



Published in final edited form as:

Psychiatry Res. 2020 December ; 294: 113505. doi:10.1016/j.psychres.2020.113505.

Internet-based hoarding assessment: the reliability and predictive validity of the internet-based Hoarding Rating Scale, Self-Report

Sara K. Nutley^a, Lyvia Bertolace^b, Luis Sordo Vieira^c, Binh Nguyen^b, Ashley Ordway^b, Heather Simpson^b, Jessica Zakrzewski^b, Monica R. Camacho^d, Joseph Eichenbaum^{d,e}, Rachel Nosheny^f, Michael Weiner^{e,f}, R. Scott Mackin^{f,g}, Carol A. Mathews^{b,*}

^a Department of Epidemiology, University of Florida Gainesville, FL, United States

^b Department of Psychiatry, University of Florida Gainesville, FL, United States

^c Department of Medicine, University of Florida Gainesville, FL, United States

^d San Francisco VA Medical Center San Francisco, California, United States

^e Department of Radiology, University of California, San Francisco San Francisco, California, United States

^f Department of Psychiatry, University of California, San Francisco San Francisco, California, United States

^g Department of Mental Health, Veterans Affairs Medical Center, San Francisco, California, United States

Abstract

The Hoarding Rating Scale – Self Report (HRS-SR) is a 5-item assessment developed to ascertain the presence and severity of hoarding symptoms. This study aimed to evaluate the validity of an online adaptation of the HRS-SR in a remote, unsupervised internet sample of 23,214 members of the Brain Health Registry (BHR), an online research registry that evaluates and longitudinally monitors cognition, medical and psychiatric health status. Convergent validity was assessed among a sub-sample of 1,183 participants who completed additional, remote measures of self-reported hoarding behaviors. Structured clinical interviews conducted in-clinic and via video conferencing tools were conducted among 230 BHR participants; ROC curves were plotted to assess the diagnostic performance of the internet-based HRS-SR using best estimate hoarding disorder (HD) diagnoses as the gold standard. The area under the curve indicated near-perfect model accuracy, and was confirmed with 10-fold cross validation. Sensitivity and specificity for distinguishing clinically relevant hoarding were optimized using an HRS-SR total score cut-off of 5. Longitudinal analyses indicated stability of HRS-SR scores over time. Findings indicate that the internet-based HRS-SR is a useful and valid assessment of hoarding symptoms, though additional research using samples with more diverse hoarding behavior is needed to validate optimal cut-off values.

*Corresponding Author: Carol A. Mathews, Telephone: 352-294-4927, Fax: 352-627-5015, carolmathews@ufl.edu.

Author Statement: We confirm that this manuscript has not been published elsewhere and is not under consideration by another journal. All authors have approved the manuscript and agree with submission to *Psychiatry Research*. Thank you for your consideration.

Keywords

Hoarding; Hoarding Disorder; Mental Disorders; Validation Study; Remote unsupervised; Internet

1. Introduction

Hoarding disorder (HD) is a psychiatric condition characterized by a persistent difficulty discarding possessions that results in a debilitating accumulation of clutter and contributes to significant impairment in daily functioning (American Psychiatric Association, 2013; Ayers et al., 2013; Chakraborty et al., 2012; Matsunaga et al., 2010; Vorstenbosch et al., 2012; Rosa et al., 2012). In the general population, the prevalence of clinically significant hoarding behavior ranges from 2% to 4% (American Psychiatric Association, 2013; Mueller et al., 2009; Samuels et al., 2008; Iervolino et al., 2009). Hoarding symptoms are often chronic and progressive, typically beginning in adolescence and increasing in severity and prevalence throughout life (Grisham et al., 2006; Cath et al., 2017). In particular, the prevalence of clinically significant hoarding appears to increase linearly by nearly 20% with every 5 years of age; among older adults, prevalence is over 6% (Cath et al., 2017).

Despite designation of hoarding disorder as a separate diagnosis in the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), factors contributing to etiology and functional consequences of this debilitating psychiatric illness are still limited, and under-diagnosis is common (Mataix-Cols & Fernández de la Cruz, 2018). Accurate diagnosis is heavily reliant on reliable assessment of hoarding behaviors and the resulting impairment in social and occupational functioning. Though traditionally diagnosed on the basis of semi-structured clinical interviews, researchers have established a number of well-validated, self-report rating scales designed to evaluate the prominent features of hoarding disorder and related symptoms and provide provisional diagnoses of HD. These self-report assessments are particularly useful for identifying HD risk in non-clinical settings where stigma is prevalent and awareness of HD as a psychiatric illness may be limited (Chasson et al., 2018).

In recent years, there has been a rapid increase in the number of psychiatric studies using unsupervised, internet-based data collection methods. Remote, self-report psychiatric measures can be obtained for large numbers of participants much more efficiently and at significantly lower cost than traditional clinical interviews. Additionally, given the ease of repeat self-report assessment, these measures are particularly well suited for longitudinal evaluation of psychiatric symptoms. In the case of hoarding disorder, the validation of brief self-report measures of hoarding behavior for unsupervised use on the internet can assist in the identification of individuals with at-risk hoarding behaviors in population-based samples. One instrument that can be adapted for remote unsupervised assessment via the internet is the Hoarding Rating Scale, Self-Report (HRS-SR), a 5-item hoarding assessment developed to ascertain the presence and severity of hoarding symptoms such as difficulty discarding, clutter, excessive acquisition, and hoarding-related stress and impairment (Tolin et al., 2010). For each HRS-SR question, scores range from 0 to 8 and are subsequently summed for a total score ranging from 0 to 40, with higher scores indicating more severe hoarding

symptomatology. The HRS-SR has been previously used in genetic, clinical, and epidemiological studies and has been proven reliable and valid for differentiating individuals with HD from those with OCD and community controls (Cath et al., 2017; Iervolino et al., 2009; Monzani et al., 2014; Iervolino et al., 2011; Ivanov et al., 2013; Perroud et al., 2011; Frost et al., 2011; Frost & Hristova, 2011). While there is not a single gold standard cut off score, previous studies have identified potential cut off formulae used to designate HD: (1) HRS-SR total score greater than 14 (Tolin et al., 2010; Faraci et al., 2018; Rodriguez et al., 2012; Turna et al., 2018); (2) HRS-SR total score greater than 17 (Mathews et al., 2014; Zilhão et al., 2016; Iervolino et al., 2009); (3) HRS-SR component scores greater than or equal to 4 (moderate) for each of the following symptoms: clutter, difficulty discarding, and either emotional distress or impairment (for a total score of >12), in line with the DSM-5 criteria for HD (Tolin et al., 2008; 2010).

Although widely used as a traditional self-report assessment, there continues to be a dearth of literature on the performance of the HRS-SR in internet-based samples. Accordingly, the present study has three primary aims. The first is to validate our newly developed internet-based and remotely-administered HRS-SR, using diagnosis of hoarding disorder based on semi-structured clinical interview as the gold standard by determining the optimal scoring method and cutoff score for identifying clinically relevant hoarding behaviors. The second aim of this study is to assess the longitudinal stability of internet-based HRS-SR scores. Our final aim is to assess the convergent validity of internet-based HRS-SR in comparison to other well-established, self-report measures of hoarding behavior. This study uses data collected from members of the Brain Health Registry (BHR), a national, online research registry that aims to identify, evaluate and longitudinally track cognition, medical and psychiatric health in adults ages 18 and older (Weiner et al., 2018).

2. Methods

2.1. Participants

The Brain Health Registry (BHR), led by researchers at the University of California, San Francisco, consists of approximately 70,000 online research participants (Brain Health Registry, 2012–2019; Weiner et al., 2018; Mackin et al., 2018). Using an ongoing recruitment model, approximately 500 new participants register for the study each month by providing their name and email address, creating a unique username and password, and providing informed consent. The sole exclusion criteria for registration is age less than 18 years. Currently, the BHR has participants from all 50 states, with over-representation of participants on the coasts and particularly in California, where the BHR was begun and recruitment efforts have been the strongest (~30% of the current sample).

