Research paper

Symptom severity and disability in psychiatric disorders: The U.S. Collaborative Psychiatric Epidemiology Survey

CrossMark

R. García-Velázquez[a,*], M. Jokela[a], T.H. Rosenström[a,b]

[a] Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Finland
[b] Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway

A B S T R A C T

Background: While most psychiatric diagnoses are based on simple counts of symptoms, some symptoms may be sign of a more severe mental syndrome than others. This calls for validated estimates of the relative severity specific symptoms imply within a disorder. We focused on four diagnostic disorders: Manic Episode (ME), Major Depressive Episode (MDE), Post-traumatic Stress Disorder (PTSD) and Generalized Anxiety Disorder (GAD). Symptom-specific severity parameters were estimated, and validated by examining their association with levels of self-reported disability in daily activities over and above the number of symptoms.
Methods: Data from the cohort study of the U.S. Collaborative Psychiatric Epidemiology Surveys (CPES) was used, which comprises the National Comorbidity Survey Replication, National Survey of American Life, and the National Latino and Asian American Study. The four analytic datasets included respondents who endorsed disorder-specific pre-screening symptoms according to the World Mental Health Survey Initiative's version of the Composite International Diagnostic Interview. Disability was measured using the WHO Disability Assessment Schedule. Item Response Theory and Tobit models were implemented.
Results: For ME, PTSD, and GAD (not MDE) symptom severity based on psychometric Item Response Theory predicted disability outcomes after adjusting for symptom count. For PTSD, symptom count was not associated with disability.
Limitations: The analytic sample for each psychiatric disorder was based on a pre-selection stemming from index criteria (e.g. sadness or pleasure loss for MDE), which implies that our results are only generalizable to those individuals at risk rather than for the entire population. Additionally, we acknowledge that the use of uni-dimensional models is only one of the several options to model psychopathological constructs.
Conclusions: The same number of symptoms may be related to different levels of disability, depending on the specific symptoms from which the person suffers. Diagnostic procedures and treatment decisions may benefit from such additional information without extra costs.

1. Introduction

Diagnostics of psychiatric disorders are based on the number of symptoms as defined in diagnostic manuals. In most instances, the diagnosis does not discriminate between the specific symptoms that the person suffers from. That is, two people with the same diagnosis may have only partially overlapping sets of symptoms. However, the different symptom combinations may contain information on the clinical severity of the person's disorder, over and above the simple number of symptoms (Fried and Nesse, 2015a). Such knowledge could be used for more informative diagnosis of psychiatric disorders.

According to classic psychometrics, *construct validity* pertains to the effective measurement of the intended theoretical construct (here, a

specific mental disorder) and to appropriate inferences based on it (Cronbach and Meehl, 1955; Loevinger, 1957). The *structural* component of construct validity implies that the symptoms indexing the construct occur and coexist consistently as its representatives. The *external* component of construct validity implies that the measure is associated with relevant external criteria not directly used to measure the construct, that is, with independently measured outcomes.

Mental disorders have been conceptualized as *harmful dysfunctions* (Wakefield, 1992), and for most psychiatric diagnoses of the DSM, the disorder must be related to *clinically significant disturbance, distress, or disability* (American Psychiatric Association, 2013). Thus, the purpose of a valid psychiatric diagnosis is to identify functionally impaired individuals as opposed to non-impaired individuals. Such a purpose also

establishes impairment as a fundamental criterion in external validity studies. If specific symptoms carry information about the clinical impairment associated with the disorder over and above the simple count of symptoms, it would be possible to use this information to sharpen diagnosis and more effectively orient therapeutic targets. Symptom-level analyses of psychiatric disorders suggest that specific symptoms may be differently related to psychosocial functioning (Fried and Nesse, 2014, 2015a; Fried et al., 2015; Tweed, 1993; Wakefield and Schmitz, 2017a, 2017b). However, the performance of specific diagnostic criteria with respect to clinical measures of disability has not been examined in detail across the most disabling psychiatric disorders, such as Major Depression or Post-traumatic Stress Disorder.

Continuous dimensions, rather than categories, seem to best explain the structure of psychiatric symptoms, suggesting that disorders reflect underlying continuous variables (Edens et al., 2006; Haslam et al., 2012; Marcus et al., 2006, 2008). Models of Item Response Theory (IRT; de Ayala, 2013; Reckase, 2009) link each symptom to a threshold on an underlying disorder continuum based on the distribution of symptom occurrence. In modeling psychiatric disorders, the *psychometric severity* parameter (known as *difficulty* parameter in ability testing, or *b* parameter in general) determines the level of the underlying disorder continuum that is required for an individual to endorse a symptom with a 50% probability (Reise and Waller, 2009; Thomas, 2011). After determining the psychometric severity of specific symptoms, it is possible to examine whether these severity estimates are related to external indicators of functional disability, including difficulties in self-care, life activities or cognition (Üstün, 2010). Fig. 1 displays the rationale of our approach.

