Relationship of a Big Five personality questionnaire to the symptoms of affective disorders

Zoe Wilks[a], Adam M. Perkins[a,*], Andrew Cooper[b], Bartlomiej Pliszka[a], Anthony J. Cleare[a], Allan H. Young[a]

[a] Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

[b] Department of Psychology, Goldsmiths, University of London, London, United Kingdom

Research Paper

*Corresponding author at: Department of Psychological Medicine, Institute of Psychiatry, Psychology & Neuroscience, King's College London & South London and Maudsley NHS Foundation Trust, Denmark Hill, Camberwell, London SE5 8AZ

E-mail address: adam.perkins@kcl.ac.uk (A. Perkins).

Abstract

Online assessments allow cost-effective, large-scale screening for psychiatric vulnerability (e.g., university undergraduates or military recruits). However, conventional psychiatric questionnaires may worsen mental health outcomes due to overmedicalizing normal emotional reactions. Personality questionnaires designed for occupational applications could circumvent this problem as they utilise non-clinical wording and it is well-established that personality traits influence susceptibility to psychiatric illness. Here we present a brief, free-to-use occupational personality questionnaire, and test its sensitivity to symptoms of Bipolar Disorder (BD) and Major Depressive Disorder (MDD) in an online sample. Our study used a cross-sectional, self-report design to assess the relationship between self-reported symptoms of affective disorders and scores on the personality dimensions of openness, conscientiousness, extraversion, agreeableness and neuroticism. We used SEM to compare affective symptoms in 8,470 individuals (mean age 25.6 $\pm$ 7.0 years; 4,717 male) with scores on an online adaption of the TSDI, a public-domain 'Big Five' personality questionnaire. ROC curve analyses assessed cut off scores for the best predictors of overall vulnerability to affective disorders (represented by a composite screening score). Neuroticism was the most robust predictor of QIDS-16 depression symptoms and MDQ Hypomania symptoms (β = 0.68 and 0.39 respectively, $p < .0001$). Extraversion was the most robust predictor of HCL-16 Hypomania symptoms ($\beta = 0.34$, $p < .0001$). ROC curve analyses suggest if the TSDI was used for screening in this sample, neuroticism cut offs of approximately 58 for men and 70 for women would provide the most useful classification of overall vulnerability to affective disorders.

Relationship of a Big Five personality questionnaire to the symptoms of affective disorders

Keywords:
Personality; Big Five model; Trait Self-Description Inventory; Affective disorders; Depression; Hypomania

1. Introduction

Affective disorders comprise the largest class of psychiatric illness and place a significant disease burden on the global population. For example, the Global Burden of Disease Study found that in 2017, depressive disorders were the third most common cause of disability (IHME, 2018). Interventions which reduce the prevalence of affective disorders are therefore likely to have significant societal and economic benefits. Psychiatric screening is a potentially useful method of flagging individuals at high risk of developing an affective disorder before they fall ill so that they can be offered early, low-cost interventions that prevent conversion into more costly, fully-fledged affective illness.

Screening programmes are, however, generally recognised as a challenging and complex aspect of medicine and may even be counterproductive (Kagee et al., 2013). For example, a study of 88,235 US soldiers returning from Iraq found that their mental health worsened between the Post-Deployment Health Assessment (PDHA) and the Post-Deployment Health Re-Assessment (PDHRA) that took place approximately six months later (Milliken *et al.* 2007). Evidence of this type has led to concerns that: "All screening programmes do harm; some do good as well, and, of these, some do more good than harm at reasonable cost." (Gray *et al.* 2008). The mechanism by which screening might worsen mental health is unknown but one possibility is that by presenting negative emotions with inappropriate context, normal emotional reactions come to be viewed as abnormal (Durà-Vilà et al., 2013).

Researchers seeking to reduce the prevalence of affective disorders may therefore benefit from some means of screening for susceptibility to affective disorders in large cohorts, but without using screening tools that present psychiatric symptoms in inappropriate contexts to the individuals being screened. One way to do this derives from personality research since personality traits influence susceptibility to affective illness. For example, neuroticism predicts risk of major depressive disorder (MDD) 25 years later (Kendler *et al.* 2006), extraversion and openness are genetically related to bipolar disorder (BD; Lo *et al.*

2016; Matsumoto *et al.* 2018) and antidepressant treatments have been shown to reduce neuroticism and increase extraversion (Tang *et al.* 2009).

Occupational personality questionnaires are specially worded for use in non-clinical populations and may therefore provide a rapid, safe and economical means of screening for susceptibility to affective disorders at a population level. However, occupational personality questionnaires are typically published commercially and thus require costly per-administration royalty payments, rendering them prohibitively expensive for large-scale use, such as online screening for risk of affective disorders in undergraduates.

