Original article

Latent class analysis of the health of the nation outcome scales: A comparison of Swiss and English profiles and exploration of their predictive utility

P. Maddison [a,*], P. Golay [b], S.J. Muncer [c]

[a] *Currency and tariff development, Tees Esk & Wear Vallleys NHS Foundation Trust, Flatts Lane, Middlesbrough, Normanby, UK*
[b] *Department of psychiatry, Lausanne university hospital (CHUV), Lausanne, Switzerland*
[c] *Department of clinical psychology, university of Teesside, Middlesbrough, UK*

1. Introduction

The Health of the nation outcome scales (HoNOS) [1] were designed to measure the health and social functioning of adults with severe mental health problems. They form part of the English mental health minimum data set and are recommended by the department of health and are part of the attempt to develop "payment by results" (PbR) for mental health [2]. They are also widely used in Australia, New Zealand and Canada [3,4], and have also been used in Europe [5]. Although they are widely used there are still questions about their psychometric validity and their ability to predict anything useful.

Originally, it was claimed that the HoNOS contained 12 separate scales which were independent of each other [1]. However, it has been shown clearly that the scales are not independent [6], which is also supported by the numerous studies that have shown that it contains four or more factors [7–12]. Unfortunately, the four or more factors that have been revealed are not always the same. The Speak factor structure [10,11] has the most evidence in support

as it has been found most often and with studies that use the largest and most representative samples. However, although this model appears to be the best it is still far from satisfactory in terms of fit, and it has also been shown that the fit is worst for the most common type of mental health problems [11]. Indeed, the proponents of the four factor model have proposed that it may be better to not use all of the items in HoNOS and instead concentrate on a two-factor model which contains two factors one measuring Depression and the other a measure of Social and cognitive problems [13,14]. Overall, it would be fair to say that there are still some questions to be answered about the psychometric properties of HoNOS.

Research into the ability of HoNOS to predict health care costs also produces mixed results. HoNOS has been used in Australia and New Zealand as part of a casemix classification system [15,16] which found associations with cost. However, it is difficult to work out the precise role of HoNOS and in particular what it adds to the predictive success, particularly as diagnosis was also used to define clusters. A more direct test of the predictive validity of HoNOS in Canada found that total baseline HoNOS score was significantly associated with in and outpatient service use including admissions, bed days, and psychiatric contacts [4]. Furthermore, an attempt to adopt a case mix approach in Germany used HoNOS as

---

one of the variables in assigning groups and overall explained 17% of the duration of stay in hospital. Again, total score on HoNOS was used as a predictor [17].

More recently, however, HoNOS total score was found not to be a useful predictor of mental health service costs in a sample of patients with common mental health problems, and indeed only the "self injury" item showed any relationship [18]. Golay et al. [19] also found that overall, the HoNOS items had weak predictive validity for duration of stay in hospital, re-hospitalization and also time before re-hospitalization. However, they conducted a latent class analysis on their HoNOS scores to reveal five distinct profiles of patients. These classes were significantly associated with different durations of hospitalization, and also the re-hospitalization variables.

It is, therefore, possible that HoNOS might have a role to play in prediction and this might be best explored by developing profiles based on latent class analysis. The aim of the present study is to investigate this possibility on a large sample of mental health patients who have had an inpatient stay during their care within a large mental health provider in the North East of England. The study will also compare the model derived from the English data to the model found in the Swiss data to assess the consistency of the approach across locations and samples. Only data relating to Working aged adult and older person services were included in the current study. At the same time, these results can be compared to other methods of scoring HoNOS by using the total score, the item scores and the various factor scores.

2. Method

2.1. Sample characteristics

Tees Esk & Wear Valleys NHS Foundation Trust (TEWV) is a large mental health provider in the North East of England. Within TEWV, the HoNOS is routinely rated at key points during a patient's care, including at the point of admission to an inpatient ward. The data used in the current study included HoNOS ratings for all patients who had an inpatient stay between October 2011 and October 2013.

