Title

Self-reporting of psychiatric illness in an online patient registry is a good indicator of the existence of psychiatric illness

Permalink

https://escholarship.org/uc/item/886872zf

Authors

Sordo Vieira, Luis
Nguyen, Binh
Nutley, Sara K
et al.

Publication Date

2022-07-01

DOI

10.1016/j.jpsychires.2022.03.022

Peer reviewed

Self-reporting of psychiatric illness in an online patient registry is a good indicator of the existence of psychiatric illness

Luis Sordo Vieira[a,b], **Binh Nguyen**[b], **Sara K. Nutley**[c], **Lyvia Bertolace**[b], **Ashley Ordway**[b], **Heather Simpson**[b], **Jessica Zakrzewski**[b], **Marie E. Jean Gilles**[b], **Rachel Nosheny**[f], **Michael Weiner**[d,e,f], **R. Scott Mackin**[d,f], **Carol A. Mathews**[b,*]

[a]Department of Medicine, University of Florida, Gainesville, FL, United States

[b]Department of Psychiatry, University of Florida, Gainesville, FL, United States

[c]Department of Epidemiology, University of Florida, Gainesville, FL, United States

[d]San Francisco VA Medical Center, San Francisco, CA, United States

[e]Department of Radiology, University of California, San Francisco, San Francisco, CA, United States

[f]Department of Psychiatry and Behavioral Sciences, University of California, San Francisco, San Francisco, CA, United States

Abstract

Online registries offer many advantages for research, including the ability to efficiently assess large numbers of individuals and identify potential participants for clinical trials and genetic studies. Of particular interest is the validity and utility of self-endorsement of psychiatric disorders in online registries, which, while increasingly more common, remain understudied. We thus assessed the comparability of prevalence estimated from self-endorsement of psychiatric disorders in one such registry, the Brain Health Registry (BHR) to prevalence computed from large US-based epidemiological studies and the degree to which BHR participants report psychiatric disorders consistently. We also examined the concordance between self-report and clinically determined diagnoses of various DSM-5 psychiatric disorders in a subset of participants who underwent direct assessments and identified possible reasons for discordance.

Rates of self-reported psychiatric disorders were closest to previously reported population prevalence rates when endorsed at multiple timepoints, and accuracy was at least 70% for all except Hoarding Disorder as compared to the clinical diagnoses. Clinical data suggested that self-endorsement of a given psychiatric diagnosis was indicative of the presence of a closely related condition, although not necessarily for the specific disorder, with the exception of major

---

depressive disorder, panic disorder, and hoarding disorder, which had high positive predictive values (85%, 73%, 100%, respectively).

We conclude that self-reporting of psychiatric conditions in an online setting is a fair indicator of psychopathology, but should be accompanied by more in-depth interviews if using data from a participant for a specific disorder.

1. Introduction

Online patient or research registries, which for the purposes of this work, are defined as a web-based "collection of information about individuals, usually focused around a specific diagnosis or condition" (U.S. Department of Health and Human Services, 2022) offer the potential to accelerate clinical research, including facilitating recruitment for in-person studies, clinical trials, genetic and other large-scale studies, and also provide the ability to gather longitudinal data at a large scale. The nature of the data in such registries is often self-reported. Because of the increased scale, self-reports of psychiatric disorders collected via the internet are increasingly being used in genetic studies in psychiatry. For example, one study used online self-reported psychiatric diagnoses combined with consumer genetics testing to identify genetic risk loci for self-reported major depression (Hyde et al., 2016), and studies of additional psychiatric disorders using this method are ongoing. However, despite the advantages of such registries, the validity and reliability of online self-report of psychiatric disorders have not been adequately evaluated.

Psychiatric disorders, which are not based on laboratory findings but rather on a set of symptom criteria may be particularly susceptible to problems of validity in the context of self-reporting. Patients might mistakenly self-report certain disorders because they do not understand the diagnostic criteria, or because of overlap of symptoms between disorders. Moreover, problems with recall might significantly impede accurate self-report (Andrews et al., 1999; Wells and Horwood, 2004).

Concordance of self-report and claims data/medical records has been previously analyzed for various non-psychiatric medical conditions, and results vary depending on condition (St. Clair et al., 2017; van den Akker et al., 2015; Wolinsky et al., 2015, 2014). Less work has been done regarding psychiatric illness, and the data regarding the utility of self-reported diagnoses are mixed. For example, one study validated 74% of self-reported previous diagnoses of depression using the Structured Clinical Interview for DSM-IV (SCID-I) (Sanchez-Villegas et al., 2008), while another found low positive predictive value and sensitivity of self-reported depression compared to cutoff-scoring using the Beck Depression Inventory-II (Zarghami et al., 2020). A study of mental health service utilization in the general population found discrepancies between self-reported and actual usage with a tendency towards under-reporting of mental health service utilization (Drapeau et al., 2011). However, the work that has been published to date does indicate that self-report questions, even if imperfect, may be useful (Sanchez-Villegas et al., 2008; Stuart et al., 2014). For example, in a study of an Australian population, 61% of participants who were found to have depression using the SCID also self-identified as having depression using the single question "Have you ever suffered from depression?" (Stuart et al., 2014).

Understanding the validity of self-reported data is critical in determining the utility of such data. For example, low validity can decrease the statistical power for genome-wide association studies (Tung et al., 2011). However, assessing the validity of self-reported psychiatric disorders is challenging for various reasons, including the lack of gold-standard biomarkers or objective measures of psychiatric illness. Thus, in the absence of large cohorts with positive self-reporting for all psychiatric disorders together with clinically validated diagnoses and/or other symptom-based scales, we reasoned that a reasonable approach that could highlight flagrant problems with either self-selection bias or the validity of the questions themselves is to compare prevalence estimates computed from the registries to known population-based prevalences.

The goals of this study were thus to: 1) determine the degree to which participants in an online registry report lifetime history of psychiatric disorders consistently over repeated questionnaire administrations, 2) compare the rates of psychiatric disorders from a national online registry (the Brain Health Registry; BHR) to reported population prevalences, thus getting a population-level sense of the utility of the BHR data and 3) in a subset of participants with both online, self-report, and in-clinic diagnoses, compare self-report to diagnoses ascertained through structured clinical interviews.

