

Article originally published in *Journal of Affective Disorders*. The final publication is available at Elsevier via <http://dx.doi.org/10.1016/j.jad.2016.09.014>

# Reconsidering the definition of Major Depression based on Collaborative Psychiatric Epidemiology Surveys

Tom Rosenström (PhD)<sup>\*,1</sup> and Markus Jokela (PhD)<sup>1</sup>

<sup>1</sup>Institute of Behavioural Sciences, University of Helsinki, Finland

\*Correspondence:

*Institute of Behavioural Sciences, University of Helsinki, Finland.*

*Siltavuorenpenger 1 A (PO Box 9), 00014, Helsinki, Finland.*

E-mail: [tom.rosenstrom@helsinki.fi](mailto:tom.rosenstrom@helsinki.fi)

## Abstract

**Background:** Diagnostic definitions for depressive disorders remain a debated topic, despite their central role in clinical practice and research. We use both recent evidence and nationally representative data to derive an empirically-based modification of DSM-IV/-5 Major Depressive Disorder (MDD).

**Method:** A modified MDD diagnosis was derived by analyzing data from Collaborative Psychiatric Epidemiology Surveys, a multistage probability sample of adults ( $n = 20\ 013$ ; age  $\geq 18$  years) in coterminous USA, Alaska and Hawaii. The old and the newly suggested MDD definitions were compared for their associated disability (WHO Disability Assessment Schedule and number of disability days in past month), suicide attempt, and other covariates.

**Results:** Our data-driven definition for major depression was "lack of interest to all or most things" plus four other symptoms from the set {weight gain, weight loss, insomnia, psychomotor retardation, fatigue, feelings of worthlessness, diminished ability to think/concentrate, suicidal ideation/attempt}. The new definition captured all the disability implied by MDD and excluded cases that showed no greater disability than the general population nor increased risk of suicide attempts. The lifetime prevalence of the new diagnosis was 14.7% (95% CI = 14–15.4%) of the population, slightly less than for the old definition (16.4%; CI = 15.4–17.3%).

**Limitations:** Only conservative modifications of MDD could be studied, because of restrictions in the symptom data.

**Conclusions:** With only small adjusting, the new definition for major depression may be more clinically relevant than the old one, and could serve as a conservative replacement for the old definition.

*Keywords:* Diagnostic definitions; data mining; disability; nosology; validity; clinical relevance

## Introduction

Clinical definitions for depressive disorders remain a debated topic, despite the high prevalence and burden of disability of these disorders. Recent studies have investigated the depressive disorders at the level of individual symptoms instead of composite clinical definition of Major Depressive Disorder (MDD)<sup>1</sup> (Bringmann et al., 2015; Cramer et al., 2012; Fried and Nesse, 2015; Keller et al., 2007; Keller and Nesse, 2005; Lux and Kendler, 2010; Oquendo et al., 2004). The symptom-level analysis is attractive for basic research because the definition of the MDD "syndrome" is not well established empirically (Haslam et al., 2012; Keller et al., 2007; Loo et al., 2012; Lux and Kendler, 2010; Solomon et al., 2001). At the same time, clinical practice requires criteria by which to assign individuals to treatment groups, and there is some evidence to support a temporal clustering of symptoms which is consistent with sudden phase transitions characteristic to syndromic states (Hosenfeld et al., 2015; van de Leemput et al., 2014). Accordingly, there is a challenge to provide empirically based answers to the question "when does depression become a mental disorder" (Maj, 2011a, 2011b). Treatment assignment based on single symptoms may be infeasible and reliable biomarkers identifying depression are lacking, implying that symptom combinations need to be considered in both research and treatment. This paper draws from the new symptom-level findings and uses United States national-level estimates to derive an empirically-based recommendation for a more clinically salient definition

---

<sup>1</sup> **Abbreviations:** MDD = DSM-IV Major Depressive Disorder; MD = an alternative definition for Major Depression to be derived; CPES = Collaborative Psychiatric Epidemiology Surveys; WHODAS = World Health Organization's Disability Assessment Schedule; WMH-CIDI = World Mental Health Survey Initiative's version of the Composite International Diagnostic Interview

for the Major Depression (MD; we drop the "Disorder" to distinguish between the suggested new ideas and the old definition, and to encompass both episodes and disorder).

The current Major Depressive Disorder (MDD) diagnosis requires the presence of at least one of the two core symptoms: (1) depressed mood and/or (2) markedly diminished interest or pleasure in all, or almost all, activities (American Psychiatric Association, 2013). However, it can be difficult to define depressed mood or sadness without referring to some other symptoms. For example, according to DSM-5 depressed mood is indicated "by either subjective report (e.g., feels sad, empty, hopeless) or observation made by others (e.g., appears tearful)" (American Psychiatric Association, 2013). But, how does one know when one is feeling sad? In April 18th 2016, Wikipedia defined sadness as "emotional pain associated with, or characterized by feelings of disadvantage, loss, despair, helplessness, disappointment and sorrow", which seems many things rather than one. In more quantitative terms, we have previously found that, after taking a sadness item into account, other self-report items provided little predictive value for the WHO-CIDI diagnosis of MDD (Rosenström et al., 2015); this is an expected finding when the "sadness" item implicitly implies multiple other symptoms. In a prospective network analysis of depressive symptoms, Bringman et al. (2015) showed that sadness has a high "indegree" but a low "outdegree" and "betweenness" in relation to other symptoms, suggesting that (statistically) it mostly summarizes other prevailing symptoms instead of predicting them. Those with a lot of problems (symptoms) are likely to become sad (endorse the symptom in future), but the sad who currently lack the other problems are relatively unlikely to get lot of problems in the future.

Among the depressive symptoms, anhedonia (low positive affect) has been found to be relatively specific to depression e.g. when comparing depression, anxiety, and schizophrenia,

whereas "depressed mood" represents "a mixture of relatively high NA [negative affect] and moderately low PA [positive affect]", therefore being less specific to depression and more related to general distress (Clark and Watson, 1991; Joiner et al., 2003). Ideally, constructing a data-driven definition should start from the most elementary (specific) components available rather than use variables that already are *a priori* given compositions of many elements. Furthermore, sadness is frequently seen as a normal, adaptive response to loss (Kleinman, 2012; Wakefield and First, 2012; Wakefield and Schmitz, 2013). In contrast, lack of interest in "all, or almost all, activities" should reliably intervene with goal-directed behaviors. Therefore, anhedonia may also be intrinsically more disabling than sadness.

Based on the findings cited above, we take only "diminished interest in all, or almost all, activities" as the core feature of depression in our analysis, as learning models of depression have done (Griffiths et al., 2014; Trimmer et al., 2015). For ease of reference, we call this "lack of interest" also as "anhedonia", although strictly speaking, it refers to lack of pleasure. We then study the following empirical questions: How do the other depressive symptoms distribute in the anhedonic population and what would be a sensible definition of depressive disorder based on that distribution? How much this new definition of depression overlaps with the old one? Which one of the definitions, the old or the new, is more clinically relevant? A diagnostic definition that implies higher level of disability, longer episodes, and greater probability of suicide attempts than an alternative definition is an example of comparatively "clinically relevant" diagnosis. These questions are analyzed here using a representative population sample of psychiatric symptoms in the United States, the World Health Organization's (WHO's) Collaborative Psychiatric Epidemiology Surveys (CPES) (Alegria et al., 2015).

## Methods

### Sample and procedures

CPES data joins together three multi-stage area probability samples, the National Comorbidity Survey Replication (NCS-R), the National Study of American Life (NSAL), and the National Latino and Asian American Study of Mental Health (NLAAS). Collection of the samples were funded by the National Institute of Mental Health and they were selected using the sampling frames and sample selection procedures that are common to the University of Michigan Survey Research Center's National Sample design, and they shared essential scientific objectives and survey instrumentation for mental health diagnostics. The joint sample design and sampling methods for the CPES data have been previously described (Heeringa et al., 2004) and the data was available to us via the Inter-university Consortium for Political and Social Research (ICPSR) service (Alegria et al., 2015). The joint sample represents all adults (i.e., age 18 years or more) residing in households in coterminous United States, Alaska and Hawaii, excluding institutionalized persons and those living on military bases (NCS-R and NSAL also excluded non-English speakers).