In total, 2325 HoNOS records were identified. Of which, 1279 were male (55%) and 1046 were female (45%), with a mean age of 40.84 years (SD = 13.16). In line with PbR developments in the UK, TEWV has adopted a patient classification system that groups patients based on their level of need. As part of this system, all patients are allocated to a "Super class" that summarises overarching disorder types into non-psychotic, psychosis and organic. The current sample was classified using the super class system as follows: 49.59% had a non-psychotic disorder (encompassing mood, anxiety, obsessive-compulsive, eating, and dissociative disorders), 48.73% had a psychosis disorder (encompassing schizophrenia, schizotypal, delusional and bi-polar disorders), 0.56% had an organic disorder (encompassing Alzheimer-s, vascular and frontotemporal dementia, unspecified and symptomatic disorders) and 1.12% had an undisclosed disorder.

2.2. Data extraction

HoNOS ratings were recorded, stored and extracted using TEWV's electronic patient record system. All of the HoNOS assessments were rated within 2 weeks of admission to the inpatient ward. For patients who had multiple inpatient stays during the 2-year period, only the first HoNOS assessment for each patient was used (subsequent HoNOS assessments were not considered). Only HoNOS assessments that contained a full set of valid scores (scales 1 to 12 rated 0 to 4) were used. In addition to the HoNOS records, other variables of interest were extracted for each

patient and included: the total length of the inpatient stay (captured in days); the date distance between discharge and re-admission (for those patients who had multiple inpatient stays, this captured the number of days between their discharge and re-admission) and the total number of inpatient stays (within the 2-year period of the extract, the total number of times a patient had an inpatient spell).

2.3. Statistical analysis

To verify the existence of specific patients HoNOS' profiles, a latent class analysis (LCA) was conducted on all 12 scales. LCA mainly differs from cluster analysis because it is model based, allows covariates (i.e. relating the class membership to external variables of interest) and classification uncertainty (i.e. for each patient a probability of class membership is given for each class). HoNOS items were dichotomized into "no serious problem" (scores 0, 1 and 2) and "severe problem" (scores 3 and 4) to reduce the number of model parameters and facilitate model estimation [19]. The best solution was determined using the Bayesian Information Criterion coefficient which balance model fit and model complexity (i.e. number of parameters [20]). A Lo-Mendell-Rubin Adjusted Likelihood Ratio Test and a Parametric Bootstrapped Likelihood Ratio Test were performed in order to determine whether a solution with one less class could present a similar degree of adjustment. The relationship between class-membership and distal outcomes (length of stay, time between discharge and re-admission and total number of inpatient stays) was estimated using a 3-step latent class regression model with the Lanza method for continuous or categorical distal variables [21,22]. With this approach the latent class analysis is first performed without being influenced by covariates. The second step is to record probabilities of class membership for each participant. The third and final step is to introduce the auxiliary variable in the model and to evaluate its relationship with class membership while taking classification uncertainty into account. Finally, in order to compare the Swiss and English classification, the model parameters were fixed according to the values of the Swiss LCA model, except for latent class means which were freely estimated. This allowed us to classify English participants according to the pre-determined Swiss model.

Correlation analyses were performed to assess the relationship between the HoNOS total score and the observed outcomes and the HoNOS factor scores and the observed outcomes. The factor structure used comprised emotional well-being (non-accidental self-injury, problems with depressed mood and other mental or behavioural problems), social well-being (problems drinking or drug taking, problems with relationships, problems with living conditions, problems with occupation and activities), personal well-being (cognitive problems, physical illness or disability, problems with activities of daily living and problems with occupation and activities) and severe disturbance (overactive, aggressive, disruptive or agitated behaviour and problems associated with hallucinations and delusions) [10].

All statistical tests were two-tailed and significance was determined at the 0.05 level. All statistical analyses were performed with the Mplus statistical package version 7.4 and IBM SPSS version 22.

3. Results

3.1. Latent class profile analysis

Characteristics from one to eight classes LCA are presented in Table 1. No model presented high entropy. The four-class solution was preferred on the basis of its lowest BIC and clinical interpretability. For the sake of parsimony, it was verified whether a solution

Table 1
Characteristics of the 1-8 class latent class analysis solutions.

