal, Canada;* [3]*School of Rehabilitation, University of Montreal, Montreal, Canada;* [4]*Department of Behavioral Sciences, Québec University in Abitibi-Temiscamingue, Rouyn, Canada;* [5]*Douglas Hospital Research Centre, Montreal, Canada*

Abstract. Psychometric properties of Client's questionnaire from the Canadian version of the Wisconsin Quality of Life-Index (CaW-QLI) were assessed with two groups, 89 English (E) and 94 French (F) individuals with schizophrenia or schizoaffective disorder. Sub-samples of 40 E and 36 F were re-interviewed within a 2-week period. Spearman correlations-SC between each domain and CaW-QLI global Score ranged from 0.39 to 0.76, while inter-domain correlations were low, confirming the multidimensional properties of the scale. Cronbach's α (internal consistency) were 0.78 (E) and 0.70 (F) for the CaW-QLI global scores and, from 0.45 to 0.88 among seven of eight domains. Test–retest (Concordance Correlation Coefficient-CCC) ranged from 0.36 to 0.80 among the domains, and from 0.80 (E) and 0.85 (F) between CaW-QLI global scores. Regarding convergent validity, SC between CaW-QLI global score and Spitzer's QOL-Index were 0.72 (E) and 0.58 (F). As hypothesized, there were higher correlations between CaW-QLI global scores (E and F) and SF-36 scales related to mental health than those related to physical health. Minor changes in the scoring are proposed to enhance face and content validity.

Introduction

In the last 15 years, there has been a growing research on the quality of life of individuals with severe psychiatric disorders such as schizophrenia. Multidimensional and psychometrically-sound tools to assess quality of life (QOL) of this population are urgently needed for use: 1) in the community-based programs being implemented; 2) in clinical trials of new drugs; 3) in clinical research on innovative and standard treatments; 4) in clinical interventions to assess clients' needs and guide clinical treatment decisions; 5) in policy making related to the development of services and deployment of resources [1–4].

Since 1980, a number of scales have been developed to measure QOL in persons with severe mental illness [1, 5–9]. However, most scales are time consuming, require trained interviewers, and/or have psychometric properties that have not been adequately evaluated. In a recent literature review of 11 instruments assessing QOL in psychiatry, Nieuwenhuizen [8] indicated that three instruments, 1) the QOL Scale [10]; 2) the QOL Interview [11] and 3) the Lancashire QOL Profile [12], have been comprehensively evaluated on the grounds of their psychometric properties. According to Revicki [13], the most frequently used instruments in clinical trials of antipsychotic medications are Lehman's QOL Interview and Heinrichs' QOL Scale.

Scales developed more recently have attempted to address some of the limitations of the first generation instruments, such as length of administration, main focus on research or limited scope. In a recent publication, Zissi [14] proposed a 'Meditational' model of QOL to link patient's subjective and self related concepts as a theoretical construct in a population of the chronically mentally ill.

The Wisconsin Quality of Life Index (W-QLI), [15], although a fairly new tool, seems to be quite promising. The W-QLI was developed when a state Medicaid agency asked the authors to design an instrument to be used in cost-benefit studies of clozapine [2]. It has incorporated most of the characteristics that should be inherent of a good QOL clinical and research tool. It is increasingly being incorporated into clinical studies [1].

The original validation study was conducted on 40 patients [15] to assess the construct validity. The relationship was established between the functioning level and the BPRS score. The patient's score on daily life functioning showed little correlation with symptoms, as expected. The test–retest reliability was

examined with 10 participants on a 3–10 day period, with percentages of agreement varying from 0.82 to 0.87 for each domain and total score. The convergent validity for the client's questionnaire has been assessed with the Spitzer's QL-Index (Pearson correlation, 0.91) and the Spitzer's Uniscale (Pearson correlation, 0.68). Given the current popularity for the W-QLI, and the limited knowledge of its psychometric properties, further validation studies in different settings are of significant interest.

The W-QLI has been modified to fit the characteristics of the Canadian population, the universal Canadian health system, and community and social services in Canada [16]. Apart from these modifications mainly in the background information section, the Wisconsin has been kept the same.