We hypothesized that most psychiatric illnesses would be consistently self-reported, and that consistent self-reporting of a given disorder might represent a more reliable phenotype than endorsement at any single timepoint. We also hypothesized that prevalence estimates from self-report would be more consistent with reported population prevalence for common and well-established psychiatric diagnoses (e.g., MDD), as compared to diagnoses that are less well-understood and/or less well-established (e.g., hoarding disorder (HD)). Lastly, we hypothesized that self-report for all psychiatric disorders compared to clinician diagnoses would suffer from false negatives (e.g. participants not being aware of having a particular diagnosis), but that the validity of a positive self-reporting would differ depending on psychiatric disorder.

2. Materials and methods

The present work was approved by the Universities of Florida and California San Francisco Institutional Review Boards. Informed consent was obtained from all participants.

2.1. Participants

Participants for this study included individuals ages 18+ who participated in the Brain Health Registry (BHR) (Weiner et al., 2018) as of December 2020, who reported their race, gender, and education, and had complete psychiatric history data in the medical history questionnaire. The BHR includes participants from all 50 US states, with approximately 500 new participants joining monthly (https://www.brainhealthregistry.org) (Nutley et al., 2020), and was created to advance the study of brain aging. It collects online self-reported data on medical and psychiatric history, family history, symptom-based instruments assessing quality of life, sleep quality, depression symptomatology, and hoarding symptomatology, and measures of subjective and objective neurocognitive functioning (Weiner et al., 2018). In this study, we focused on psychiatric history, which includes self-reported diagnoses

of current/past major depressive disorder, obsessive-compulsive disorder, specific/social phobia, hoarding disorder, generalized anxiety disorder, post-traumatic stress disorder, panic disorder, and bipolar disorder, all answered in a "yes" or "no" fashion.

2.2. Comparison with formal clinical assessments

A subset (N = 235) of BHR participants also participated in semi-structured clinical interviews as part of a larger study on hoarding symptoms, recruited as controls (113/235) or as possible HD (122/235) based on their total scores on the Hoarding Rating Scale, Self-Report (HRS-SR) (see (Nutley et al., 2020) for details). Participants underwent clinical interview conducted by telephone or teleconference by trained clinical personnel using the Mini International Neuropsychiatric Interview (MINI) (Sheehan et al., 1998), the Yale-Brown Obsessive Compulsive Scale (Y-BOCS) (Goodman et al., 1989), the Structured interview for Hoarding Disorder (SIHD) (Nordsletten et al., 2013), the Savings Inventory-Revised (SI-R) (Frost et al., 2004), and the UCLA Hoarding Severity Scale (UHSS) (Saxena et al., 2015).

Assignment of diagnoses was performed using a best estimate/consensus procedure as previously described (Nutley et al., 2020). Participants were assigned diagnoses of either "definite", "probable" or "not present", with probable indicating that the participant did not meet full diagnostic criteria according to the available data, but that sufficient data were available to suggest that the diagnosis was likely to be present (for example, all symptoms met, but the evaluators suspected under-reporting of impairment or distress). For this study, we grouped "present" and "probable" into a single positive diagnosis. When there were sufficient data to assign a probable or definite diagnosis, it was determined that participants met criteria for that diagnosis. We grouped dysthymia and MDD into a single category termed depression. Similarly, in addition to examining these disorders separately, anxiety disorders were grouped into a single category that included generalized anxiety disorder (GAD), panic disorder (PD), social phobia (SP), and/or agoraphobia and termed anxiety disorder. Dysthymia and agoraphobia are not assessed in the BHR, thus, in this sample, the BHR anxiety disorder category includes GAD, PD and/or specific or social phobia, and depression represents MDD only. In the BHR, the term "any psychiatric illness" represents the self-reporting of GAD, PD, social or specific phobia, PTSD, MDD, bipolar disorder, OCD, or HD. In the clinically diagnosed group, the term "any psychiatric illness" represents the presence of GAD, PD, social phobia, agoraphobia, PTSD, MDD, dysthymia, bipolar disorder (BD), OCD or HD.

2.3. Comparison with prevalence estimates in previous epidemiological studies

We estimated the lifetime prevalence (cases/population size) of the psychiatric conditions assessed in the BHR based on self-report for participants who had completed the medical history module at least twice using four adjudication methods: Endorsement of a psychiatric disorder (either current or past history) at any timepoint (Any), at every timepoint that the participant completed (All), at more than half the completed timepoints (Most), and temporally consistent endorsement (e.g. if a participant endorsed lifetime history at some point, they endorsed it at all subsequent timepoints afterwards) (Temporal consistency).

As comparators, we used the two largest epidemiological cross-sectional studies of mental health in the US adult population: the National Comorbidity Survey Replication (NCS-R) and the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). As HD did not become a formal diagnosis until 2013, prevalence estimates for this disorder are not reported in the NCS-R or NESARC. Thus, we used previous studies of prevalence of HD in adult populations (Ayers, 2017; Postlethwaite et al., 2019) and compared these with prevalences computed using 1) self-reporting of HD, and 2) the Hoarding Rating Scale-Self Report (HRS-SR), a well validated scale used to detect pathological hoarding (Tolin et al., 2010).

- The National Comorbidity Survey Replication (NCS-R) surveyed prevalence of psychiatric disorders in 9,282 English-speaking adults in the contiguous United States, using an adapted version of the World Mental Health Composite International Diagnostic Interview (WMH-CIDI, (Alonso et al., 2002)), collected between 2001 and 2003 using DSM-IV hierarchical criteria (Kessler et al., 2004; Kessler and Merikangas, 2004, p. 0) and updated in July 2007. Some conditions, such as OCD, were only assessed in a subset of the NCS-R sample. For this study, we used prevalence estimates and two times the standard error reported on the NCS website as a measure of uncertainty of the prevalence estimate. We used the published age categories: young adult (18–29), adult (30–44), middle age (45–59), and older adult (60+).

- The National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) surveyed approximately 43,000 American adults (wave 1, 2001–2002) for prevalence of alcohol and drug use, as well as several common psychiatric disorders using DSM-IV criteria via the Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV) (Grant et al., 2003; Hasin and Grant, 2015). The second wave re-interviewed approximately 34,000 participants 3 years after the first wave. For this study, we used lifetime prevalence estimates as well as two times standard errors as a measure of uncertainty of the prevalence estimate for panic disorder, generalized anxiety disorder (GAD), post-traumatic stress disorder (PTSD, assessed in wave 2), MDD, bipolar I and bipolar II (reported as a combined rate), and age categories (young adult (18–29), adult (30–44), middle age (45–64), and older adult (65+)) as reported in a review of multiple NESARC-based studies (Hasin and Grant, 2015) and cited studies (Compton et al., 2007; Grant et al., 2005a, 2005b; Pietrzak et al., 2011).