The Trait Self-Description Inventory (TSDI) is an exception, as it is a public domain occupational personality questionnaire developed in the 1990s as a joint venture between the U.S.A. and U.K. military (Collis & Elshaw, 1998). The TSDI measures the well-established 'Big Five' model of personality which derives from lexical rather than psychiatric research and consists of openness, conscientiousness, extraversion, agreeableness and neuroticism (Digman, 1990; Costa & McCrae, 1992; Goldberg, 1993). Big Five self-report questionnaires have been safely used in occupational settings for decades (e.g., Barrick *et al.* 2001) but nevertheless have clinical utility (e.g., Kendler et al. 2011), suggesting the TSDI has promise as a free-to-use, non-medicalising, psychiatric screening tool. But the TSDI has never been tested for psychiatric sensitivity. In this study we aimed to address this gap in the literature by testing the sensitivity of the TSDI to self-reported affective disorders symptoms in a large online sample.

2. Methods

2.1. Study Sample

Subjects were 8,470 volunteers (4,717 male, 3753 female) recruited via advertisements on social media websites Imgur and Facebook, as well as on the King's College London circular email system. To maximise potential participation, the only exclusion criterion was being less than 18 years of age. The mean $\pm$ SD of age was 25.6 $\pm$ 7.0 years. The Ethics Committee of King's College London approved the study and all participants gave informed consent.

2.2. Study measures

2.2.1. Personality and Mood assessment

The Trait Self-Description Inventory (TSDI) was developed as a 172-item paper and pencil self-report questionnaire measuring the Big Five personality dimensions of openness, conscientiousness, extraversion, agreeableness and neuroticism (Collis & Elshaw, 1998). It

was recently converted into a 50-item online version (Patrick *et al.* 2018). In twenty eight of the fifty items, participants use a seven-point response to indicate the extent to which an adjective such as "organized" or "considerate" is characteristic of them. In the remaining twenty two items, participants use a nine point response to indicate the extent to which they agree with statements such as "I tend to get upset easily" or "I like to keep all my belongings neat and organised". Each dimension of the Big Five in this questionnaire comprises ten items but not all have an equal split between adjectives and statements. Furthermore, to minimise social desirability, six out of ten extraversion items are reverse scored (e.g., "I am a very shy person"). This preponderance of reverse scored items typically produces mean extraversion scores that are negative (e.g., -2.2; Patrick et al., 2018) which can lower face validity and cause confusion during interpretation. The Hypomania Check List (HCL-16; Forty *et al.* 2010) measures symptoms of hypomania and is an abbreviated version of the HCL-32 (Angst *et al.* 2005). The HCL-16 has sensitivity (83%) and specificity similar to that of the HCL-32 (71%). The 16 items in the HCL-16 were summed to produce a total score for this scale. For the ROC analyses reported below, a participant received a positive screening result if they had a total score of 12 or higher. This equated to being more than 1 SD unit above the mean in this sample.

The Mood Disorder Questionnaire (MDQ; Hirschfeld *et al.* 2000) measures hyperthymia and hypomania with good sensitivity (73%) and specificity (90%) and was administered as a second screening tool for more severe bipolar spectrum disorders. The MDQ assesses the lifetime experience of 13 symptoms using a yes/no response format. If the participant endorses at least two of these symptoms, they are asked if they have experienced these concurrently using a yes/no response format. Lastly, participants are asked to what extent these symptoms have caused them problems using a four-point scale from 'No problem' to 'Serious problem'. To derive a continuous score for the analyses reported below, we summed all 15 of these items to produce an MDQ total. A participant received a positive screening result on the MDQ if they responded 'yes' to at least 7 of the symptom items, and responded both 'yes' to the item asking if they had experienced these symptoms concurrently and reported moderate or serious problems as a result of these symptoms.

The Quick Inventory of Depressive Symptomatology (QIDS-16; Rush *et al.* 2003) measures depressive symptoms in the past seven days and was adapted from the 30-item Inventory for Depressive Symptomatology (IDS-30; Rush *et al.* 1996). A total score for the QIDS-16 was calculated using the scoring instructions provided by Rush et al. (2003); total scores can range from 0 to 27. A participant received a positive screening result on the QIDS-

16 if they had a total score of 16 or greater. Using previously published suggested cut offs, this corresponds to severe, or very severe, depression (Reilly et al., 2015; Rush et al., 2003).

2.2.2. Procedure

Study advertisements directed volunteers to our custom-built web-portal which presented and scored the questionnaires (for a demonstration version, see www.measureyourpersonality.com, using access code 57562353) There was no monetary incentive for participation however in return for their cooperation each completer received a detailed personality profile explaining how their personality scores compare to the average and providing advice on how to manage their personality.

2.3 Statistical analysis

The Statistical Package for Social Sciences (SPSS) version 24 (IBM, New York, USA) was used to calculate the descriptive statistics and Pearson's *r* correlations. Correlations were used to measure associations between personality scores and affective disorders symptoms. Correlations greater than .10 were regarded as nontrivial on the basis that a measure which accounts for 1% of variance is regarded as having utility in large applied samples (Schmidt & Hunter, 1998).

To control for the effect of shared variance between the five personality traits, a series of structural equation models (SEMs) were constructed with the five personality traits simultaneously predicting the outcome variables. For each SEM, the measurement model for the personality traits was constructed by allowing each item from the TDSI to load on the specified *a priori* personality trait only, thus creating five latent variable personality traits as predictors in each model. Given the only small to moderate overlap between the three outcome variables and given the apparent differential relationships between the personality traits and MDQ Hypomania and HCL-16 Hypomania apparent from the correlations, separate SEMs were run for each outcome. In each of the three SEMs, the items belonging to each outcome measure were specified to load on a single latent variable. Lastly, in each SEM gender was also included as a control predictor variable. All SEM analyses were undertaken using MPlus 7 (Muthen & Muthen, 2015).

