Fall 4-30-2023

Logistic Regression With Machine Learning Sheds Light on the Problematic Sexual Behavior Phenotype

Shui Jiang

Keanna Wallace

Esther Yang

Leslie Roper

Garima Aryal

See next page for additional authors

OPEN

Logistic Regression With Machine Learning Sheds Light on the Problematic Sexual Behavior Phenotype

Shui Jiang, MSc, Keanna Wallace, BSc, Esther Yang, MSc, Leslie Roper, MSc, Garima Aryal, MSc, Dawon Lee, MD, Rohit J. Lodhi, MD, PhD, MRCPsych, FRANZCP, Randolph Arnau, PhD, Rick Isenberg, MD, Bradley Green, PhD, David Wishart, PhD, and Katherine J. Aitchison, BM BCh, PhD, FRCPsych

Objectives: There has been a longstanding debate about whether the mechanisms involved in problematic sexual behavior (PSB) are similar to those observed in addictive disorders, or related to impulse control or to compulsivity. The aim of this report was to contribute to this debate by investigating the association between PSB, addictive disorders (internet addiction, compulsive buying), measures associated with the construct known as reward deficiency (RDS), and obsessive-compulsive disorder (OCD).

Methods: A Canadian university Office of the Registrar invited 68,846 eligible students and postdoctoral fellows. Of 4710 expressing interest in participating, 3359 completed online questionnaires, and 1801 completed the Mini-International Neuropsychiatric Interview. PSB was measured by combining those screening positive (score at least 6) on the Sexual Addiction Screening Test—Revised Core with those self-reporting PSB. Current mental health condition(s) and childhood trauma were measured by self-report. OCD was assessed by a combination of self-report and Mini-International Neuropsychiatric Interview data.

Results: Of 3341 participants, 407 (12.18%) screened positive on the Sexual Addiction Screening Test—Revised Core. On logistic regression, OCD, attention deficit, internet addiction, a family history of PSB, childhood trauma, compulsive buying, and male gender were associated with PSB. On multiple correspondence analysis, OCD appeared to cluster separately from the other measures, and the pattern of data differed by gender.

Conclusions: In our sample, factors that have previously been associated with RDS and OCD are both associated with increased odds of PSB. The factors associated with RDS appear to contribute to a separate data cluster from OCD and to lie closer to PSB.

Key Words: addictive disorders, compulsivity, reward deficiency syndrome, obsessive-compulsive disorder, logistic models

(*J Addict Med* 2023;17: 174–181)

A theory of hypersexuality with dependence was first proposed by Orford[1] in 1978. The term sex addiction was used by Carnes[2] in 1983, 1990, and 1991 and by Goodman[3] in 1998. Continuation of a sexual behavior despite adverse consequences and/or distress caused or worsened by the sexual behavior has been a consistent theme in the literature,[1,3,4] despite differences of opinion as to whether the condition should be classified as an addiction or as an impulse control disorder. Diagnostic criteria proposed by Carnes[4] in 2005 led to sex addiction being included in the *Diagnostic and Statistical Manual of Mental Disorders* (Third Edition Revised) (302.87). However,

From the Department of Medical Genetics, Faculty of Medicine and Dentistry, College of Health Sciences, University of Alberta, Edmonton, AB, Canada (SJ, KW, EY, LR, GA, DL, KJA); Department of Psychiatry, Faculty of Medicine and Dentistry, College of Health Sciences, University of Alberta, Edmonton, AB, Canada (KW, EY, LR, GA, DL, RJL, KJA); Department of Emergency Medicine, Faculty of Medicine and Dentistry, College of Health Sciences, University of Alberta, Edmonton, AB, Canada (EY); Department of Medicine, Faculty of Medicine and Dentistry, College of Health Sciences, University of Alberta, Edmonton, AB, Canada (DL); Department of Psychiatry, University of Western Ontario, London, ON,Canada (RJL); Department of Psychiatry, Schulich School of Medicine & Dentistry, Western University, London, ON, Canada (RJL); Department of Psychiatry, College of Medicine, University of Saskatchewan, Saskatoon, Saskatchewan, Canada (RJL); School of Psychology, University of Southern Mississippi, Hattiesburg, MS (RA); Psychological Counselling Services, Ltd., Scottsdale, AZ (RI); Department of Psychology and Counseling, University of Texas at Tyler, Tyler, TX (BG); Department of Biological Sciences, Faculty of Science, College of Natural and Applied Sciences, University of Alberta, Edmonton, AB, Canada (DW); Neuroscience and Mental Health Institute, University of Alberta, Edmonton, AB, Canada (KJA); Women and Children's Health Research Institute, University of Alberta, Edmonton, AB, Canada (KJA); and Division of Clinical Sciences, Psychiatry Section, Northern Ontario School of Medicine, Sudbury, ON, Canada (KJA).

Received for publication April 30, 2022; accepted July 23, 2022.

Supplemental digital content is available for this article. Direct URL citation appears in the printed text and is provided in the HTML and PDF versions of this article on the journal's Web site (www.journaladdictionmedicine.com).

Supported by an Alberta Centennial Addiction and Mental Health Research Chair and transitional funding (to KJA) and the American Foundation for Addiction Research (to KJA). Several scholarships, including a University of Alberta Dean's Doctoral Student Award 2021–2022, a University of Alberta Medical Science Graduate Program Scholarship 2020, a University of Alberta Doctoral Recruitment Scholarship 2017, and an Alberta Graduate Excellence Scholarship 2020 provided additional funding for the first author (SJ). A Fulbright-Canada-Palix Foundation grant to Dr. Patrick Carnes facilitated his contributions to study design.

KJA has received 2 research grants from Janssen Inc., Canada (fellowship grants for trainees). The analysis and data interpretation were conducted independently.