BHR participants are invited by email to semi-annually complete online cognitive tests and comprehensive self-report questionnaires related to medical history and symptomology, health behaviors, quality of life factors, cognitive function, and memory. In addition to the core questions, modules targeting specific areas of interest have been added over time. The Hoarding and Clutter module was added to the BHR in February 2017 to gather information about lifetime diagnoses of HD and obsessive-compulsive related disorders, and hoarding symptoms including an online version of the HRS-SR (Supplemental Table 1), which was

adapted from the Hoarding Rating Scale (HRS). As BHR participants are re-contacted every six months and asked to complete the questionnaires with respect to their current symptoms, many participants in the dataset have internet-based HRS-SR data from multiple time points. For all analyses using a single time point, participants' most recent data from the Hoarding and Clutter module was collected.

Of 70,000 participants enrolled in the BHR as of December 2019, 23,214 individuals (75.5% female; mean age = 60.7 years, SD = 13.0 years) had completed the Hoarding and Clutter module at the time of this study. The majority of participants identified as white (88.3%) and a large number obtained a college (39.9%) or graduate (41.4%) degree.

2.2 Subsample 1: Additional Self-Report Hoarding Assessments

Of 23,214 participants included in this analysis, 1,183 individuals (80.6% female, mean age = 61.2 years) were contacted by the research team for additional assessments. These participants completed two additional self-report measures of hoarding behavior and severity, the Saving Inventory, Revised (SI-R; Frost et al., 2014) and the UCLA Hoarding Severity Scale (UHSS; Saxena et al., 2007) as a part of a larger study on the relationship between hoarding and functional disability. Both surveys were administered remotely in an unsupervised setting using an electronic survey distribution tool.

For this component of the study, BHR participants with HRS-SR scores of 14 or greater (at any BHR data collection time point), as well as those with self-reported hoarding disorder were recruited as cases ($n=176$). Those whose most recent HRS-SR score equaled 0 were identified as controls ($n=854$). A small number of individuals with most recent HRS-SR scores between 1 and 14 were also recruited for participation and were identified as intermediate participants ($n=153$).

The SI-R is a 23-item self-report questionnaire that measures problems with clutter, difficulty discarding, excessive acquisition, and hoarding-related impairment (Frost et al., 2014). Each item is scored on a 5-point scale; the total possible score is 92. The SI-R is widely used in hoarding research, demonstrates good test-retest reliability, and has been shown to reliably discriminate between individuals with HD and both community and clinical controls (Kellman-McFarlane et al., 2019). A cutoff score between 40 and 42, depending on the study, is often used to identify individuals with clinically significant hoarding.

The UHSS is a 10-item, semi-structured interview that assesses the frequency of hoarding symptoms over a one-week period and evaluates predictors of functional impairment among those with HD (i.e. perfectionism, indecisiveness, and procrastination; Saxena et al., 2007; Saxena et al., 2015). Individual items are scored on a 5-point scale and are summed for a total score ranging from 0 to 40, with higher scores indicating a higher prevalence of hoarding symptoms. Although traditionally clinician-administered, the UHSS was self-administered in this study as has been done in our previous studies (Chou et al., 2018; Archer et al., 2019).

2.3. Subsample 2: Clinical Interview

Two hundred and thirty participants in subsample 1 underwent a semi-structured clinical diagnostic interview. Participants for this component of the study were randomly selected from three groups: (1) case ($n=121$, 52.6%), (2) control ($n=108$, 47.0%), and (3) intermediate ($n=1$, 0.4%) groups in subsample 1. Interviews were conducted in-person or via telephone and were approximately one to three hours in length. Clinicians trained in the assessment of hoarding disorder and other psychiatric conditions for research purposes conducted interviews using three well-validated psychiatric assessments: The Mini International Neuropsychiatric Interview (MINI), the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS), and the Structured Interview for Hoarding Disorder (SIHD).

The MINI (English version 7.0.2) is a brief, structured diagnostic interview designed to evaluate 17 of the most common psychiatric health conditions included in the DSM-5 and the 10th revision of the International Classification of Diseases (ICD-10; Sheehan et al., 1998). In this study, the MINI was administered to the participant by the clinician and the item-level data were subsequently used to identify major DSM-5 Axis I diagnoses.

The Y-BOCS is a well-validated measure of obsessive compulsive symptoms and impairment and severity caused by them. The Y-BOCS version used in this study consists of a symptom checklist of 69 common obsessions and compulsions and a 10-item scale that rates time occupied by obsession and compulsions, related impairment and distress, and perceived degree of control over thoughts and behaviors (Goodman et al., 1989). Administered using a semi-structured interview, symptoms are organized in 15 theme-based categories related to obsessions and compulsions that are scored as absent, currently present, and present in the past but not currently present. Individual severity/impairment items are scored on a scale of 0 (no symptoms) to 4 (extreme symptoms) and summed for a total score ranging from 0 to 40, with separate subtotals for obsessions and compulsions. The Y-BOCS was used in this study to determine lifetime and current diagnoses of obsessive compulsive disorder (OCD).

The SIHD is a semi-structured assessment of core HD diagnostic criteria (DSM-5) and corresponding HD specifiers (Nordsletten et al., 2013). Assessment subsections reflect each of the disorder's six core features; the first question of each section is close-ended and subsequent items invite interviewees to elaborate. Using skip rules, the SIHD also aids clinicians in identifying other possible causes of hoarding (i.e. OCD, ASD). In this study, the SIHD was used to diagnose lifetime and current hoarding disorder and to assess the degree of insight the participant had into their symptoms.

2.4. Clinical Best Estimate Procedure

A best estimate (BE) procedure using information obtained from the clinical interviews (MINI, Y-BOCS, and SIHD) combined with self-report data from the SI-R and UHSS scales was used to assign clinical diagnoses that were used as the gold standard in this analysis. HRS-SR scores were not used to assign clinical diagnoses. Best estimators (seven clinicians with research expertise who did not conduct interviews and were blinded to the case/control status of the participant as well as their HRS-SR scores) reviewed clinical interview

documents and questionnaires as well as interviewer notes. Lifetime diagnoses were assigned for Hoarding Disorder as well as other psychiatric diagnoses (not part of the present analysis) using DSM-5 criteria.

For each lifetime diagnosis, best estimators gave a rating of “not present”, “probable”, or “definite”, and provided information related to age of onset and severity as appropriate. A “probable” diagnosis was coded if the interviewee did not meet full diagnostic criteria, but information gathered from questionnaires and interviewer notes suggested a diagnosis was likely to be present.

For all interviewees, one best estimate was initially conducted. If all psychiatric diagnoses (including HD) were determined to be “not present”, a second estimator was not assigned, and the interviewee was determined a control. If any diagnoses were rated as “probable” or “definite”, a second estimator was assigned. Upon completion of two best estimator ratings, discrepancies were discussed as appropriate to reach consensus. If consensus was not achieved, a third best estimator was assigned and a consensus was reached between the three clinicians.