Receiver Operating Characteristic (ROC) curve analyses were undertaken to examine how each of the five personality traits concurrently predicted the screening results for the MDQ, HCL-16 and the QIDS-16. The thresholds for each outcome measure used to classify the sample as a positive or negative screening result were outlined above. A composite

screening score was also calculated for each participant by summing the number of positive screening classifications on the three outcome measures, with scores ranging from 0 to 3 (i.e. from no positive screening outcomes to a positive screening outcome on all three measures). This summed screening score was then dichotomised to use ROC curves to contrast those with three positive screening outcomes versus all other participants. Youden's *J* index (Youden, 1950) was used to derive optimal cut offs and test performance indicators for each personality trait in relation to each outcome measure. Sensitivity indicates the ability of the test to correctly detect true positives, specificity indicates the ability of the test to correctly detect true negatives, Positive Predictive Value (PPV) refers to the accuracy of a positive test result and Negative Predictive Value (NPV) refers to the accuracy of a negative test result. Area under the ROC curve (AUC) was calculated for all personality traits in relation to each outcome. In general terms, AUCs should be above at least 0.60 to demonstrate the predictor has acceptable discriminative utility in relation to the outcome, although one should also bear in mind the relative cost of administering the predictor and the intended purpose of the testing. All ROC curve analyses were undertaken using the pROC package in R (Robin et al., 2011).

3. Results

Descriptive Statistics

Table 1 displays means, standard deviations and correlations for the Big Five personality dimensions and affective disorders symptoms. Due to the large sample size, there were significant mean differences between the sexes for all personality scores. The size of the differences were in the range of one or two points except for neuroticism and agreeableness, in which women scored higher than men by 8.83 and 4.4 points respectively (the total range was 74 points for neuroticism and 59 points for agreeableness). Nine personality-symptom correlations were large enough to be nontrivial, eight of which were present in both sexes, indicating that they are robust. QIDS-16 depression correlated positively with neuroticism (.561) and openness (.108) but negatively with extraversion (-.190) and conscientiousness (-.184). MDQ hypomania correlated positively with neuroticism (.287) and openness (.128) but negatively with conscientiousness (-.134). Finally, HCL-16 hypomania correlated positively with extraversion (.285).

Table 1 about here

Structural Equation Modelling Analyses

In the first SEM, the personality traits were modelled to predict the QIDS16 depression outcome. This structural model fit the data reasonably well overall, χ^2 (2124) = 82319, p < .0001; RMSEA = 0.067, CFI = 0.86. Neuroticism was the most robust predictor of depression, with a standardised path weight of 0.68, p < .0001. Openness, β = 0.11, p < .0001, conscientiousness, β = -0.14, p < .0001, and agreeableness, β = -0.07, p < .0001, also had significant, albeit modest, relationships with depression. Extraversion and gender did not significantly predict depression. In the second SEM, the personality traits were modeled to predict the MDQ Hypomania outcome. This structural model fit the data reasonably well overall, χ^2 (2059) = 68060, p < .0001; RMSEA = 0.062, CFI = 0.88. Neuroticism was the most robust predictor of MDQ Hypomania, with a standardised path weight of 0.39, p < .0001. All other predictor variables were also significant predictors of the outcome, although of a lesser magnitude than neuroticism: openness, β = 0.12, p < .0001, conscientiousness, β = -0.09, p < .0001, agreeableness, β = -0.06, p< .0001, extraversion, β = 0.16, p < .0001, and gender, β = -0.13, p< .0001. In the final SEM, the personality traits were modeled to predict the HCL-16 Hypomania outcome. This structural model fit the data reasonably well overall, χ^2 (2124) = 68987, p < .0001; RMSEA = 0.061, CFI = 0.88. Extraversion was the most robust predictor of HCL-16 Hypomania, with a standardised path weight of 0.34, p < .0001. All other predictor variables, with the exception of conscientiousness, were also significant predictors of the outcome, although of a much lesser magnitude than extraversion: openness, β = 0.12, p < .0001, agreeableness, β = 0.08, p< .0001, neuroticism, β = 0.04, p < .0001, and gender, β = 0.08, p< .0001.

Receiver Operating Characteristic Curve Analyses

Given the observed sex difference across the personality traits, particularly for neuroticism, the ROC curve analyses were undertaken and reported here separately by sex. The proportion of male and female participants receiving a positive screening outcome for the affective disorder measures can be seen in Tables 2 and 3, respectively. For the composite screening outcome measure, 2% of the overall sample had a positive screen on all three measures (N = 166), 8.4% of the sample had a positive screen on two of the outcome measures (N = 714), 25.5% of the sample had a positive screen on one of the outcome measures (N = 2160), and 64.1% of the sample had no positive screening (N = 5430). Tables 2 and 3 show the

proportion of the sample who had a positive screen on all three outcomes for males and females, respectively.

