SciMedCentral

Journal of Addiction Medicine and Therapy

Review Article

Mokken Scaling analysis of Revised Clinical Interview Schedule (CIS-R) psychiatric status rating scales in a nationally representative sample: the 2007 Adult Psychiatric Morbidity Survey of England

Stochl J[1,2,3]* and Tim Croudace[4]

[1]Department of Psychiatry, University of Cambridge, UK
[2]Cameo Early Intervention Services, Cambridgeshire and Peterborough NHS Foundation Trust, UK
[3]Department of Kinanthropology, Charles University, Czech Republic
[4]Mental Health and Addiction Research Group (MHARG) HYMS and Department of Health Sciences Alcuin College, University of York, UK

*Corresponding author
Stochl J, Department of Psychiatry, University of Cambridge, Addenbrookes Hospital, Box 189, Cambridge, CB2 0QQm, UK, Tel: +447587146299; Email: js883@cam.ac.uk

Submitted: 01 August 2013

Accepted: 23 August 2013

Published: 28 August 2013

Keywords
- Structured clinical interview
- Psychometrics
- Clinical interview schedule -Revised
- Nonparametric item response theory
- Mokken scale analysis

Abstract

This study aims to investigate double monotonicity of Revised Clinical Interview Schedule (CIS-R) psychiatric status rating scales data from 2007 Adult Psychiatric Morbidity Survey (APMS) within the framework of Mokken models. Results show that the items of the scale are sufficiently unidimensional in the general population for the CIS-R responses to be scalable according to broad Mokken principles. These do not require recourse to the parametric models for item response function curves typical of most applications of IRT in patient reported outcome measures research (PROMs). Our illustrative results provide an exemplar of the method. The methods are however more widely relevant for phenotype work in clinical and behavioural research, and so should appeal to those who work on addictions or in clinical medicine.

INTRODUCTION

Psychiatric research and clinical practice are influenced by diagnostic classification systems or their revisions. Psychometric evaluation of structured clinical interview data is therefore of methodological interest, as well as conceptually relevant to debates over parsimonious description of disorders, for example in terms of dimensional versus categorical perspectives, or some alternative "hybrid approaches" that are suggested by methodologists. Indeed an invigorated literature on clinical classification and diagnosis is emerging at the interface of clinical research and information systems for structured / sequenced data collections. In this regard it is useful to be able to study clinical symptom data on psychiatric status in unselected populations, such as community samples of adults living in ordinary households. In such populations there are healthy, unwell but not help-seeking, and unhealthy treated individuals not just those with clinical problems or some established degree of risk.

Das-Munshi et al. [1] reported on the symptoms of anxiety and depression in such a UK population survey, using symptoms captured by a comprehensive clinical interview schedule. Their analysis utilized a clinical structured interview that has been applied frequently in psychiatric epidemiology - the CIS-R which has been used repeatedly by the Office of National Statistics

(formerly the Office of Population Census and Surveys) and also by clinical trialists to measure baseline psychiatric symptoms (usually for anxious/depressive morbidity). The CIS–R enables the generation of ICD–10 diagnoses and as the authors indicate: "…unusually, the CIS–R includes a comprehensive assessment of anxiety and depressive symptoms regardless of whether the respondent meets criteria for a specific disorder. It is therefore particularly well suited for exploration of sub-threshold conditions such as mixed anxiety and depressive disorder" [1]. Their aim was to seek evidence for a notion of mixed anxiety and depressive disorder (MADD) based on clinical information on signs and symptoms initially used to make DSM diagnoses. MADD was described by the authors (in the manner of a hypothesis) as both "a provisional diagnosis in ICD-10 and DSM-IV", and involving "the presence of both anxiety and depressive symptoms". Our motivation is more methodological and is oriented from a "psychometric epidemiology" perspective.

Using data from the same survey (the third in the series, conducted in 2007) we sought to evaluate the strength of these psychiatric status measures in terms of a simple psychometric scaling model, for measuring the severity of common mental disorder symptoms (morbidity). We do not consider any of the screening questionnaires that were also included, since they do not offer clinical ratings, but comprise traditional self report questionnaire screens e. g. for alcohol. The APMS survey relied on the CIS-R to assess common mental disorder in the first ("phase one") interview. The APMS reports used CIS-R data as the basis for diagnosis, based on the range of non-psychotic symptoms that it assesses (over the period of the past week). This use of the CIS-R is typical, with responses used to generate both an overall score and for diagnoses (after application of relevant clinical algorithms). Our analysis relates to the overall scoring approach. We sought to operationalise a dimension of anxiety and depressive morbidity severity. This involves the idea of a single dimension for the CIS-R symptoms, as well as scaling criteria.