The purpose of this study was to validate the client's questionnaire of the CaW-QLI in both, English and French language by assessing: 1) its face, content and construct validity; 2) its internal consistency; 3) its test–retest reliability within a 2-week time frame; 4) its convergent and discriminant validity, using the Spitzer's QL-Index and Uniscale (patient/client version) [17], and the MOS SF-36 (SF-36) [18]. Due to the comparability between the American and the Canadian versions of the W-QLI, the conclusions from this study are also relevant for the original W-QLI.

Method

Description of the Wisconsin Quality of Life Index

The W-QLI is the only scale currently available which solicits information about the client, from three perspectives: directly from the client, from a professional care provider, and from a family member (whenever possible). The dimensions covered in the client's W-QLI are the following: satisfaction level with different life domains, occupational activities, psychological well-being, symptoms, physical health, social relations/support, finances, and activities of daily living. Some of these subscales include existing scales: the satisfaction with the Life Domains Scale [19]; the Bradburn's Affect Balance Scale [20]; the outcome scale related to frequency and type of social contact of the International Pilot Study of Schizophrenia [21]; and the Spitzer's QL-Index and Uniscale (patient/ client version) [17].

Each domain is scored individually. A global QOL score can then be calculated by summing the scores of each domain and dividing them by the number of domains. The clients are also able to weight their perception of importance of each domain at the end of the scale. This is an unique property of the W-QLI which produces a global weighted score.

A further attractive feature of this scale is that the client, provider, and family member questionnaires contain sections requesting the respondent to indicate three treatment-related goals, the relative importance of each goal (on a 10-point scale), and the degree to which the respondent believes each goal has been achieved.

The modifications to the original W-QLI have been done in the English and simultaneously in the French translation. This French translation was produced following the procedures developed by Flaherty [22] and by Guillemin and his colleagues [23]. The intention of these procedures is to produce a version which is semantically and conceptually equivalent to the original version. Both the English- and French-Canadian editions of the questionnaires were approved by the original team of authors from the University of Wisconsin–Madison.

Study procedures

Participants
Individuals between the ages of 18 and 65 years, with a diagnosis of schizophrenia or schizo-affective disorder, according to the Diagnosis and Statistical Manual of Mental Disorders (DSM-IV criteria), diagnosed by certified psychiatrists and had been treated through mental health services for at least 24 months prior to their participation. Individuals needed to be physically healthy and able to give informed consent. This study was approved by the Research and Ethics Committees of the four participating hospitals.

Participants were recruited in two Canadian centres, Halifax (Nova Scotia), and Montreal (Quebec), from outpatient services, i.e., rehabilitation programs, socio-professional services or residential community resources. Clinicians were approached to identify and recruit participants who met the study's inclusion criteria. Those who agreed to participate were interviewed by the research assistant who explained the nature of the study. Participants read and signed comprehensive consent forms to carry out the interview, and a chart audit.

Data collection
Although, the W-QLI was designed as a self-administred instrument the data collection was conducted by trained interviewers who provided assistance if needed. Ten dollars was given to each client as an honorarium for participation.

Sub-samples of 36 English-speaking and 39 French-speaking clients were re-interviewed within a 2-week period. In Nova Scotia the sub-sample was randomized, but in Quebec, those re-tested, were the first ones to agree to be re-interviewed.

Chart audits from the clients' charts were done to collect demographic and clinical data not included in the CaW-QLI. The average administration time of the client's CaW-QLI was 25 min.

Table 1. Means, standard deviations of the CaW-QLI domains and Spearman correlations with the CaW-QLI global score, the Spitzer uniscale and Spitzer-QLI (Total sample = 183 patients)