Tables 2 and 3 show the test performance indicators for each of the five personality traits in relation to the screening results on each of the MDQ, HCL-16, QIDS-16 and the composite screening score for males and females, respectively. For the MDQ, neuroticism was the only trait with an AUC above 0.60 for males and females, and it had reasonable test sensitivity and specificity for both groups. For the HCL-16, extraversion was the only trait with an AUC above 0.60, and it also had reasonable test sensitivity and specificity. Neuroticism was a good predictor of the QIDS-16 for males and females, with an AUC of 0.81 for males and 0.77 for females, and with good test sensitivity and specificity across both groups. Extraversion also had an AUC value above 0.60 for the QIDS-16 for males and females. Lastly, for the composite screening measure, neuroticism had an AUC value of 0.77 for males and 0.76 for females, and good test sensitivity and specificity. Conscientiousness also had an AUC above 0.60 for this outcome measure for males and females. Figure 1 shows the ROC curve plots for the key predictor personality trait for each of the four screening outcome measures for both males and females. The thresholds used to generate the test indicators shown in Tables 2 and 3 are also shown in Figure 1. All the thresholds charted in Figure 1 show sex differences, with the largest being a 17-point difference between the neuroticism cut off scores for males and females on the composite screening score (Figure 1D). It should be noted for the composite screening score that using a higher threshold for the males of 58 on neuroticism also resulted in reasonable test sensitivity and specificity, with values of 0.66 and 0.75, respectively.

Tables 2, 3 and Figure 1 about here

---

4. Discussion

These data showed that a brief, public domain, occupational personality questionnaire, namely the TSDI, is sensitive to affective disorder symptoms when used in a large-scale online context. Of these symptom measures, QIDS-16 depression scores showed the strongest relationship to personality, displaying nontrivial correlations with four out of the Big Five personality dimensions. The largest two depression-personality correlations were a positive correlation with neuroticism and a negative correlation with extraversion. These findings fit

the historic notion that neurotic introverts are especially vulnerable to depression (e.g., Gray, 1970). The two smaller depression-personality correlations were a negative correlation with conscientiousness and a positive correlation with openness. The former fits longitudinal research that portrays conscientiousness as a general index of health (Friedman & Martin, 2011) and the latter fits the notion that high scores on openness to fantasy confer increased risk of depression (Carillo *et al.* 2001). SEMs, controlling for the overlap between the personality traits, largely followed the relationships found with the bivariate correlations between the traits and the outcome measures.

The relationship between hypomania symptoms and personality scores was less clear cut. It is plausible that hypomania should correlate positively with extraversion, but we found that MDQ hypomania scores were unrelated to extraversion in this sample. However, against expectations, MDQ hypomania correlated positively with neuroticism and openness, echoing the correlation pattern found between personality and QIDS-16 depression. This suggests that MDQ hypomania captures some depression variance, an idea that is backed up by the .440 correlation between the two measures. The smaller correlation of .097 between the QIDS-16 scores and the HCL-16 scores suggests the latter is a purer measure of hypomania. This notion is backed up by the nontrivial positive correlation between extraversion and HCL-16 scores, which fitted our expectations.

ROC curve analyses were then used to examine how well the traits predicted screening outcomes on the MDQ, HCL-16 and QIDS-16. Using cut offs based on previously published work with these affective disorder scales, between approximately 10-20% of the sample were categorised with an at-risk profile for each of these outcomes. We also created a composite screening score by summing the number of positive screening outcomes for each participant. 1.4% of males and 2.6% of females had a positive screening outcome on all three outcome measures. The ROC curve analyses showed that neuroticism was the most robust predictor of the screening outcomes for the MDQ, QIDS-16 and the composite screening score, with AUC values between 0.66 and 0.81 for each outcome. The HCL-16 screening outcome was not particularly well-predicted by any of the personality traits, although extraversion had an AUC value above 0.60 for males and females. For the MDQ, QIDS-16 and composite score, the optimal cut off score for neuroticism was between 49 and 52 for males, and between 60 and 69 for females. In broad terms, we could say that individuals in the general community scoring higher on neuroticism than the cut off values in this range (considering gender) are at-risk for affective disorders and may benefit from preventative interventions that address this risk.

Regarding the practicalities of using the online version of the TSDI as a psychiatric screening tool in large populations, such as an annual undergraduate intake cohort at a major university, the simplest option is to use scores on neuroticism as it displayed higher AUC values than any of the other Big Five personality dimensions. The negligible cost of administering the online version of the TSDI plus the inexpensive, benign nature of preventative interventions (e.g., a weekly meeting with a personal tutor or student counsellor) suggests organisations would be justified in experimenting with cut off scores for neuroticism in their specific population of interest order to minimise the number of vulnerable people who slip through the net yet also minimise the number of low risk people who are mistakenly offered interventions.