Send correspondence to Katherine J. Aitchison, BM BCh, PhD, FRCPsych, Departments of Psychiatry and of Medical Genetics, College of Health Sciences, Faculty of Medicine and Dentistry, University of Alberta, 116 St & 85 Ave, Edmonton, AB T6G 2R3, Canada. E-mail: kaitchis@ualberta.ca.

ISSN: 1932-0620/23/1702-0174

DOI: 10.1097/ADM.0000000000001078

despite draft proposed criteria for the inclusion of sex addiction and internet addiction in *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition) (*DSM-5*),[5] the only behavioral addiction currently included in the main body of *DSM-5* (American Psychiatric Association 2013) is gambling disorder (312.31), which was first included in *DSM* as an impulse control disorder.[6] In the *International Statistical Classification of Diseases*, compulsive sexual behavior disorder was proposed for inclusion as an impulse control disorder in 2014,[7] which occurred in 2018.[8] However, the scientific discussion about whether compulsive sexual behavior disorder could constitute a manifestation of a behavioral addiction was acknowledged,[9] and it was predicted (as in the case of problematic gambling)[6] that understanding would evolve as research elucidated the phenomenology and neurobiological underpinnings of the condition.

Sex addiction is defined as a model for understanding egodystonic hypersexuality characterized by preoccupation (behavioral salience) with and narrowing of the repertoire of sexual activity, an inability to stop (i.e., compulsion aspect) or decrease in both internal (e.g., fantasy) and external behaviors despite unwanted consequences, tolerance (here defined as resulting in increased frequency, duration, or risk associated with behaviors), and withdrawal (e.g., dysphoric mood when the behavior is stopped), with reinstatement of behaviors to compensate for the dysphoric mood.[5] Compulsive sexual behavior disorder is characterized by a persistent pattern of failure to control intense, repetitive sexual impulses or urges, resulting in repetitive sexual behavior over an extended period (e.g., 6 months or more) that causes marked distress or impairment in personal, family, social, educational, occupational, or other important areas of functioning.[8] The term problematic sexual behavior (PSB) is used because it focuses on a potentially broader range of behaviors and their problematic outcomes rather than on the heterogeneous etiology and pathology.[10] PSB is characterized by repetitive sexual behaviors, and it is associated with uncontrolled sexual urges/impulses and distress,[11] and social and functional impairments.[8,11] It has been estimated that the prevalence of PSB is 3% to 6% in adults,[4,12–14] with higher frequencies in specific populations. Reid[15] reported that 19% of college men met the criteria for PSB, whereas Giordano and Cecil[16] found that 11.1% of men and women in an undergraduate sample met these criteria. Cashwell et al.[17] reported that, of 379 undergraduates, 21.2% of men and 6.7% of women screened positive, that is, scored in a range indicating that they should be offered further assessment for PSB.

Various risk factors have been associated with PSB, including those previously associated with the construct known as reward deficiency syndrome (RDS) (e.g., substance[18] and nonsubstance use disorders,[19] attention-deficit/hyperactivity disorder [ADHD],[20] and personality disorder),[21] other psychiatric disorders,[22] and childhood trauma.[23] Mick and Hollander[24] hypothesized that impulsivity initiated the early stage of PSB, with compulsivity being involved in repetitive behaviors and hence in the maintenance of PSB. Owing to the classification of PSB within *International Statistical Classification of Diseases, Tenth Revision*, as an impulse control disorder, we specifically included obsessive-compulsive disorder (OCD) in our study. The magnitude of the contribution of the various different postulated mechanisms and whether PSB encompasses various syndromes with different etiologies are not known at present. Moreover, systematic data on the prevalence of PSB and associated sociodemographic factors across diverse populations, including nontreatment seekers, remain to be provided.[8]

In light of the above, this paper aims to further explore the prevalence, sociodemographic features, and nosology of the construct known as PSB in a diverse sample of adults in postsecondary education. We hypothesized that, first, screening positive for factors previously associated with RDS (e.g., internet addiction, compulsive buying, nicotine dependence, and pathological gambling), and second, OCD and childhood trauma would be associated with PSB in this sample derived from a Canadian postsecondary education institution.

METHODS

The study inclusion criteria were as follows: at least 18 years of age, undergraduates, graduates, postdoctoral fellows, and recently convocated students registered in at least 1 course in the preceding year, except for those who had completely withdrawn after registration without any reasons to not return in the next academic term or not be on campus (such as suspension), and being able to answer in English. A Canadian university Office of the Registrar invited 68,846 eligible students and postdoctoral fellows by email (all available students registered 2016–2019). Students interested in participating emailed the study team. These students were then sent an email invitation to review the participant information and complete consent online, followed by the screening measures (hosted by the Qualtrics platform), with a subsequent email inviting them to complete the Mini-International Neuropsychiatric Interview (MINI) version 7.0.2 online.

Measures

The Sexual Addiction Screening Test—Revised (SAST-R) is a 45-item screening tool comprising several subscales designed to detect potential hypersexuality of an addictive nature. The 20-item SAST-R Core subscale was used as a screener for the characteristics associated with PSB with positive screen indicated by a score of 6 or greater (the cutoff for the general population).[25] Each question is answered in a binary manner (yes/no = 0/1). Sexual activities deemed to be problematic by participants in the past 30 days were also separately assessed by asking the following question: "Within the last 30 days, how often have you participated in activities of a sexual nature that you would regard as problematic?"