Following Das-Munshi we binary scored the CIS-R psychopathological data and evaluated whether the symptoms endorsed at >1 score thresholds form a single unidimensional continuum, when evaluated using non-parametric item response theory, by performing a Mokken scaling analysis. Our analysis embraces the utility of a simple, coarsened scaling approach to the latent ordering of survey members into strata of increasing likelihood of the presence of common mental disorder.

METHODS

Sample description

APMS 2007 was a cross-sectional survey of adults which provides data on the prevalence of treated mental health problems and untreated psychiatric disorder in the population of England. Our secondary analysis for this study utilized data from the Adult Psychiatric Morbidity Survey as archived for secondary analysis through the UK Data Service [2]. The dataset comprised 7,403 individuals – 3,197 (43.2%) males of mean age 50.8 (sd=18.4), and 4,206 (56.8%) females of mean age 51.4 (sd=18.8)). In APMS the CIS-R was applied to enumerate the prevalence of common mental disorder. Survey interviewers used a computer assisted method to adminster the CIS-R as part of the first phase of a multi-stage design (we do not consider the second stage clinical interview data in this report). CIS-R, by assessing psychiatric status of individuals, in terms of core psychopathology offers a structured clinical assessment sufficient to yield diagnoses of common mental disorders and these have been reported on extensively by the APMS team. CIS-R was distinct from other instrumentation used in the first phase, which comprised more conventional questionnaire-based screening instruments for mental disorders.

Measures

The clinical assessment (Revised Clinical Interview Schedule CIS-R [3]) consists of psychiatric status ratings of 14 symptoms (here-after referred to as items) addressed by subscales including symptom scores for fatigue, sleep problems, irritability, worry and depression (both depressed mood and depressive thoughts). To simplify analysis we recoded the responses to binary data to identify presence versus absence of symptoms, since we were interested in the simple dimensional summary of coarsely categorized common mental disorder (CMD) morbidities. The summary ratings for the included symptoms are listed in Table 1 where the prevalence of the symptoms is reported as the percentage who scored 0 (indicating symptoms absent or mild) versus 1 (indicating symptoms present). This was achieved using thresholds for dichotomization used by previous studies, to achieve some principled level of clinical consistency and validity.

Analysis

Mokken scaling techniques are a useful tool for researchers who wish to construct unidimensional tests or use questionnaires forms comprising multiple binary or polytomous items [4]. The stochastic cumulative scaling model offered by this approach is ideally suited when the intention is to score an underlying latent attribute by simply summing item responses, here sums of "symptoms present". It can assist in the determination of the dimensionality of tests or scales without recourse to factor analysis, but can also complement such parametric analysis, in some cases. It also enables consideration of reliability, without reliance on the internal consistency reliability estimates provided by Cronbach's alpha coefficient.

Mokken models belong to a broad class of statistical models called non-parametric item response theory (abbreviated often to NIRT models). The major advantage of NIRT over more commonly used item response models, such as the Rasch latent trait model, is that they relax many of the quite strong assumptions about the non-linear behaviour of response probabilities that are invoked by the members of the family of parametric IRT models [5]. This broader, more inclusive approach enables items with increasing but less regularly formed curve shapes to be entertained for the item response function (IRF) and still included in a model on these properties alone.

In this study we investigate if CIS-R psychopathology data conform to the rather strong assumptions of the *double monotonicity model* (DMM) introduced to the psychometric literature by Mokken [6]. This requires unidimensionality checking by computation of item and scale Loevinger´s

Table 1: Descriptive statistics for CIS-R binary data: univariate (diagonal) and bivariate symptom endorsement frequencies (entries below diagonal) and proportions (entries above diagonal).

	Somatic	Fatigue	Concentration/ forgetfulness	sleep disturbance	Irritability	Worry about physical health	Depressed mood	Depressive thoughts	Worry	Anxiety	Phobic anxiety	Panic attacks	Compulsive behaviours	Obsessional thoughts
Somatic	470 (6.3%)	5%	2.6%	4.2%	2.9%	2%	2.7%	2.5%	3.6%	2.1%	1.3%	1.1%	0.8%	1.1%
Fatigue	369	2186 (29.5%)	8.3%	17.2%	10.4%	5.7%	8.8%	7.7%	12.1%	6.3%	3.7%	2.3%	2.4%	3.4%
Concentration/ forgetfulness	195	616	762 (10.3%)	6.8%	5.3%	2.8%	4.8%	4.6%	6%	3.4%	2.2%	1.7%	1.4%	1.9%
sleep disturbance	311	1272	506	2375 (32.1%)	9.6%	5.1%	7.7%	6.9%	11.6%	5.8%	3.6%	2.1%	2.4%	3.7%
Irritability	211	773	391	707	1216 (16.4%)	3.1%	5.9%	5.7%	8.6%	4.3%	2.8%	1.8%	1.9%	2.6%
Worry about physical health	151	420	209	378	231	560 (7.6%)	3.4%	3%	4.3%	2.4%	1.6%	1.2%	1%	1.6%
Depressed mood	202	655	358	568	437	249	903 (12.2%)	6.6%	7.2%	4.4%	2.4%	1.9%	1.5%	2.3%
Depressive thoughts	185	567	337	513	419	225	491	723 (9.8%)	7.1%	4.3%	2.5%	1.9%	1.6%	2.3%
Worry	270	896	442	856	639	320	535	525	1403 (19%)	6.7%	3.2%	2.2%	2.3%	3.1%
Anxiety	159	467	253	433	319	179	326	322	494	634 (8.6%)	1.9%	1.7%	1.4%	1.9%
Phobic anxiety	96	276	161	265	204	120	177	186	238	143	401 (5.4%)	1.3%	1.1%	1.2%
Panic attacks	78	173	127	154	133	92	138	144	163	126	95	209 (2.8%)	0.8%	0.9%
Compulsive behaviours	59	179	100	175	140	74	109	120	169	102	81	62	292 (3.9%)	0.9%
Obsessional thoughts	80	252	142	274	190	117	168	169	231	137	89	63	63	406 (5.5%)