Using data from three large studies of psychiatric epidemiology (Alegria et al., 2015; Pennell et al., 2004), we estimated psychometric symptom severity using IRT models and examined their associations with the external validity criterion of self-reported disability, controlling for symptom count. We focused on four disorders – Manic Episode (ME), Major Depressive Episode (MDE), Post-traumatic Stress Disorder (PTSD) and Generalized Anxiety Disorder (GAD)—which constitute the four most impairing disorders among those examined by Druss et al. (2009) as rated by the Sheehan Disability Scale.

2. Methods

2.1. Data

The Collaborative Psychiatric Epidemiological Surveys (CPES) comprised three multi-stage area probability samples conducted between 2001 and 2003: the National Comorbidity Survey Replication (NCS-R, N = 9282), the National Survey of American Life (NSAL, N = 6082), and the National Latino and Asian American Study (NLAAS, N = 4649).
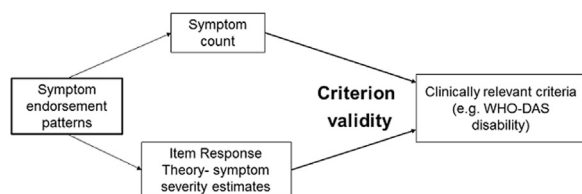
Fig. 1. Validation of Item Response Theory-derived symptom severity as referred to a criterion. Footnote: Once a set of symptoms has been defined (e.g. according to DSM-5), symptom endorsement data can be used to (a) derive symptom counts, which generally determine whether a diagnosis is fulfilled (e.g. Generalized Anxiety Disorder requires compulsory criterion plus three associated symptoms), and (b) obtain information on symptom characteristics (i.e. Item Response Theory-based severity). Symptom count and severity estimates can be studied in association with a clinically relevant variable such as disability (criterion validity). Using data at the symptom-level maximizes the use of information and brings insights on single symptom functioning. The utility of Item Response Theory and the validity of its assumptions is probed by the ability of latent severity estimates to predict the criterion variable over and above symptom count.

The data was available to us via the Inter-university Consortium for Political and Social Research service (Alegria et al., 2015). The joint sample is representative of adults (i.e., age 18 years or more) residing in households in coterminous United States, Alaska and Hawaii, excluding institutionalized persons and those living on military bases (NCS-R and NSAL also excluded non-English speakers). Comparable methodologies were used, including using trained lay interviewers to conduct interviews primarily in person. The average response rate of the CPES is 72.7%. Details of each survey can be found elsewhere (Heeringa et al., 2004; Pennell et al., 2004). The final CPES sample includes 20,013 individuals (8550 men and 11,463 women).

From the overall dataset, a different subsample was extracted for each disorder (GAD, ME, MDE, PTSD) containing individuals who answered to the specific diagnostic interview after having endorsed pre-screening symptoms. Our samples were a total of N = 4214 for ME, N = 4152 for MDE, N = 3128 for PTSD, and N = 3610 for GAD.

2.2. Measures

2.2.1. Psychiatric symptoms

The respondents were interviewed according to the World Mental Health Survey Initiative's version of the Composite International Diagnostic Interview (WMH-CIDI), which is a modified version of the original WHO-CIDI (Kessler and Üstün, 2004). Both WMH-CIDI and the other CPES questions were administered using a computer-assisted interview. The presence of a symptom was determined exactly as in the DSM-IV part of the ICPSR documentation for the diagnostic algorithms (Alegria et al., 2015).

2.2.2. Functional impairment

The World Health Organization Disability Assessment Schedule (WHODAS) was used to measure disability as defined in the International Classification of Function, Disease and Health. Scores are a product of frequency and severity of problems (none, mild, moderate, severe) that respondents reported experiencing in the past 30 days, and are normalized to values ranging from 0 to 100, where higher numbers indicate worse functioning. We used a single overall score based on the average disability over the domains of functioning: cognition, mobility, self-care, social interaction, role functioning, participation (Üstün, 2010).

As another indicator of validity based on clinical relevance, we used the item "Beginning yesterday and going back 30 days, how many days out of the past 30 were you totally unable to work or carry out your normal activities because of problems with either your physical health, your mental health, or your use of alcohol or drugs?" as a variable called "disability days" from now on. Due to highly similar findings, results using disability days are shown only in the Supplement. Both outcomes were assessed at a general level, and thus were not disorder-specific.

2.3. Statistical analysis

2.3.1. Item Response Theory for estimating psychometric severity

For each of the four disorders, a series of unidimensional IRT models were estimated. This procedure was chosen because the disorders are typically treated as single diagnostic syndromes, implying unidimensionality. Each set of diagnostic criteria was analyzed for *essential unidimensionality* to prevent interpretation of overly biased models (Reise and Rodriguez, 2016). Since diagnostic definitions do not make a distinction in terms of background variables, we chose to keep the paradigm as comparable as possible and hence estimated a general severity parameter per criterion (i.e. no gender- or age-stratified analyses were conducted).

The diagnostic criteria composed of several single symptoms were analyzed by splitting them (for instance MDE A.4 *change in sleep,* which contains *insomnia or hypersomnia*). Those criteria that were necessary