Measures used to screen for internet addiction, compulsive buying, personality disorders, ADHD, nicotine dependence, and pathological gambling were as follows: the Internet Addiction Test (IAT),[26] the Richmond Compulsive Buying Scale (RCBS),[27] the 8-item Standardized Assessment of Personality—Abbreviated Scale as a Self-Administered Screening Test (SA-SAPAS),[28] the full Adult ADHD Self-Report Scale (ASRS) version 1.1,[29] the Fagerström Test for Nicotine Dependence,[30] and the *DSM-5* Pathological Gambling Diagnostic Form,[31] respectively. Sex, sexual orientation, a family history of domestic violence or of sex addiction, having received a diagnosis of OCD or any other mental health condition(s), and childhood trauma were additionally collected by self-report. The MINI was used to output psychiatric disorders by *DSM-5* criteria.

Data Analysis

Demographic Data and Logistic Regression

Data were analyzed by STATA (Stata/SE 16) and R (version 3.6.3) after dropping the missing data. (Supplemental Digital Content, http://links.lww.com/JAM/A387). The PSB variable was created by combining those screening positive on the SAST-R Core with those endorsing self-reported PSB, after subtracting the two SAST-R Core questions referred to below. Childhood trauma (275 of 3347 [8.22%]), a family history of domestic violence (387 of 3288 [11.77%]), and SAST-R Core question 1 (which enquires about a history of sexual abuse in childhood or adolescence) were combined as childhood trauma. Self-reported family history of sex addiction ("yes"/"no") data were combined with SAST-R Core question 2 (which enquires about parental trouble with sexual behavior) to create a new family history of PSB variable. IAT was regrouped as "yes" (n = 1174, the mild, moderate, and severe groups) and "no" (n = 2160, the normal group).

Differences in data distributions between those with and without PSB were assessed by Pearson χ^2 test for categorical variables and by nonparametric testing (Mann-Whitney U test). Tetrachoric (for binary variables) and Kendall τ (for categorical variables) correlations were used for data correlation. SA-SAPAS, OCD, ASRS, IAT, childhood trauma, a family history of PSB, and gender were used as covariates in logistic regression analyses with PSB as the dependent variable. Machine learning (ML) logistic regression as well as conventional was used to enhance the robustness of the models and to address the data distribution imbalance (a relatively low frequency of PSB). Sensitivity analyses were conducted to ascertain whether using the unadjusted SAST-R Core, childhood trauma, and a family history of PSB as the dependent resulted in similar predictors for the logistic regression (Supplemental Digital Content, Supplementary Figs. 3 and 4, http://links.lww.com/JAM/A375). All reported P values are unadjusted for multiple testing.

ML Approaches

The "Haven" and "dplyr" packages, and the "forcats" and "creditmodel" packages in R were used for data input, encoding and splitting the data set (into 80% training and 20% testing), respectively. The Synthetic Minority Oversampling Technique from the "Caret" package was used to address data imbalance in the training set.[32] Leave-one-out cross-validation in the "Caret" package was used for resampling in the training set.[33] Logistic regression in the "caret" package was used for regression. Accuracy, area under the receiver operating characteristic (ROC) curve, sensitivity, specificity, and F score were used to evaluate the models. The "pROC" package was used for the area under the ROC curve.

Multiple Correspondence Analysis

Multiple correspondence analysis (MCA) was used for analyzing the pattern of relationships between multiple categorical variables in terms of how they are geometrically positioned in relation to latent dimensions. MCA was conducted on OCD, RCBS, SA-SAPAS, ASRS, IAT, and childhood trauma, stratifying by gender (women, n = 2084; men, n = 1012). Because a family history of PSB was correlated with the combined childhood trauma variable (tetrachoric $\rho = 0.69$, $P < 0.001$, n = 3019; Fig. 1; Supplemental Digital Content, Supplementary Table 3,

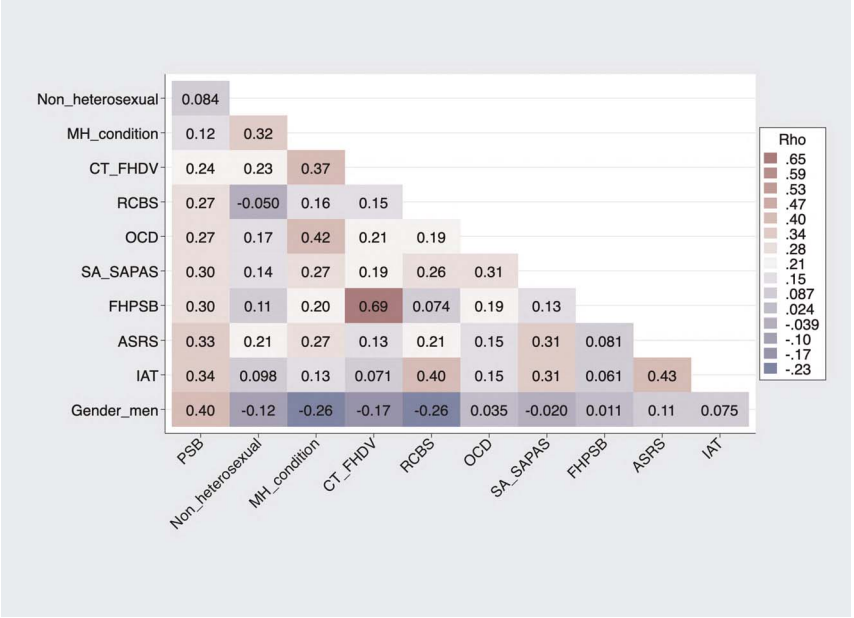

FIGURE 1. Correlation matrix for PSB with demographic and clinical variables. *P* values are provided in Supplemental Digital Content, Supplementary Table 3, http://links.lww.com/JAM/A375. MH_condition indicates mental health condition; CT_FHDV, childhood trauma and a family history of domestic violence (including physical, verbal/emotional and sexual childhood trauma and SAST-R Core question 1); FHPSB, a family history of problematic sexual behavior (including a family history of sexual addiction and SAST-R Core question 2).

http://links.lww.com/JAM/A375), these variables were combined as a supplementary childhood trauma-related variable. The "FactoMineR," "factoextra," and "GDAtools" packages in R were used for analysis and data visualization.