CPES is a "complex sample design" that allows for unbiased estimation of population statistics for the United States of America, but this requires survey-weighted estimation (Heeringa et al., 2004; Lumley, 2010). The CPES weights are products of a weight for unequal probability of selection, a weight for nonresponse, and a weight for post-stratification. The unequal probability of selection per selected individual results from a four-stage sampling process: a primary stage sampling of U.S. Metropolitan Statistical Areas and counties, followed by second-stage sampling of area segments, a third-stage sampling of housing units, and a fourth-

stage of random selection of eligible respondents from the housing units. The sample is post-stratified to a  $11 \times 12$  grid of population totals for geographic domain by race/ancestry (Vietnamese, Filipino, Chinese, All other Asian, Cuban, Puerto Rican, Mexican, All other Hispanic, Afro-Caribbean, African-American, White, All Other). The final CPES sample includes 20 013 individuals (8550 men and 11 463 women).

## Assessment instruments

The selected respondents were interviewed according to the World Mental Health Survey Initiative's version of the Composite International Diagnostic Interview (WMH-CIDI), which is a modified version of the original WHO-CIDI (Kessler and Üstün, 2004). Both WMH-CIDI and the other CPES questions were administered using a computer-assisted interview (Alegria et al., 2015). Although the WMH-CIDI allows for both 12-month and lifetime diagnoses, for simplicity, we concentrate on the data on lifetime diagnoses and on the presence *versus* absence of the depressive symptoms. A sensitivity analysis is conducted for a central disability outcome using 12-month diagnoses obtained by additionally requiring positive endorsement on the CPES variable "V00928" (an episode of being sad/or/discouraged/or/uninterested and having other problems during the past 12 months).

The presence of a symptom was determined exactly as in the "DSM-IV Major Depressive Episode" part of the ICPSR documentation for diagnostic algorithms (Alegria et al., 2015). For example, presence of anhedonia was inferred when the participant had answered positively to either one of the two diagnostic questions. The first question was "Think about the period of [(several days/two weeks)] or longer during that episode when your [(sadness/or/discouragement/or/lack of interest)] and other problems were most severe and

frequent. During that period of [(several days/two weeks)], did you lose interest in almost all things like work and hobbies and things you like to do for fun?". The second one was "Think about the period of [(several days/two weeks)] or longer during that episode when your [(sadness/or/discouragement/or/lack of interest)] and other problems were most severe and frequent. Did you feel like nothing was fun even when good things were happening?". Also the diagnosis of Major Depressive Episode was exactly as described in the diagnostic algorithm, with the exception that we did not use the bereavement exclusions in the E section of the document. The criterion has been highly controversial (Bolton et al., 2016; Kendler et al., 2008; Mojtabai, 2011; Shear et al., 2011; Wakefield and First, 2012; Wakefield and Schmitz, 2013; Zisook et al., 2007), and the focus of this study is on improving the diagnostic symptom combination. Nevertheless, it is possible that the removal of sadness from the MD criteria herein renders the bereavement exclusion a non-problem. Exclusions for episodes due to physical illness and/or lack of clinically significant impairment were in place, but mania and/or psychosis hierarchy exclusions were not used.

## **Covariates**

In addition to WMH-CIDI, we use the retrieved responses to the questions "Beginning yesterday and going back 30 days, how many days out of the past 30 were you totally unable to work or carry out your normal activities because of problems with either your physical health, your mental health, or your use of alcohol or drugs?" This variable ("V04452" in CPES codebook) is called "disability days" herein. In addition to the disability days, we use data on the World Health Organization's Disability Assessment Schedule (WHODAS). CPES data includes six variables that score disability from 0 (none) to 100 (max. disability) in the WHODAS domains of Cognition (the



variable "V08813"), Mobility ("V08814"), Self care ("V08812"), Social interaction ("V08818"), Role functioning or time out of role ("V08815"), and Participation or role impairment ("V08816") (Alegria et al., 2015). Herein, the WHODAS variable was the average of these six normalized disability indicators.

Disability days and WHODAS variables are indicators of clinical relevance. Another previously used indicator of relevance is persistence of the condition, which reflects temporal burden implied by the diagnosis (@ Wakefield and Schmitz, 2013). Very short episodes of a disorder may not be clinically relevant because the implied disability is so short-lived (cf. common cold) and because potential treatments might not work fast enough to provide benefits during the lifespan of the disorder. Therefore, we studied the length of the worst (variables "V00858" and "V00859"), the last (most recent; variables "V00861" and "V00862"), and the longest ("V00948" and "V00949") episode of being "sad/or/discouraged /or/uninterested" while also having "some of the other problems most of the day nearly every day".

As outcomes of depression diagnoses in Logistic regression, we studied the subjects' interview reports of having attempted suicide previously (value 1 in the variable "V02044"), having a low education (12 or less education years in age 21 or more; constructed from variables "V08172" and "V07306", resp.), being not married nor cohabiting (value 1 in "V08759"), feeling that one is a burden to others due to health problems (problems interfered with activities of close ones a lot or extremely; value 4 or 5 in "V06811"), being unhappy when taking all things together (value 4 in "V06870"), and having a depressed relative (value 1 or bigger in "V06929"). All these variables are potentially associated with depression based on well-known previous research, and here they provide some additional characterization of different diagnostic definitions.

## Statistical methods

CPES project distributes "pseudo" primary sampling unit and strata information (very similarly to NHANES III multistage-design) (Lumley, 2010)). Jointly these information are sufficient to provide good approximations to sampling errors via Taylor Series Linearization, as applied here (Binder, 1983; Lumley, 2010). Specifically, we used the statistical software R version 3.2.2 (2015-08-14), supplemented with the Survey package version 3.30-3 (Lumley, 2010, 2004). Survey-design objects were specified by inserting CPES variables "SECLUSTR" and "SESTRAT" to the "id" and "strata" arguments of the "svydesign" function, respectively (argument "nest" = TRUE). The "CPESWTSH" variable was the input to the "weights" argument, implying that we used all available cases in the CPES data; this is in contrast to the weights "CPESWTLG" that would select for the subset of NCS-R participants who participated in both Part I and Part II of the NCS-R (note that in CPES User Guide this latter variable is named as "CPESWGTS", but such variable cannot be found from the codebook or the data).

All analyses were based on survey-weighted estimates, as computed by the relevant methods in the Survey package. Specifically, all means, standard deviations, covariances, and correlations are survey-weighted population estimates, and the (varimax-rotated) exploratory factor analysis (Lawley and Maxwell, 1971) was based on the survey-weighted product-moment covariance matrix and maximum weighted likelihood estimation ("svyfactanal" function in the Survey package). Parallel analysis for number of factors to retain was used, (Horn, 1965; Humphreys and Montanelli Jr, 1975) as in other recent structural investigations (Fried et al., 2016; Uher et al., 2008), and implemented using "nFactors" R package, version 2.3.3, and by comparing the survey-weighted correlation matrix eigenvalues to mean eigenvalues of the

null/reference (no-correlations) distribution. Eigenvalues capture the amount of variation along principal components and are typically arranged in a descending order. The parallel-analysis reference distribution is generated using uncorrelated items and therefore the eigenvalues resulting from it are affected only by sampling error, not any 'underlying' factor. Thus, the real-data eigenvalues exceeding the reference distribution suggest underlying factors, the number of which corresponds to the running number of the last descending-order eigenvalue that still exceeds the equivalent reference. Parallel analysis has been found to perform well, and it is often chosen because classic (likelihood-ratio) tests for factor number have been found biased (Hayashi et al., 2007).

We use exploratory factor analysis strictly as a method for rough exploration of symptom correlations instead of modeling latent constructs; the suggested diagnostic measures were viewed as formative tools, as recommended in recent psychometric and philosophically oriented psychiatric research (Fried et al., 2016; Kendler, 2015). In a way of sensitivity analysis, or robustness analysis, we also provide an oblique bifactor rotation in the supplementary material using "GPArotation" R package, version 2014.11-1 (Jennrich and Bentler, 2012).

The quasi-Poisson regression coefficients are survey-weighted estimates. Poisson regression is an appropriate regression method for an outcome variable that is a count (Gelman and Hill, 2007), as for the number of disability days, and the quasi-Poisson specification was used as per instructions in the Survey package's help page for the "svyglm" function; it is also the proper specification for taking into account possible non-Poisson dispersion in the outcome variable (Gelman and Hill, 2007).

## Results

### A data-driven definition for MD

Altogether 82% of those who suffer from sadness or anhedonia, suffer from both. Overall, the prevalence of the other symptoms is at least as high in the anhedonic as in the sad people (Supplementary Table S1). Given these data and the arguments in the Introduction, the decision to limit the core symptoms to anhedonia seems reasonable and will be in effect from this point onwards.