	Domains scores				Spearman correlation with CaW-QLI global score	
	English		French		English	French
	Mean	SD	Mean	SD	r	r
CaW-QLI domains (−3 to +3)						
Satisfaction level	1.04	1.32	0.78	1.34	0.56*	0.70*
Occupational activities	0.73	1.62	0.82	1.38	0.65*	0.51*
Psychological well-being	0.68	1.57	0.87	1.33	0.76*	0.64*
Symptoms/attitudes	1.44	1.28	1.74	1.05	0.67*	0.50*
Physical health	0.69	1.50	0.67	1.41	0.70*	0.50*
Social relations/support	0.95	1.55	1.02	1.47	0.70*	0.62*
Finances	0.17	1.52	0.30	1.45	0.64*	0.60*
Activities of daily living	1.06	1.67	1.04	1.75	0.39*	0.43*
CaW-QLI global score	0.85	0.92	0.91	0.80	–	–
Spitzer-QL index (0 to 10)	6.74*	2.37	7.66*	2.02	0.72*	0.58*
Spitzer uniscale (1 to 10)	7.18	2.31	6.69	1.91	0.36*	0.56

*$p < 0.01$.

Data analysis

Descriptive statistics were used to report the two clinical groups in terms of socio-demographic and clinical data. Construct validity was assessed using Spearman correlations between domains, and between domains and the global score of the CaW-QLI. Internal consistency was calculated for all subscales using Cronbach's α. The test–retest reliability was determined by using Concordance Correlation Coefficient-CCC [24]. To test the convergent validity, Spearman correlations were used between the Spitzer's QL-Index and Uniscale, the SF-36 and the global CaW-QLI score. Regarding the SF-36, our hypothesis was that it would be a higher correlation with the mental health and role emotion subscales than the other subscales.

Results

Participants' characteristics

The total population sample from the two urban centres was 183, including 89 English-speaking clients (E) and 94 French-speaking clients (F). Most of the participants were caucasian and 83.1% were single. *t*-test and χ^2-tests between the English- and French-speaking clients indicated that the French-speaking participants were younger (mean 39.5 years vs. 44.4; $t = -3.40$, $p < 0.01$), and that their mean level of education was higher (mean 11.0 years vs. 9.6; $t = 3.23$, $p < 0.01$) than for English-speaking participants. In the French sample there were more males (18.9% vs. 67.4%, $\chi^2 = 4.12$, $p < 0.05$), more participants living alone (45.7% vs. 23.6%, $\chi^2 = 9.55$, $p < 0.01$) and receiving welfare (92.6% vs. 80.9%, $\chi^2 = 4.74$, $p < 0.05$) than in the English sample

(Table 1). The living arrangements were different ($\chi^2 = 31.73$, $p < 0.01$) for the French and English samples. More French participants were living in apartment/home (41.5% vs. 22.5%), in rooming houses (10.6% vs. 5.6%), foster homes (11.7% vs. 3.4%), and supervised apartments (10.6% vs. 3.4%). On the other hand, more English participants were living in group homes (29% vs. 14.9%), and institutions (12.4% vs. 1.1%). There was no difference in terms of diagnosis or sub-types of schizophrenia between the two samples, 51.9% had schizophrenia, paranoid type; 9.3% undifferentiated, and 4.9% residual type. Only 14.2% had schizo-affective disorder.

Domains and global quality of life scores

Table 1 reports the means and standard deviations in the eight life domains for both English and French participants[a]. No statistically significant differences were observed between the two groups for the CaW-QLI, but there was one for the Spitzer's QL-Index ($t = 2.69$, $p < 0.01$). According to this data, both groups of clients rated the symptoms domain as the highest (E: 1.44; F: 1.74) and the finances domain as the lowest (E: 0.17; F: 0.30).

Construct validity

The CaW-QLI is based upon the theory that QOL is a complex multidimensional construct. This implies that each domain can measure a distinct aspect of QOL which, at the same time, is related to the same

[a] The scores in each item of the W-QLI range from −3 to +3. The scores in the Spitzer's QL-Index range from 0 to 10 and Spitzer's Uniscale from 1 to 10.

Table 2. Cronbach's α and concordance correlation coefficient between test and retest for every domain and the CaW-QLI global score

	# Items	Cronbach's α		Concordance Correlation Coefficient: Test–retest	
		English (n = 89)	French (n = 94)	English (n = 36)	French (n = 39)
CaW-QLI domains					
Satisfaction level	10	0.86	0.88	0.74	0.51
Occupational activities	3	0.33	0.08	0.71	0.66
Psychological well-being	10	0.73	0.63	0.53	0.69
Symptoms/attitudes	5	0.81	0.73	0.72	0.83
Physical health	3	0.63	0.63	0.47	0.65
Social relations/support	6	0.75	0.82	0.68	0.60
Finances	3	0.45	0.58	0.57	0.80
Activities of daily living	6	0.64	0.56	0.76	0.36
Global score	8	0.78	0.70	0.80	0.85

underlying construct. If the domains are distinct from one another, the correlations between domains should be lower than the ones between each domain and the global score. On the other hand, if these dimensions pertain to the same underlying concept, one should expect significant correlations between each domain and the global score.