2.5. Statistical Analysis

Analyses were conducted using SAS 9.4. Confidence intervals were calculated using a level of 95% confidence. Pearson’s correlation coefficient was used to evaluate convergent validity between the internet-based HRS-SR and the SI-R and UHSS in subsample 1. Pearson correlation coefficients were interpreted using recommended criteria: 0.0–0.3 negligible; 0.31–0.5 low; 0.51–0.7 moderate; 0.71–0.9 high; 0.91–1.0 high positive (Mukaka, 2012). ROC curves were plotted to assess the diagnostic performance of the internet-based HRS-SR using the Saving Inventory – Revised (cut-off of 42) and the UCLA Hoarding Severity Scale (cut-off of 20) as the gold standard. In subsample 2, Cohen’s kappa was used to examine the agreement between a best estimate diagnosis of HD and whether an individual met criteria for HD based on a DSM-5 diagnostic proxy. Kappa values were interpreted using the following criteria: 0.0–0.2 none; 0.21–0.39 minimal; 0.40–0.59 weak; 0.60–0.79 moderate; 0.80–0.90 strong; 0.91–1.0 almost perfect (McHugh, 2012). ROC curves were plotted to assess the diagnostic performance of the internet-based HRS-SR using best estimate consensus HD diagnoses as the gold standard. For all ROC curves, optimal HRS-SR cut-points were determined using Youden’s index (J), a measure that identifies the maximum potential effectiveness of an assessment (Roupp et al., 2008). Values range from 0 to 1, with higher values indicating better diagnostic accuracy. Ten-fold cross validation was used to evaluate model robustness. Data were randomly divided into ten equal subgroups to be used as testing samples. The remaining 90% of the data were used to train ten separate logistic regression models which used internet-based HRS-SR to predict in-clinic best estimate consensus HD diagnoses.

Reliable change—Given that a single BHR time point was used to determine case status, and many participants completed the internet-based HRS-SR at multiple time points, we further examined the stability of HRS-SR scores over time. First, Pearson’s correlation coefficient was used to assess the relationship between HRS-SR total scores for each pair of

adjacent time points. The Reliable Change Index (RCI; Jacobson et al., 1984) was used to identify clinically meaningful change in HRS-SR scores between adjacent time points. To calculate RCI, the difference between each pair of adjacent time points was determined and the result was divided by the standard error of the difference in scores (i.e.

$$RCI = \frac{(HRS_{post} - HRS_{pre})}{SE_{diff}}$$

When the RCI exceeded 1.96, it was determined that the

difference in HRS-SR scores was greater than what would be expected by chance variation alone, and the temporal change in HRS-SR score (between adjacent time points) was considered clinically meaningful. Descriptive statistics were then used to determine the minimum change in HRS-SR score needed to denote a clinically meaningful change.

3. Results

In comparison to BHR participants who did not complete the *Hoarding and Clutter* module, a higher proportion of individuals included in this analysis were female (75.5% vs 71.0%, $X^2(1, N=70233) = 156.7, p < .0001$), white (88.3% vs 78.3%, $X^2(1, N=66514) = 1006.7, p < .0001$), college educated (71.8% vs 72.1%, $X^2(1, N=68913) = 888.5, p < .0001$), and of older age ($M=60.7, SD=13.0$ vs. $M=55.2, SD=14.6; t(65352) = -47.3, p < 0.0001$). A detailed comparison between the study sample and the total BHR population is outlined in Table 1.

Among 23,214 BHR participants who completed the Hoarding and Clutter elective module, HRS-SR scores ranged from 0 to 40, with a mean score of 3.86 ($SD=5.90$). For all questions, item scores ranged from 0 to 8, with mean scores between 0 and 1. For each item, more than 60% of the sample indicated no symptoms (score of 0). Using pre-existing cut-point criteria used for classifying clinically relevant hoarding behaviors that represent likely diagnoses of hoarding disorder: 1,606 participants (6.9%) had an HRS-SR total score greater than 14; 1,096 participants (4.7%) had an HRS-SR total score greater than 17; and 862 participants (3.7%) had HRS-SR component scores greater than or equal to 4 (moderate) for each of the following symptoms: clutter, difficulty discarding, and either emotional distress or impairment. Figure 1 shows the distribution of HRS-SR total scores in the entire BHR sample, and in subsamples 1 and 2.

3.1. Convergent Validity

Among 1,183 individuals included in subsample 1, mean HRS-SR scores for the control group were 0.14 ($SD=0.35, n=854$), scores for the intermediate group were 5.93 ($SD=4.09, n=153$), and scores for the case group were 21.47 ($SD=5.23, n=176$). Convergent validity between the internet-based HRS-SR and the SI-R ($r=0.887, p<0.0001$) and UHSS ($r=0.880, p<0.0001$) in subsample 1 ($n=1,183$) was high, indicating a strong relationship between measured constructs. Among those with a BE diagnosis of HD (i.e. subsample 2, $n=115$), convergent validity between the internet-based HRS-SR and the SI-R ($r=0.661, p<0.0001$) and UHSS ($r=0.592, p=0.0009$) was lower, although still significant. Among participants with a BE diagnosis of no HD ($N=115$), similar patterns emerged for the SI-R ($r=0.660, p<0.0001$), although convergent validity between the internet-based HRS-SR and UHSS was low and insignificant ($r=0.369, p=0.1942$). For individuals in subsample 2 ($n=230$),

convergent validity between the SI-R and UHSS was moderate for both those with ($r=0.626$ $p=0.0004$) and without HD ($r=0.562$ $p=0.0361$).

ROC curves (not shown) were plotted to assess the diagnostic performance of the internet-based HRS-SR using the self-reported and remotely-obtained Saving Inventory – Revised and UCLA Hoarding Severity Scale as the gold standard ($n=230$). Cutoff values of ()42 and ()20 were used to classify clinically significant hoarding behavior using the SI-R and UHSS, respectively. The empirical ROC curve plotted sensitivity against (1-specificity) for different cutoff values of total HRS-SR scores. Both models (e.g., using the SI-R and the UHSS cutoffs) suggested that sensitivity and specificity (accurate identification of true positives and negatives, respectively) were optimized using an HRS-SR total score cut-off of () 9 (SI-R sensitivity: 0.974, specificity: 0.933 Youden's J: 0.907; UHSS sensitivity: 0.962, specificity: 0.923 Youden's J: 0.884). AUC values indicate near perfect model accuracy (SI-R AUC: 0.979; UHSS AUC: 0.974).

3.2 Validity of HRS-SR scores

Of 230 interviewees, 108 (47.0%) received a BE consensus of “definite”, 7 (3.0%) received a BE consensus of “probable”, and 115 (50.0%) received a BE consensus of “not present.” For the purposes of this analysis, ratings of “probable” and “definite” were reclassified as one category. Two (0.9%) controls received a best estimate diagnosis of hoarding disorder based on the clinical interview and SI-R and UHSS data; both individuals scored between 0 and 2 on the web-based HRS-SR. Eight (6.6%) cases and one intermediate (100.0%) received a best estimate diagnosis of no HD. For hoarding diagnoses, inter-rater agreement was high ($\kappa=0.98$, 95% CI= [0.96, 1.00]). Two participants required three best estimates – in both cases, consensus was achieved with agreement between the second and third best estimators. Those with a best estimate diagnosis of HD scored significantly higher on the HRS-SR ($M=18.5$, $SD=7.2$ vs. $M=1.4$, $SD=3.9$; $t(228) = -22.2$, $p < 0.0001$), SI-R ($M=47.5$, $SD=12.9$ vs. $M=9.3$, $SD=8.9$; $t(227) = -26.1$, $p < 0.0001$), and UHSS ($M=21.1$, $SD=6.3$ vs. $M=4.1$, $SD=4.4$; $t(227) = -23.9$, $p < 0.0001$) than those without. The distribution of scores for individual HRS-SR items among those with and without a BE HD diagnosis are depicted in Figure 2.

ROC curves were plotted to assess the diagnostic performance of the internet-based HRS-SR using in-clinic best estimate consensus HD diagnoses as the gold standard (Figure 3). The model suggested that sensitivity and specificity differentiating those with HD from those with no HD were optimized using an HRS-SR total score cut-off value of () 5 (sensitivity=0.97, specificity=0.93; Table 2). The area under the curve (AUC: 0.972) indicates near perfect model accuracy. In the ten-fold cross validation, all 10 models showed similar results and suggested that sensitivity and specificity were optimized using an HRS-SR total score cut-off value of () 5. Use of this cutoff in the testing samples yielded high predictive ability, with one relative model outlier that nevertheless correctly classified the majority of participants with a sensitivity of 90% and specificity of 85% at a cutoff of () 5 (data not shown).