While all the symptoms are correlated in the general population, many were virtually uncorrelated in the anhedonic population (see Supplementary Table S2). Parallel analysis indicated that four factors are needed to summarize the symptom data among the anhedonic population (Figure 1). The symptoms that represent two conditions (e.g., having insomnia or hypersomnia) are split into two separate symptoms, because there are no *a priori* reasons to believe that they represent similar etiologies (also previous studies have endorsed this choice) (Loo et al., 2012).

We conducted an explanatory factor analysis using survey-weighted covariance matrix and four latent dimensions (Table 1). The three first factors appeared to capture the self-contradictory aspects of the DSM-5 depressive symptoms (e.g., one cannot simultaneously be both insomnic and hypersomnic), being clear methods (or 'nuisance') factors as expected. However, explicit modeling of these factors may decrease bias in the remaining factors. Indeed, the final factor resembled the familiar concept of depressive syndrome: almost all symptoms loaded positively on it, except psychomotor agitation and hypersomnia (loading < 0.1). The

Supplementary Table S3 verifies that a similar interpretation held for an oblique bi-factor rotation. Bi-factor rotation re-interprets the factor model by trying to enforce one general factor and several non-overlapping secondary factors. However, hypersomnia and psychomotor retardation still had very weak loadings onto the general factor despite this attempt of enforcing a general syndrome for all the symptoms. The cut-off value 0.1 for a 'meaningful' factor loading, which is the default of the base package of R, may appear liberal in conventional psychometric terms. However, we are examining correlations conditioned on having anhedonia. Because all the traditional depressive symptoms are correlated with each other, the conditioning attenuates the correlations and factor loadings. Thus, what conventionally might be seen a liberal cut-off actually reflects our choice to be conservative in leaving out traditionally used depressive symptoms from the new definition.

To define MD both faithfully for the data and with minimal changes compared to DSM-5, we thus define the present MD construct as **anhedonia plus four other symptoms from the set {weight gain, weight loss, insomnia, psychomotor retardation, fatigue, feelings of worthlessness, diminished ability to think/concentrate, suicidal ideation/attempt}**. In other words, we leave out sadness as discussed in the Introduction, and psychomotor agitation and hypersomnia because, in the anhedonic population, they did not significantly load onto any shared factor with most of the other symptoms. We use anhedonia (a core symptom) plus four other symptoms (i.e., total of five) as the requirement for diagnosis, in an analogy to DSM-5. As in the old definition for MDE, the symptoms needed to be present during the same two-week period and the exclusions for episodes due to physical illness and/or lack of clinically significant impairment were in place (these requirements had little effect on the associations between diagnosis and disability; not shown).

## Prevalence and demographics of MD in the United States

Having found an alternative syndromal definition for MD, we proceeded to estimate its prevalence and associated disability. The lifetime prevalence of MD was 14.7% (95% CI = 14–15.4%) of the population, whereas the prevalence of the usual DSM-IV MDD was 16.4% (95% CI = 15.4–17.3%). Despite the similar prevalences, MD does not capture the exact same people as MDD, as evident from Table 2.

The Supplementary Table S4 displays a summary of the sample. Overall, those who would get a diagnosis of DSM-IV MDD without getting the (newly suggested) diagnosis of MD are closer to the general population in most characteristics than the MD and MDD groups. On average, those diagnosed with MDD but no MD have 1.87 symptoms less than those with both (Cohen's  $d = 1.89$ ,  $t = 35.7$ ,  $p < 0.001$ ).

## Disability associated with MD versus MDD

Among the two definitions, DSM-IV MDD and our new MD, the MD is much more strongly associated with disability days than MDD (Table 3). In fact, those DSM-IV MDD diagnoses that did not simultaneously satisfy the criteria for MD did not have a statistically significant effect on the number of disability days (Table 3). Furthermore, those who satisfied MDD criteria without satisfying the MD criteria did not significantly differ from the rest of the population in WHODAS score (Cohen's  $d = 0.015$ ,  $t = 0.323$ ,  $p = 0.747$ ), whereas those having MD without MDD did differ ( $d = 0.755$ ,  $t = 6.98$ ,  $p < 0.001$ ; Supplementary Table S4). In addition to Table 3, we checked analogous results for the group having both MD and MDD, finding that predicted disability is

between MDD and MD as expected (coefficient 0.573, CI = 0.323–0.823). Similarly, we verified that the average WHODAS score increased through the sub-groups having MDD without MD, MDD, MDD and MD, MD, and MD without MDD (3.297, 4.954, 5.557, 5.979, 8.09, resp.), showing a gradient of increasing disability when moving from DSM-5 MDD to the newly suggested MD definition.

Regarding effect sizes, the quasi-Poisson regression coefficient of 0.457 for the DSM-IV MDD diagnosis implies a 58% higher number of disability days per month in comparison to those not satisfying the MDD criteria. In comparison, the quasi-Poisson regression coefficient of 0.833 for the MD diagnosis implies a 130% higher number of disability days per month in comparison to those not satisfying the MD criteria. In other words, the rate effect of MD on disability days is 2.24 times that of the DSM-IV MDD (see Supplementary Table S4 for group averages).

The 12-month diagnoses imply even more disability than the lifetime diagnoses, because the disability days outcome refers to the past month (Table 3). The highest increase in disability days was associated with the MD diagnosis without MDD. This condition increased the number of disability days by 312%. The qualitative pattern of findings was similar to that of the lifetime diagnoses (Table 3). The rest of the paper will discuss lifetime diagnoses only.

## **Correlates of MD versus MDD**

We now continue the characterization of the MDD sufferers without MD, the MDD sufferers, the MD sufferers, and the MD sufferers without MDD, by assessing different types of disability and by using other clinically relevant variables than disability. Figure 2 shows how the diagnostic groups are represented in the WHODAS sub-scales and episode length, further characterizing the above

findings on disability, whereas the Figure 3 shows odds ratios for several other binary outcomes and for a few common DSM-IV disorders other than depressive disorders.

Most notably, MD sufferers are characterized by more impairment in the WHODAS domains than in the MDD sufferers, and a higher likelihood of previous suicide attempt. It is possible to show by explicit test within a multivariate logistic regression model e.g. that MDD without MD implies a decreased risk of suicide (OR = 0.371,  $p < 0.001$ ), whereas the other forms of depression in general imply an increased risk (OR = 2.35,  $p < 0.001$ ). Of course, this is readily evident from the Figure 3. In contrast, "unhappiness", feeling that one is a "burden to others", familiarity, and marriage status did not show marked differences across the diagnostic groups. MD sufferers without MDD were distinguished from the other groups by their higher likelihood of having low educational status, which may be related to their higher cognitive disability and time out of role (e.g. studying). Although we did not find drastic differences in episode length, those who had MD without MDD had longer longest depressive episode than those who had MDD without MD ( $d = 0.215$ ,  $t = 2.43$ ,  $p = 0.017$ ), and a similar but statistically non-significant trend was observed for the worst and the last episodes ( $d = 0.091$ ,  $t = 1.74$ ,  $p = 0.084$ , and  $d = 0.244$ ,  $t = 1.37$ ,  $p = 0.173$ , respectively).

MD had a higher level of comorbidity with other DSM-IV disorders than MDD, even after adjusting for the WHODAS score (Figure 3). Differences in comorbidity were most clear for substance use disorders, attention deficit disorder and panic disorder, but not so much for anxiety or post-traumatic stress. For example, MDD and MD had clearly non-overlapping confidence intervals for the attention deficit disorder (Figure 3); in that case, 7.15% (95% CI = 5.52–8.79%) of MDD cases had a lifetime comorbid attention deficit disorder, whereas the same



estimate for MD was 11% (95% CI = 8.85–13.1%). Altogether 52.9% (95% CI = 50.5–55.4%) of the MDD cases had had some of the other diagnoses studied in the Figure 3 during their lifetime, whereas 61.2% (95% CI = 58.9–63.5%) of the MD cases satisfied the criteria for some of the other non-depression lifetime diagnoses.

## Discussion

Here we propose an alternative, data-driven definition for MD that is similar in spirit to the well-known DSM-IV definition of MDD, but based on less conceptually overlapping symptom descriptions. The new MD is slightly less prevalent in the population than the familiar MDD, but significantly more associated with disability and suicide attempts than the MDD. In fact, those who satisfied MDD criteria without satisfying the MD criteria did not significantly differ from the rest of the US population in terms of disability days or WHODAS score. In this sense, the new MD definition is more pertinent to the question "when does depression become a disorder" than the old MDD definition (Maj, 2011a, 2011b; Wakefield and Schmitz, 2013).