2.4. Assessments of hoarding disorder in the BHR

To compare self-reported hoarding to our previously validated cutoff criteria using HRS-SR scores (Nutley et al., 2020), we assessed a subset of participants in the BHR who completed the medical history module and the HRS-SR at least twice. Likely HD was defined as a mean HRS-SR score of 14 or higher, which is widely accepted as a clinically significant threshold (Tolin et al., 2010). We also compared self-reported rates of HD to rates determined using an HRS-SR cutoff of 10, which we have previously identified as a reliable cutoff score to identify problematic hoarding (Nutley et al., 2020).

2.5. Statistical analyses

We used the R statistical software, version 3.6.3 (R Core Team, 2018), using the tidyverse 1.3.0 package and the epiR package v 2.0.19. Prevalence estimates and 95% confidence intervals (Clopper-Pearson method) were calculated for each self-report psychiatric diagnosis in the BHR, first for the entire sample, by age group, and by age and gender. We used the R Caret Package v.6.0–86 confusion matrix function to compare self-report to best estimate diagnoses. For each diagnosis we calculated the sensitivity, specificity, percent correctly classified (accuracy), positive predictive value (PPV), and negative predictive value (NPV) using the clinical best-estimate diagnosis as the gold-standard, and using the consensus self-report based on the Any and Most methods. Cohen's kappas were computed for measuring agreement between BHR participant and clinical interview using the Most and Any methods.

3. Results

3.1. Participant demographics

44,156 individuals (33,305 females, 10,851 males) reported their race, gender, and education, and completed the psychiatric disorders history in the medical history module of the BHR at least once (mean age (SD): females 57.3 years (13.1), males 61 years (13.7), (ages 90+ at baseline are recorded as age 90) and 78.9% reported completing a 2-year college degree or higher. Of these, 25,037 completed the module at least twice (88.9% non-Latino white, 18,800 females (mean age (SD) = 58.5 (12.4)), 6,237 males (mean age (SD) = 62.6 years (12.3))). There was no significant difference in completion rates between males and females. Participants who identified solely as white were more likely to complete the medical module at least twice compared to other/multiple races (58.7% vs 44.8%; $\chi^2(1, N = 44156) = 417.83$, $p < 0.001$). See Table 1 for full demographics.

3.2. Consistency of self-report

We first assessed the consistency of self-reported psychiatric diagnoses in the BHR by examining the percentage of people who endorsed a given disorder at baseline who also self-reported the same condition most or all of the time (Table 2). Relative consistency (Most) of self-report was high (over 80%) and complete consistency (All) was moderate (over 60%) for most disorders, with the clear exception of HD, where consistency was 50%.

We then assessed the proportion of people who endorsed a condition in a temporally consistent manner (if they endorsed at one timepoint, they endorsed at all subsequent timepoints) compared to those who endorsed at any timepoint. The highest proportion of temporally consistent self-reported diagnoses was bipolar disorder (78%), followed by PTSD (73%), depression (70%), GAD (67%), OCD (63%), panic disorder (57%) and hoarding disorder (45%).

3.3. Comparison of self-reported prevalences to population estimates

We next compared the estimated lifetime prevalences using the four different methods (Any, Temporal Consistency, All, Most) to prevalences reported in the NCS-R and NESARC.

Self-report by a participant at >50% (Most) and all timepoints (All) as well as temporally-consistent self-report yielded rates that were the most comparable to both the NCS-R and NESARC, with the exception of GAD, which had a higher prevalence rate in the BHR than in the population-based studies regardless of method (Fig. 1).

As the BHR sample is skewed towards older adulthood, and thus may not be directly comparable to prevalence estimates derived from the entire age distribution, we repeated the analyses, stratifying by age group (Fig. 2) and by age and gender (Supplementary Table S3). Lifetime prevalence estimates for the middle age group (ages 45–59) in the BHR were very similar to those reported in the NCS-R and, to a somewhat lesser extent, in the NESARC. Again, the Any method produced higher estimates whereas the other three methods yielded rates resembling population rates, with the exceptions of GAD.

3.4.   Lifetime prevalence of hoarding disorder in the BHR

15,880 of the 25,037 subjects (63.4%) who completed the medical history module at least twice also completed the Hoarding Rating Scale-Self Report (HRS-SR) at least twice. Lifetime prevalence of self-reported HD was substantially lower than the published population estimates (2.5% (Postlethwaite et al., 2019), 6% for ages 55+ (Cath et al., 2017)) regardless of estimation method (Most: 0.5%; Any: 1.7%). In contrast, the prevalence of HD derived using HRS-SR cutoff scores was closer to the previously published estimate for older adults. 883 (5.6%) subjects had a mean HRS score of >=14, indicating likely HD (Tolin et al., 2010). Of 87 subjects who self-reported HD using the Most method, 81 (93%) had a mean HRS of 10 or higher, which is the cutoff that we identified in our previous validation study (Nutley et al., 2020). Of the 279 subjects who self-reported HD using the Any method, 193 (69%) had a mean HRS of 10 or higher.

3.5.   Comparison of self-reported diagnoses with clinical consensus data

We next examined the validity of self-report by calculating accuracy (percent correctly classified), positive and negative predictive value (PPV, NPV), sensitivity, and specificity using the Most and Any methods from cross-tabulation compared to consensus diagnoses in the 235 participants (demographics can be found in supplementary material) who underwent direct clinical assessment (Table 3), as well as Cohen's kappas (Table 4).

Of the 109 participants who endorsed at least one non-HD psychiatric condition most of the time, 94 (86%) met clinical criteria for at least one non-HD psychiatric condition. Accuracy was at least 70% for all diagnoses except HD (Table 3). Specificity was high for all individual diagnoses, and above 90% for all but GAD, while sensitivity was low to moderate (0.07–0.59) across diagnoses using the Most method. Positive predictive value (PPV) was good to high (>0.60) for HD, depression, and panic disorder, as well as for any psychiatric disorder and for any anxiety disorder. The low accuracy rate for HD (56%) and relatively low accuracy rate for depression (74%) were mediated by false negatives. Cohen's kappas were highest for depression and lowest for OCD (Table 4).