In such endeavours we emphasise that it is crucial to specify different cut off scores for men and women as the point of most accurate classification of psychiatric problems occurs in men at a lower point on the neuroticism scale than women, reflecting the well-established finding that men on average score lower on neuroticism than women (Patrick et al., 2018). In the case of the AUC values on the composite score, the sex difference in neuroticism cut offs is 17 points (52 for men versus 69 for women; Figure 1D). Whilst statistically the optimal point to maximise test sensitivity and specificity, this option most likely overestimates the number of men who are at risk due to the small number of men who scored three on the composite measure combined with its discontinuous, categorial nature. A more practicable cut off for men in this sample would be located at the lower peak on the blue line which is almost as effective a classifier as the first peak but corresponds to approximately 58 on the neuroticism scale. Such a cut off would more closely resemble that produced using the traditional method for identifying extreme scorers in psychometric studies, which is to flag up individuals who scoring one SD or more above the mean on neuroticism (in this case 61.6 for men and 70.3 for women). On this basis, if this sample were a cohort of undergraduates, we would select neuroticism cut off scores of 58 for men and 70 for women, as they would capture approximately equal degrees of proneness to psychiatric difficulties for men and women relative to the general population, even though the absolute scores are different.

As the Big Five model of personality stems from lexical rather than psychiatric research, there is no reason why the personality scores obtained in this study should relate to affective disorders symptoms, unless they represent a genuine association. Moreover, the patterning of the personality-symptom associations resembled that seen in previous studies, suggesting that our personality questionnaire is doing more than detecting noise variance.

The main limitation of our study is that our sample comprised users of social media such as Facebook and Imgur. For this reason, the results can only be viewed as explaining the relationship between personality domains and affective disorders symptoms in such individuals. However, as this is a growing sector of the population, it could be argued such participants are increasingly representative of the general public. As mentioned earlier, the HCL-16 was not entirely well-predicted by any of the five personality traits, so it may be worth examining the role of lower-order personality facets when seeking to predict this outcome.

We conclude that the TSDI personality questionnaire can be used in large samples (e.g., in university students) to identify individuals with a personality profile that confers elevated risk of developing an affective disorder. More specifically, we showed that screening for scores on neuroticism above a range of around 58 for men and 70 for women provides good predictive validity for outcome measures of affective disorders, particularly depression. This is particularly important given the pressing need to address the increasingly rising tide of depressive and related disorders in younger cohorts (Kessler *et al.* 2003) which makes it important to be able to screen for affective disorders on a large scale, but without causing overmedicalisation problems. Our results suggest the TSDI provides a cost-effective method for doing this.

References

Angst, J., Adolfsson, R., Benazzi, F., Gamma, A., Hantouche, E., Meyer, T.D., Skeppar, P., Vieta, E., Scott, J., 2005. The HCL-32: towards a self-assessment tool for hypomanic symptoms in outpatients. Journal of Affective Disorders, 88, 217-233.

Barrick, M.R., Mount, M.K., Judge, T.A., 2001. Personality and performance at the beginning of the new millennium: What do we know and where do we go next?. International Journal of Selection and Assessment, 9, 9-30.

Carrillo, J.M., Rojo, N., Sánchez-Bernardos, M.L., Avia, M.D., 2001. Openness to experience and depression. European Journal of Psychological Assessment, 17, 130-136.

Collis, J. M., Elshaw, C.C., 1998. The development of the trait self-description inventory: Results of US/UK collaboration. TTCP/HUM/98-001.

Costa, Jr., P.T., McCrae, R.R., 1992. Revised NEO personality inventory (NEO-PI-R) and NEO five-factor (NEO-FFI) inventory professional manual. PAR: Odessa, FL.

Digman, J.M., 1990. Personality structure: Emergence of the five-factor model. Annual Review of Psychology, 41, 417-440.

Durà-Vilà, G., Littlewood, R. and Leavey, G., 2013. Depression and the medicalization of sadness: Conceptualization and recommended help-seeking. International Journal of Social Psychiatry, 59(2), pp.165-175.

Friedman, H., Martin, L.R., 2011. The longevity project: surprising discoveries for health and long life from the landmark eight-decade study. Hay House: Carlsbad, CA.

Forty, L., Kelly, M., Jones, L., Jones, I., Barnes, E., Caesar, S., Fraser, C., Gordon-Smith, K., Griffiths, E., Craddock, N., Smith. D.J., 2010. Reducing the Hypomania Checklist (HCL-32) to a 16-item version. Journal of Affective Disorders, 124, 351-356.

Goldberg L.R., 1993. The structure of phenotypic personality traits. American Psychologist, 48, 26-34.

Gray, J.A., 1970. The psychophysiological basis of introversion-extraversion. Behaviour Research and Therapy, 8, 249-266.

Gray, J.M., Patnick, J., Blanks, R.G., 2008. Maximising benefit and minimising harm of screening. BMJ: British Medical Journal, 336, 480.

Hirschfeld, R.M., Williams, J.B., Spitzer, R.L., Calabrese, J.R., Flynn, L., Keck, Jr., P.E., Lewis, L., McElroy, S.L., Post, R.M., Rapport, D.J., Russell, J.M. 2000. Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire. American Journal of Psychiatry, 157, 1873-1875.

Institute for Health Metrics and Evaluation (IHME). Findings from the Global Burden of Disease Study 2017. Seattle, WA: IHME, 2018.

Kagee, A., Tsai, A.C., Lund, C. and Tomlinson, M., 2013. Screening for common mental disorders in low resource settings: reasons for caution and a way forward. International health, 5(1), pp.11-14.