Lower diagonal = counts; upper diagonal = percentages

Table 2: Mokken scaling analysis of CIS-R psychiatric symptoms: summary of results.

Symptom	Item H	Standard error of item H	Number of significant monotonicity violations	Number of significant IIO violations
Somatic	0.34	0.01	0	0
Fatigue	0.54	0.01	0	0
Concentration/forgetfulness	0.42	0.01	0	0
Sleep disturbance	0.46	0.01	0	0
Irritability	0.41	0.01	0	0
Worry about physical health	0.35	0.01	0	0
Depressed mood	0.45	0.01	0	0
Depressive thoughts	0.50	0.01	0	0
Worry	0.50	0.01	0	0
Anxiety	0.42	0.01	0	0
Phobic anxiety	0.35	0.02	0	0
Panic attacks	0.51	0.02	0	0
Compulsive behaviours	0.29	0.02	0	0
Obsessional thoughts	0.31	0.01	0	0

scalability coefficients [7], checking that IRFs are monotone non-decreasing functions of the hypothesized latent common mental disorder severity trait, and ensuring that IRFs for the relevant psychopathology symptom items do not intersect. If these assumptions are met then a) the APMS survey respondents can be ordered with respect to their latent severity of mixed anxiety and depressive morbidity on the basis of binary-recoded sum score, and b) the scores derived from use of the CIS-R then have achieve the desirable property of invariant item ordering. Invariant item ordering (IIO) allows the researcher to order items according to their commonality/rarity (or as more commonly expressed, by their prevalence). This property helps researchers to communicate useful features of the hierachical ordering of clinical diagnostic interview psychopathology items that might inform the content or order of future clinical testing or enquiry about more severe psychopathology in related clinical areas. Scales with IIO also have several other features described fully elsewhere [4]. The software that was used for the psychometric analyses in this study is freeware, R [8] version 3 and commands from user-written package "mokken" [9] thereby enabling easy replication by other researchers who wish to explore the method on their own data or instruments.

RESULTS

After listwise deletion of missing data 7,403 survey respondents CIS-R data were analysed. The prevalence of the binary recoded symptoms, and their pair-wise endorsement proportions are shown in Table 1 using descriptive statistics for binary data (including co-occurrence of pairs of symptoms). The most common experienced symptoms in the APMS include sleep disturbance (32.1%), fatique (29.5%), and worry (19%), the rarest are panic attacks (2.8%) and compulsive behaviours (3.9%). The most common pairs of symptoms (co-occurring) were a combination of fatique and sleep disturbance (17.2%), fatique and worry (12.1%), and sleep disturbance and worry (11.6%). These are substantially lower than the highest prevalence symptoms individually (Table 1).

Table 2 shows summary results from the Mokken scaling analysis of binary CIS-R data. For all psychopathology items except "compulsive behaviours", the scalability coefficients (denoted in the column Item H) exceeded 0.3, an indication that such symptoms capture a single dimension of common mental disorder severity, when conceptualized as endorsed symptoms (present or absent at the threshold value operationalised by the score recoding >1 rating). Compulsive behaviours fit least well among the other CIS-R items for common mental disorder. However the estimate of 0.29 is only slightly below the recommended range of item H values.

The scalability estimate for these CIS-R symptoms (all symptoms) was 0.43 a sufficiently large value that suggests the scale is of medium "strength" according to criteria suggested by Sijtsma & Molenaar [5]. Furthermore, there were no violations of either monotonicity nor IIO for any of the CIS-R items suggesting that all symptoms meet the monotonicity assumption and IIO. Finally, a value for the score reliability was estimated at 0.84 (moderately satisfactory) (Table 2).