As expected, each domain demonstrated significant Spearman correlations with the global score (Table 1). For seven of the eight domains, these correlations were over 0.56 for the English and over 0.50 for the French clients. The correlations between domains were low, and significantly lower in some areas. They ranged from 0.07 (satisfaction level and activities of daily living) to 0.59 (psychological well-being and symptoms) in the English sample and from 0.01 (occupational activities and activities of daily living) to 0.52 (physical health and symptoms) in the French sample.

Internal consistency

The Cronbach's α for the whole instrument were 0.78 and 0.70 for the English and French clients respectively (Table 2). For seven of the eight domains, they ranged from 0.45 to 0.86 (E) and from 0.56 to 0.88 (F). For the domain of occupational activities, they were respectively 0.33 (E) and 0.08 (F). It has to be noted that the α in the domains of psychological well-being and activities of daily living were obtained after the item from the Spitzer's QL-Index was removed from the respective subscale.

Test–retest reliability

Concordance correlation coefficient ranged from 0.47 to 0.76 among the QOL domains in the English population, whereas it ranged between 0.36 and 0.83 in the French one (Table 2). The test–retest for the global score were 0.80 (E) and 0.85 (F).

Table 3. Spearman correlations between CaW-QLI global score and MOS SF-36 dimensions

	English	French
Physical functioning	0.15	0.27*
Role physical	0.40**	0.39**
Bodily pain	0.36**	0.25
General health	0.52**	0.41**
Vitality	0.41**	0.43**
Social functioning	0.38**	0.51**
Role emotional	0.51**	0.46**
Mental health	0.53**	0.59**

$*p < 0.05;$ $**p < 0.01.$

Convergent and discriminant validity

Convergent validity was assessed using the Spearman correlations between the CaW-QLI global score, the Spitzer's QL-Index and Uniscale. The correlations between the CaW-QLI global score and the Spitzer's QL-Index were 0.72 and 0.58 for the English and French clients, and 0.36 and 0.56 with the Uniscale (Table 1).

Discriminant validity was assessed by comparing the correlation between the CaW-QLI global score and the SF-36 subscales (Table 3). As expected, the correlations were higher with dimensions of the SF-36 related to mental health (E: 0.53; F: 0.59) and role emotional (E: 0.51; F: 0.46), and lower with the dimensions related to physical functioning (E: 0.15; F: 0.27), bodily pain (E: 0.36; F: 0.25) and role physical (E: 0.40; F: 0.39).

Discussion

The purpose of this study was to contribute to the further validation of English and French versions of the CaW-QLI. The results showed that the internal consistency for the whole scale and for four of the

domains (satisfaction level, symptoms attitude, psychological well-being, social relation/support) was good. On the other hand, the α in the finance (3 items), activity of daily living (6 items), and physical health (3 items) domains were lower (from 0.45 to 0.64). According to Martinez Arias [25] and Gulliksen [26], two factors influence the variability of the internal consistency: the sample size and the number of items. The larger the sample, the higher the α. By the same token, the larger the number of items, the higher the α. They suggest that scales that have 10 or more items should have α scores of 0.70 or higher, but a 4-item scale, an α around 0.50 is acceptable[b]. The domain of occupational activities showed a very low α, particularly the French version. This could be explained by the fact that this domain has only three items, and the scoring for two items is inconsistent, despite addressing similar areas (example: doing less activities is scored −3 in one item, and 0 in the other). In order to improve the internal consistency, the scoring should be revised.

The test–retests reliability for the global scores are good (F and E). The test–retests were acceptable for most of the domains; however, the physical health (E), psychological well-being (E), and activities of daily living (F) domains were less stable.