Kendler, K.S., Gatz, M., Gardner, C.O., Pedersen, N.L., 2006. Personality and major depression: a Swedish longitudinal, population-based twin study. Archives of General Psychiatry, 63, 1113-1120.

Kendler, K.S., Myers, J., Reichborn- Kjennerud, T., 2011. Borderline personality disorder traits and their relationship with dimensions of normative personality: A web- based cohort and twin study. Acta Psychiatrica Scandinavica, 123, 349-59.

Kessler, R.C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K.R., Rush, A.J., Walters, E.E., Wang, P.S. 2003. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA, 289, 3095-3105.

Matsumoto., Y., Suzuki, A., Shirata, T., Takahashi, N., Noto, K., Goto, K., Otani, K. 2018. Implication of the DGKH genotype in openness to experience, a premorbid personality trait of bipolar disorder. Journal of Affective Disorders, 238, 539-541.

Milliken, C.S., Auchterlonie, J.L., Hoge, C.W., 2007. Longitudinal assessment of mental health problems among active and reserve component soldiers returning from the Iraq war. JAMA, 298, 2141-2148.

Muthén, L.K. and Muthén, B.O., 2015. Mplus User's Guide Seventh Edition ed. Los Angeles: CA: Muthén & Muthén.

Patrick, F., Young, A., Williams, S.C.R., Perkins, A.M., 2018. Pre-screening of clinical trials volunteers using an online personality questionnaire. Neuropsychiatric Disease and Treatment. In press.

Reilly, T.J., MacGillivray, S.A., Reid, I.C. and Cameron, I.M., 2015. Psychometric properties of the 16-item Quick Inventory of Depressive Symptomatology: a systematic review and meta-analysis. Journal of psychiatric research, 60, pp.132-140.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.C. and Müller, M., 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC bioinformatics, 12(1), p.77.

Rush, A.J., Gullion, C.M., Basco, M.R., Jarrett, R.B., Trivedi, M.H., 1996. The inventory of depressive symptomatology (IDS): psychometric properties. Psychological Medicine, 26, 477-486.

Rush, A.J., Gullion, C.M., Basco, M.R., Jarrett, R.B. and Trivedi, M.H., 1996. The inventory of depressive symptomatology (IDS): psychometric properties. Psychological medicine, 26(3), pp.477-486.

Rush, A.J., Trivedi, M.H., Ibrahim, H.M., Carmody, T.J., Arnow, B., Klein, D.N., Markowitz, J.C., Ninan, P.T., Kornstein, S., Manber, R., Thase, M.E., 2003. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS-16), clinician rating (QIDS-16-C), and self-report (QIDS-16-SR): a psychometric evaluation in patients with chronic major depression. Biological Psychiatry, 54, 573-583.

Schmidt, F.L., Hunter, J.E., 1998. The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. Psychological Bulletin, 124, 262-274.

Tang, T.Z., DeRubeis, R.J., Hollon, S.D., Amsterdam, J., Shelton, R., Schalet, B., 2009. Personality change during depression treatment: a placebo-controlled trial. Archives of General Psychiatry, 66, 1322-1330.

Youden, W.J., 1950. Index for rating diagnostic tests. Cancer, 3(1), pp.32-35.

Table 1. Means, standard deviations and correlations for entire sample as well as males and females only, for each of the Big Five personality dimensions and measures of affective symptoms.

Variable	Overall Mean (SD)	Male	Female	1	2	3	4	5	6	7	8
1. Openness	46.4(11.6)	47.8(11.4)	44.7(11.7)	-	.017	-.044**	.090**	.047**	.096**	**.128****	**.108****
2. Conscientiousness	46.3(11.0)	45.5(10.7)	47.4(11.2)	.036*	-	.017	.125**	-.104**	-.037**	**-.134****	**-.190****
3. Extraversion	-10.7(15.3)	-11.9(15.2)	-9.2(15.4)	-.023	.018	-	.134**	-.285**	**.285****	-.003	**-.205****
4. Agreeableness	43.9(11.6)	41.9(11.8)	46.3(10.9)	.120**	.123**	.114**	-	.080**	**.101****	-.036**	-.012**
5. Neuroticism	49.4(16.5)	45.5(16.1)	54.33(15.7)	.087**	-.119**	-.340**	.040**	-	.054**	**.287****	**.584****
6. HCL-16	8.6(2.9)	8.3(2.8)	8.9(2.9)	**.109****	-.042**	**.305****	**.102****	.022	-	.356**	.086**
7. MDQ	8.4(4.1)	8.6(4.0)	8.2(4.2)	**.117****	**-.116****	.015	-.014	**.307****	.358**	-	.433**
8. QIDS-16	9.7(5.68)	9(5.5)	10.6(5.8)	**.130****	**-.196****	**-.233****	-.036**	**.577****	.065**	.422**	-

				-	.018	-	.134**	-.285**	**.285****	-.003	**-.205****
				-.046**	-	.001	-	.080**	**.101****	-.036**	-.012**
				.116**	.097**	.127**	-	-	.054**	**.287****	**.584****
				.082**	-.146**	-.296**	.020	-	-	.356**	.086**
				.109**	-.049**	**.249****	.066**	.043**	-	.358**	.433**
				.132**	**-.149****	-.016	-.047**	**.311****	.366**	-	.422**
				.127**	**-.215****	**-.203****	-.043*	**.568****	.086**	.466**	-

Note: Correlations for whole sample (n = 8,470) in upper right of matrix, correlations for males (n = 4,717, upper) and females (n = 3,753, lower) in lower left of matrix. Personality/symptom correlations over .10 are shown in bold font.