RESULTS

Of 4710 interested in participating, 237 withdrew from the study, 3359 completed online questionnaires using the Qualtrics platform, and 1801 completed the MINI (Supplemental Digital Content, Supplementary Fig. 1, http://links.lww.com/JAM/A375).

Demographic Data

The mean score on the SAST-R Core scale among all participants was 2.23 (SD, 2.73), with 12.18% (407 of 3341) of participants scoring above the threshold for screening positive in the general population.[25] The frequency of PSB was 16.53% (532 of 3219). There was a significant difference in the distribution of the following variables by those with and without PSB: sex, sexual orientation, OCD, RCBS, a family history of PSB, SA-SAPAS, current mental health condition(s), childhood trauma, ASRS, and IAT (Table 1; Supplemental Digital Content, Supplementary Table 2, http://links.lww.com/JAM/A375).

TABLE 1. Distribution of Demographic and Clinical Variables by Problematic Sexual Behavior (PSB) Group

Categories	PSB "Yes" 532/3219 (16.53%)	PSB "No" 2687/3219 (83.47%)	Total Number	*P* Value
SAST-R Core score	6.58 ± 3.22 (N = 532)	1.38 ± 1.56 (N = 2687)	3219	$P < 0.001$*
Age	22.32 ± 5.13 (N = 532)	22.34 ± 5.05 (N = 2686)	3218	$P = 0.81$*
Gender				
Women	232 (10.82%)	1913 (89.18%)	2145	$P < 0.001$†
Men	290 (28.27%)	736 (71.73%)	1026	
Total	522 (16.46%)	2649 (83.54%)	3171	
Sexual Orientation				
Heterosexual	420 (15.85%)	2230 (84.15%)	2650	$P = 0.034$†
Non-heterosexual	110 (19.50%)	454 (80.50%)	564	
Total	530 (16.49%)	2684 (83.51%)	3214	
OCD				
No	490 (15.88%)	2596 (84.12%)	3086	$P < 0.001$†
Yes	40 (33.33%)	80 (66.67%)	120	
Total	530 (16.53%)	2676 (83.47%)	3206	
RCBS				
No	471 (15.54%)	2559 (84.46%)	3030	$P < 0.001$†
Yes	61 (32.28%)	128 (67.72%)	189	
Total	532 (16.53%)	2687 (83.47%)	3219	
FHPSB				
No	453 (15.38%)	2492 (84.62%)	2945	$P < 0.001$†
Yes	71 (34.63%)	134 (65.37%)	205	
Total	524 (16.63%)	2626 (83.37%)	3150	
SA-SAPAS				
No	425 (14.76%)	2454 (85.24%)	2879	$P < 0.001$†
Yes	107 (31.47%)	233 (68.53%)	340	
Total	532 (16.53%)	2687 (83.47%)	3219	
Mental health condition(s)				
No	399 (15.35%)	2200 (84.65%)	2599	$P < 0.001$†
Yes	131 (21.87%)	468 (78.13%)	599	
Total	530 (16.57%)	2668 (83.43%)	3198	
Childhood trauma				
No	339 (13.94%)	2092 (86.06%)	2431	$P < 0.001$†
Yes	185 (25.45%)	542 (74.55%)	727	
Total	524 (16.59%)	2634 (83.41%)	3158	
ASRS				
No	321 (12.97%)	2153 (87.03%)	2474	$P < 0.001$†
Yes	211 (28.32%)	534 (71.68%)	745	
Total	532 (16.53%%)	2687 (83.47%)	3219	
IAT				
Normal (No)	241 (11.50%)	1855 (88.50%)	2096	$P < 0.001$†
Mild, moderate or severe (Yes)	291 (25.91%)	832 (74.09%)	1123	
Total	532 (16.53%)	2687 (83.47%)	3219	

*Mann-Whitney U test.

†Pearson's χ^2 test.

OCD indicates obsessive-compulsive disorder (including self-reported and MINI-diagnosis); RCBS, Richmond Compulsive Buying Scale; FHPSB, a family history of problematic sexual behavior (including a family history of sex addiction and SAST Core question 2); SA-SAPAS, Standardized Assessment of Personality–Abbreviated Scale as a Self-Administered Screening Test; Childhood trauma, childhood trauma (including physical, verbal/emotional and sexual childhood trauma and SAST Core question 1) and a family history of domestic violence; ASRS, the Adult ADHD Self-Report Scale (v1.1); IAT, the Internet Addiction Test.

Correlational analyses showed that PSB was correlated with nonheterosexual sexual orientation, current mental health condition(s), childhood trauma, RCBS, OCD, SA-SAPAS, a family history of PSB, ASRS, IAT, and gender (men) (Fig. 1; Supplemental Digital Content, Supplementary Table 3, http://links.lww.com/JAM/A375). Interestingly, a family history of PSB was strongly correlated with childhood trauma (tetrachoric $\rho = 0.69$, $P < 0.001$, n = 3019), and both were correlated with PSB. Of those responding "yes" to a family history of sex addiction, 32.65% (16 of 49) screened positive on the SAST-R Core. Our data also showed that (*a*) PSB was correlated with unadjusted SAST-R Core (tetrachoric $\rho = 0.97$, $P < 0.001$, n = 3219) and self-reported PSB (tetrachoric $\rho = 0.70$, $P < 0.001$, n = 3271), and (*b*) self-reported OCD was correlated with OCD MINI diagnosis (tetrachoric $\rho = 0.47$, $P < 0.001$, n = 1785; Supplemental Digital Content, Supplementary Table 4, http://links.lww.com/JAM/A375).