Cohen's kappa was used to examine the agreement between a best estimate lifetime diagnosis of HD and whether an individual met criteria for HD based on a summation of

HRS-SR scores for the three items that assess the core DSM-5 HD criteria: clutter, difficulty discarding, and impairment or distress (Table 3). As excessive acquisition is not a core DSM-5 criterion for HD, it was not included in this analysis. As the cut-points increased, Kappa values, sensitivity, and negative predictive value decreased, while specificity and positive predictive value (PPV) remained unchanged. Using these criteria, sensitivity, negative predictive value, and kappa values were optimized when HD was classified using a summation of HRS-SR component scores indicating mild or greater symptomatology (i.e. reporting clutter, difficulty discarding, and either emotional distress or impairment scores 1 (total component score of 3). Specificity and positive predictive value were optimized using component scores of 3 (total component score of 9).

3.3 Reliability of HRS-SR Scores

Of 230 participants with clinical interviews, 197 (85.7%) completed the internet-based HRS-SR at multiple time points. Figure 4 depicts the longitudinal trajectory of HRS-SR scores for each participant, stratified by the number of time points completed. For most participants, scores remained consistent across time points, irrespective of best estimate HD. Pearson's correlation coefficient confirmed that HRS-SR scores were highly correlated between adjacent time points (Time 1 and 2: $r=0.88$; Time 2 and 3: $r=0.91$; Time 3 and 4: $r=0.88$; Time 4 and 5: $r=0.90$; Time 5 and 6: $r=0.92$).

A temporal mean of HRS-SR scores was computed for each individual. The majority of those without BE HD had an average HRS-SR score less than 5 (93.9%) and nearly all participants with HD (98.3%) had an average score greater than or equal to 5, indicating that a mean cut-off value of () 5 reasonably differentiated those with and without HD.

Among those who completed the HRS-SR at multiple time points, ROC curves using means of all available HRS-SR scores (rather than most recent score) to predict best estimate consensus HD were plotted. The model suggested that sensitivity and specificity were optimized using an HRS-SR total score cut-off value of () 6 (sensitivity=0.98, specificity=0.95; AUC: 0.973).

A total of 797 observations from the 197 participants with multiple time points were next used to evaluate meaningful change in HRS-SR scores over time (567 total time point pairs; i.e. time points 1 and 2, time points 2 and 3, etc.). The reliable change index suggested that a 4-point increase or decrease in HRS-SR score between subsequent time points indicated meaningful change in hoarding symptoms. Meaningful change in HRS-SR scores was observed between 30.6% of subsequent time point pairs ($n=168$). However, only 5.5% ($n=11$) of participants had HRS-SR scores that crossed the cut-off score of 5 suggested by the ROC curve analysis. Even fewer participants (2.5%, $n=5$) had total scores that increased (or decreased) from ($<$) 5 to () 10 (or vice versa) (that is, from unaffected to likely affected, based on the cutoffs identified from the ROC analysis in the clinical sample and in the larger subsample with additional hoarding data, respectively).

4. Discussion

The findings of this analysis indicate that the internet-based HRS-SR is a valid and temporally stable diagnostic assessment of hoarding behavior and a useful instrument for evaluating the presence of clinically relevant hoarding symptoms (e.g., likely hoarding disorder) in comparison to diagnoses from clinicians and other self-report assessments. Our results indicate high concordance between lifetime HD diagnoses as determined by clinical interviews, other self-report measures of hoarding symptoms, and the internet-based HRS-SR.

Our primary aim was to validate an internet-based and remotely-administered version of the HRS-SR by determining the optimal scoring method and cutoff score for identifying clinically relevant hoarding behavior. Our findings indicate that assessment of hoarding symptoms using the internet-based HRS-SR may be most accurate when HRS-SR total score, rather than component subscales, is used to quantify symptom presence and severity. In particular, the internet-based HRS-SR differentiated individuals with and without a lifetime diagnosis of HD with optimal sensitivity and specificity using a cut-off value of 5 on the total score measure. We note that this cut-off value is much lower than what has been suggested in previous HRS-SR validation studies (Tolin et al., 2010; Tolin et al., 2018), and may be most useful for identifying individuals who do not have hoarding disorder, rather than identifying those who do. The low cutoff value in the clinical sample is likely due to the fact that only one participant who underwent clinical assessment had a total score in the mid-range of the Hoarding Rating Scale (between 1 and 14), creating a nearly bimodal distribution of HRS-SR scores and limiting our ability to determine with confidence the relationship between scores in the 1–14 range and lifetime hoarding diagnoses as clinically assessed. It is likely that the ROC analysis conducted in the clinical sample is thus not optimal for distinguishing individuals with lifetime HD from those with mild, subclinical hoarding behaviors in the intermediate range of scores. Until further work is done to determine the relationship between mid-range HRS-SR scores and clinical diagnoses of HD, we suggest using a cut-off value of ()5, as identified in this study, to distinguish individuals who are not likely to have HD (e.g., those that score 5 on the HRS-SR) rather than to identify likely HD.

The results of our other analyses, specifically the DSM-5 based scoring and the ROC curves using SI-R and UHSS cutoffs as the gold standard, are perhaps more useful for identifying HD, and point to a cutoff score of 10 ()9. We suggest that this cutoff score is more appropriate than the lower cutoff for use in identifying individuals who have clinically significant hoarding symptoms, and are likely to meet clinical criteria for HD. However, given the limitations discussed above, we suggest that a conservative approach would be to use this cutoff to determine clinically significant hoarding or subclinical/probable HD, and to continue to use the previously validated cutoff of 14 to identify definite HD, until further work is done to evaluate the psychometric properties of the internet-based HRS-SR using a more robust range of initial scores. A cutoff of 14 has been previously recommended by Tolin and colleagues (2010) and has been used in a number of genetic and epidemiological studies to successfully differentiate individuals with HD from those with OCD and community controls (Cath et al., 2017; Iervolino et al., 2009; Monzani et al., 2014; Iervolino

et al., 2011; Ivanov et al., 2013; Perroud et al., 2011; Frost et al., 2011; Frost & Hristova, 2011). Examination of the prevalence of likely HD in the BHR sample also supports this suggestion. Using this cut-off, the population prevalence of HD among 23,214 BHR participants was 6.9%. Given the older mean age of the sample, and the potential for selection bias which accompanies the elective nature of questionnaires in the BHR (i.e. the Hoarding and Clutter module may be of particular interest to BHR participants with hoarding behavior), we would expect the population prevalence of HD to be slightly greater than 6% (as observed in previous epidemiological investigation of hoarding symptoms among older adults; Cath et al., 2017). Accordingly, a cutoff value of () 14 to distinguish clinically significant hoarding behavior seems appropriate, while a cutoff of 10 () 9 gives a population prevalence for HD of 13.3% in this sample, which is higher than has been reported in previous samples, again suggesting that this cutoff, while useful for identifying problematic hoarding, may in fact be too lenient to identify true HD with certainty.

Our second aim was to assess the temporal stability of internet-based HRS-SR scores. Our findings indicate that the majority of participants have HRS-SR scores remain consistent between time points (~6 months), irrespective of best estimate HD diagnosis. Meaningful variation in HRS-SR scores was suggested to occur when participants experienced a 4-point increase or decrease in total score between adjacent time points. We found that participants' HRS-SR scores changed by 4 points or more approximately 30% of the time, suggesting a meaningful and detectable change in hoarding symptom severity. However, the meaning of a 4 point change in scores is quite different when starting from a low score (e.g., a change from 5 to 9 or 10), than when starting from a higher score (e.g., a change from 15 to 19 or 20), particularly in light of the strong right skew of the scores in the total sample. In the former case, it could be argued that a change of 4 points indicates that the respondent may be at risk for developing symptoms of hoarding disorder, while in the latter, a change of 4 points most likely indicates a relative worsening of symptoms in someone who is currently suffering from hoarding disorder. We note that scores crossed a threshold from unaffected to possibly affected (i.e., from <5 to >5) or from unaffected to likely affected (i.e., from <5 to >9) only a very small proportion of the time (5.5% and 2.5%, respectively). Though future studies are needed to assess the validity of reliable change indices and evaluate attributes associated with reliable deterioration or improvement in hoarding behavior, this index, when used in combination with the cutoff scores identified here, provides a baseline for progress monitoring in population-based samples.