* p < .05. ** p < .01.

Table 2. Results from the ROC curve analyses for the MDQ, HCL-16, QIDS-16 and the composite screening score for male participants

TSDI Scale	Cutoff Score	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV
MDQ positive screen (N = 713/15.1%) versus control (N = 4004)						
Openness	51.5	**0.60 (0.57-0.62)**	0.54	0.61	0.20	0.88
Conscientiousness	41.5	**0.57 (0.55-0.60)**	0.47	0.65	0.19	0.87
Extraversion	-19.5	0.51 (0.49-0.53)	0.38	0.65	0.16	0.85
Agreeableness	44.5	0.51 (0.49-0.54)	0.50	0.54	0.16	0.86
Neuroticism	48.5	**0.66 (0.64-0.69)**	0.65	0.59	0.22	0.90
HCL-16 positive screen (N = 595/12.6%) versus control (N = 4122)						
Openness	51.5	**0.58 (0.56-0.60)**	0.53	0.60	0.16	0.90
Conscientiousness	49.5	**0.53 (0.50-0.55)**	0.68	0.37	0.13	0.90
Extraversion	-6.5	**0.64 (0.62-0.67)**	0.53	0.68	0.19	0.91
Agreeableness	37.5	**0.54 (0.52-0.57)**	0.74	0.33	0.14	0.90
Neuroticism	45.5	**0.52 (0.50-0.55)**	0.56	0.50	0.14	0.89
QIDS-16 positive screen (N = 660/14%) versus control (N = 4057)						
Openness	47.5	**0.59 (0.57-0.61)**	0.67	0.48	0.17	0.90
Conscientiousness	43.5	**0.59 (0.57-0.61)**	0.56	0.58	0.18	0.89
Extraversion	-18.5	**0.64 (0.62-0.66)**	0.56	0.65	0.21	0.90
Agreeableness	52.5	0.50 (0.48-0.53)	0.22	0.82	0.16	0.87
Neuroticism	51.5	**0.81 (0.79-0.83)**	0.78	0.69	0.29	0.95
Composite score positive screen (N = 67/1.4%) versus control (N = 4650)						
Openness	45.5	**0.58 (0.52-0.65)**	0.78	0.39	0.02	0.99
Conscientiousness	42.5	**0.60 (0.53-0.67)**	0.60	0.59	0.02	0.99
Extraversion	-19.5	0.52 (0.44-0.59)	0.40	0.65	0.02	0.99
Agreeableness	48.5	0.48 (0.41-0.56)	0.36	0.68	0.02	0.99
Neuroticism	51.5	**0.77 (0.72-0.82)**	0.81	0.63	0.03	0.99

Note: PPV = Positive Predictive Value; NPV = Negative Predictive Value. Significant AUC values are shown in bold font.

Table 3. Results from the ROC curve analyses for the MDQ, HCL-16, QIDS-16 and the composite screening score for female participants

TSDI Scale	Cutoff Score	AUC (95% CI)	Sensitivity	Specificity	PPV	NPV
MDQ positive screen (N = 632/16.8%) versus control (N = 3121)						
Openness	45.5	**0.57 (0.54-0.59)**	0.58	0.52	0.20	0.86
Conscientiousness	39.5	**0.58 (0.55-0.60)**	0.36	0.77	0.24	0.86
Extraversion	16.5	0.50 (0.47-0.52)	0.95	0.06	0.20	0.83
Agreeableness	40.5	0.51 (0.49-0.54)	0.29	0.75	0.19	0.84
Neuroticism	59.5	**0.67 (0.65-0.69)**	0.62	0.63	0.26	0.89
HCL-16 positive screen (N = 688/18.3%) versus control (N = 3065)						
Openness	46.5	**0.56 (0.54-0.59)**	0.56	0.56	0.22	0.85
Conscientiousness	38.5	**0.55 (0.52-0.57)**	0.29	0.79	0.24	0.83
Extraversion	-9.5	**0.62 (0.59-0.64)**	0.65	0.54	0.24	0.87
Agreeableness	54.5	**0.53 (0.50-0.55)**	0.29	0.77	0.22	0.83
Neuroticism	59.5	**0.54 (0.52-0.57)**	0.47	0.60	0.21	0.84
QIDS-16 positive screen (N = 798/21.3%) versus control (N = 2955)						
Openness	45.5	**0.58 (0.54-0.59)**	0.58	0.53	0.25	0.82
Conscientiousness	44.5	**0.60 (0.58-0.62)**	0.52	0.63	0.27	0.83
Extraversion	-5.5	**0.61 (0.59-0.63)**	0.73	0.45	0.26	0.86
Agreeableness	35.5	0.51 (0.49-0.54)	0.20	0.85	0.27	0.80
Neuroticism	62.5	**0.77 (0.76-0.79)**	0.65	0.75	0.41	0.89
Composite score positive screen (N = 99/2.6%) versus control (N = 3654)						
Openness	44.5	**0.59 (0.54-0.65)**	0.69	0.47	0.03	0.98
Conscientiousness	38.5	**0.66 (0.60-0.72)**	0.47	0.78	0.06	0.98
Extraversion	3.5	0.51 (0.45-0.56)	0.27	0.77	0.03	0.98
Agreeableness	54.5	0.47 (0.41-0.53)	0.25	0.76	0.03	0.97
Neuroticism	68.5	**0.76 (0.71-0.80)**	0.58	0.80	0.04	0.99