Regression Models

Logistic Regression

Outliers were removed from the logistic regression (n = 47, with 8 categorized as "yes" to having PSB; Supplemental Digital Content, Supplementary Fig. 2, http://links.lww.com/JAM/A375) if their Pearson standardized residual had an absolute value of more than 2,[34] or Hosmer-Lemeshow $\Delta\hat{\beta}$ more than 4.[35] The following variables were associated with PSB: OCD, ASRS, IAT, a family history of PSB, SA-SAPAS, childhood trauma, RCBS, and gender (men) (odd ratios and *P* values are provided in Fig. 2A).

ML Logistic Regression

ML logistic regression showed that RCBS, SA-SAPAS, OCD, a family history of PSB, ASRS, childhood trauma, IAT, and gender (men) were associated with PSB (odd ratios and *P* values are provided in the Fig. 2B). A validation of the aforementioned model was undertaken. The validation model was trained with adjusted SAST-R Core (n = 3141) and tested with self-reported PSB (n = 3145, odd ratios and *P* values are provided in the Fig. 2C). Sensitivity analyses were performed using unadjusted SAST-R Core and a family history of sexual addiction (Supplemental Digital Content, Supplementary Figs. 3 and 4, http://links.lww.com/JAM/A375).

Multiple Correspondence Analysis

The MCA results showed that, in both women and men, the first 2 dimensions accounted for more than 50% of explained variance (scree plots are provided in Supplemental Digital Content, Supplementary Fig. 5, http://links.lww.com/JAM/A375). PSB showed slightly different data clustering patterns by gender in these dimensions. In dimensions 1 and 2, PSB was closer to SA-SAPAS, IAT, and ASRS (and RCBS in women) but not OCD (or RCBS in men). Interestingly, RCBS showed a stronger association with OCD in men than with other variables. In addition, OCD (and RCBS in men) showed higher representation in dimension 2 than in dimension 1. In dimensions 4 and 5, in women, PSB showed an equivalent representation in dimensions 4 and 5 and displayed a stronger association with SA-SAPAS. In men, PSB was only well represented in dimension 5 and locating further away from the other variables (Figs. 3, 4; Supplemental Digital Content, Supplementary Figs. 6 and 7, http://links.lww.com/JAM/A375).

DISCUSSION

The frequency of PSB, as defined by screening positive on the SAST-R Core, was 8.49% in the women and 19.7% in the men in our sample, consistent with prior reports in this population.[15–17] Factors consistently associated with PSB were sex, ADHD, personality disorder, OCD, internet addiction and compulsive buying, childhood trauma, and a family history of PSB. Our results also showed that the self-reported data were correlated with the screening tool (in the case of the SAST-R and PSB) or data via MINI (Supplemental Digital Content, Supplementary Table 4, http://links.lww.com/JAM/A375).

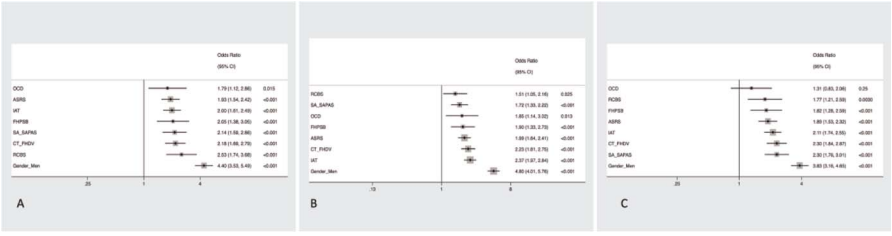

FIGURE 2. Forest plots of PSB logistic regression analysis. A, By conventional logistic regression. B, By ML logistic regression. C, Validation ML logistic regression model. A, The log likelihood of the model was −1149.98 (n = 3049, after removing outliers, with a pseudo $R^2 = 0.16$, $P < 0.00001$). The model passed the link test with a significant *P*(hat) ($P < 0.001$) and a nonsignificant *P*(hatsq) ($P = 0.12$). The probability of the Hosmer-Lemeshow χ^2 test was insignificant ($P = 0.64$), suggesting a good model fit. The mean VIF and the condition number were 1.08 and 6.95, respectively. The model accuracy, area under the ROC curve, sensitivity, specificity, and *F* score were 84.32%, 77.37%, 14.46%, 98.19%, and 0.23, respectively. B, Machine learning logistic regression was conducted on the data set without removing the outliers (training, n = 2477; testing, n = 619; train to test ratio, 0.80). The model accuracy, area under the ROC curve, sensitivity, specificity, and *F* score were 74.64%, 73.90%, 53.55%, 79.76%, and 0.45, respectively. C, The validation model was trained using adjusted PSB, n = 3141, and was tested by monthly PSB, n = 3145. The model accuracy, area under the ROC curve, sensitivity, specificity, and *F* score were 75.04%, 74.26%, 59.11%, 76.97%, and 0.34, respectively. FHPSB includes a family history of problematic sexual behavior (including a family history of sex addiction and SAST-R Core question 2); CT_FHDV, childhood trauma and a family history of domestic violence (including physical, verbal/emotional and sexual childhood trauma and SAST-R Core question 1); VIF, variance inflation factor.