Although our alternative definition of MD is partly data-driven, it is clear that it stands firmly on the shoulders of the old MDD definition. This conservative modification is not accidental. While it is important to study the depressive symptoms as dynamic systems (Bringmann et al., 2015; Cramer et al., 2012) and search for biomarkers (Kapur et al., 2012), neither of the approaches currently promises fast advances. High-dimensional nonlinear systems are difficult to analyze both theoretically and empirically, and some of the biomarkers for depression have turned out to be relevant only for sickness behaviors but not cognitive symptoms of depression (Jokela et al., 2015) or suffer from a very low predictive value (Fried and

Kievit, 2015). Neurovegetative symptoms associated with sickness behaviors have been found less strongly associated with clinically relevant characteristics of depression than cognitive symptoms (Lux and Kendler, 2010). Another practical argument for gradual progress is that the clinical community has shown resistance to drastic changes in nosology (Kendler, 2013). Furthermore, the recently endorsed "coherence theory of truth" in context of psychiatric nosology stresses that a new proposal should fit well with other things we know (Kendler, 2015), and such a fit is more difficult to demonstrate for definitions that have less to do with the existing pool of knowledge. We hope that the present definition of MD is both a conservative modification and a change for the better, and serves to keep the scientific method—a process of successive revisions—running despite the apparent standstill.

According to a longitudinal symptom-network analysis, sadness, a central part of the core symptom of "depressed mood" in the DSM-IV and DSM-5 definitions of MDD, does not appear to play a key role in the emergence of multiple depressive symptoms (Bringmann et al., 2015). Moreover, sadness may not directly imply pathology (Kleinman, 2012; Maj, 2011b; Wakefield and Schmitz, 2013). Further still, terms that reflect "depressed mood" (e.g., sad or blue) appear less specific to depression in comparison to anhedonia (Clark and Watson, 1991; Joiner et al., 2003). This study demonstrated that a definition without the sadness-including symptom better captures clinically relevant cases of depression. This new definition retained the other core symptom, anhedonia, in its original role, which is in line with the models of goal-directed behavior that have recently evoked much interest regarding psychiatric applications (Griffiths et al., 2014; Maia and Frank, 2011; Trimmer et al., 2015); for example, Griffiths et al. conclude: "Anhedonia, a common symptom in depression, maps closely onto deficits within outcome

valuation circuitry, and is the clearest example of how problems with reward value lead to reductions in goal-directed action" (Griffiths et al., 2014).

In this paper, we defined both the old and the new depression diagnosis without the bereavement exclusion criterion, because it has been very controversial (Bolton et al., 2016; Kendler et al., 2008; Kleinman, 2012; Mojtabai, 2011; Shear et al., 2011; Wakefield and Schmitz, 2013), and because it is not included in the most recent criteria (American Psychiatric Association, 2013). Also, the discussion on the bereavement criterion is not central for the present paper. We were interested in finding a better symptom combination for rapid clinical diagnosing of depression based on A criteria of DSM-IV and -5, and bereavement is not considered a depressive symptom in itself. Of course, future research could investigate whether or not our modified definition has any special role in that debate.

## **Limitations**

A major strength of the empirical estimates provided in this study is that they are representative of the entire US population. In contrast, generalizability to the other populations is a limitation and a topic for further study. However, there is evidence for the same underlying structure of major depression despite significant cultural differences (Kendler et al., 2015), and depressive disorders have been described throughout the written history (Beck, 1967). Likewise, we mainly restricted our analysis to lifetime diagnoses and future studies could explore 12-month diagnoses, or even immediate reactions to life events. It was established, however, that the findings on disability days followed the same qualitative pattern with the same implications for diagnostic revisions both for those with lifetime depression diagnosis and for those with episode also within the past 12 months (Table 3). In addition, the present study draws from the data

collected for prevalence estimation within the old diagnostic system, and the paper is therefore limited to the variables that were perceived as relevant for the DSM-IV system. However, clearly there was still room for exploration in these data, as we could derive the novel MD definition.

Although the proposed new diagnostic definition for MD has the undeniable strength of excluding MDD cases without observable disability, the new definition may have the drawback of increased psychiatric comorbidity. At first sight, this is an obvious outcome of the higher overall level of disability in MD cases than in the MDD cases, but the finding persisted despite adjusting for the WHODAS disability score. Given these preliminary findings, future studies could investigate the novel MD definition in the wider context of the total diagnostic system.

Although not a limitation regarding the comparison between MD and MDD, the present study is generally limited by the fact that full symptom data was available only from individuals who had endorsed either anhedonia or sadness criterion, or both. This does not limit analyses in the anhedonic population, and our MD is automatically excluded if one does not belong to that sub-population. Therefore, diagnostic information was available for the entire sample. However, it is possible (even likely) that explorative analyses using wider range of symptoms and other etiologic information will eventually achieve even better diagnostic constructs in comparison to the present ones. We hope the present analysis encourages such efforts in representative large datasets, and also point out that researchers have frequently worked with a larger range of symptoms than is available in DSM-5 (Beck, 1967).

In summary, this study found strong support for re-defining the major depression as "lack of interest to all or most things" plus four other symptoms from the set {weight gain, weight loss, insomnia, psychomotor retardation, fatigue, feelings of worthlessness, diminished ability to

think/concentrate, suicidal ideation/attempt}. This new definition captured all the disability implied by DSM-IV Major Depressive Disorder and excluded cases that show no greater disability than the general population. The analysis was representative of the population of the United States.

## References

Alegria, M., Jackson, J.S., Kessler, R.C., Takeuchi, D., 2015. Collaborative Psychiatric Epidemiology Surveys (CPES), 2001-2003 [United States]. ICPSR20240-v8. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2015-12-09.

<http://doi.org/10.3886/ICPSR20240.v8>.

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. American Psychiatric Association, Arlington, VA.

Beck, A.T., 1967. Depression: Clinical, Experimental, and Theoretical Aspects. University of Pennsylvania Press, Philadelphia, Pennsylvania.

Binder, D.A., 1983. On the variances of asymptotically normal estimators from complex surveys. *Int. Stat. Rev.* 51, 279–292. doi:[10.2307/1402588](https://doi.org/10.2307/1402588)

Bolton, J.M., Au, W., Chateau, D., Walld, R., Leslie, W.D., Enns, J., Martens, P.J., Katz, L.Y., Logsetty, S., Sareen, J., 2016. Bereavement after sibling death: a population-based longitudinal case-control study. *World Psychiatry* 15, 59–66. doi:[10.1002/wps.20293](https://doi.org/10.1002/wps.20293)

Bringmann, L.F., Lemmens, L.H.J.M., Huibers, M.J.H., Borsboom, D., Tuerlinckx, F., 2015. Revealing the dynamic network structure of the Beck Depression Inventory-II. *Psychol. Med.* 45, 747–757.

doi:[10.1017/S0033291714001809](https://doi.org/10.1017/S0033291714001809)

Clark, L.A., Watson, D., 1991. Tripartite model of anxiety and depression: psychometric evidence and taxonomic implications. *J. Abnorm. Psychol.* 100, 316–336.

Cramer, A.O.J., Borsboom, D., Aggen, S.H., Kendler, K.S., 2012. The pathoplasticity of dysphoric episodes: differential impact of stressful life events on the pattern of depressive symptom inter-correlations. *Psychol. Med.* 42, 957–965. doi:[10.1017/S003329171100211X](https://doi.org/10.1017/S003329171100211X)

Fried, E.I., van Borkulo, C.D., Epskamp, S., Schoevers, R.A., Tuerlinckx, F., Borsboom, D., 2016.

Measuring depression over time ... or not? Lack of unidimensionality and longitudinal measurement invariance in four common rating scales of depression? *Psychol. Assess.* Epub ahead of print, 1–14. doi:[10.1037/pas0000275](https://doi.org/10.1037/pas0000275)

Fried, E.I., Kievit, R.A., 2015. The volumes of subcortical regions in depressed and healthy individuals are strikingly similar: a reinterpretation of the results by Schmaal et al. *Mol. Psychiatry*, 21, 724–725. doi:[10.1038/mp.2015.199](https://doi.org/10.1038/mp.2015.199)

Fried, E.I., Nesse, R.M., 2015. Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. *BMC med.* 13, 72. doi:[10.1186/s12916-015-0325-4](https://doi.org/10.1186/s12916-015-0325-4)

Gelman, A., Hill, J., 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, New York.