Number of classes	Size of the smallest class	Entropy	BIC
1	2325	N/A	28217.193
2	479	0.696	27442.476
3	360	0.670	26816.105
4	202	0.655	26784.535
5	122	0.672	26788.572
6	99	0.640	26806.444
7	90	0.646	26852.677
8	56	0.653	26897.719

BIC: Bayesian information criterion.

with one less class could present a similar degree of adjustment. The Lo-Mendell-Rubin Adjusted Likelihood Ratio Test indicated a three class only solution could be acceptable ($P = 0.480$) but the Parametric Bootstrapped Likelihood Ratio Test indicated that the four class solution was preferable ($P < 0.001$). Given simulation studies have shown that the Parametric Bootstrapped Likelihood Ratio Test performed better than the Lo-Mendell-Rubin Adjusted Likelihood Ratio Test [23] the four-class solution was finally selected.

The characteristics of the four-class solution are presented in Fig. 1. The first class (9% of the sample) consisted of patients with a profile of depression with major social problems. The second class (13% of the sample) consisted of patients presenting with a psychosis disorder. The third class (40% of the sample) contained patients presenting with a profile of major depression with self-injury risk. The fourth and final class (38% of the sample) consisted of patients presenting with a profile of mild psychosis.

A cross-tabulation of the data was performed that assessed the relationship between the most likely four latent classes and the three PbR "Super Class" classifications. Reassuringly, there was a significant association between the latent classes and the Super Class allocation ($\chi^2[9] = 760.308$, $P = < 0.001$). Seventy-seven percent of the patients classified in the non-psychotic super class were allocated to latent classes one and three (encompassing

profiles of depression with major social problems and major depression with self-injury risk) and 79% of the patients classified in the psychosis super class were allocated to latent classes two and four (encompassing profiles of psychosis and mild psychosis). The small numbers of organic patients were spread across latent classes one, two and four.

3.2. Analysis of class membership and observed outcomes

Latent class membership was significantly associated with different inpatient durations ($\chi^2[3] = 226.360$, $P = < 0.001$). Longest to shortest inpatient durations were associated with class membership as follows: class two (Psychosis; mean = 47.29 days, SE = 3.89), class three (major depression with self-injury risk; mean = 36.74 days, SE = 1.69); class four (mild psychosis; mean = 16.37 days, SE = 0.54), and class one (depression with major social problems; mean = 11.58 days, SE = 0.99).

Latent class membership was significantly associated with time distances between discharge and re-admission ($\chi^2[3] = 199.295$, $P = < 0.001$). Interestingly, the latent class memberships that indicated the shortest hospital durations were associated with quicker hospitalisation re-admission rates. Shortest to longest re-admission rates for class membership were as follows: class one (mean = 32.67 days, SE = 8.73), class four (mean = 140.14 days, SE = 10.31), class three (mean = 194.13 days, SE = 12.21), and finally class two (mean = 313.39 years, SE = 25.11).

Class membership was significantly associated to the number of inpatient admissions that patients experienced over the 2-year period ($\chi^2[3] = 12.595$, $P = 0.006$). However, pairwise comparisons of the latent classes indicated that the difference between classes was minimal and that not all classes were significantly different from one another. Class one (mean = 1.246 stays, SE = 0.059) was significantly different to classes two (mean = 1.406 stays, SE = 0.050) and three (mean = 1.426 stays, SE = 0.031) but no other significant differences were apparent.
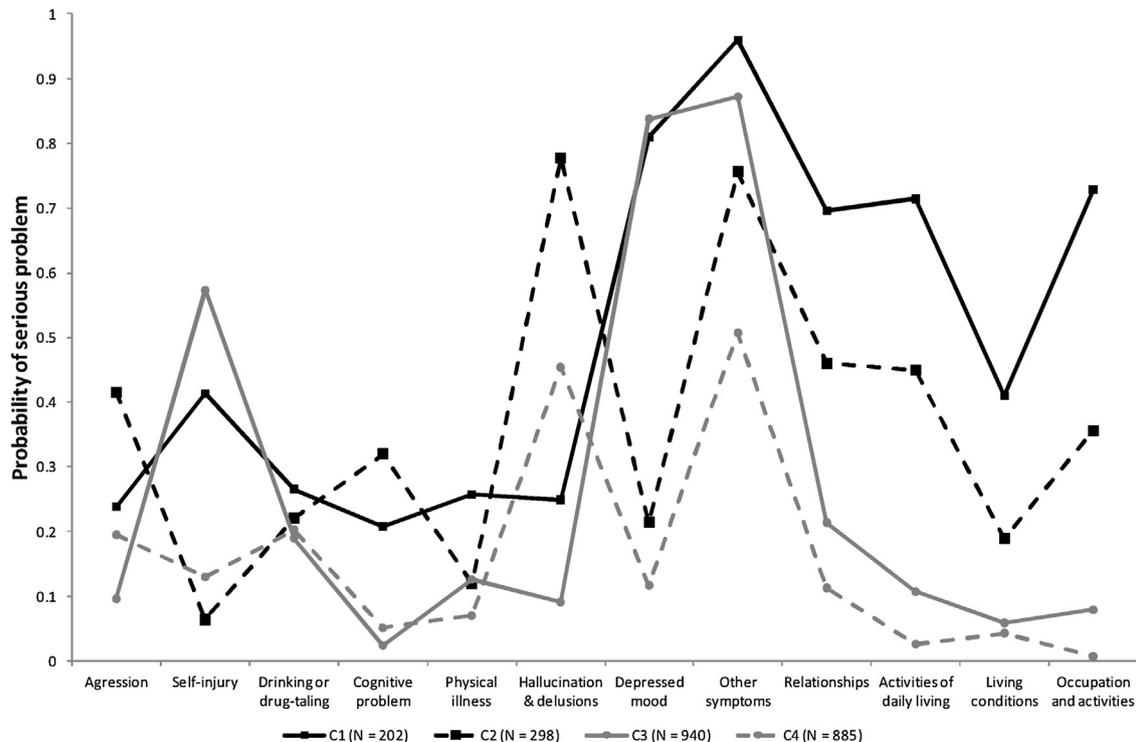
Fig. 1. Probability of having a serious problem in each item of the HoNOS within four latent classes. C1: depression with major social problems ($n = 202$); C2: psychosis ($n = 298$); C3: major depression and self-injury risk ($n = 940$); C4: mild psychosis ($n = 885$).