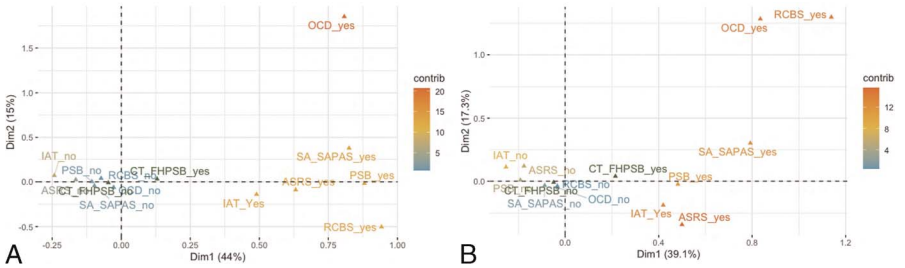

FIGURE 3. Results of MCA (dimensions 1 and 2) in women (A) and men (B). The maps of data binarized into 2 categories (yes/no, for each variable) in dimensions 1 and 2 in women and men are shown. CT_FHPSB includes childhood trauma and a family history of problematic sexual behavior.

Model validation (training with SAST-R Core subtracting the first 2 questions and testing with self-reported PSB) resulted in very similar predictors (Figs. 2B, C). This means that (1) the original dependent variable (PSB measured by SAST-R Core subtracting the first 2 questions and adding in the self-report data) performs similarly to using SAST-R Core subtracting the first 2 questions as the dependent variable, and (2) splitting the original data set into training and test gives comparable results to using only self-reported PSB as the dependent variable in the test set. A sensitivity analysis (Supplemental Digital Content, Supplementary Figs. 3 and 4, http://links.lww.com/JAM/A375) was undertaken using the original variables, namely, unadjusted SAST-R Core, childhood trauma, and a family history of sex addiction. The logistic regression model using these variables (pseudo R^2 = 0.13) revealed that RCBS, ASRS, IAT, OCD, childhood trauma, and gender (men) remained significant. A family history of sex addiction was the only variable that became nonsignificant (P = 0.17), probably because of the lower frequency of endorsement of this variable (n = 49) compared with that of SASTC2 (endorsement, 214 of 3268 [6.55%]). Using the new PSB variable, including childhood trauma and a family history of PSB but excluding sexual orientation as well as current mental health condition(s) improved the model fit slightly in the logistic regression model (e.g., sensitivity: unadjusted 77.20% vs. adjusted 79.70%; F score: unadjusted 0.42 vs. adjusted 0.45). The ML logistic regression models showed higher sensitivity and F scores compared with the conventional logistic regression.

The variables previously associated with the RDS construct, such as SA-SAPAS, ASRS, and IAT, were associated with PSB. In the final logistic (Fig. 2) and ML logistic regressions (Fig. 2), RCBS was associated with PSB. Of note, there were gender differences regarding RCBS: RCBS was negatively correlated with men (Fig. 1). In addition, the distribution of PSB differed significantly between OCD-positive and -negative groups (P < 0.001, Table 1), and PSB was correlated with OCD (ρ = 0.27, P < 0.001; Fig. 1; Supplemental Digital Content, Supplementary Table 3, http://links.lww.com/JAM/A375). Our regression results showed that OCD was associated with PSB (OR of 1.79, P = 0.015, and OR of 1.85, P = 0.013 for conventional and ML logistic regression, respectively; sensitivity models: OR of 2.28, P < 0.001, and OR of 1.73, P = 0.033, respectively).

The goal of the MCA is to investigate the complex relationships between these variables and PSB, without making any prior assumptions. MCA considers the geometric definition of the data rather than the statistical definition. Our MCA results showed that all the predictors displayed associations with PSB in the first 2 dimensions, with OCD (and RCBS in men) displaying a weaker association with PSB (Figs. 3A, B; Supplemental Digital Content, Supplementary Fig. 6, http://links.lww.com/JAM/A375). These findings are consistent with those in our regression (Fig. 2) and correlation analyses (Fig. 1). Although the other 3 dimensions did not achieve higher eigenvalues, they still accounted for around 40% of the explained variance. Therefore, to prevent possible loss of information, we retained the data up to dimension 5 (Fig. 4). In dimensions 4 and 5, PSB had a good quality of representation and a relatively high contribution to the dimensions (Supplemental Digital Content, Supplementary Tables 5 and 6, http://links.lww.com/JAM/A375). Data from dimensions 4 and 5 showed that there were gender differences regarding the association/correlation patterns among the data; for example, in women, PSB and SA-SAPAS were grouped relatively closely, but in men, PSB and SA-SAPAS were clustered far apart (Figs. 4A, B; Supplemental Digital Content, Supplementary Fig. 7, http://links.lww.com/JAM/A375). Interestingly, PSB exhibited a weaker association with all other variables in men in dimension 5, implying that there may be a subset of PSB that is not explained by the hypotheses of RDS or of OCD.

Limitations

Limitations of this study include the following. First, the self-selection recruitment strategy may have impacted sample characteristics; hence, our findings should be interpreted with caution and require replication. In addition, although all eligible participants were given an equal chance to participate in the research, data collected using a "voluntary response" method may not be representative. However, of note, the frequency of PSB in our data set was comparable with that seen in other studies of adults in postsecondary education as mentioned previously.[12,17] Moreover, the purpose of participation in this study as outlined to the students was to act as "a *healthy control subject* in a three-phase study about genes associated with sexual addiction and related conditions." Given this, it could be expected that the frequency of problematic behaviors and addictions and related conditions (the reward deficiency construct) might be expected to be lower, not higher, than the general population frequencies