Griffiths, K.R., Morris, R.W., Balleine, B.W., 2014. Translational studies of goal-directed action as a framework for classifying deficits across psychiatric disorders. *Front. Syst. Neurosci.* 8.

doi:[10.3389/fnsys.2014.00101](https://doi.org/10.3389/fnsys.2014.00101)

Haslam, N., Holland, E., Kuppens, P., 2012. Categories versus dimensions in personality and psychopathology: a quantitative review of taxometric research. *Psychol. Med.* 42, 903–920.

doi:[10.1017/S0033291711001966](https://doi.org/10.1017/S0033291711001966)

Hayashi, K., Bentler, P.M., Yuan, K.-H., 2007. On the likelihood ratio test for the number of factors in exploratory factor analysis. *Struct. Equ. Modeling* 14, 505–526.

doi:[10.1080/10705510701301891](https://doi.org/10.1080/10705510701301891)

Heeringa, S.G., Wagner, J., Torres, M., Duan, N., Adams, T., Berglund, P., 2004. Sample designs and sampling methods for the Collaborative Psychiatric Epidemiology Studies (CPES). *Int. J. Meth. Psychiatric Res.* 13, 221–240. doi:[10.1002/mpr.179](https://doi.org/10.1002/mpr.179)

Horn, J.L., 1965. A rationale and test for the number of factors in Factor Analysis. *Psychometrika* 30, 179–185.

Hosenfeld, B., Bos, E.H., Wardenaar, K.J., Conradi, H.J., van der Maas, H.L.J., Visser, I., de Jonge, P., 2015. Major depressive disorder as a nonlinear dynamic system: bimodality in the frequency distribution of depressive symptoms over time. *BMC psychiatry* 15, 222. doi:[10.1186/s12888-015-0596-5](https://doi.org/10.1186/s12888-015-0596-5)

Humphreys, L.G., Montanelli Jr, R.G., 1975. An investigation of the parallel analysis criterion for determining the number of common factors. *Multivariate Behav. Res.* 10, 193–205.

- Jennrich, R.I., Bentler, P.M., 2012. Exploratory bi-factor analysis: the oblique case. *Psychometrika* 77, 442–454. doi:[10.1007/s11336-012-9269-1](https://doi.org/10.1007/s11336-012-9269-1)
- Joiner, T.E., Brown, J.S., Metalsky, G.I., 2003. A test of the tripartite model's prediction of anhedonia's specificity to depression: patients with major depression versus patients with schizophrenia. *Psychiatry Res.* 119, 243–250. doi:[10.1016/S0165-1781\(03\)00131-8](https://doi.org/10.1016/S0165-1781(03)00131-8)
- Jokela, M., Virtanen, M., Batty, G.D., Kivimäki, M., 2015. Inflammation and specific symptoms of depression. *JAMA Psychiatry* 73, 87–88. doi:[10.1001/jamapsychiatry.2015.1977](https://doi.org/10.1001/jamapsychiatry.2015.1977)
- Kapur, S., Phillips, A.G., Insel, T.R., 2012. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol. Psychiatry* 17, 1174–1179. doi:[10.1038/mp.2012.105](https://doi.org/10.1038/mp.2012.105)
- Keller, M.C., Neale, M.C., Kendler, K.S., 2007. Association of different adverse life events with distinct patterns of depressive symptoms. *Am. J. Psychiatry* 164, 1521–1529. doi:[10.1176/appi.ajp.2007.06091564](https://doi.org/10.1176/appi.ajp.2007.06091564)
- Keller, M.C., Nesse, R.M., 2005. Is low mood an adaptation? Evidence for subtypes with symptoms that match precipitants. *J. Affect. Disord.* 86, 27–35. doi:[10.1016/j.jad.2004.12.005](https://doi.org/10.1016/j.jad.2004.12.005)
- Kendler, K.S., 2015. Toward a limited realism for psychiatric nosology based on the coherence theory of truth. *Psychol. Med.* 45, 1115–1118. doi:[10.1017/S0033291714002177](https://doi.org/10.1017/S0033291714002177)
- Kendler, K.S., 2013. A history of the DSM-5 scientific review committee. *Psychol. Med.* 43, 1793–1800. doi:[10.1017/S0033291713001578](https://doi.org/10.1017/S0033291713001578)
- Kendler, K.S., Aggen, S.H., Li, Y., Lewis, C.M., Breen, G., Boomsma, D.I., Bot, M., Penninx, B.W.J.H., Flint, J., 2015. The similarity of the structure of DSM-IV criteria for major depression in depressed



women from China, the United States and Europe. *Psychol. Med.* 45, 1945–1954.

doi:[10.1017/S0033291714003067](https://doi.org/10.1017/S0033291714003067)

Kendler, K.S., Myers, J., Zisook, S., 2008. Does bereavement-related Major Depression differ from Major Depression associated with other stressful life events? *Am. J. Psychiatry* 165, 1449–1455.

doi:[10.1176/appi.ajp.2008.07111757](https://doi.org/10.1176/appi.ajp.2008.07111757)

Kessler, R.C., Üstün, T.B., 2004. The World Mental Health (WMH) Survey Initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *Int. J. Meth. Psychiatric Res.* 13, 93–121. doi:[10.1002/mpr.168](https://doi.org/10.1002/mpr.168)

doi:[10.1002/mpr.168](https://doi.org/10.1002/mpr.168)

Kleinman, A., 2012. Culture, bereavement, and psychiatry. *The Lancet* 379, 608–609.

doi:[10.1016/S0140-6736\(12\)60258-X](https://doi.org/10.1016/S0140-6736(12)60258-X)

Lawley, D.N., Maxwell, A.E., 1971. *Factor Analysis as a Statistical Method*, 2nd ed. Butterworths & Co, London.

Lumley, T., 2010. *Complex Surveys: A Guide to Analysis Using R*, 1 ed. Wiley, Hoboken, N.J.

Lumley, T., 2004. Analysis of complex survey samples. *J. Stat. Softw.* 9, 1–19.

doi:[10.18637/jss.v009.i08](https://doi.org/10.18637/jss.v009.i08)

Lux, V., Kendler, K.S., 2010. Deconstructing major depression: a validation study of the DSM-IV symptomatic criteria. *Psychol. Med.* 40, 1679–1690. doi:[10.1017/S0033291709992157](https://doi.org/10.1017/S0033291709992157)

Maia, T.V., Frank, M.J., 2011. From reinforcement learning models to psychiatric and neurological disorders. *Nat. Neurosci.* 14, 154–162. doi:[10.1038/nn.2723](https://doi.org/10.1038/nn.2723)

Maj, M., 2011a. Refining the diagnostic criteria for major depression on the basis of empirical evidence. *Acta Psychiatr. Scand.* 123, 317. doi:[10.1111/j.1600-0447.2011.01680.x](https://doi.org/10.1111/j.1600-0447.2011.01680.x)

Maj, M., 2011b. When does depression become a mental disorder? *Br J Psychiatry* 199, 85–86. doi:[10.1192/bjp.bp.110.089094](https://doi.org/10.1192/bjp.bp.110.089094)

Mojtabai, R., 2011. Bereavement-related depressive episodes: Characteristics, 3-year course, and implications for the DSM-5. *Arch. Gen. Psychiatry* 68, 920–928. doi:[10.1001/archgenpsychiatry.2011.95](https://doi.org/10.1001/archgenpsychiatry.2011.95)

Oquendo, M.A., Barrera, A., Ellis, S.P., Li, S., Burke, A.K., Grunebaum, M., Endicott, J., Mann, J.J., 2004. Instability of symptoms in recurrent major depression: a prospective study. *Am. J. Psychiatry* 161, 255–261. doi:[10.1176/appi.ajp.161.2.255](https://doi.org/10.1176/appi.ajp.161.2.255)

Rosenström, T., Elovainio, M., Jokela, M., Pirkola, S., Koskinen, S., Lindfors, O., Keltikangas-Järvinen, L., 2015. Concordance between Composite International Diagnostic Interview and self-reports of depressive symptoms: a re-analysis. *Int. J. Meth. Psychiatric Res.* 24, 213–225. doi:[10.1002/mpr.1478](https://doi.org/10.1002/mpr.1478)

Shear, M.K., Simon, N., Wall, M., Zisook, S., Neimeyer, R., Duan, N., Reynolds, C., Lebowitz, B., Sung, S., Ghesquiere, A., Gorscak, B., Clayton, P., Ito, M., Nakajima, S., Konishi, T., Melhem, N., Meert, K., Schiff, M., O'Connor, M.-F., First, M., Sareen, J., Bolton, J., Skritskaya, N., Mancini, A.D., Keshaviah, A., 2011. Complicated grief and related bereavement issues for DSM-5. *Depress. Anxiety* 28, 103–117. doi:[10.1002/da.20780](https://doi.org/10.1002/da.20780)

Solomon, A., Haaga, D.A.F., Arnow, B.A., 2001. Is clinical depression distinct from subthreshold depressive symptoms? A review of the continuity issue in depression research. *J. Nerv. Ment. Dis.* 189, 498–506.