Table 2
Crosstabulation of English most likely class and Swiss most likely class allocation.

English most likely class	Swiss most likely class							
	Anxiety and depression with psychotic symptoms and social problems		Drug or alcohol use with no other problems		Major depression with no social problems but self-harm risk		Psychotic symptoms with no social problems	
	n	%	n	%	n	%	n	%
Depression with major social problems	121	59.90	1	0.50	80	39.60		0.00
Psychosis	110	36.91	21	7.05	17	5.70	150	50.34
Major depression with self-injury risk	7	0.74	134	14.26	799	85.00		0.00
Mild psychosis		0.00	476	53.79	5	0.56	404	45.65

Table 3
Exploration of the HoNOS (total score and factors) with other variables.

Correlations			
HoNOS level	Length of stay	Total number of inpatient stays	Distance between discharge and re-admission
Total score	0.013	0.000	−0.004
Emotional well-being	−0.211[b]	−0.034	−0.145[b]
Social well-being	0.007	0.00	−0.005
Personal well-being	0.147[b]	−0.001	0.045[a]
Severe disturbance	0.138[b]	0.052[a]	0.122[b]

[a] Correlation is significant at the 0.05 level.
[b] Correlation is significant at the 0.01 level.

3.3. Comparison of Swiss and English HoNOS profiles

The English profiles showed some similarities but also some differences with the Swiss profiles. The raw data was crosstabulated to assess how the English classifications mapped onto the Swiss classifications (Table 2). The classes reflecting depressive disorders showed a high level of agreement across both the English and the Swiss classes indicating similarities across samples. For instance, 85% of the patients classified in the English class 3 (major depression with self-injury risk) were also classified in the Swiss class 4 (major depression with no social problems but self-harm risk). Similarly, 60% of the patients classified in the English class 1 (depression with major social problems) were also classified in the Swiss class 1 (anxiety and depression with psychotic symptoms and social problems). However, the classes reflecting the psychotic disorders indicated less agreement across the English and the Swiss classes suggesting that casemixes relating to psychotic disorders may be more sample specific.

3.4. Total and factor score correlations

The correlation between HoNOS total and factor scores and the observed outcome measures are given in Table 3. It is clear that total HoNOS score has no significant relationship with any of the observed outcome measures. Only the severe disturbance factor has a significant positive relationship with all three-outcome measures, although this is a weak relationship, in particular with total number of inpatient stays. Emotional well-being has a significant negative relationship with all three outcome measures. Of particular interest is the negative relationship between emotional well-being and length of stay and between discharge and readmission. This suggests that perhaps patients with poor scores on EWB are being released from hospital too soon.