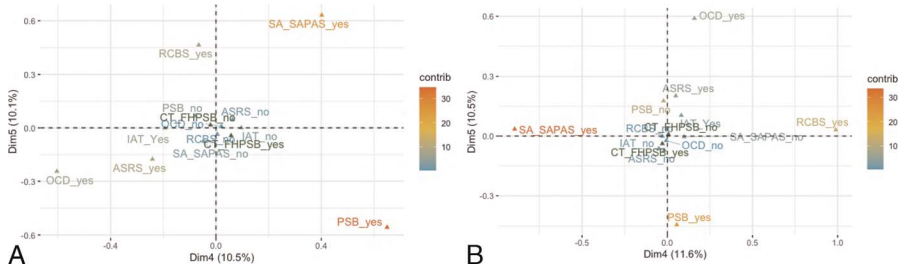

FIGURE 4. Results of MCA (dimensions 4 and 5) in women (A) and men (B). The maps of data binarized into 2 categories (yes/no, for each variable) in dimensions 4 and 5 in women and men are shown. CT_FHPSB includes childhood trauma and a family history of problematic sexual behavior.

of these. Therefore, if there is any bias in the sample, it might trend toward the absence of an association between PSB and RDS. Subsample sizes were another limiting factor. Specifically, because the frequency of endorsement of OCD was relatively low (120 of 3206 [3.74%]), although it appears on the MCA that PSB clusters more closely to the measures of reward deficiency than to OCD, this should not be overinterpreted. For sexual orientation and sex, the smaller groups were combined in the former case and dropped in the latter. For the former, bisexual (366 of 3371 [10.86%]), homosexual (129 of 3371 [3.83%]), and other (91 of 3371 [2.70%]) were regrouped as "nonheterosexual sexual orientation." Because of the small numbers in the gender minority groups (trans and other, 47 of 3375 [1.39%]), only men and women were included in the logistic regression analysis. Another potential limitation is that we dropped missing values. However, because the proportion of missing data was relatively small (5% or less, descriptions of missing data are provided in the Supplemental Digital Content, Supplementary Tables 7 and 8, http://links.lww.com/JAM/A375), this should not be problematic.[36] Although various methods may be used to impute Although various methods may be used to impute data, some such methods (e.g., monotone imputation or chained equations) may introduce bias and perfect (inaccurate) predictions, as the methods themselves involve the fitting of a regression model to a categorical outcome.[37] We also dropped 2 measures, the Fagerström Test for Nicotine Dependence (15 of 3332 [0.45%]) and the Pathological Gambling Diagnostic Form (76 of 3332 [2.28%]), because of very low endorsement rates (which were lower than in another university student sample).[38] Another limitation is that logistic regression is sensitive to high correlations between predictor variables. In our data set, the correlation coefficients were less than 0.80,[39] and the mean variance inflation factor and the condition number were at acceptable levels (variance inflation factor, <10; condition number, <10) in our final models,[40] indicating that multicollinearity was not a significant concern. A power calculation using STATA used the parameters for OCD (the weakest predictor, $P_1 = 0.16$, $P_2 = 0.24$, pseudo $R_2 = 0.16$) that, for an α of 0.05, a sample size of only 264 would give a power of 0.8. Finally, the proportion of the variance (pseudo $R^2 = 0.16$, Fig. 2) explained by our analysis was relatively low. Of note, the variables that we included in our models were driven by the prior literature. Future research could include feature selection techniques, alternative ML models, and consideration of genetic and other biological variables. Validation through additional independent data sets with overlapping variables would also be advantageous.

CONCLUSIONS

In summary, we have identified associations between PSB in adults in our sample derived from individuals in postsecondary education and the following: compulsive buying, personality disorder, OCD, ADHD, childhood trauma, a family history of PSB, internet addiction, and gender (men). Our results suggest not only that measures associated with RDS are associated with PSB but also that OCD is a vulnerability factor. Our data are consistent with different subgroups within PSB with slightly somewhat contrasting etiological mechanisms by sex.

ACKNOWLEDGMENTS

The authors would like to acknowledge the contribution of Dr. Patrick Carnes to the study design, and volunteer input of Pennie Carnes, Grace Li, Hana Graham, and Chanelle Martens to data collection. SJ would like to thank Afia Anjum from DW's lab for her assistance with R. There was no pharmaceutical or industry support for the work reported herein.

REFERENCES

1. Orford J. Hypersexuality: implications for a theory of dependence. *Br J Addict Alcohol Other Drugs*. 1978;73(3):299–210.
2. Carnes P. *Don't Call It Love: Recovery from Sexual Addiction*. New York, NY: Bantam Books; 1991:439.
3. Goodman A. Sexual addiction: terminology and theory. *Psychiatr Times*. 1998;15(7):22–26.
4. Carnes P. *Sexual Addiction*. 8 ed. vol I. Philadelphia, PA: Lippincott, Williams & Wilkins; 2005.
5. Carnes PJ, Hopkins TA, Green BA. Clinical relevance of the proposed sexual addiction diagnostic criteria: relation to the Sexual Addiction Screening Test-Revised. *J Addict Med*. 2014;8(6):450–461.
6. Rosenthal RJ. Inclusion of pathological gambling in DSM-III, its classification as a disorder of impulse control, and the role of Robert Custer. *Int Gambl Stud*. 2020;20(1):151–170.
7. Rehm J, Probst C, Kraus L, et al. The addiction concept revisited. In: Reframing Addiction: Policies, Processes and Pressures. Barcelona, Barcelona, Spain: The ALICE RAP Project, 2014:103–117.
8. Kraus SW, Krueger RB, Briken P, et al. Compulsive sexual behaviour disorder in the ICD-11. *World Psychiatry*. 2018;17(1):109–110.
9. Griffiths MD. Compulsive sexual behaviour as a behavioural addiction: the impact of the internet and other issues. *Addiction*. 2016;111(12):2107–2108.
10. Joannides P. The challenging landscape of problematic sexual behaviors, including "sexual addiction" and "hypersexuality". *New Directions in Sex Therapy: Innovations and Alternatives*. London, United Kingdom: Routledge/Taylor & Francis Group; 2012:69–83.