The analysis indicated a good construct validity. As expected, each domain demonstrated higher correlations with the global CaW-QLI score than with any other domain. These inter-domain correlations are comparable partly with Becker's data who reported the lowest correlations between finances and activities of daily living (0.07) and highest between occupational activities and social relations (0.70) (results not available for global score).

As for the convergent validity, the correlations between the Spitzer's QL-Index/Uniscale, and the W-QLI global score were lower than those obtained by Becker and her team [14], which were respectively 0.90 (with the Index) and 0.68 (with the Uniscale). The correlations between the CaW-QLI global score and the scores of SF-36 subsclaes were higher with domains related to mental health than those related to physical health. This seems to indicate that the W-QLI captures features regarding QOL and mental health.

Regarding face and, the lowest score attributed to the financial domain was congruent with previous findings [27], however the highest score associated with the symptoms domain was intriguing. In reviewing the scoring procedures from the original authors, it was discovered that some items of this domain were over-rated[c]. For instance, in the item

that measures suicidality, a score of 0 equals "feeling like killing yourself frequently". So, even if this score of 0 is theoretically neutral, it reflects something negative in this example. The same applies to some items of the occupational activities, physical health domains and to the Spitzer's QL-Index items. In this case, when re-scoring the original scores of 0, 1 and 2 in −3 to +3, the original score of 0 becomes −3, the score 1 becomes 0, and 2 becomes +3. The scoring procedures of these three domains should be addressed.

A recommendation is to change the scoring of each item that produces an over scoring by recalibrating the response choice so that a zero score reflects a neutral state.

The domains included in the CaW-QLI were chosen because they were identified as important for the QOL and because they had been shown to be good predictors of outcome in schizophrenia. The fact that some existing scales were included within a domain helps to improve the content validity. However, in the case of the Spitzer's QL-Index, this could be questionable. This scale was originally developed for people with cancer, and the best expected improvement was a revert to previous level of functioning, despite the impact of the illness or the treatment. The situation is different in outcome studies with individuals with schizophrenia, where an improvement in comparison with the baseline condition is expected following pharmacotherapy and/or psychosocial rehabilitation. It could be problematic, in the context of outcome studies, if the highest score is attributed to the 'usual' level of functioning. Consequently, it is recommended to remove the five questions of the Spitzer's QL-Index (which are distributed in five of the eight domains) from the scoring of the domains. If clinicians/researchers want to keep the Spitzer inside the W-QLI, it is recommended to score the Spitzer's QL-Index separately according to its original scoring formula [15]. Our strategy is to replace the Spitzer items by equivalent questions designed to measure neutral points, and to capture improvement and deterioration.

In conclusion, the CaW-QLI has unique characteristics not available in existing QOL scales. Among them are that: 1 – Solicits information from clients, care providers and families; 2 – Satisfaction scores are weighted by the level of importance according to the client; 3 – Collects useful clinical information to guide clinicians in a more client driven treatment objectives. (This was confirmed in our study by most of the participants and care givers who were interested in using it as a clinical tool.)

These characteristics can explain its rapid popularity and justify the efforts to address the psychometric weaknesses outlined in this paper, in order to enhance its future as possible goal standard outcome measure for clinical and research purposes.

[b] According to these criteria, seven out of eight subscales in the CaW-QLI have good to acceptable internal consistency.

[c] An item or a scale is over-rated when a positive score does not equal a positive aspect, or a score of 0 does not equal neutrality.