Our final aim was to assess the convergent validity of the internet-based HRS-SR in comparison to well-established, self-report measures of hoarding behavior. Our findings suggested a strong relationship between the HRS-SR and SI-R, as well as the HRS-SR and the UHSS. Given the similarity between the three measures of hoarding behavior and the brevity of the HRS-SR, researchers may prefer to use the HRS-SR in remote assessment of HD, where time limitations or participant burden may be of increased concern.

4.1. Limitations

The present analysis is limited by use of a fairly homogenous sample predominated by white, educated females of older age. Internet-based data collection methods may preclude

individuals without reliable computer or internet access from participating in this study. Additionally, as a result of our participant selection process, a comparatively small number of participants who underwent additional clinical assessment had a total score in the mid-range of the Hoarding Rating Scale (between 1 and 14); findings may not represent the diverse spectrum of individuals endorsing hoarding behavior and may not adequately distinguish individuals with lifetime HD from those with mild, subclinical hoarding behaviors when considering a more diverse ranges of scores.

Furthermore, clinical assessment procedures limited the ability to objectively assess hoarding symptoms through outside observation and consider alternative psychiatric or medical diagnoses. Thus, the internet-based HRS-SR is sensitive for identifying hoarding symptoms that are distressing or impairing, but cannot substitute for a clinical diagnosis to exclude other causes of hoarding behavior.

Lastly, we were unable to delineate the exact duration of time between completion of the HRS-SR, SI-R/UHSS, and clinical interviews. Though it is known that all clinical interviews were conducted within one year of completing the Hoarding and Clutter module, a small number of participants may have repeated the online assessment after the clinical interview was conducted. However, our findings indicate that HRS-SR scores are temporally stable and that recommended cut-off criteria are comparable when predicting HD using either a subject's most recent or temporal mean score. Accordingly, it is unlikely that assessment chronology introduced bias.

4.2. Conclusions

These analyses suggest that the internet-based HRS-SR is a useful, temporally stable, and valid assessment of hoarding symptoms. In this sample, an HRS-SR total score of 5 or greater was both sensitive and specific for differentiating individuals who met clinical DSM-5 criteria for lifetime hoarding disorder from those who did not. As this cut-off value is much lower than what has been suggested in previous HRS-SR validation studies, and our additional analyses identified a cutoff score of 9, we suggest that HRS-SR scores of 5 or lower may be best thought of as an indication of exclusion of likely HD, whereas scores of 10 or higher may indicate those with likely HD. However, until additional research is done to confirm the relationship between clinical diagnosis and HRS-SR scores in the mid-range (e.g., between 10–14), we recommend that if clinical interviews are not an option for assigning HD diagnoses, researchers follow more conservative guidelines, and use HRS-SR cutoff scores of 14 or higher to identify individuals with likely HD and scores of 10 or higher to identify individuals with subclinical symptoms or those who are at-risk for HD. We also note that work still remains to be done to determine the relationship between HRS-SR scores and other causes of hoarding behavior, including dementia, mild cognitive impairment, and other psychiatric illnesses.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding Statement: This work was supported by the National Institutes of Health (grant number R01MH117114).

Conflicts of 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. Dr. Mackin has received research funding from the National Institutes of Health and Janssen Research and Development LLC.

References

- American Psychiatric Association. 2013. Diagnostic and statistical manual of mental disorders, 5th ed. Arlington, VA.
- Archer CA, Moran K, Garza K, Zakrzewski JJ, Martin A, Chou CY, Uhm SY, Chan J, Gause M, Salazar M and Plumadore J, 2019. Relationship between symptom severity, psychiatric comorbidity, social/occupational impairment, and suicidality in hoarding disorder. *Journal of Obsessive-Compulsive and Related Disorders*, 21,158–164.
- Ayers CR, Saxena S, Golshan S and Wetherell JL, 2010. Age at onset and clinical features of late life compulsive hoarding. *International Journal of Geriatric Psychiatry: A journal of the psychiatry of late life and allied sciences*, 25(2), 142–149.
- Brain Health Registry (BHR). 2012–2019. <http://www.brainhealthregistry.org> (accessed 12 March 2020).
- Cath DC, Nizar K, Boomsma D and Mathews CA, 2017. Age-specific prevalence of hoarding and obsessive compulsive disorder: a population-based study. *The American Journal of Geriatric Psychiatry*, 25(3), 245–255. [PubMed: 27939851]
- Chakraborty V, Cherian AV, Math SB, Venkatasubramanian G, Thennarasu K, Mataix-Cols D and Reddy YJ, 2012. Clinically significant hoarding in obsessive-compulsive disorder: results from an Indian study. *Comprehensive psychiatry*, 53(8), 1153–1160. [PubMed: 22796017]
- Chasson GS, Guy AA, Bates S and Corrigan PW, 2018. They aren't like me, they are bad, and they are to blame: A theoretically-informed study of stigma of hoarding disorder and obsessive-compulsive disorder. *Journal of obsessive-compulsive and related disorders*, 16, 56–65.
- Chou CY, Mackin RS, Delucchi KL and Mathews CA, 2018. Detail-oriented visual processing style: Its role in the relationships between early life adversity and hoarding-related dysfunctions. *Psychiatry research*, 267, 30–36. [PubMed: 29883858]
- Faraci P, Perdighe C, Del Monte C and Salianni AM, 2018. Reliability, validity and factor structure of the Compulsive Acquisition Scale (CAS). *Clinical Neuropsychiatry: Journal of Treatment Evaluation*, 15(1), 42–49.
- Frost RO and Hristova V, 2011. Assessment of hoarding. *Journal of clinical psychology*, 67(5), 456–466. [PubMed: 21351103]
- Frost RO, Hristova V, Steketee G and Tolin DF, 2013. Activities of daily living scale in hoarding disorder. *Journal of obsessive-compulsive and related disorders*, 2(2), 85–90. [PubMed: 23482436]
- Frost RO, Steketee G and Grisham J, 2004. Measurement of compulsive hoarding: saving inventory-revised. *Behaviour research and therapy*, 42(10), 1163–1182. [PubMed: 15350856]
- Frost RO, Steketee G and Tolin DF, 2015. Comorbidity in hoarding disorder. *Focus*, 13(2), 244–251.
- Frost RO, Tolin DF, Steketee G and Oh M, 2011. Indecisiveness and hoarding. *International Journal of Cognitive Therapy*, 4(3), 253–262.
- Frost RO, Tolin DF, Steketee G, Fitch KE and Selbo-Bruns A, 2009. Excessive acquisition in hoarding. *Journal of anxiety disorders*, 23(5), 632–639. [PubMed: 19261435]
- Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, Heninger GR and Charney DS, 1989. The Yale-Brown obsessive compulsive scale: I. Development, use, and reliability. *Archives of general psychiatry*, 46(11), 1006–1011. [PubMed: 2684084]
- Grisham JR, Frost RO, Steketee G, Kim HJ and Hood S, 2006. Age of onset of compulsive hoarding. *Journal of anxiety disorders*, 20(5), 675–686. [PubMed: 16112837]