11. Dickenson JA, Gleason N, Coleman E, et al. Prevalence of distress associated with difficulty controlling sexual urges, feelings, and behaviors in the United States. *JAMA Netw Open*. 2018;1(7):e184468.
12. Carnes PJ, Green BA, Merlo LJ, et al. PATHOS: a brief screening application for assessing sexual addiction. *J Addict Med*. 2012;6(1):29–34.
13. Carnes PJ. *Facing the Shadow: Starting Sexual and Relationship Recovery: A Gentle Path to Beginning Recovery from Sex Addiction*. Carefree, AZ: Gentle Path Press; 2015.
14. Klein V, Rettenberger M, Briken P. Self-reported indicators of hypersexuality and its correlates in a female online sample. *J Sex Med*. 2014;11(8):1974–1981.
15. Reid RC. Differentiating emotions in a sample of men in treatment for hypersexual behavior. *J Soc Work Pract Addict*. 2010;10(2):197–213.
16. Giordano AL, Cecil AL. Religious coping, spirituality, and hypersexual behavior among college students. *Sex Addict Compuls*. 2014;21(3):225–239.
17. Cashwell CL, Giordano A, Lewis TA, et al. Using the PATHOS questionnaire for screening sexual addiction among college students: a preliminary exploration. *Sex Addict Compuls*. 2015;22(2):154–166.
18. Deneke E, Knepper C, Green BA, et al. Comparative study of three levels of care in a substance use disorder inpatient facility on risk for sexual addiction. *Sex Addict Compuls*. 2015;22(2):109–125.
19. Carnes PJ. The sexual addiction screening process. *Clin Manag Sex Addict*. 2019;21–39.
20. Blankenship R, Laaser M. Sexual addiction and ADHD: is there a connection? *Sex Addict Compuls*. 2004;11(1–2):7–20.
21. Ballester-Arnal R, Castro-Calvo J, Giménez-García C, et al. Psychiatric comorbidity in compulsive sexual behavior disorder (CSBD). *Addict Behav*. 2020;107:106384.
22. Grant JE, Lust K, Chamberlain SR. Body dysmorphic disorder and its relationship to sexuality, impulsivity, and addiction. *Psychiatry Res*. 2019;273:260–265.
23. London S, Quinn K, Scheidell JD, et al. Adverse experiences in childhood and sexually transmitted infection risk from adolescence into adulthood. *Sex Transm Dis*. 2017;44(9):524–532.
24. Mick TM, Hollander E. Impulsive-compulsive sexual behavior. *CNS Spectr*. 2006;11(12):944–955.
25. Carnes P, Green B, Carnes S. The same yet different: refocusing the Sexual Addiction Screening Test (SAST) to reflect orientation and gender. *Sex Addict Compuls*. 2010;17(1):7–30.
26. Young KS. Internet addiction: the emergence of a new clinical disorder. *Cyberpsychol Behav*. 1998;1(3):237–244.
27. Ridgway NM, Kukar-Kinney M, Monroe KB. An expanded conceptualization and a new measure of compulsive buying. *J Consum Res*. 2008;35(4):622–639.
28. Merlhiot G, Mondillon L, Vermeulen N, Basu A, Mermillod M. Adaptation and validation of the standardized assessment of personality—Abbreviated Scale as a Self-Administered Screening Test (SA-SAPAS). *J Psychol Psychother*. 2014;4:(6):1–9
29. Kessler RC, Adler L, Ames M, et al. The World Health Organization Adult ADHD Self-Report Scale (ASRS): a short screening scale for use in the general population. *Psychol Med*. Feb 2005;35(2):245–256.
30. Heatherton TF, Kozlowski LT, Frecker RC, et al. The Fagerström test for nicotine dependence: a revision of the Fagerstrom tolerance questionnaire. *Br J Addict*. 1991;86(9):1119–1127.
31. Rennert L, Denis C, Peer K, et al. DSM-5 gambling disorder: prevalence and characteristics in a substance use disorder sample. *Exp Clin Psychopharmacol*. 2014;22(1):50–56.
32. Benedan L, Monti GS. Predicting the risk of gambling activities in adolescence: a case study. Berlin, Brandenburg, Germany: Springer; 2019:47–57. doi:https://dx.doi.org/10.1007/978-3-030-51222-4_5
33. Nwanganga F, Chapple M. Practical Machine Learning in R. Hoboken, NJ: John Wiley & Sons; 2020.
34. Chen CY, Yang HCP, Chen CW, Chen TH. Diagnosing and revising logistic regression models: effect on internal solitary wave propagation. *Eng Comput*. 2008;25(2):121–139.
35. Hosmer DW Jr, Lemeshow S, Sturdivant RX. *Applied logistic regression*. vol 398. Hoboken, NJ: John Wiley & Sons; 2013.
36. Momeni A, Pincus M, Libien J. Imputation and missing data. In: *Introduction to Statistical Methods in Pathology*. Gewerbestrasse, Cham, Switzerland: Springer International Publishing, 2018:185–200.
37. White IR, Daniel R, Royston P. Avoiding bias due to perfect prediction in multiple imputation of incomplete categorical variables. *Comput Stat Data Anal*. 2010;54(10):2267–2275.
38. Öncel SY, Gebizlioğlu ÖL, Alioğlu FA. Risk factors for smoking behavior among university students. *Turk J Med Sci*. 2011;41(6):1071–1080.
39. Shrestha N. Detecting multicollinearity in regression analysis. *Am J Appl Math Stat*. 2020;8(2):39–42.
40. Kim JH. Multicollinearity and misleading statistical results. *Korean J Anesthesiol*. 2019;72(6):558–569.