514

References

1. Awad AG, Voruganti LNP, Heslegrave RJ. Measuring quality of life in patients with schizophrenia. Pharmacoeconomics 1997; 11(1): 32–47.
2. Becker M, Diamond R. New developments in quality of life measurement in schizophrenia. In: Katsching H, Freeman H, Sartorius N (eds), Quality of life in mental disorders. Chichester, England: John Wiley & Sons, 1997: 119–133.
3. Katschnig H, Freeman H, Sartorius N, eds. Quality of life in mental disorders. Chichester, England: John Wiley & Sons, 1997.
4. Ware JE, Gandek B. The SF-36 Health Survey: Development and use in mental health research and the IQOLA project. Int J Ment Health 1994; 23: 49–73.
5. Atkinson MJ, Zibin S. Quality of life measurement among persons with chronic mental illness: A critique of measures and methods. Ottawa: Health Promotion and Programs Branch, Health Canada, 1996.
6. Lehman AF, Burns BJ. Severe mental illness in the community. In: B Spilker (ed), Quality of life assessments in clinical trials. New York: Raven Press Ltd., 1990: 357–362.
7. Lehman AF. Instruments for measuring quality of life in mental illnesses. In: Katschnig H, Freeman H, Sartorius N (eds), Quality of life in mental disorders. Chichester, England: John Wiley & Sons, 1997: 80–94.
8. Nieuwenhuizen CV, Schene AH, Boevink WA, Wolf JK, et al. Measuring the quality of life of clients with severe mental illness: A review of instruments. Psychiat Rehabilitation J 1997; 20(4): 33–41.
9. Voruganti L, Heslegrave RJ, Awad AG. Quality of life measurement during antipsychotic drug therapy of schizophrenia. J Psychiatry Neurosci 1997; 22(4): 267–274.
10. Heinrichs DW, Hanlon TE, Carpenter Jr WT. The quality of life scale: An instrument for rating the schizophrenia deficit syndrome. Schizophrenia Bull 1984; 10(3): 388–398.
11. Lehman AF. A quality of life interview for the chronically mentally ill. Eval Program Plann 1988; 11(1): 51–62.
12. Oliver JPJ. The social care directive: Development of a quality of life profile for use in community services for the mentally ill. Soc Work & Soc Sci Rev 1991; 3(1): 5–45.
13. Revicki DA. Methods of Pharmacoeconomic evaluation of psychopharmacologic therapies for patients with schizophrenia. J Psychiat Neurosci 1997; 22(4): 256–266.
14. Zissi A, Barry MM, Cochrane R. A mediational model of quality of life for individuals with severe mental health problems. Psychological Medicine 1998; 28, 1221–1230.
15. Becker M, Diamond R, Sainfort FA. A new patient focused index for measuring quality of life in persons with severe and persistent mental illness. Quality of Life Research 1993; 2(4): 239–251.
16. Diaz P, Mercier C. An evaluation of the Wisconsin Quality of Life Questionnaires for clinical application and research in Canada. Quality of Life Newsletter 1996; 16(2): 11–12.
17. Spitzer WO, Dobson AJ, Hall J, et al. Measuring the quality of life of cancer patients. A concise QL-Index for use by physicians. J Chron Disease 1981; 34: 585–597.
18. Ware JE, Snow KK, Kosinski M, Gandek B. SF-36: Health survey manual and interpretation guide. Boston: New England Medical Center, The Health Institute, 1993.
19. Baker F, Intagliata J. Quality of life in the evaluation of community support systems. Eval Program Plann 1982; 5: 69–79.
20. Bradburn NM. The structure of psychological well-being. Chicago: Aldine, 1969.
21. Strauss JS, Carpenter WT. Prediction of outcome in schizophrenia. Part II. Arch Gen Psychiat 1974; 31: 37–42.
22. Flaherty J. Developing instruments for cross-cultural psychiatric research. J Nerv Ment Dis 1986; 46(176): 5.
23. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. J Clini Epidemi 1993; 46: 1417–1432.
24. Lin L. A Concordance correlation coefficient to evaluate reproducibility. Biometrics 1989 Mar, 45(1): 225–268.
25. Martinez Arias R. Psicometria: Teoria de los Tests Psicologicos y Educativos. Madrid: Editorial Sintesis, SA, 1995.
26. Gulliksen H. Theory of mental test. New York: John Wiley, 1950.
27. Mercier C. Improving the quality of life of people with severe mental disorders. In: Romney DM, Brown R, Fry P (eds), Improving the quality of life in normal and disabled populations. Boston: Kluwer Academic Publishers, 1994: 165–192.

Address for correspondence: Pablo Diaz, Schizophrenia and Continuing Care Program, Centre for Addiction and Mental Health, Clarke Site, 250 College Street, Room 729A, Toronto, ON MST1R8, Canada
Tel: (416) 979-4747 Ext 2387; Fax: (416) 979-6936