Trimmer, P.C., Higginson, A.D., Fawcett, T.W., McNamara, J.M., Houston, A.I., 2015. Adaptive learning can result in a failure to profit from good conditions: implications for understanding depression. *Evol. Med. Public Health* 2015, 123–135. doi:[10.1093/emph/eov009](https://doi.org/10.1093/emph/eov009)

Uher, R., Farmer, A., Maier, W., Rietschel, M., Hauser, J., Marusic, A., Mors, O., Elkin, A., Williamson, R.J., Schmael, C., Henigsberg, J., Perez, J., Mendlewicz, J., Janzing, J.G.E., Zobel, A., Skibinska, M., Kozel, D., Stamp, A.S., Bajs, M., Placentino, A., Barreto, M., McGuffin, P., Aitchison, K.J., 2008. Measuring depression: comparison and integration of three scales in the GENDEP study. *Psychol. Med.* 38, 289–300.

van de Leemput, I.A., Wichers, M., Cramer, A.O.J., Borsboom, D., Tuerlinckx, F., Kuppens, P., van Nes, E.H., Viechtbauer, W., Giltay, E.J., Aggen, S.H., Derom, C., Jacobs, N., Kendler, K.S., van der Maas, H.L.J., Neale, M.C., Peeters, F., Thiery, E., Zachar, P., Scheffer, M., 2014. Critical slowing down as early warning for the onset and termination of depression. *Proc. Natl. Acad. Sci. U S A* 111, 87–92. doi:[10.1073/pnas.1312114110](https://doi.org/10.1073/pnas.1312114110)

van Loo, H.M., de Jonge, P., Romeijn, J., Kessler, R.C., Schoevers, R.A., 2012. Data-driven subtypes of major depressive disorder: a systematic review. *BMC Med.* 10, 156. doi:[10.1186/1741-7015-10-156](https://doi.org/10.1186/1741-7015-10-156)

Wakefield, J.C., First, M.B., 2012. Validity of the bereavement exclusion to major depression: does the empirical evidence support the proposal to eliminate the exclusion in DSM-5? *World Psychiatry* 11, 3–10. doi:[10.1016/j.wpsyc.2012.01.002](https://doi.org/10.1016/j.wpsyc.2012.01.002)

Wakefield, J.C., Schmitz, M.F., 2013. When does depression become a disorder? Using recurrence rates to evaluate the validity of proposed changes in major depression diagnostic thresholds.

World Psychiatry 12, 44–52. doi:[10.1002/wps.20015](https://doi.org/10.1002/wps.20015)

Zisook, S., Shear, K., Kendler, K.S., 2007. Validity of the bereavement exclusion criterion for the diagnosis of major depressive episode. World Psychiatry 6, 102–107.

## Tables

**Table 1. Factor loadings of symptoms in the anhedonic population**

	<b>Factor 1</b>	<b>Factor 2</b>	<b>Factor 3</b>	<b>Factor 4</b>
Weight loss	-0.049	<b>0.981</b>	-0.008	<b>0.173</b>
Weight gain	0.054	<b>-0.608</b>	-0.021	<b>0.132</b>
Insomnia	<b>-0.777</b>	0.093	0.024	<b>0.215</b>
Hypersomnia	<b>0.994</b>	-0.030	-0.007	0.075
Psychomotor retardation	-0.060	0.035	<b>-0.299</b>	<b>0.362</b>
Psychomotor agitation	-0.057	0.038	<b>0.995</b>	0.016
Fatigue	0.063	-0.025	<b>-0.113</b>	<b>0.351</b>
Feelings of worthlessness or guilt	0.010	-0.019	0.033	<b>0.363</b>
Diminished ability to think or concentrate	-0.078	-0.009	0.007	<b>0.441</b>
Suicidal ideation	-0.068	0.019	0.013	<b>0.270</b>

Note: Varimax-rotated loadings from a 4-factor Exploratory Factor Analysis in the anhedonic population (sample n = 3443, estimated population total = 40 818 803). Loadings above 0.10 in absolute value are highlighted with bold font.

**Table 2. A contingency table for MDD and MD**

	<b>No MD</b>	<b>Yes MD</b>	<b>Marginal</b>
<b>No MDD</b>	168 957 187 (81.0%)	5 494 157 (2.6%)	174 451 344 (83.6%)
<b>Yes MDD</b>	8 831 969 (4.2%)	25 267 305 (12%)	34 099 274 (16.4%)
<b>Marginal</b>	177 789 156 (85.2%)	30 761 462 (14.8%)	208 550 618 (100%)

Note: A contingency table of estimated MDD (old definition) and MD (data-driven definition) population totals. Respective percentages are given in the parentheses. Notice that the sum of marginals is not exactly equal with the estimated population total, 209 500 125 (SE = 10 569 621) individuals, because of the sampling error in the distinct estimates (sum of domain estimates for cell totals vs. direct estimate of population total). The survey-weighted estimates are based on a sample of 19 820 participants with diagnostic data available.

**Table 3. Quasi-Poisson regression coefficients when predicting number of disability days**

Variable	Model							
	I	II	III	IV	V	VI	VII	VIII
Age	0.01*** (0.01, 0.01)	0.01*** (0.01, 0.02)	0.01*** (0.01, 0.01)	0.01*** (0.01, 0.01)	0.01*** (0.01, 0.02)	0.01*** (0.01, 0.02)	0.01*** (0.01, 0.01)	0.01*** (0.01, 0.01)
Male sex	-0.12 (-0.29, 0.05)	-0.08 (-0.25, 0.09)	-0.17* (-0.32, -0.01)	-0.16* (-0.32, -0.004)	-0.10 (-0.26, 0.06)	-0.08 (-0.23, 0.08)	-0.16* (-0.32, -0.01)	-0.16* (-0.32, -0.01)
MDD	0.46*** (0.27, 0.64)				0.98*** (0.76, 1.20)			
MD		0.83*** (0.56, 1.11)				1.28*** (1.05, 1.51)		
MDD w/o MD			-0.10 (-0.45, 0.25)				0.01 (-0.43, 0.45)	
MD w/o MDD				1.16*** (0.90, 1.41)				1.42*** (1.15, 1.68)
Constant	0.09 (-0.17, 0.35)	-0.08 (-0.39, 0.22)	0.26* (0.04, 0.48)	0.14 (-0.10, 0.37)	-0.0002 (-0.26, 0.26)	-0.13 (-0.39, 0.14)	0.25* (0.02, 0.48)	0.15 (-0.08, 0.38)
Observations	15 565	15 571	15 565	15 571	15 565	15 571	15 565	15 571

\*p<0.05; \*\*p<0.01; \*\*\*p<0.001

Note: Quasi-Poisson regression coefficients when predicting number of disability days per month. Four models are shown, depending on one of the independent variables being MDD (Major Depressive Disorder; Model I) diagnosis, MD (the newly-suggested Major Depression; Model II) diagnosis, MDD without MD (Model III), or MD without MDD (Model IV). The columns 5–8 (Models V to VIII) show the same models put for 12-month instead of lifetime diagnoses. The 95% confidence intervals are in parentheses.

## Figure captions

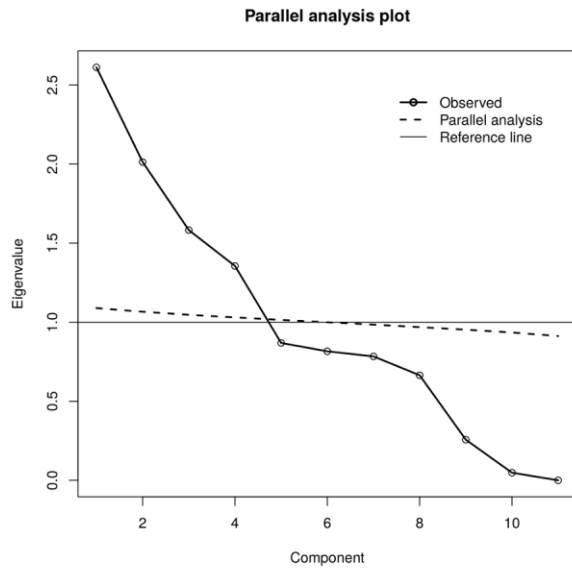
*Figure 1: Parallel analysis plot in the anhedonic population (Sample  $n = 3443$ , corresponding estimated population total = 40 818 803). Parallel analysis line shows the expected eigenvalues under no correlations, but taking sampling variance into account. Reference line at the value 1 shows the expectation without any sampling variance. The estimated number of factors is the number of descending-order eigenvalues above the parallel-analysis line (i.e., 4).*