- Iervolino AC, Perroud N, Fullana MA, Guipponi M, Cherkas L, Collier DA and Mataix-Cols D, 2009. Prevalence and heritability of compulsive hoarding: a twin study. *American Journal of Psychiatry*, 166(10), 1156–1161.
- Iervolino AC, Rijdsdijk FV, Cherkas L, Fullana MA and Mataix-Cols D, 2011. A multivariate twin study of obsessive-compulsive symptom dimensions. *Archives of general psychiatry*, 68(6), 637–644. [PubMed: 21646580]
- Ivanov VZ, Mataix-Cols D, Serlachius E, Lichtenstein P, Anckarsäter H, Chang Z, Gumpert CH, Lundström S, Långström N and Rück C, 2013. Prevalence, comorbidity and heritability of hoarding symptoms in adolescence: a population based twin study in 15-year olds. *PloS one*, 8(7).
- Jacobson NS and Traux P, 1991. Clinical significance: A statistical approach to defining the meaningful change in psychotherapy research. *J Consult Clin Psychol*, 59(1), 12–19. [PubMed: 2002127]
- Kellman-McFarlane K, Stewart B, Woody S, Ayers C, Dozier M, Frost RO, Grisham J, Isemann S, Steketee G, Tolin DF and Welsted A, 2019. Saving inventory–Revised: Psychometric performance across the lifespan. *Journal of affective disorders*, 252, 358–364. [PubMed: 30999092]
- Mackin RS, Insel PS, Truran D, Finley S, Flenniken D, Nosheny R, Ulbright A, Comacho M, Bickford D, Harel B and Maruff P, 2018. Unsupervised online neuropsychological test performance for individuals with mild cognitive impairment and dementia: results from the Brain Health Registry. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 10, 573–582.
- Mataix-Cols D and Fernández de la Cruz L, 2018. Hoarding disorder has finally arrived, but many challenges lie ahead. *World Psychiatry*, 17(2), 224–225. [PubMed: 29856565]
- Mathews CA, Delucchi K, Cath DC, Willemsen G and Boomsma DI, 2014. Partitioning the etiology of hoarding and obsessive-compulsive symptoms. *Psychological medicine*, 44(13), 2867–2876. [PubMed: 25066062]
- Matsunaga H, Hayashida K, Kiriike N, Nagata T, & Stein DJ (2010). Clinical features and treatment characteristics of compulsive hoarding in Japanese patients with obsessive-compulsive disorder. *Cns Spectrums*, 15(4), 258–266. [PubMed: 20414175]
- McHugh ML, 2012. Interrater reliability: the kappa statistic. *Biochemia medica: Biochemia medica*, 22(3), 276–282. [PubMed: 23092060]
- Monzani B, Rijdsdijk F, Harris J and Mataix-Cols D, 2014. The structure of genetic and environmental risk factors for dimensional representations of DSM-5 obsessive-compulsive spectrum disorders. *JAMA psychiatry*, 71(2), 182–189. [PubMed: 24369376]
- Mueller A, Mitchell JE, Crosby RD, Glaesmer H, & de Zwaan M 2009. The prevalence of compulsive hoarding and its association with compulsive buying in a German population-based sample. *Behaviour Research and Therapy*, 47(8), 705–709. [PubMed: 19457476]
- Mukaka M, 2012. Statistics Corner: A guide to appropriate use of. *Malawi Medical Journal*, 24(3), 69–71. [PubMed: 23638278]
- Nordsletten AE, de la Cruz LF, Pertusa A, Reichenberg A, Hatch SL and Mataix-Cols D, 2013. The structured interview for hoarding disorder (SIHD): development, usage and further validation. *Journal of Obsessive-Compulsive and Related Disorders*, 2(3), 346–350.
- Perroud N, Guipponi M, Pertusa A, Fullana MA, Iervolino AC, Cherkas L, Spector T, Collier D and Mataix-Cols D, 2011. Genome-wide association study of hoarding traits. *American Journal of Medical Genetics, Part B*, 156(2), 240.
- Rodriguez CI, Herman D, Alcon J, Chen S, Tannen A, Essock S and Simpson HB, 2012. Prevalence of hoarding disorder in individuals at potential risk of eviction in New York City: a pilot study. *The Journal of nervous and mental disease*, 200(1).
- Rosa AC, Diniz JB, Fossaluza V, Torres AR, Fontenelle LF, De Mathis AS, da Conceição Rosário M, Miguel EC and Shavitt RG, 2012. Clinical correlates of social adjustment in patients with obsessive-compulsive disorder. *Journal of psychiatric research*, 46(10), 1286–1292. [PubMed: 22800713]
- Ruopp MD, Perkins NJ, Whitcomb BW and Schisterman EF, 2008. Youden Index and optimal cut-point estimated from observations affected by a lower limit of detection. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 50(3), 419–430.

- Samuels JF, Bienvenu OJ, Grados MA, Cullen B, Riddle MA, Liang KY, ... & Nestadt G 2008. Prevalence and correlates of hoarding behavior in a community-based sample. *Behaviour research and therapy*, 46(7), 836–844. [PubMed: 18495084]
- Saxena S, Ayers CR, Dozier ME and Maidment KM, 2015. The UCLA hoarding severity scale: development and validation. *Journal of affective disorders*, 175, 488–493. [PubMed: 25681559]
- Saxena S, Brody AL, Maidment KM and Baxter LR Jr, 2007. Paroxetine treatment of compulsive hoarding. *Journal of psychiatric research*, 41(6), 481–487. [PubMed: 16790250]
- Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R and Dunbar GC, 1998. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of clinical psychiatry*.
- Tolin DF, Frost RO and Steketee G, 2010. A brief interview for assessing compulsive hoarding: the Hoarding Rating Scale-Interview. *Psychiatry research*, 178(1), 147–152. [PubMed: 20452042]
- Tolin DF, Frost RO, Steketee G and Fitch KE, 2008. Family burden of compulsive hoarding: results of an internet survey. *Behaviour research and therapy*, 46(3), 334–344. [PubMed: 18275935]
- Tolin DF, Frost RO, Steketee G, Gray KD and Fitch KE, 2008. The economic and social burden of compulsive hoarding. *Psychiatry research*, 160(2), 200–211. [PubMed: 18597855]
- Tolin DF, Gilliam CM, Davis E, Springer K, Levy HC, Frost RO, Steketee G and Stevens MC, 2018. Psychometric properties of the hoarding rating scale-interview. *Journal of obsessive-compulsive and related disorders*, 16, 76–80 [PubMed: 31544015]
- Turna J, Patterson B, Simpson W, Pullia K, Khalesi Z, Kaplan KG and Van Ameringen M, 2018. Prevalence of hoarding behaviours and excessive acquisition in users of online classified advertisements. *Psychiatry research*, 270, 194–197. [PubMed: 30261409]
- Vorstenbosch V, Hood HK, Rogojanski J, Antony MM, Summerfeldt LJ and McCabe RE, 2012. Exploring the relationship between OCD symptom subtypes and domains of functional impairment. *Journal of Obsessive-Compulsive and Related Disorders*, 1(1), 33–40.
- Weiner MW, Nosheny R, Camacho M, Truran- Sacrey D, Mackin RS, Flenniken D, Ulbricht A, Insel P, Finley S, Fockler J and Veitch D, 2018. The Brain Health Registry: An internet- based platform for recruitment, assessment, and longitudinal monitoring of participants for neuroscience studies. *Alzheimer's & Dementia*, 14(8), 1063–1076.
- Wootton BM, Diefenbach GJ, Bragdon LB, Steketee G, Frost RO and Tolin DF, 2015. A contemporary psychometric evaluation of the Obsessive Compulsive Inventory—Revised (OCI-R). *Psychological assessment*, 27(3), 874. [PubMed: 25664634]
- Zilhão NR, Smit DJ, Boomsma DI and Cath DC, 2016. Cross-disorder genetic analysis of tic disorders, obsessive-compulsive, and hoarding symptoms. *Frontiers in psychiatry*, 7, 120. [PubMed: 27445875]