DISCUSSION

This study aimed to investigate some basic psychometric properties of CIS-R within the framework of nonparametric item response theory. To accomplish this, we checked whether the CIS-R meets the assumptions of double monotonicity Mokken model. The results show that the scale can be considered as unidimensional measure of common mental disorder, because (with exception of item "compulsive behavior"), all items have the Loevinger scalability coefficients above the widely recommended value of 0.3. Compulsive behavior seems to be less correlated with other items suggesting that this symptom may be less comorbid with respect to other symptoms covered by CIS-R. However, Loevinger scalability coefficient for this item (0.29) is only slightly below the recommended criteria and therefore we prefer to keep this item in the scale. There are no violations of monotonicity for any CIS-R items which indicates that as the CIS-R sumscore increases, the mean score of each item increases as well.

Finally, there are no intersections of characteristic curves of items. Together with unidimensionality, monotonicity, and local independence (discussed below), this ensures that CIS-R meets assumptions of double monotonicity model and that the scale has IIO property. Therefore, the respondents can be safely ordered with respect to severity of common mental disorder by using the sumscore of the CIS-R items (i. e. number of symptoms).

Several aspects of our scaling results deserve comment, since to our knowledge this is the first application of non-parametric psychometric item response theory models to CIS-R symptoms. For example, Mokken models formally require an additional assumption of local independence of item responses. The local independence assumption states that an individual's response to an item is not influenced by his or her responses to the other items in the same test/scale [5]. Within the framework of Mokken models, this assumption is difficult to check empirically, although some statistical methods have been developed for this purpose [10-13]. Therefore we do not address this issue here apart from the statement that the conceptualization of the CIS-R scale does not arise worries about local dependence of items.

Some additional limitations of our analyses must also be appreciated. First, we used dichotomous scoring of CIS-R and did not exploit the full ordinal range of the items. Our results may not extend to ordinal scoring of CIS-R. Further, we did not incorporate the weighted survey design as such methodology has not been developed for Mokken models. Relationships between the Mokken analysis model and some restricted latent class models allow for some further scrutiny of the correspondence between our results and those from latent structure analysis of CIS-R. We plan to return to this correspondence in a future piece of work across APMS samples.

ACKNOWLEDGEMENT

JS was partly supported by the Czech Ministry of Education (research project MSM 0021620864).

REFERENCES

1. Das-Munshi J, Goldberg D, Bebbington PE, Bhugra DK, Brugha TS, Dewey ME, et al. Public health significance of mixed anxiety and

depression: beyond current classification. Br J Psychiatry. 2008; 192: 171-177.

2. National Centre for Social Research and University of Leicester, Adult Psychiatric Morbidity Survey, 2007, UK Data Archive: Colchester, Essex.

3. Brugha TS, Bebbington PE, Jenkins R, Meltzer H, Taub NA, Janas M, et al. Cross validation of a general population survey diagnostic interview: a comparison of CIS-R with SCAN ICD-10 diagnostic categories. Psychol Med. 1999; 29: 1029-1042.

4. Stochl J, Jones PB, Croudace TJ. Mokken scale analysis of mental health and well-being questionnaire item responses: a non-parametric IRT method in empirical research for applied health researchers. BMC Med Res Methodol. 2012; 12: 74.

5. Sijtsm K, Molenaar IW. Introduction to nonparametric item response theory. Measurements Methods for the Social Sciences, ed. RM Jaeger. Vol. 5. 2002, London: Sage Publications. 170.

6. Mokken RJ. A theory and procedure of scale analysis1971. The Hague: Mouton. 354.

7. Loevinger J. A systematic approach to the construction and evaluation of tests of ability. Psychol Monogr. 1947; 61.

8. Dean CB, Nielsen JD. Generalized linear mixed models: a review and some extensions. Lifetime Data Anal. 2007; 13: 497-512.

9. van der Ark, Mokken RJ (Version 2.7.5): An R package for Mokken scale analysis [computer software]. 2013.

10. Douglas J, Rim Kim Hae, Brian Habing, Furong Gao. Investigating local dependence with conditional covariance functions. J Educ Behav Stat. 1998; 23: 129-151.

11. Ip EH. Adjusting for information inflation due to local dependency in moderately large item clusters. Psychometrika. 2000; 65: 73-91.

12. Ip EH. Testing for local dependency in dichotomous and polytomous item response models. Psychometrika. 2001; 66: 109-132.

13. Hoskens M, De Boeck P. A parametric model for local item dependencies among test items. Psychol Methods. 1997; 2: 261-277.

Cite this article

Stochl J, Croudace T (2014) Mokken Scaling analysis of Revised Clinical Interview Schedule (CIS-R) psychiatric status rating scales in a nationally representative sample: the 2007 Adult Psychiatric Morbidity Survey of England. J Addict Med Ther 2(1): 1005.