Due to the high false positive rates we observed for GAD, OCD, PTSD, bipolar disorder, and social/specific phobia, and the high false negative rates for HD, we examined the rates and patterns of false positives and false negatives for self-report diagnoses in the BHR (see

supplementary material). We hypothesized that false positives could be explained in the majority of cases by either the presence of a commonly comorbid or otherwise confounding condition (e.g., OCD vs HD), or by the presence of subclinical symptomatology. Although these analyses are based on a small number of participants (for example, five false positives for OCD), we observed common patterns that could explain or partly explain the endorsement of conditions. For example, twelve of the fifteen patients who self-reported PTSD but did not meet criteria for PTSD endorsed lifetime trauma plus re-experiencing and/or depression or anxiety (symptomatic trauma). The five participants who self-reported OCD but did not meet criteria all endorsed hoarding obsessions, and also endorsed other obsessions and/or recurrent thoughts but did not meet severity criteria. GAD self-reporting could to a significant extent be explained by the presence of depression and/or another anxiety disorder (See supplementary material).

4. Discussion

The major findings of this study were 1) the consistency of endorsement of psychiatric disorders over repeated assessments varies across psychiatric conditions, with the highest rates in bipolar disorder, depression and PTSD, and hoarding disorder being the least consistently endorsed condition; 2) with the exception of GAD, which had higher prevalences in the BHR, and HD, which had lower prevalence, rates of psychiatric disorders in the BHR were similar when compared to population-based studies; 3) overall accuracy of self-report diagnoses is generally high, but positive predictive value was more variable when compared to diagnoses based on clinical assessment; and 4) self-reporting of a specific disorder may not always accurately reflect a clinical diagnosis of that disorder, but is generally indicative of related psychopathology.

4.1. Consistency of endorsement over repeated assessments

Studies suggest that implausible inconsistencies across longitudinal studies (endorsing history of condition and then not endorsing upon re-interview) are fairly common in self-report of chronic conditions, including in mental health (Cigolle et al., 2018; Jensen et al., 2019). For example, almost 50% of respondents who self-reported having a mental health condition that lasted less than 6 months answered that they have never had that condition upon re-interview 4 years later (Jensen et al., 2019). Temporal inconsistency in the reporting of traumatic events is also common (Hepp et al., 2006). In a cohort of participants of age 25, only 44% of participants with a previous diagnosis of depression before age 21 recalled key depressive symptoms (Wells and Horwood, 2004). Together with the previous literature, the rates of temporally consistent self-reporting in this study suggest that consistency is a key consideration in psychiatric health surveys, and important to consider in retrospective studies, with some disorders having more inconsistent patterns of response (such as HD and panic disorder).

4.2. Comparison of self-reporting prevalence with published prevalence

Estimated prevalence derived using repeated endorsement were similar to previously published prevalences, with the exception of GAD and HD. In contrast, endorsement at any single timepoint resulted in inflated prevalence estimates, suggesting high rates of false

positives. To our knowledge, this is the first study that has compared prevalences of multiple psychiatric disorders using different methods of adjudication from patterns of response to large epidemiological studies. It is notable that GAD and HD were the clear outliers in our analyses, with GAD consistently over-reported and HD consistently under-reported compared to population prevalences and to direct clinical assessment. We hypothesize that the high rates of self-reported GAD in the BHR might be due to confounding of GAD with other anxiety disorders and/or overlapping symptomatology with depression (Zbozinek et al., 2012). Indeed, 60% of participants with GAD also self-reported either depression or another anxiety disorder, and the observed rates of GAD in the BHR (16–25%) were more similar to published prevalence for any anxiety disorder than to those for GAD alone.

In contrast, we hypothesize that HD continues to be under-recognized as a unique psychiatric disorder. Although a small number of participants with clinically ascertained HD endorsed OCD, which is a related condition, in the BHR, the vast majority did not endorse having either OCD or HD (supplementary material). Given the low rate of self-reported OCD, it is unlikely that the low self-reported rates of HD are exclusively due to confounding with OCD, but rather are more likely caused by a lack of insight/education about HD on the part of the participant and on the part of their treating clinicians. Moreover, the relatively recent addition of HD to the fifth edition of the DSM (American Psychiatric Association, 2013) and the 11th edition of the International Classification of Diseases might be a reason for under-endorsement of HD in online registries. However, positive self-reporting of HD was extremely accurate in the clinically ascertained subset.

We focused on positive predictive value (PPV) as a key measure of validity. We argue that although accuracy is important for epidemiological-style analyses, PPV is a more important metric for the purposes of identifying participants for clinical trials, genetic studies or similar research. False negatives cannot be avoided with simple yes/no questions of lifetime history. As expected, simple yes/no questions of psychiatric illness have moderate to poor sensitivity (Stuart et al., 2014; Zarghami et al., 2020).

That said, our comparison of self-reported psychiatric illness to formal clinical assessment suggests that self-reporting of psychiatric illness in the BHR is an indicator of the existence of psychopathology more generally, and that in many cases, the presence of a false positive may signal the existence of a subclinical or related disorder. The PPV for the self-reporting of any psychiatric diagnosis other than HD was>=0.85, suggesting that self-reported psychiatric history in online registries is a useful screen for the presence of psychiatric illness in general, and for some specific disorders in particular, but also that additional assessment measures may be needed. For example, the positive predictive value was high for HD, depression, and panic disorder in this sample, and thus self-reporting might be a sufficient screening tool for these disorders. The observed positive predictive value for depression (85%) using a single question in the BHR surpassed the previously observed values of 61.9% (Stuart et al., 2014), and 74.2% (Sanchez-Villegas et al., 2008). For other disorders, such as PTSD and OCD, self-reporting may be a better indicator of subclinical symptoms (trauma plus re-experiencing symptoms or MDD/anxiety or subclinical obsessive-compulsive or hoarding symptoms), although our results are limited by a small sample size for these disorders.

Based on these findings, we argue that in the BHR at least, self-report of a given psychiatric illness, if not indicative of a clinically diagnosable disorder, does appear to identify a construct that is related to the psychiatric illness of interest and thus may also be valuable as screening questions. For some disorders, such as GAD, HD, and possibly also bipolar disorder, the addition of questions aimed at elucidating presence of symptomatology related to the disorder of interest (e.g., the HRS-SR or the GAD-7(Spitzer et al., 2006)) might be beneficial.

This study represents one of the first to examine the validity of self-reported mental health history, in particular self-reported HD, in a large online registry. It has several limitations, however. First, the BHR population is largely older, highly educated, and female and under-represents minoritized ethnocultural groups, and thus is not representative of the US population, whereas epidemiological studies seek to find a representative set of the population—thus differences in demographics could account for the observed differences in prevalences, and have a larger effect than adjudication methodology. Second, it is impossible to validate all of the BHR self-reported data at an individual level, and HD was over-represented (~50%) in the subset of participants who underwent clinical assessments, and thus the prevalence of other psychiatric disorders in this sample was not representative of the population at large. Third, we have not yet replicated the findings in this paper in other online registries of mental health. It is possible, although arguably unlikely, that overreporting of self-reported psychiatric conditions is specific to the BHR. For example, the nature of the BHR might attract participants who are concerned about their cognition and mental health. As the availability of registries increases, this will ultimately need to be closely examined as every registry will have its own type of bias. Finally, we could not determine the reasons for observed false negatives.

5. Conclusion

The results of this study indicate that self-reported mental health history in online registries is, when used thoughtfully, a valuable resource. One future suggestion for online registries that include a mental health self-report component is to include, particularly for those disorders with lower accuracy of self-report diagnoses, well-validated symptom-based instruments to complement the self-reported history, as was done for HD (Nutley et al., 2020). Moreover, when longitudinal assessments are available, the consistency of self-report should carefully be assessed, as we expect that consistent self-report might be a better indicator of the psychopathology being assessed. Lastly, although we conclude that self-reporting of psychiatric conditions in an online setting is a fair indicator of psychopathology, it should be accompanied by more in-depth interviews if using data from a participant for a specific disorder.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Declaration of competing interest

Dr. Mathews has received royalties from W.W. Norton and Company, and travel funding from the Tourette Association of America. She has received research funding from the National Institutes of Health, the International Obsessive Compulsive Foundation, the Tourette Association of America, and the Patient Centered Research Outcomes Institute. She is an unpaid member of several advisory boards, including the Tourette Association of America, the International OCD Foundation, and the Family Foundation for OCD Research. Dr. Mackin has received research funding from the National Institutes of Health and Janssen Research and Development LLC. Dr. Sordo Vieira has received funding from the American Mathematical Society, the US National Science Foundation, and the National Institutes of Health. Dr. Nosheny has received research funding from the National Institutes of Health and Genentech, Inc., and travel funding from Mount Sinai Medical Center. Dr. Simpson has received royalties from the Tourette Association of America for consultation and speaking engagements. Dr. Weiner has served on Advisory Boards for Eli Lilly, Cerecin/Accera, Roche, Alzheon, Inc., Merck Sharp & Dohme Corp., Nestle/Nestec, PCORI/PPRN, Dolby Family Ventures, National Institute on Aging (NIA), Brain Health Registry and ADNI. Dr. Weiner serves on the Editorial Boards for Alzheimer's & Dementia, TMRI and MRI. He has provided consulting and/or acted as a speaker/lecturer to Cerecin/Accera, Inc., BioClinica, Nestle/ Nestec, Roche, Genentech, NIH, The Buck Institute for Research on Aging, FUJIFILM-Toyama Chemical (Japan), Garfield Weston, Baird Equity Capital, University of Southern California (USC), Cytox, and Japanese Organization for Medical Device Development, Inc. (JOMDD) and T3D Therapeutics. Dr. Weiner holds stock options with Alzheon, Inc., Alzeca, and Anven. Dr. Weiner receives support for his work from the National Institutes of Health, Department of Defense, Patient-Centered Outcomes Research Institute, the California Dept. of Public Health, The University of Michigan, Siemens, Biogen, Hillblom Foundation, the Alzheumer's Association, the State of California, Johnson and Johnson, Kevin and Connie Shanahan, GE, VUmc, Australian Catholic University, The Stroke Foundation and the Veterans Administration.

Abbreviations:

BHR Brain Health Registry

References

Alonso J, Ferrer M, Romera B, Vilagut G, Angermeyer M, Bernert S, Brugha TS, Taub N, Mccolgen Z, Girolamo GD, Polidori G, Mazzi F, Graaf RD, Vollebergh WAM, Buist-Bowman MA, Demyttenaere K, Gasquet I, Haro JM, Palacín C, Autonell J, Katz SJ, Kessler RC, Kovess V, Lépine JP, Arbabzadeh-Bouchez S, Ormel J, Bruffaerts R, THE ESEMeD/MHEDEA 2000 INVESTIGATORS, 2002. The European study of the epidemiology of mental disorders (ESEMeD/ MHEDEA 2000) project: rationale and methods. Int. J. Methods Psychiatr. Res 11, 55–67. 10.1002/ mpr.123. [PubMed: 12459795]

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders, fifth ed. American Psychiatric Association. 10.1176/appi.books.9780890425596.

Andrews G, Anstey K, Brodaty H, Issakidis C, Luscombe G, 1999. Recall of depressive episode 25 years previously. Psychol. Med 29, 787–791. 10.1017/S0033291799008648. [PubMed: 10473305]

Ayers CR, 2017. Age-specific prevalence of hoarding and obsessive compulsive disorder: a population-based study. Am. J. Geriatr. Psychiatr 25, 256–257. 10.1016/j.jagp.2016.12.001.

Cath DC, Nizar K, Boomsma D, Mathews CA, 2017. Age-specific prevalence of hoarding and obsessive compulsive disorder: a population-based study. Am. J. Geriatr. Psychiatry Off. J. Am. Assoc. Geriatr. Psych 25, 245–255. 10.1016/j.jagp.2016.11.006.

Cigolle CT, Nagel CL, Blaum CS, Liang J, Quiñones AR, 2018. Inconsistency in the self-report of chronic Diseases in panel surveys: developing an adjudication method for the health and retirement

study. J. Gerontol. B Psychol. Sci. Soc. Sci 73, 901–912. 10.1093/geronb/gbw063. [PubMed: 27260670]

Compton WM, Thomas YF, Stinson FS, Grant BF, 2007. Prevalence, correlates, disability, and comorbidity of DSM-IV drug abuse and dependence in the United States: results from the national epidemiologic survey on alcohol and related conditions. Arch. Gen. Psychiatr 64, 566. 10.1001/archpsyc.64.5.566. [PubMed: 17485608]

Drapeau A, Boyer R, Diallo FB, 2011. Discrepancies between survey and administrative data on the use of mental health services in the general population: findings from a study conducted in Qúebec. BMC Publ. Health 11, 837. 10.1186/1471-2458-11-837.

Frost RO, Steketee G, Grisham J, 2004. Measurement of compulsive hoarding: saving inventory-revised. Behav. Res. Ther 42, 1163–1182. 10.1016/j.brat.2003.07.006. [PubMed: 15350856]

Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, Heninger GR, Charney DS, 1989. The Yale-Brown obsessive compulsive scale. I. Development, use, and reliability. Arch. Gen. Psychiatr 46, 1006–1011. 10.1001/archpsyc.1989.01810110048007. [PubMed: 2684084]

Grant BF, Dawson DA, Stinson FS, Chou PS, Kay W, Pickering R, 2003. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. Drug Alcohol Depend. 71, 7–16. 10.1016/s0376-8716(03)00070-x. [PubMed: 12821201]

Grant BF, Hasin DS, Stinson FS, Dawson DA, June Ruan W, Goldstein RB, Smith SM, Saha TD, Huang B, 2005a. Prevalence, correlates, co-morbidity, and comparative disability of DSM-IV generalized anxiety disorder in the USA: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Psychol. Med 35, 1747–1759. 10.1017/S0033291705006069. [PubMed: 16202187]

Grant BF, Stinson FS, Hasin DS, Dawson DA, Chou SP, Ruan WJ, Huang B, 2005b. Prevalence, correlates, and comorbidity of bipolar I disorder and axis I and II disorders: results from the national epidemiologic survey on alcohol and related conditions. J. Clin. Psychiatr 66, 1205–1215. 10.4088/jcp.v66n1001.

Hasin DS, Grant BF, 2015. The national epidemiologic survey on alcohol and related conditions (NESARC) waves 1 and 2: review and summary of findings. Soc. Psychiatr. Psychiatr. Epidemiol 50, 1609–1640. 10.1007/s00127-015-1088-0.

Hepp U, Gamma A, Milos G, Eich D, Ajdacic-Gross V, Rössler W, Angst J, Schnyder U, 2006. Inconsistency in reporting potentially traumatic events. Br. J. Psychiatry 188, 278–283. 10.1192/bjp.bp.104.008102. [PubMed: 16507971]

Hyde CL, Nagle MW, Tian C, Chen X, Paciga SA, Wendland JR, Tung JY, Hinds DA, Perlis RH, Winslow AR, 2016. Identification of 15 genetic loci associated with risk of major depression in individuals of European descent. Nat. Genet 48, 1031–1036. 10.1038/ng.3623. [PubMed: 27479909]

Jensen HAR, Davidsen M, Christensen AI, Ekholm O, 2019. Inconsistencies in self-reported health conditions: results of a nationwide panel study. Int. J. Publ. Health 64, 1243–1246. 10.1007/s00038-019-01287-0.

Kessler RC, Berglund P, Chiu WT, Demler O, Heeringa S, Hiripi E, Jin R, Pennell B-E, Walters EE, Zaslavsky A, Zheng H, 2004. The US national comorbidity survey replication (NCS-R): design and field procedures. Int. J. Methods Psychiatr. Res 13, 69–92. 10.1002/mpr.167. [PubMed: 15297905]

Kessler RC, Merikangas KR, 2004. The national comorbidity survey replication (NCS-R): background and aims. Int. J. Methods Psychiatr. Res 13, 60–68. 10.1002/mpr.166. [PubMed: 15297904]

Nordsletten AE, Fernández de la Cruz L, Pertusa A, Reichenberg A, Hatch SL, Mataix-Cols D, 2013. The structured interview for hoarding disorder (SIHD): development, usage and further validation. J. Obsessive-Compuls. Relat. Disord 2, 346–350. 10.1016/j.jocrd.2013.06.003.

Nutley SK, Bertolace L, Sordo Vieira L, Nguyen B, Ordway A, Simpson H, Zakrzewski J, Camacho MR, Eichenbaum J, Nosheny R, Weiner M, Mackin RS, Mathews CA, 2020. Internet-based hoarding assessment: the reliability and predictive validity of the internet-based Hoarding Rating Scale, Self-Report. Psychiatr. Res 294, 113505 10.1016/j.psychres.2020.113505.

Pietrzak RH, Goldstein RB, Southwick SM, Grant BF, 2011. Prevalence and Axis I comorbidity of full and partial posttraumatic stress disorder in the United States: results from wave 2 of the national epidemiologic survey on alcohol and related conditions. J. Anxiety Disord 25, 456–465. 10.1016/j.janxdis.2010.11.010. [PubMed: 21168991]

Postlethwaite A, Kellett S, Mataix-Cols D, 2019. Prevalence of Hoarding Disorder: a systematic review and meta-analysis. J. Affect. Disord 256, 309–316. 10.1016/j.jad.2019.06.004. [PubMed: 31200169]

R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Sanchez-Villegas A, Schlatter J, Ortuno F, Lahortiga F, Pla J, Benito S, Martinez-Gonzalez MA, 2008. Validity of a self-reported diagnosis of depression among participants in a cohort study using the Structured Clinical Interview for DSM-IV (SCID-I). BMC Psychiatr. 8, 43. 10.1186/1471-244X-8-43.

Saxena S, Ayers CR, Dozier ME, Maidment KM, 2015. The UCLA hoarding severity scale: development and validation. J. Affect. Disord 175, 488–493. 10.1016/j.jad.2015.01.030. [PubMed: 25681559]

Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC, 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J. Clin. Psychiatr 59 (Suppl. 20), 22–33 quiz 34–57.

Spitzer RL, Kroenke K, Williams JBW, Löwe B, 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch. Intern. Med 166, 1092. 10.1001/archinte.166.10.1092. [PubMed: 16717171]

St Clair P, Gaudette É, Zhao H, Tysinger B, Seyedin R, Goldman DP, 2017. Using self-reports or claims to assess disease prevalence: it's complicated. Med. Care 55, 782–788. 10.1097/MLR.0000000000000753. [PubMed: 28617703]

Stuart AL, Pasco JA, Jacka FN, Brennan SL, Berk M, Williams LJ, 2014. Comparison of self-report and structured clinical interview in the identification of depression. Compr. Psychiatr 55, 866–869. 10.1016/j.comppsych.2013.12.019.

Tolin DF, Frost RO, Steketee G, 2010. A brief interview for assessing compulsive hoarding: the Hoarding Rating Scale-Interview. Psychiatr. Res 178, 147–152. 10.1016/j.psychres.2009.05.001.