4. Discussion

Using LCA methods, we extracted four patient profiles from our data (derived from scoring patterns on HoNOS individual items). Entropy was moderate indicating heterogeneity within class remained after the classification process. Despite the relative simple structure, potentially interesting inter-individual differences may remain within each profile. The profiles were interpreted as representing four casemix types and are described as depression with major social problems, psychosis, major depression with self-injury and mild psychosis. The classes demonstrated relatively good concordance with the high level PbR "Super class" classification system in England. However, the results differ slightly from the paper by Golay et al. where five patient profiles were apparent. This may suggest that the Swiss sample was more heterogeneous because one extra class was needed in order to adequately reflect the patients' profile.

A comparison of the Swiss and English profiles indicated that there are some strong likenesses between the classes, particularly in those profiles that reflect depressive disorders. However, the psychotic profiles were not so strongly related and displayed a more dispersed mapping across the classes. Furthermore, the Swiss class 2 (drug or alcohol use with living condition and occupation problems) was not present in the English data. While the Swiss sample did not include records from Lausanne Alcohol withdrawal unit, records from Lausanne university hospital's psychiatry department included 13.9% hospital stays from the substance use withdrawal unit. This may explain why a specific class may have emerged in the Swiss sample.

In understanding the differences between the Swiss classes and English classes, it is important to consider the service configurations across both locations. The Lausanne university hospital's psychiatry department's catchment area is about 620′000. A total of 329 adult psychiatric beds are available on three geographical sites. Diverse outpatient settings are proposed. A specialized outpatient early psychosis program [24] and assertive community treatment teams are [25] also available and work in close collaboration with inpatient units. Regarding treatment options, personality disorder and substance abuse disorder could be treated within both inpatient and outpatient settings.

TEWV is a large mental health trust that provides a range of mental health, learning disability and eating disorder services across a large geographical area. A total of 1000 beds are available across 22 sites, providing services to patients requiring acute inpatient care, intensive care, challenging behaviour, 24 hour nursed care or rehabilitation services. Inpatient services in TEWV

form part of a planned and integrated whole system approach to care which is delivered in conjunction with community services. Although inpatient facilities are available for patients presenting with personality disorder and substance disorder, admissions are avoided wherever possible, with most treatment delivered in a community setting.

Using a LCA method for the HoNOS data indicated greater utility when looking at the relationship with observed outcomes than using the traditional HoNOS total/factor score. The HoNOS total and factor scores bared little to no relation with the observed outcomes. However, the classification system identified significant associations between class membership and inpatient duration and class membership and time distance to re-admission. The results indicated that those in the psychosis class had a significantly longer inpatient stay than all of the other classes (an average of 47 days). This was approximately four times longer than the classification with the shortest stay (depression with major social problems). The results also identified a scenario where those patients in class one (depression with major social problems) had the shortest hospital duration but were re-admitted significantly quicker than any other class. However, the data did not indicate a strong relationship between the classes and the total number of inpatient stays observed.

The present study has some limitations. Fixing model parameters in order to use a pre-defined model to classify observations from another sample assume they come from the same population which is not the case here because participants, although being all psychiatric patients, came from different countries. Future studies based on multigroup latent-class analysis could further enable the analysis of latent structures of observed categorical variables across two or more groups and better highlight generality and specificity of each sample.

5. Conclusion

Taken altogether, the results suggest that the method of generating HoNOS profile based classes using LCA, as proposed by Golay et al. could be very useful. The results mainly suggest "generality" of the classes across locations and samples but that there is some degree of "specificity" (i.e. the best solution is not always the same depending on the location). The classes were particularly useful when examining their relationship with distal outcomes despite the fact that the HoNOS total and factor scores added little to no value.

Disclosure of interest

The authors declare that they have no competing interest.

References

[1] Wing JK, Beevor AS, Curtis RH, Park SBG, Hadden S, Burns A. Health of the nation outcome scales (HONOS): research and development. Br J Psychiatry 1998;172:11–8.

[2] Department of Health. Mental health payment by results guidance for 2013-14. Available at https://www.gov.uk/government/uploads/attachment/Mental-Health-PbR-Guidance-for-2013-2014.pdf (accessed 8 August 2013).