Highlights

- The internet-based Hoarding Rating Scale is a valid measure of hoarding symptoms.
- Hoarding scores are relatively stable between adjacent time points.
- Scores of 5 indicate low-risk of clinical or subclinical hoarding

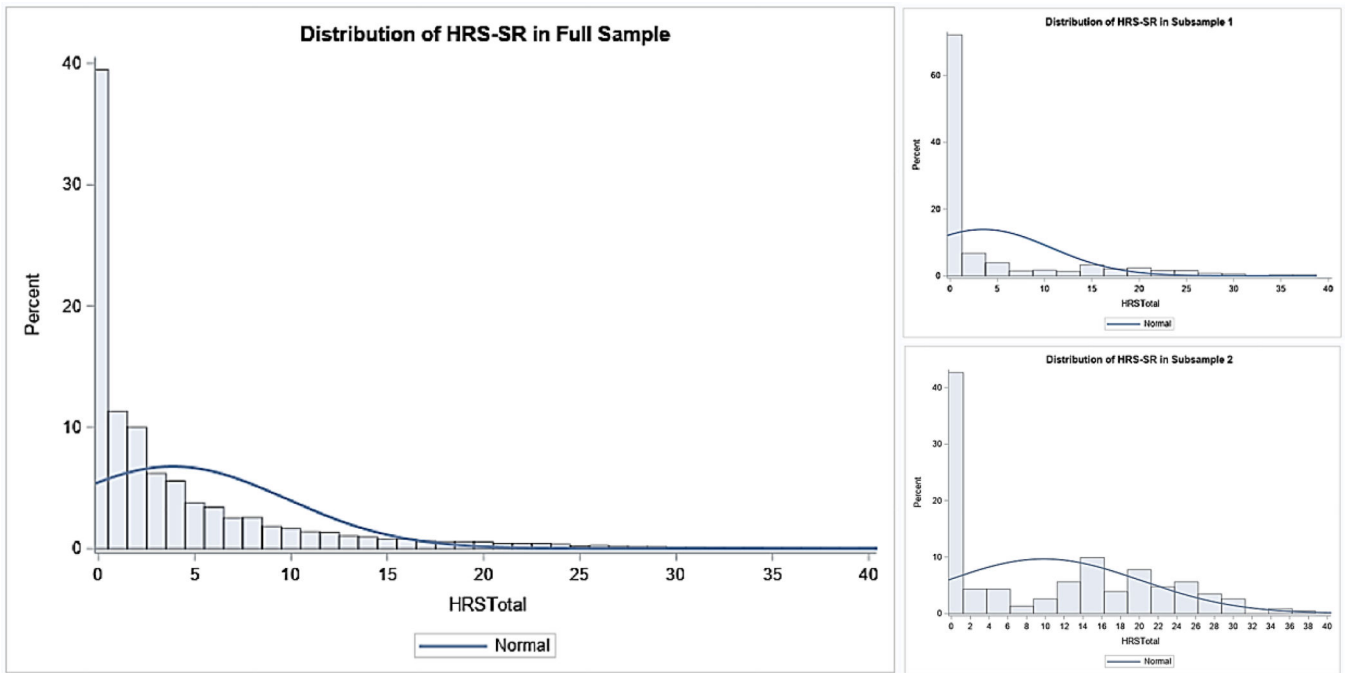


Figure 1:
Distribution of HRS-SR total scores among BHR participants, sub-sample 1, and sub-sample 2

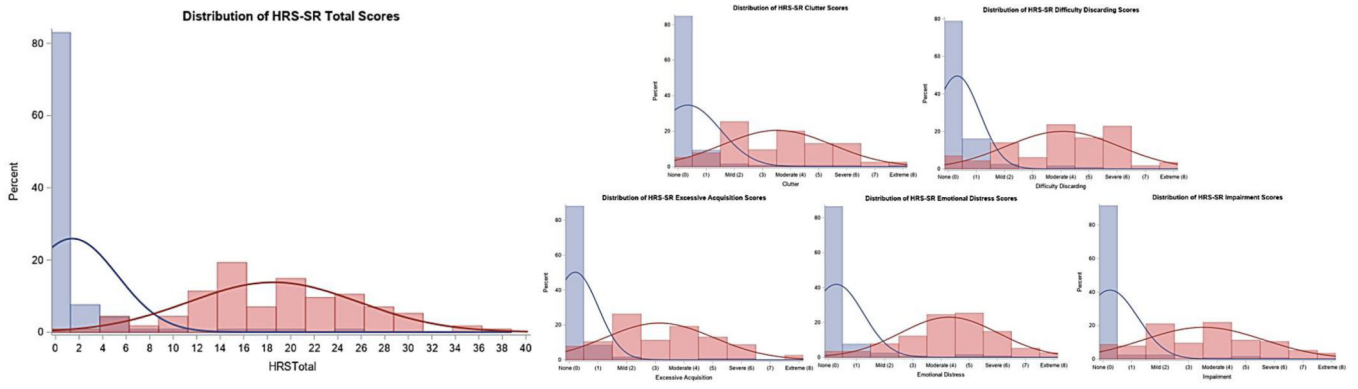


Figure 2:
Distribution of HRS-SR scores, by best estimate hoarding disorder diagnosis
Red histograms represent individuals with best estimate HD and blue histograms represent individuals without best estimate HD

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

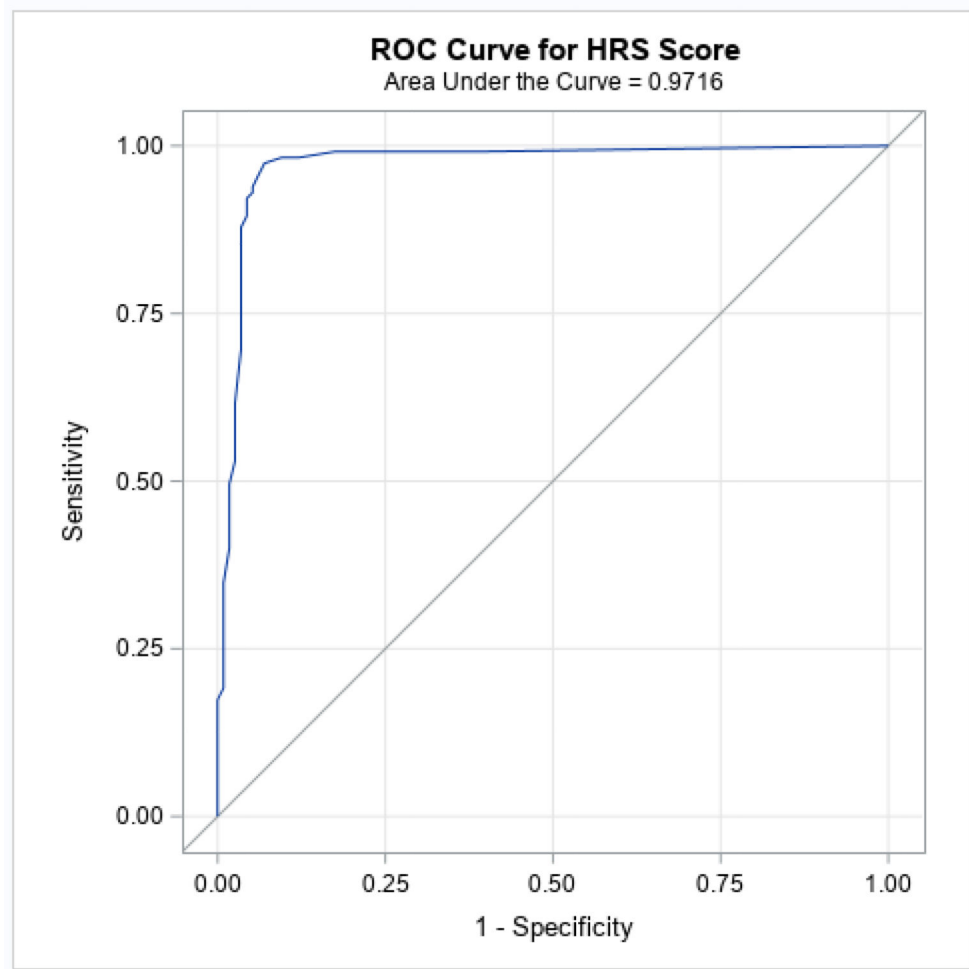


Figure 3:
ROC curve for different cutoff values of total HRS-SR scores predicting BE HD