Tung JY, Do CB, Hinds DA, Kiefer AK, Macpherson JM, Chowdry AB, Francke U, Naughton BT, Mountain JL, Wojcicki A, Eriksson N, 2011. Efficient replication of over 180 genetic associations with self-reported medical data. PLoS One 6, e23473. 10.1371/journal.pone.0023473. [PubMed: 21858135]

U.S. Department of Health and Human Services, 2022, January 4. List of Registries. National Institutes of Health. Retrieved. https://www.nih.gov/health-information/nih-clinical-research-trials-you/list-registries. (Accessed 28 February 2022). from.

van den Akker M, van Steenkiste B, Krutwagen E, Metsemakers JFM, 2015. Disease or no disease? Disagreement on diagnoses between self-reports and medical records of adult patients. Eur. J. Gen. Pract 21, 45–51. 10.3109/13814788.2014.907266.

Weiner MW, Nosheny R, Camacho M, Truran-Sacrey D, Mackin RS, Flenniken D, Ulbricht A, Insel P, Finley S, Fockler J, Veitch D, 2018. The Brain Health Registry: an internet-based platform for recruitment, assessment, and longitudinal monitoring of participants for neuroscience studies. Alzheimers Dement. 14, 1063–1076. 10.1016/j.jalz.2018.02.021. [PubMed: 29754989]

Wells JE, Horwood LJ, 2004. How accurate is recall of key symptoms of depression? A comparison of recall and longitudinal reports. Psychol. Med 34, 1001–1011. 10.1017/S0033291703001843. [PubMed: 15554571]

Wolinsky FD, Jones MP, Ullrich F, Lou Y, Wehby GL, 2015. Cognitive function and the concordance between survey reports and medicare claims in a nationally representative cohort of older adults. Med. Care 53, 455–462. 10.1097/MLR.0000000000000338. [PubMed: 25793268]

Wolinsky FD, Jones MP, Ullrich F, Lou Y, Wehby GL, 2014. The concordance of survey reports and medicare claims in a nationally representative longitudinal cohort of older adults. Med. Care 52, 462–468. 10.1097/MLR.0000000000000120. [PubMed: 24714584]

Zarghami M, Taghizadeh F, Moosazadeh M, Kheradmand M, Heydari K, 2020. Validity of self-reporting depression in the Tabari cohort study population. Neuropsychopharmacol. Rep 40, 342–347. 10.1002/npr2.12138. [PubMed: 32951353]

Zbozinek TD, Rose RD, Wolitzky-Taylor KB, Sherbourne C, Sullivan G, Stein MB, Roy-Byrne PP, Craske MG, 2012. Diagnostic overlap of generalized anxiety disorder and major depressive disorder in a primary care sample. Depress. Anxiety 29, 1065–1071. 10.1002/da.22026. [PubMed: 23184657]
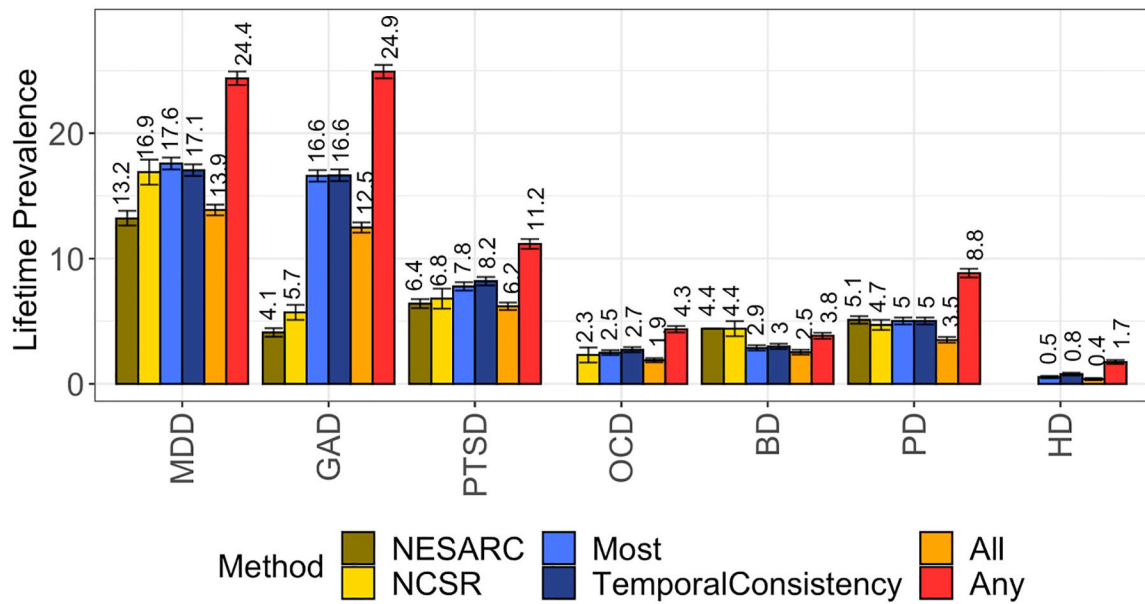
Fig. 1.
(Prevalence from adjudication methods vs NCSR and NESARC)