[3] Page AC, Hooke GR, Rutherford EM. Measuring mental health outcomes in a private psychiatric clinic: health of the Nation outcome scales and medical outcomes short form SF-36. Aust NZ J Psychiatry 2001;35:377–81.

[4] Kisely S, Campbell L, Cartwright J, Cox M, Campbell J. Do the health of the Nation outcome scales measure outcome. Can J Psychiatry 2010;55:431–8.

[5] Lovaglio PG, Monzani E. Health of the nation outcome scales evaluation in a community setting population. Qual Life Res 2012;21:1643.

[6] Williams B, Speak B, Hay P, Muncer S. An evaluation of the independence of the Health of the nation outcome scales. Australas Psychiatry 2014;22:473–6.

[7] McClelland R, Trimble P, Fox M, Stevenson MR, Bell B. Validation of an outcome scale for use in adult psychiatric practice. Int J Qual Health Care 2000;9:98–105.

[8] Newnham EA, Harwood KE, Page AC. The subscale structure and clinical utility of the health of the nation outcome scale. J Mental Health 2009;18:326–34.

[9] Preston NJ. The Health of the nation outcome scales: validating factorial structure and invariance across two health services. Aust NZ J Psychiatry 2000;34:512–9.

[10] Speak B, Hay P, Muncer S. HoNOS – Their utility for payment by results in mental health. Int J Health Care Qual Assur 2015;28:115–28.

[11] Speak B, Muncer S. Factorial structure of the Health of the nation outcome scales: An ordinal confirmatory factor analysis using a national sample of clinician ratings in England. Int J Ment Health Nurs 2016;25:87–98.

[12] Trauer T. The structure of the Health of the nation outcome scales (HoNOS). J Mental Health 1999;8:499–509.

[13] Muncer S, Speak B. Mokken scale analysis and confirmatory factor analysis of the Health of nation outcome scales. Pers Individ Diff 2016;94:272–6.

[14] Muncer S, Speak B. Confirmatory factor analysis of a two-scale model of the Health of the nation outcome scales (HoNOS) across diagnostic categories. Psychiatry Res 2017;12–4.

[15] Burgess P, Pirkis J, Buckingham W, Eagar K, Solomon S. Developing a casemix classification for specialist mental health services. Casemix Quarterly 1999;1:4–20.

[16] Eagar K, Gaines P, Burgess P, Green J, Bower A, Buckingham B, et al. Developing a New Zealand casemix classification for mental health services. World Psychiatry 2004;3:172–7.

[17] Andreas S, Dirmaier J, Harfst T, Kawski S, Koch U, Schulz H. Development and evaluation of a case group concept for inpatients with mental disorders in Germany: using self-report and expert-rated instruments. Eur Psychiatry 2009;24:105–11.

[18] Twomey C, Prina AM, Baldwin DS, Das-Munshi J, Kingdon D, Koeser L, et al. Utility of the Health of the nation outcome scales (HoNOS) in predicting mental health service costs for patients with common mental health problems: historical cohort study. Plos One 2016. http://dx.doi.org/10.1371/journal.pone 0167103.

[19] Golay P, Basterrechea L, Conus P, Bonsack C. Internal and predictive validity of the French health of the nation outcome scales: need for future directions. Plos One 2016;11(8):1–13.

[20] Schwarz G. Estimating the dimension of a model. Ann Stat 1978;6(2):461–4.

[21] Lanza ST, Tan X, Bray BC. Latent class analysis with distal outcomes: a flexible model-based approach. Struct Equ Modeling 2013;20(1):1–26.

[22] Asparouhov T, Muthén B. Auxiliary variables in mixture modeling: three-step approaches using M plus. Struct Equ Modeling 2014;21(3):329–41.

[23] Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. Struct Equ Modeling 2007;14(4):535–69.

[24] Baumann PS, Crespi S, Marion-Veyron R, Solida A, Thonney J, Favrod J, et al. Treatment and early intervention in psychosis program (TIPP-Lausanne): implementation of an early intervention programme for psychosis in Switzerland. Early Interv Psychiatry 2013;7(3):322–8.

[25] Alameda L, Golay P, Baumann P, Morandi S, Ferrari C, Conus P, et al. Assertive outreach for "difficult to engage" patients: a useful tool for a subgroup of patients in specialized early psychosis intervention programs. Psychiatry Res 2016;239:212–9.