Note: PPV = Positive Predictive Value; NPV = Negative Predictive Value. Significant AUC values are shown in bold font.

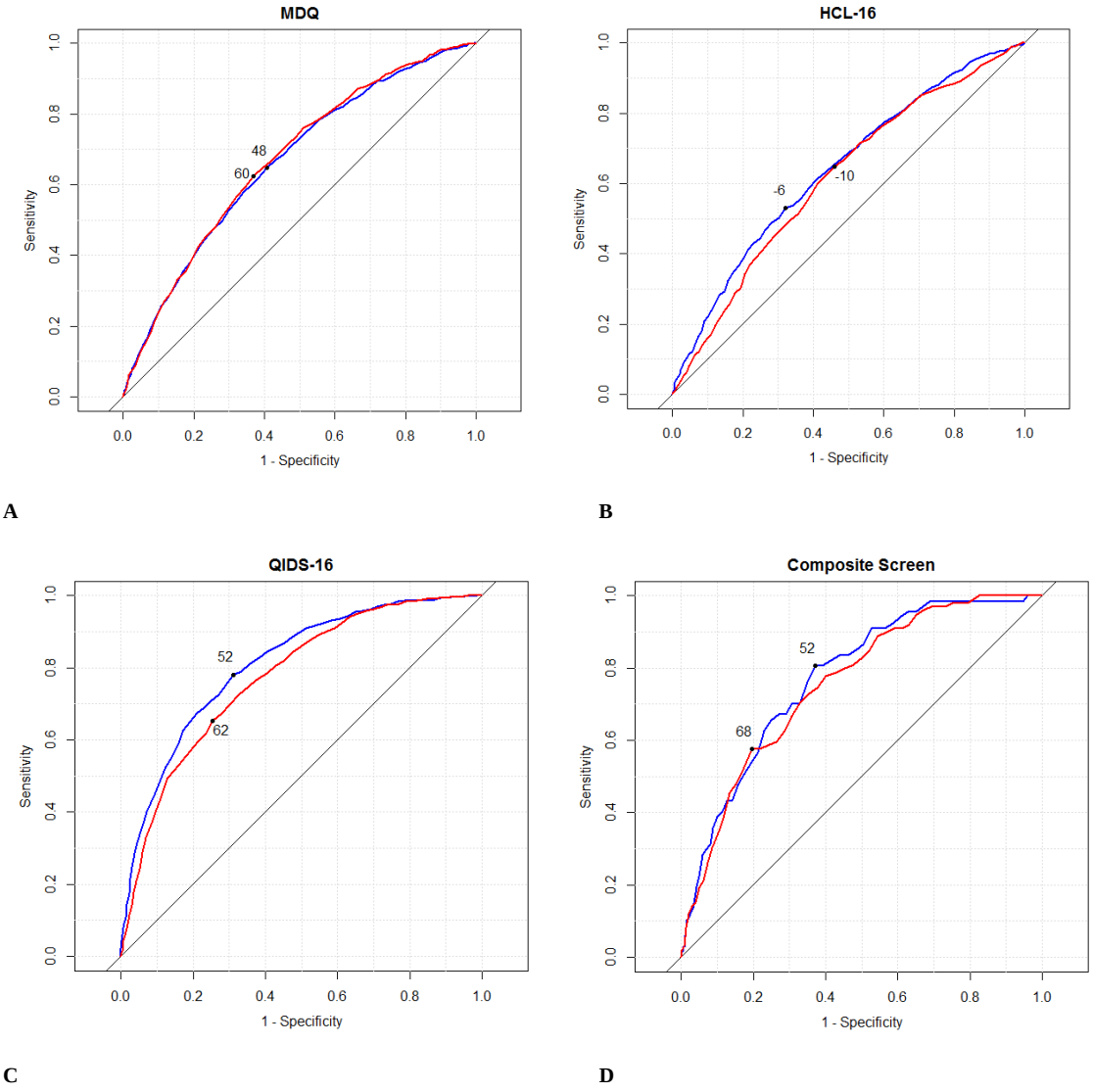

Figure 1. Receiver Operating Characteristic curves for males (in blue) and females (in red) showing neuroticism predicting the MDQ outcome (A), Extraversion predicting the HCL-16 outcome (B), neuroticism predicting the QIDS-16 outcome (C), and neuroticism predicting the composite screening outcome (D) .