Prevalence estimates of various psychiatric conditions based on different methodologies in the BHR vs NESARC and NCS-R. BD= Bipolar Disorder, MDD = Major Depressive Disorder, GAD = Generalized Anxiety Disorder, OCD=Obsessive-Compulsive Disorder, PD=Panic Disorder, PTSD=Post-Traumatic Stress Disorder. OCD was not assessed in the NESARC. HD was not assessed in either NESARC or NCS-R.
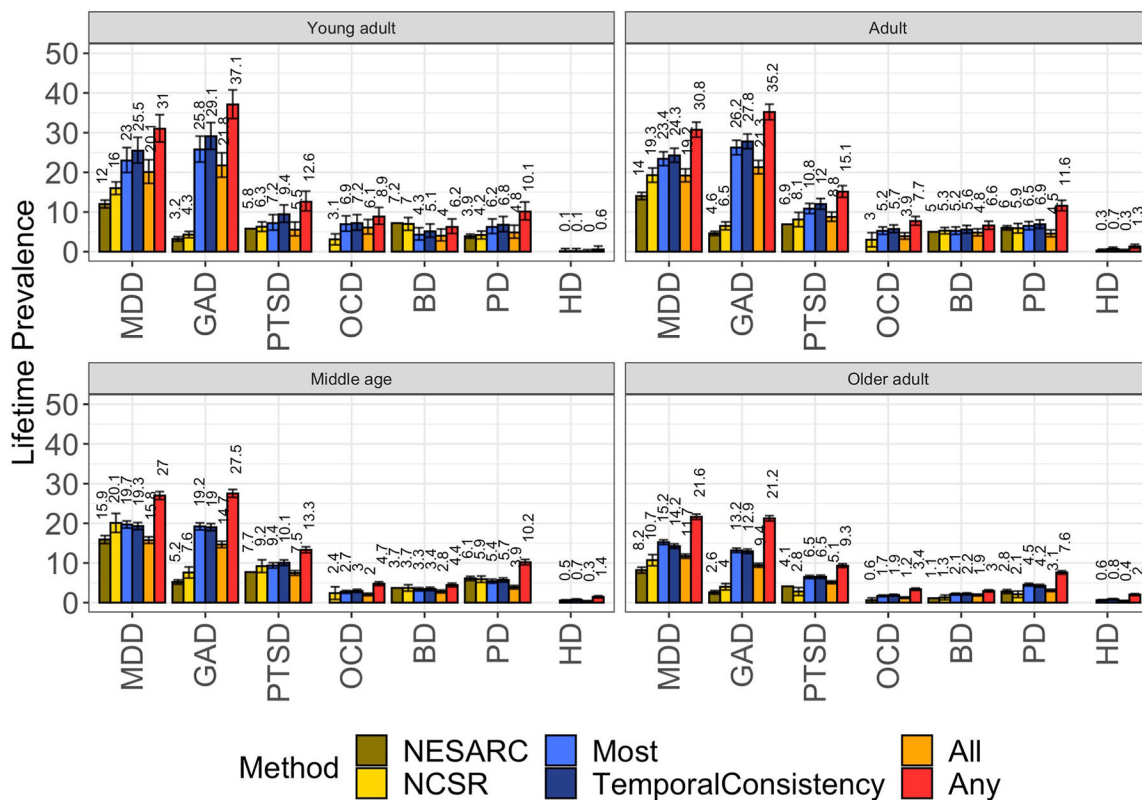
Fig. 2.

(Prevalence from adjudication methods vs NCSR and NESARC stratified by age)

Prevalence estimates based on the All, Most and Any Method based on NCS-R and NESARC populations stratified by age. OCD was not assessed in the NESARC. HD was not assessed in NESARC nor in NCS-R.

Table 1

Demographics of participants in the Brain Health Registry. No post-secondary degree indicates that the highest level of education by participant is grammar school, high school, or some college. Post-HS degree indicates that the highest level of education is a 2 year degree, 4 year degree, master degree, doctoral degree or professional degree.

	Overall, N = 44,156[a]	Completion of BHR medical module		p-value[b]
		At least twice, N = 25,037[a]	Only once, N = 19,119[a]	
Age, years	60 (51, 67)	61 (53, 68)	59 (48, 67)	<0.001
Race				<0.001
Multiple/Other	6,224 (14%)	2,788 (11%)	3,436 (18%)	
Non-Latino White	37,932 (86%)	22,249 (89%)	15,683 (82%)	
Gender				0.060
Female	33,305 (75%)	18,800 (75%)	14,505 (76%)	
Male	10,851 (25%)	6,237 (25%)	4,614 (24%)	
Education				<0.001
No post-secondary degree	9,320 (21%)	4,422 (18%)	4,898 (26%)	
Post-HS degree	34,836 (79%)	20,615 (82%)	14,221 (74%)	

[a]Median (IQR); n (%).

[b]Wilcoxon rank sum test; Pearson's Chi-squared test.

Table 2

(Consistency of self-reporting per condition after endorsing at baseline using the Most and All Method).

	Cases	Most	All
MDD	4751	4010 (84%)	3475 (73%)
GAD	4466	3677 (82%)	3124 (70%)
PTSD	2012	1725 (86%)	1549 (77%)
OCD	718	537 (75%)	470 (65%)
BD	772	677 (88%)	630 (82%)
PD	1474	1076 (73%)	877 (59%)
HD	219	109 (50%)	94 (43%)

MDD = Major Depressive Disorder, GAD = Generalized Anxiety Disorder, PTSD= Post-Traumatic Stress Disorder, OCD=Obsessive-Compulsive Disorder, BD=Bipolar Disorder, PD=Panic Disorder, HD= Hoarding Disorder.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

(Sensitivity and specificity, based on Most and Any methods).

Condition	Method	Prevalence	Accuracy	Sensitivity	Specificity	PPV	NPV
Any psychopathology	Most	0.69	0.70	0.63	0.86	0.91	0.50
	Any		0.77	0.74	0.83	0.91	0.58
Any non-HD psychopathology	Most	0.60	0.74	0.67	0.84	0.86	0.63
	Any		0.80	0.80	0.80	0.85	0.73
Any anxiety	Most	0.35	0.72	0.50	0.84	0.63	0.76
	Any		0.73	0.65	0.77	0.60	0.80
HD	Most	0.50	0.56	0.11	1.00	1.00	0.53
	Any		0.60	0.19	1.00	1.00	0.55
Depression	Most	0.45	0.74	0.53	0.92	0.85	0.70
	Any		0.77	0.66	0.86	0.80	0.76
Panic disorder	Most	0.15	0.87	0.23	0.98	0.73	0.88
	Any		0.87	0.49	0.94	0.59	0.91
GAD	Most	0.22	0.79	0.59	0.84	0.51	0.88
	Any		0.72	0.65	0.74	0.41	0.88
PTSD	Most	0.13	0.87	0.47	0.93	0.48	0.92
	Any		0.85	0.53	0.90	0.43	0.93
Social phobia	Most	0.09	0.91	0.14	0.98	0.43	0.92
	Any		0.88	0.24	0.94	0.28	0.93
OCD	Most	0.06	0.92	0.07	0.98	0.17	0.94
	Any		0.89	0.13	0.94	0.13	0.94
Bipolar disorder	Most	0.05	0.94	0.36	0.97	0.40	0.97
	Any		0.94	0.36	0.97	0.36	0.97

Descriptive statistics of accuracy of self-endorsement based on the most and any method vs clinical diagnoses. Prevalence refers to the proportion of cases in the clinical interviews (N = 235).

Table 4

(Cohen's kappa based on Most and Any methods).

Condition	Method	κ
GAD	Most	0.41
	Any	0.32
PD	Most	0.30
	Any	0.46
Depression	Most	0.47
	Any	0.53
OCD	Most	0.06
	Any	0.07
Bipolar disorder	Most	0.35
	Any	0.33
Hoarding disorder	Most	0.11
	Any	0.19
PTSD	Most	0.40
	Any	0.39

Cohen's kappas were computed using the Most and Any method and compared to clinical consensus data (N = 235).