*Figure 2: Disability scores in the diagnostic groups. Forest plot of average WHODAS sub-scale score and average episode length in non-depressed individuals [no MD(D)], the Major Depressive Disorder sufferers without the new Major Depression diagnosis (MDD w/o MD), the MDD sufferers, the MD sufferers, the MD sufferers without MDD (MD w/o MDD). Whiskers denote 95% confidence intervals. The episode length is not shown for the "no MD(D)" condition, because these individuals have no episodes.*

*Figure 3: Binary outcomes in the diagnostic groups. Forest plot of odds ratios for DSM-IV disorders and other binary outcomes from their regression on the Major Depressive Disorder (MDD) sufferers without the new Major Depression (MD) diagnosis (MDD w/o MD), the MDD sufferers, the MD sufferers, and the MD sufferers without MDD (MD w/o MDD). Odds ratios for the DSM-IV disorders are adjusted for WHODAS disability score to avoid inflated comorbidity due to the previously shown differences in general disability. All diagnoses are lifetime diagnoses. Whiskers denote 95% confidence intervals.*

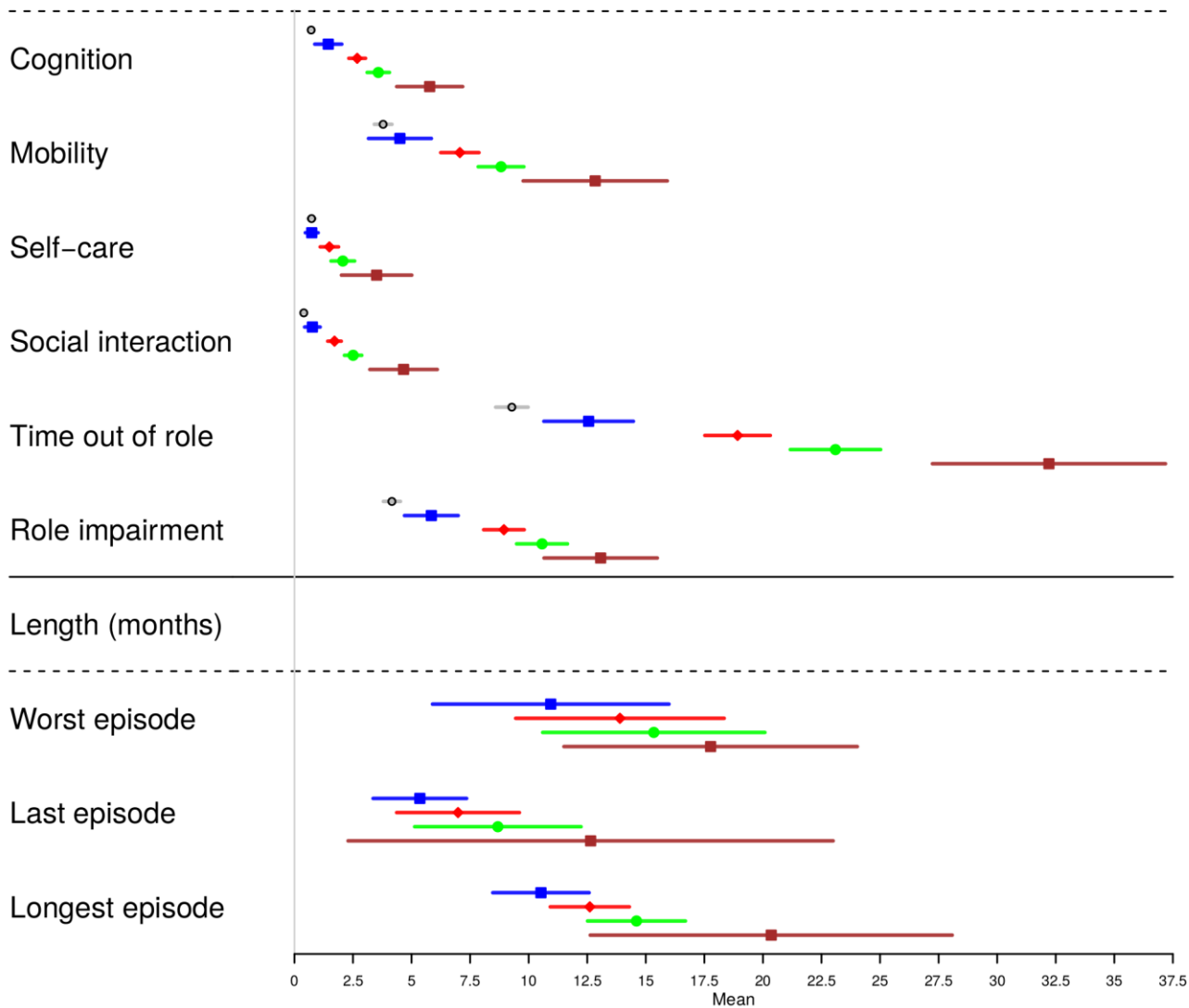


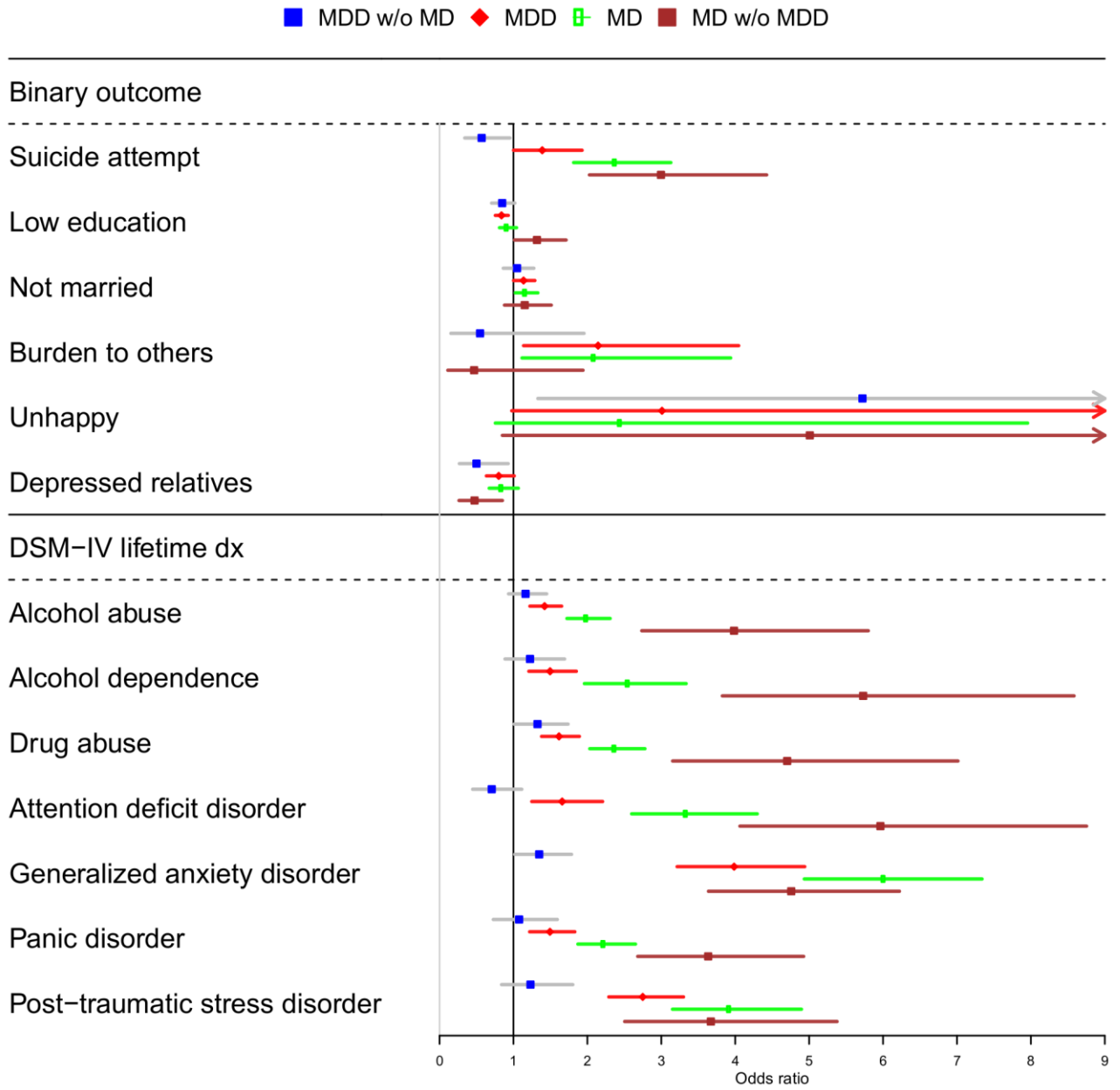
## Figures



○ no MD(D) ■ MDD w/o MD ◆ MDD ● MD ■ MD w/o MDD

WHODAS score





## Supplementary Tables:

# Reconsidering the definition of Major Depression based on Collaborative Psychiatric Epidemiology Surveys

Tom Rosenström (PhD)<sup>\*,1</sup> and Markus Jokela (PhD)<sup>1</sup>

<sup>1</sup>Institute of Behavioural Sciences, University of Helsinki, Finland

\* Correspondence:

*Institute of Behavioural Sciences, University of Helsinki, Finland.*

*Siltavuorenpenger 1 A (PO Box 9), 00014, Helsinki, Finland.*

E-mail: [tom.rosenstrom@helsinki.fi](mailto:tom.rosenstrom@helsinki.fi)

**Supplementary Table S1. Symptom prevalence**

	<b>Total prevalence</b>	<b>Prevalence in anhedonia</b>	<b>Prevalence in depressed mood</b>
Depressed mood	0.2290	0.979	1.000
Loss of interest	0.1950	1.000	0.835
Weight loss	0.1500	0.666	0.650
Weight gain	0.0318	0.143	0.135
Insomnia	0.1720	0.757	0.744
Hypersomnia	0.0349	0.156	0.149
Psychomotor retardation	0.0883	0.418	0.382
Psychomotor agitation	0.0239	0.103	0.105
Fatigue	0.1930	0.877	0.832
Feelings of worthlessness or guilt	0.0957	0.459	0.417
Diminished ability to think or concentrate	0.1980	0.890	0.856
Suicidality	0.1570	0.708	0.678
Sub-group sample size	20 013	3 443	4 000
Estimated population total	209 500 125	40 818 803	47 891 758

Note: Symptom prevalence in the general (1st column), the anhedonic (2nd col), and the sad or low-mood (3rd col) population. The two last rows show the analytic sample size and the estimate for the population total, respectively.

**Supplementary Table S2. Correlations between the depressive symptoms in anhedonic people.**

	<b>mood</b>	<b>wloss</b>	<b>wgain</b>	<b>insomnia</b>	<b>hypersomnia</b>	<b>retardation</b>	<b>agitation</b>	<b>fatigue</b>	<b>worthlessness</b>	<b>thinking</b>	<b>suicidal</b>
mood	1.000	0.140	-0.290	0.160	-0.220	-0.012	-0.062	-0.053	0.130	0.063	0.170
wloss	0.140	1.000	-0.920	0.250	-0.220	0.097	0.016	-0.072	-0.053	0.015	0.030
wgain	-0.290	-0.920	1.000	-0.220	0.170	-0.066	-0.120	0.027	0.056	-0.032	-0.068
insomnia	0.160	0.250	-0.220	1.000	-0.980	0.130	0.085	-0.057	-0.037	0.210	0.150
hypersomnia	-0.220	-0.220	0.170	-0.980	1.000	-0.110	-0.150	0.089	0.024	-0.230	-0.160
retardation	-0.012	0.097	-0.066	0.130	-0.110	1.000	-0.660	0.160	0.100	0.290	-0.037
agitation	-0.062	0.016	-0.120	0.085	-0.150	-0.660	1.000	-0.390	-0.068	-0.210	-0.034
fatigue	-0.053	-0.072	0.027	-0.057	0.089	0.160	-0.390	1.000	0.016	0.130	-0.130
worthlessness	0.130	-0.053	0.056	-0.037	0.024	0.100	-0.068	0.016	1.000	0.042	0.290
thinking	0.063	0.015	-0.032	0.210	-0.230	0.290	-0.210	0.130	0.042	1.000	-0.033
suicidal	0.170	0.030	-0.068	0.150	-0.160	-0.037	-0.034	-0.130	0.290	-0.033	1.000

Note: sample n = 3443, estimated population total = 40 818 803.

**Supplementary Table S3. Oblique bi-factor loadings in the anhedonic population**

	<b>Factor 1</b>	<b>Factor 2</b>	<b>Factor 3</b>	<b>Factor 4</b>
Weight loss	<b>0.3590</b>	<b>0.9300</b>	-0.0059	-0.0066
Weight gain	0.0118	<b>-0.6210</b>	-0.0190	-0.0231
Insomnia	<b>0.2500</b>	0.0106	0.0163	<b>0.7700</b>
Hypersomnia	0.0406	0.0047	0.0076	<b>-0.9980</b>
Psychomotor retardation	<b>0.3670</b>	-0.0296	<b>-0.2960</b>	0.0680
Psychomotor agitation	0.0119	-0.0011	<b>0.9970</b>	0.0015
Fatigue	<b>0.3400</b>	-0.0867	<b>-0.1080</b>	-0.0625
Feelings of worthlessness or guilt	<b>0.3520</b>	-0.0903	0.0382	-0.0175
Diminished ability to think or concentrate	<b>0.4330</b>	-0.0990	0.0125	0.0701
Suicidal ideation	<b>0.2700</b>	-0.0376	0.0152	0.0610

Note: Oblique bi-factor analysis in the anhedonic population. The "bifactorQ" function from the "GPArotation" R package was applied to the factor loadings and the varimax rotation matrix of the main manuscript. This rotation method tries to find a rotation where each symptom loads onto a main factor plus one and only one other factor. The results align closely with the varimax-based interpretations of the main manuscript, with the exception of factor reordering and that the bi-factor model would further excluded weight gain from depressive symptoms. Since exploratory oblique bi-factor rotation estimates are technically challenging and may involve local maxima, we erred for a conservative choice to keep the weight-gain criterion in the diagnostic set for MD.

**Supplementary Table S4. Summary of the sample and the diagnostic groups**

	<b>All</b>			<b>MD</b>		<b>MDD</b>		<b>MDD w/o MD</b>	
Sub-population total		209 500 125		30 761 462		34 099 274		8 831 969	
<b>Variable</b>	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>Mean</b>	<b>SD</b>	<b>Mean</b>	<b>SD</b>	<b>Mean</b>	<b>SD</b>
Age (years)	20013	44.6	17.4	42.2	14.4	42.8	15.1	43.9	16.2
Male	20013	0.481	0.50	0.365	0.481	0.371	0.483	0.435	0.496
WHODAS	10871	3.200	8.83	5.98	9.57	4.95	8.45	3.30	6.27
Disability days	15571	1.88	6.11	3.39	7.89	2.36	7.00	1.72	5.81
Symptom: mood	20013	0.229	0.420	0.990	0.099	0.990	0.101	0.986	0.117
Symptom: anhedonia	20013	0.195	0.396	1.000	0.000	0.894	0.308	0.589	0.492
Symptom: weight loss	20013	0.15	0.357	0.721	0.448	0.68	0.467	0.586	0.493
Symptom: weight gain	20013	0.032	0.176	0.158	0.365	0.151	0.358	0.112	0.316
Symptom: insomnia	20013	0.172	0.378	0.833	0.373	0.775	0.418	0.624	0.485
Symptom: hypersomnia	20013	0.035	0.183	0.133	0.339	0.163	0.369	0.237	0.426
Symptom: mot. retard.	20013	0.088	0.284	0.498	0.500	0.416	0.493	0.209	0.407
Symptom: mot. agit.	20013	0.024	0.153	0.105	0.306	0.100	0.301	0.118	0.322
Symptom: fatigue	20013	0.193	0.395	0.927	0.261	0.887	0.317	0.770	0.421
Symptom: worthlessness	20013	0.096	0.294	0.546	0.498	0.442	0.497	0.193	0.395
Symptom: thinking	20013	0.198	0.399	0.960	0.196	0.909	0.288	0.777	0.417
Symptom: suicidal	20013	0.157	0.363	0.790	0.407	0.712	0.453	0.535	0.499

Note: Summary of the sample and the diagnostic groups. N refers to the raw number of non-missing observations for the indicated variable overall, whereas the Mean and Standard deviation (SD) refer to survey-weighted population estimates for that variable in the diagnostic group indicated in the sup-headings of the pertinent columns. Abbreviations: "WHODAS" = World Health Organization Disability Assessment Schedule -based score; "no MD(D)" = sub-group of those not satisfying MD nor MDD criteria; "MD" = sub-group of people satisfying the (new) Major Depression criteria; "MDD" = sub-group satisfying the (old) Major Depressive Disorder criteria; "MDD w/o MD" = sub-group of people satisfying the (old) MDD criteria without satisfying the (new) MD criteria. "All" refers to the entire sample/US population.