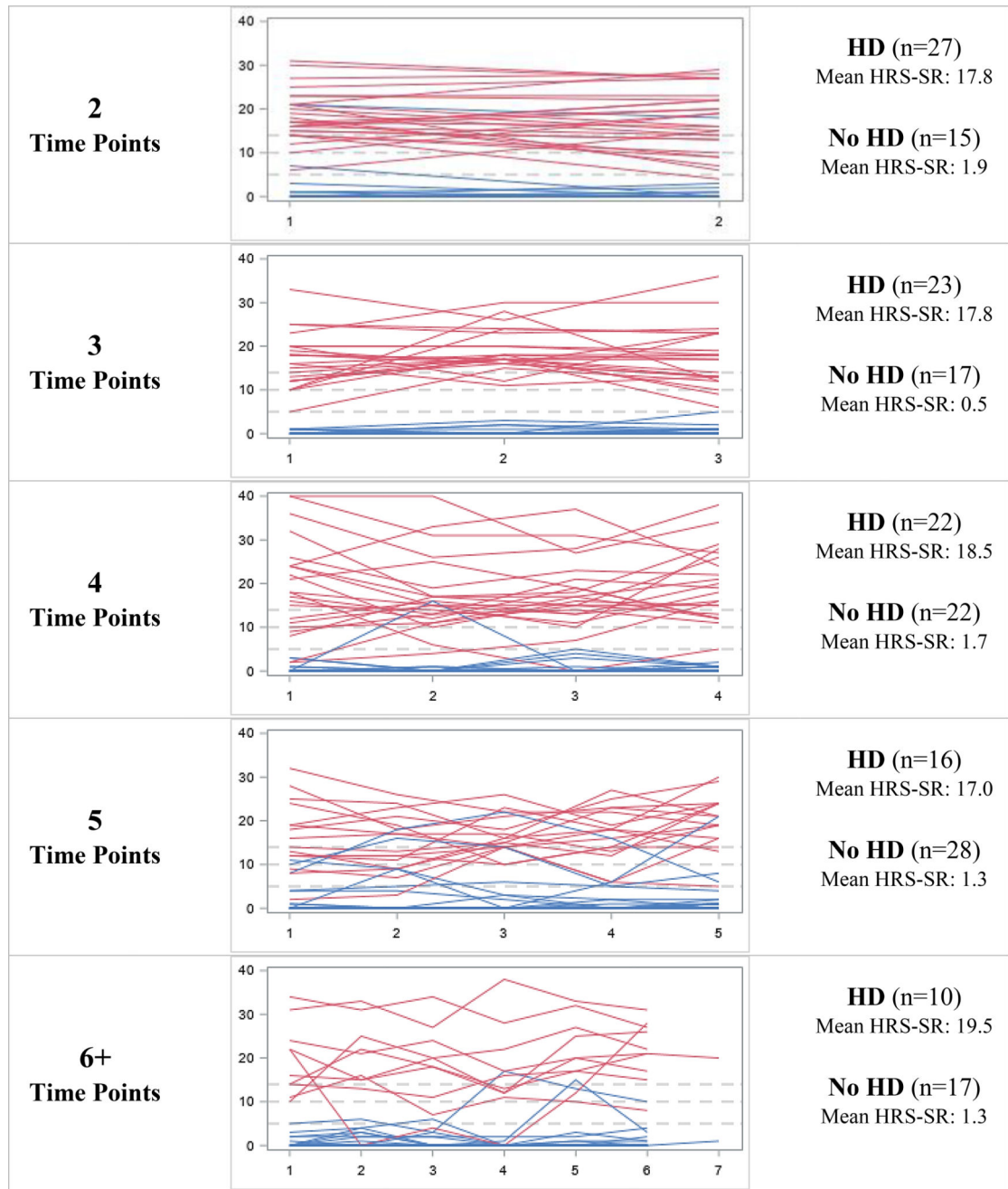


Figure 4:
 Longitudinal trajectory of HRS-SR scores, by number of time points completed and best estimate hoarding disorder diagnosis
 Red lines represent individuals with best estimate HD and blue lines represent individuals without best estimate HD

Table 1:

Characteristics of BHR Participants Completing the Hoarding and Clutter Module

	Completed Hoarding and Clutter Module N=23,214 (%)	Did Not Complete Hoarding and Clutter Module N=47,493 (%)	p	Effect Size
Gender ^a			<0.0001	0.05
Male	5694 (24.5)	13642 (29.0)		
Female	17520 (75.5)	33377 (71.0)		
Race ^b			<0.0001	-0.12
White	20311 (88.3)	24043 (78.3)		
Other	2701 (11.7)	9454 (21.7)		
Education ^c			<0.0001	0.11
Less than College	4353 (18.8)	12765 (27.9)		
College	9258 (39.9)	18216 (39.9)		
Graduate/Professional	9603 (41.4)	14718 (32.2)		
Age (mean, sd) ^d	60.7 (13.0)	55.2 (14.6)	<0.0001	0.40

Effect size was calculated using Cohen's d, Cramer's ϕ or Cramer's V method of effect size, as appropriate.

^aDid Not Complete Hoarding and Clutter Module N=47,019

^bCompleted Hoarding and Clutter Module N=23,012; Did Not Complete Hoarding and Clutter Module N=43,502

^cDid Not Complete Hoarding and Clutter Module N=45,699

^dCompleted Hoarding and Clutter Module N=22,242; Did Not Complete Hoarding and Clutter Module N=43,112

Table 2:

Sensitivity and specificity for different cutoff values of total HRS-SR scores

HRS-SR Cut-off	Sensitivity	Specificity	Youden's J
1	0.99	0.62	0.61
2	0.99	0.83	0.82
3	0.98	0.88	0.86
4	0.98	0.90	0.89
5	0.97	0.93	0.90
6	0.96	0.94	0.90
7	0.94	0.95	0.89
8	0.93	0.95	0.88
9	0.92	0.96	0.88
10	0.90	0.96	0.85
11	0.88	0.97	0.84
12	0.87	0.97	0.84
13	0.83	0.97	0.79
14	0.76	0.97	0.72
15	0.70	0.97	0.66
16	0.63	0.97	0.59
17	0.57	0.97	0.54
18	0.53	0.97	0.50
19	0.50	0.98	0.48
20	0.43	0.98	0.42

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3:

Sensitivity and specificity for different cutoff values of HRS-SR DSM-5 proxy

	Sensitivity	Specificity	PPV	NPV	Kappa
Clutter, difficulty discarding, and either emotional distress or impairment scores 1	89.6	95.7	95.4	90.2	0.90 (0.84, 0.96)
Clutter, difficulty discarding, and either emotional distress or impairment scores 2	77.4	98.3	97.8	81.3	0.78 (0.71, 0.86)
Clutter, difficulty discarding, and either emotional distress or impairment scores 3	48.7	99.1	98.25	65.9	0.49 (0.39, 0.59)
Clutter, difficulty discarding, and either emotional distress or impairment scores 4	38.3	99.1	97.8	61.6	0.39 (0.29, 0.48)
Clutter, difficulty discarding, and either emotional distress or impairment scores 5	15.7	99.1	94.7	54.0	0.16 (0.09, 0.23)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript