The varying burden of depressive symptoms across adulthood : Results from six NHANES cohorts

García-Velázquez, Regina

2019-03-01

García-Velázquez , R , Jokela , M & Rosenstrom , T H 2019 , ' The varying burden of depressive symptoms across adulthood : Results from six NHANES cohorts ' , Journal of Affective Disorders , vol. 246 , pp. 290-299 . https://doi.org/10.1016/j.jad.2018.12.059

http://hdl.handle.net/10138/325155
https://doi.org/10.1016/j.jad.2018.12.059

submittedVersion

Highlights

- Prevalence of depressive symptoms and functional impairment decoupled across adulthood.
- Age moderated the association between depressive symptoms and high impairment.
- Middle-aged adults were more likely to report high impairment.
- Three individual symptoms showed age-specific patterns of impairment.

Abstract

Background. Depressive symptoms differ from each other in the degree of functional impairment they cause. The incidence of depression varies across the adult lifespan. We examined whether age moderates the impairment caused by depressive symptoms.

Methods. The study sample (n= 21,056) was adults drawn from six multistage probability samples from the National Health and Nutrition Examination Survey series (NHANES, years 2005 to 2016) conducted in the United States using cross-sectional, representative cohorts. Depressive symptoms were assessed with the nine-item Patient Health Questionnaire (PHQ-9). We used regression models to predict high functional impairment, while controlling for sociodemographic variables and physical disorders.

Results. Age moderated the association between depressive symptoms and functional impairment: middle-aged adults perceived moderate and severe symptoms as more impairing than did others. Older adults reported slightly higher impairment due to mild symptoms. The individual symptoms of low mood, feelings of worthlessness and guilt, and concentration difficulties were more strongly related to high impairment in mid-adulthood as compared to early and late adulthood.

Limitations. Cross-sectional data allows only between-person comparisons. The PHQ-9 is brief and joins compound symptoms into single items. There was no information available concerning comorbid mental disorders. Co-occurring physical disorders were self-reported.

Conclusions. Symptoms of depression may imply varying levels of impairment at different ages. The results suggest a need for age adjustments when estimating the functional impact of depression in the general population. Additionally, they show a need for more accurate assessments of depression-related impairment at older ages. Evidence-based programs may generally benefit from symptom- and age-specific findings.

The Varying Burden of Depressive Symptoms Across Adulthood: Results from Six NHANES Cohorts.

R. García-Velázquez[a*], M. Jokela[a], T.H. Rosenström[a,b]

[a] Department of Psychology and Logopedics, University of Helsinki.

[b] Department of Mental Disorders, Norwegian Institute of Public Health.

* Correspondence author:

Department of Psychology and Logopedics, University of Helsinki.

Haartmaninkatu 3, E215. P.O. Box 21, 00014 University of Helsinki. Finland

Phone: +358 (0)40 93 73 933. E-mail: regina.garciavelazquez@helsinki.fi

Word count: 4818.

Keywords: depression; functional impairment; adulthood; symptom analysis; cross-sectional.

Depression is a common disorder that ranks among the top causes of global disability (Whiteford et al., 2013; World Health Organization, 2017). Unsurprisingly, this has motivated a large body of research aimed at characterizing depression. Epidemiological studies have shown that Major Depressive Disorder (MDD) reaches peak prevalence during midlife. This trajectory is widely documented in large cross-sectional (Jorm, 2000; Blanchflower & Oswald, 2016) and longitudinal datasets (Sutin et al., 2013; Cheng, Powdthavee, & Oswald, 2017), and particularly in western countries (Le Bon & Le Bon, 2014; Steptoe et al., 2015). However, a key question remains underexplored: Is depression equally impairing across the adult lifespan?

There is evidence that there is age-specific variance in the prevalence of not only depression sum-scores, but also of individual depressive symptoms (Hegeman, Kok, van der Mast, & Giltay, 2012; Hegeman, de Waal, Comijs, Kok, & van der Mast, 2015; Schaakxs, Comijs, Lamers, Beekman, & Penninx, 2017). It seems that younger patients diagnosed with MDD are more prone to affective and cognitive symptoms, while somatic symptoms are more common in older patients. The different age trends found in clinical samples of depressed individuals, however, cannot be directly generalized to the community at large due to the possibility of different correlation structures (Foster & Mohler-Kuo, 2018). Thus, it remains unclear whether the age patterns seen in clinical samples are found in the adult general population (i.e., along the whole spectrum of depressive symptoms).

Depressive symptoms are heterogeneous in more aspects than prevalence. According to clinical and community studies, it appears that individual depressive symptoms vary in risk factors, biomarkers, and responses to antidepressants (Fried & Nesse, 2015a; Jokela et al., 2016). Importantly, symptoms also differ in their association with functioning (Fried & Nesse, 2014). This is to some extent understandable, given that depressive symptoms expand across several domains of human experience (i.e., somatic, affective-motivational, and cognitive). Thus, studies have addressed independently how

symptom prevalence varies across age groups, and how the disabling effects of depressive symptoms vary. A wider approach is necessary to tackle the question of whether age has implications for the impairment caused by depressive symptoms.

There are reasons to think that depressive symptoms may show different patterns of impact across the adult lifespan. The first reason has to do with contextual factors: symptoms could interfere with people's daily routines differently depending on the person's life stage. Several life transitions, such as social role changes when entering adulthood, may relate to the burden caused by depressive symptoms. For instance, fatigue or concentration problems could be more taxing for younger adults who are in their early work careers and/or taking care of young children, compared to older adults for whom work life and parenthood do not impose the same demands. A second reason for looking into age patterns of depression is physical deterioration, which challenges the assessment of depression in older adults (Haigh et al., 2018). Some symptoms, such as difficulties with sleep and concentration, are well-known correlates of ageing (Mander et al., 2017) and at the same time are diagnostic criteria for MDD (APA, 2013). Additionally, chronic disease may cause symptoms similar to those of depression (Molarius & Janson, 2002). Studying the impairment associated with individual symptoms is fundamental given the interplay of comorbid chronic disease, deterioration, and different pathways to depression in ageing populations.

In order to better characterize the burden of depressive symptomatology across the adult lifespan, we (1) examined the prevalence of aggregated and individual symptoms of depression across age groups in a representative sample of the general population of the United States, and (2) inspected whether age group plays a role in the associations between symptoms of depression and difficulties with normal life activities. This provides useful information for understanding age-related needs and for accurately calibrating the evaluation of depression-related functional impairment at different ages.

We used data from the National Health and Nutrition Examination Surveys (NHANES), which consists of cross-sectional measurements of a nationally representative sample of adults in the U.S. (CDC, 2017).

Methods

Participants

Participants were from the NHANES (CDC, 2017) for 2005-2006 (n=5,334), 2007-2008 (n=5,995), 2009-2010 (n=6,360), 2011-2012 (n=5,615), 2013-2014 (n=5,924), and 2015-2016 (n=5,735) with a total of 34,963 participants who were on average 46.24 years old (sample weighted, SE=.112). The NHANES protocol selected the samples to represent the U.S. adult population. Depressive symptoms have been assessed for all the adult participants in the NHANES since 2005. The samples are new for every study cohort, and thus there are no repeated observations.

Our analytic sample included only the participants who reported having depressive symptoms, and therefore could report the impairment caused by them. Put differently, one cannot report the impairment caused by symptoms he or she does not suffer from. The reader can observe in the next section that the wording of the item measuring functional impairment was clear in this respect (i.e., *if you checked off any problems,* how difficult have *those problems* made it for you…). Thus, the structure of the questionnaire defined our analytic sample in what comes to the analysis of functional impairment. Additionally, the initial sample had missing questionnaire data in each cross-sectional cohort (between 9% and 12.5%). The size of our analytic sample was n=21,056. The descriptive information is based on the entire sample, and thus the prevalence of symptoms, for example, is to be interpreted in the context of the general population.

Measures

In the NHANES protocol depressive symptoms and functional impairment were assessed by questionnaire and by a single item, respectively, in computer-assisted personal interviews. Depressive symptoms were assessed using the nine item Patient Health questionnaire (PHQ-9). The questionnaire items query how often the participant had been suffering from concrete depressive symptoms during the last 2 weeks, each self-rated on a 4-point response scale (0=Not at all, 1=Several days, 2=More than half the days, 3=Nearly every day). The nine symptoms correspond to the MDD diagnostic criteria in the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), and are as follows: anhedonia, low mood, problems with sleep, low energy or fatigue, changes in appetite, feelings of worthlessness or guilt, concentration difficulties, psychomotor alterations, and thoughts of self-harm or death.

The level of difficulty in normal life was measured with the question "If you checked off any problems, how difficult have those problems made it for you to do your work, take care of things at home, or get along with other people?", also rated on a 4-point Likert scale (0=Not difficult at all, 1=Somewhat difficult, 2=Very difficult, 3=Extremely difficult). This single item is considered a measure of functional impairment for which there is substantial evidence of convergent validity with a number of variables measuring quality of life, functional status, and use of health care services (Kroenke & Spitzer, 2002; Spitzer et al., 2000, 1999). In our analyses, this variable was dichotomized by merging together the two lower and the two higher values and coding it as 0 or 1, the latter indicating that it was *very* or *extremely difficult* to carry out normal activities. This variable will be referred to as high functional impairment. The sum-score of the PHQ-9 questionnaire was calculated by summing the nine items together, resulting in a variable ranging between 0 and 27. *Age* was reported in years and top coded at 80 to preserve the anonymity of participants.

The covariates controlled for were as follows: *gender* (reported as "Male" or "Female"); *race/ethnicity,* recoded into four categories (1=Mexican American, 2=other Hispanic, 3=non-Hispanic white, 4=non-Hispanic black, or 5=Other including multiracial); *marital status* including six categories (1=married, 2=widowed, 3=divorced, 4=separated, 5=never married, 6=living with partner); *ratio of family income to poverty level* (dichotomized with a threshold at 1, informing whether the income of a household is below or above the poverty level); a count of self-reported *physical medical conditions* (diabetes, heart disease, stroke, pulmonary disease, and cancer); and NHANES *sampling year*. The epidemiological variables we controlled for were selected based on availability in the NHANES data and on previous literature showing that they have an association with major depression (e.g., Kessler & Essex, 1982; Moussavi et al., 2007; Haushofer & Fehr, 2014).

Statistical analyses

The scoring system of the PHQ-9 can be interpreted based on the cut-off points of 5, 10, 15, and 20, indicative of mild, moderate, moderately severe, and severe levels of depressive symptoms (Kroenke & Spitzer, 2002; Kroenke et al., 2010). These categories are used for aiding the interpretation of the results when the PHQ-9 score is examined, but the PHQ-9 score was introduced in our analyses as a ratio scale. We first conducted preliminary analyses to examine potential nonadditive effects of age and the PHQ-9 scores when predicting functional impairment.

A significant interaction was found between the PHQ-9 score and age in predicting high impairment, thus we defined four age categories to discriminate between young adulthood, midlife, and postretirement age for further study: 18 to 30 years, 31 to 50 years, 51 to 65 years, and 66 years and older. This allowed us to examine differential patterns of association between depression and impairment across life stages. We present this model as Step 1.

We performed logistic regression-based differential item functioning (DIF) analysis. DIF analyses assess whether some items (symptoms) behave differently across population subgroups (e.g., age groups) when adjusting for possible differences in the underlying "trait" (depression). We also conducted DIF analyses for gender in order to consider its potential bias on response distributions. DIF was examined by comparing nested regression models where the grouping variable (i.e., sex or age group) was included as a predictor of item responses, first only as an additive effect (known as uniform DIF) and then as an interaction effect (nonuniform DIF). An improvement in the fit from one model to another indicates DIF, since the grouping variable contributes to explaining the responses to the item. The criterion for DIF-flagging was a change in Nagelkerke's pseudo $R^2 \geq 0.02$ (Gelin & Zumbo, 2003). For a detailed review of our DIF approach see Choi (2016).

We provide different descriptive tables showing the prevalence of individual depressive symptoms according to age group. In order to test whether the differences between age groups were statistically significant, we performed chi-square tests and calculated the effect sizes with Cramer's V following the formula $\sqrt{\chi^2/n \times df}$, where χ^2 is the chi-square statistic and *df* is the minimum number of categories minus one (Cramèr, 1946). In this case *df*=3, since both age group and the PHQ-9 items had four different categories. We considered the effect size to be small when surpassing .06, to be medium moderate when at least .17, and to be large when reaching .30.

We analyzed the association of all individual symptoms with high functional impairment in a series of stepwise models. First, nine logistic regression models were fitted (Step 2), one for each of the nine PHQ-9 symptoms including its interaction with age group. These models were adjusted for all covariates, but not for other PHQ-9 symptoms. In Step 3 we estimated the fully adjusted model where all PHQ-9 symptoms and interactions with age group were collapsed into one single model, and thus

the regression estimates of the influence of individual symptoms on high functional impairment were adjusted for the presence of the other symptoms plus covariates.

We provide Akaike's information criterion (AIC), which penalizes for model complexity (Burnham & Anderson, 2002). AIC was used for comparing the models in Step 1 and 3. A lower AIC is indicative of a better fit or higher parsimony. Nagelkerke's pseudo R^2 was calculated, which is asymptotically independent of the sample size and can be interpreted as the proportion of the outcome variation explained by the predictors. Pseudo R^2 values range between 0 and 1 (Nagelkerke, 1991).

The analyses were conducted using R 3.4.0 software (R Core Team, 2017) and the packages *lordif* (Choi, 2016), and *survey* (Lumley, 2014). Sampling weights were used to achieve population-representative estimates in all the analyses presented. Sampling weights for the NHANES data (variable *wtmec2yr*) were used according to the guidelines provided by the NHANES project (CDC, 2015). A demonstration code file of our analyses can be found from the Supplementary materials.

Results

Distribution of depressive symptoms according to age group

Survey-weighted descriptive information on the sample is presented in Table 1. The distribution of the PHQ-9 severity scores (i.e., based on cut-offs) and of the functional impairment item were similar to those found in other studies using western community samples, but overall milder (Kocalevent et al., 2013; Rief et al., 2004). Depression scores rose with age, and declined approximately after age 55 (Figure 1).

The frequencies of individual symptom endorsement per age group are found in Figure 2 (see Figure S1 in the Supplement to see the distribution according to age in years). The severity of symptoms was milder compared to other general population findings, but such data are, however,

scarce for individual PHQ-9 items (Rief et al., 2004). The symptoms did not show any meaningful age-prevalence pattern. The chi-square tests of association between age group and symptom responses were statistically significant due to the large sample size, but the effect size according to Cramer's V was less than .056 for all PHQ-9 items, which is negligible. Thus, the age groups showed very similar symptom prevalence patterns.

We also inspected the subset of the sample reporting high functional impairment (n= 1,032). Chi-square tests and effect sizes revealed no age-group patterns in individual symptom prevalence (chi-square tests were statistically significant due to large sample size, but effect sizes registered values between .015 and .041, which is considered negligible). The reader can find a more detailed characterization of symptom prevalence in the subsample reporting high functional impairment in Figure S2, where the distributions are shown without the age grouping. Note this information is merely descriptive, and confounding variables may contribute to the distributional patterns.

Regression models for high functional impairment

In preliminary analyses there was an interaction effect between depression sum-score and age in predicting high functional impairment, which remained present throughout several sensitivity analyses. DIF analyses indicated that neither age group nor gender were a source of DIF in the PHQ-9 items (Supplement, Tables S1 and S2). At this point gender did not register a statistically significant effect (p-value $\leq$.05) in any of the models from Step 1 to 3, either at the aggregate level or at the symptom-specific level. From here on we will elaborate on the relationship of PHQ-9 depression sum-scores and high functional impairment across the four age groups, with no distinction between genders.

The model in Step 1 shows the statistically significant interaction effect of age group and depression sum-score on functional impairment (Table S3). The interaction can be observed in Figure 3, where the association between PHQ-9 depression sum-scores and reported functional impairment

varies in strength (i.e., slope) for the different age groups. For depression sum-scores between 1 and 9, corresponding to mild depression, older adults were slightly more likely to report high impairment than were younger adults. Starting from moderate scores onward adults aged over 65 became gradually less likely than middle-aged adults to report high impairment. For instance, at a score of 8, the probability of reporting high impairment was 6% for adults over 65 and 4% for those aged 31 to 50. At a score of 20 the probability of feeling highly impaired 53% in the age group over 65 and 67% in the 31 to 50 year age group. The coefficients and details of the models tested in every step are shown in Supplementary Table S3. The predictor sum-score alone accounted for the 85% of the predictive power in the model in Step 1 (Nagelkerke's R^2=.269 for a model with the depression sum-score as only predictor), the remainder being attributable to age and interactions plus covariates.

All separate regression models testing the effect of individual symptoms on high functional impairment showed a statically significant main effect, and also some significant interactions between the symptom and age group (Step 2). The AIC for the symptom-level, fully adjusted model (Step 3) was lower than that of the depression sum-score fully adjusted model (Step 1), which suggests that the model including the individual symptoms explains the data better that the one with the depression sum-score. The effects in Step 3 revealed that when adjusting for all other symptoms, age group moderated the relationship between high functional impairment and three symptoms: low mood, feelings of worthlessness and guilt, and concentration difficulties (interactions displayed in Figure 4). The effects of the rest of the symptoms remained uniform across age groups.

To quantify the specific contribution of each PHQ-9 symptom on high functional impairment, a series of Nagelkerke's R^2 were calculated: we calculated the model in Step 3 by excluding one symptom at a time, and then compared the proportion of change in R^2 with respect to that of the full

model (R^2=.331). Table 2 presents the percentage of loss in R^2 when each symptom was excluded. The magnitude of each unique contribution was, however, very similar and rather small.

After identifying three individual symptoms for which the association to high impairment was moderated by age group, we performed a sensitivity analysis. We calculated a sum-score for the PHQ-9 excluding these three symptoms (i.e., a sum score based on six items) and then calculated a regression model identical to that in Step 1. The interaction effects of this new sum-score variable with age were not statistically significant, suggesting that the interaction of the PHQ-9 sum-score with age is driven by the three symptoms (Supplementary Table S4). We found the same for age group and for age in years.

Discussion

The results of our analysis suggest that reported functional impairment attributed to depressive symptoms is not a simple function of aggregated depression scores. The association we found was instead moderated by age, so that moderate to severe depression was perceived as more impairing by middle-aged adults as compared to adults in early adulthood and postretirement age. Adults of postretirement age were slightly more likely to report high impairment due to mild depressive symptoms, but as scores increased from moderate to severe, the likelihood to report high impairment decreased compared to other age groups (Figure 3). The youngest age group (18-31 years) was in general less likely to report feeling high functional impairment as compared to middle-aged adults (age groups of 31 to 50 and 51 to 65 years). In addition, symptom-level analyses revealed a significant role of age in the functional impairment attributed to the specific symptoms of low mood, concentration difficulties, and feelings of worthlessness and guilt. These associations remained after adjusting for the other PHQ-9 individual symptoms and for covariates such as chronic physical disorders, marital status,

and living below the poverty threshold, which are known to relate to clinical depression (e.g., Kessler & Essex, 1982; Moussavi et al., 2007; Haushofer & Fehr, 2014).

This study implemented models both at the level of depression sum-score and individual symptoms, obtaining very similar results. The goodness of fit registered by both approaches was also very similar, with fit indices slightly favoring the symptom-level approach (Table S3). Thus, according to our results, depression sum-scores functioned efficiently as a proxy for all the individual symptoms in what comes to predicting functional impairment. However, examining symptoms individually offered a deeper understanding of depression as a syndrome because only three specific symptoms showed age-related patterns of functional impairment (Table 2). They explained the interaction effect observed between aggregated depression sum-scores and age (Table S4). This information was concealed as a single number in the sum-score.

Symptom-level research is important also because different depression scales have different content, which implies that a sum-score calculated from one scale may contain different symptoms than a sum-score derived from a different scale (Fried, 2017). The findings obtained using sum-scores of different depression scales may not be directly comparable. Symptom-level research has the potential to reveal valuable information for making informed decisions about MDD classification or treatment. For example, diagnostic criteria that are only weakly associated with relevant outcomes could be revised. Nosological theories may well be tested or developed based on symptom-level evidence. Treatment guidelines could benefit from identifying the most impairing symptoms. MDD has been pointed out as a problematic diagnosis by many researchers for being simultaneously under- and overdiagnosed, and for responding inconsistently to treatment (Fried, 2015; Lorenzo-Luaces, 2015; Maj, 2011; Wakefield & Schmitz, 2017). This may relate to the heterogeneous symptom combinations being collapsed to same diagnostic outcomes (Fried & Nesse, 2015b; Østergaard et al., 2011), which

motivates symptom-based approaches. Therefore, we analyzed functional impairment with respect individual diagnostic symptoms, not just their sums.

The impairing effects of concentration difficulties are well documented in studies on clinical depression (Gonda et al., 2015; Lam et al., 2014). This is because concentration difficulties and indecisiveness, define the DSM criterion of MDD most related to cognitive impairment. Research has found that impaired attention and executive functioning are likely trait-markers of depression, meaning that they predispose a person to MDD and remain present after remission (Lee et al., 2012). In line with this, a recent study examining working-aged patients with MDD found that concentration difficulties related to health-related quality of life independently of the severity of depression (Fattori et al., 2017). Our finding that disrupted ability to concentrate is a more impairing symptom at ages of 31 to 65 years in the general population supports the hypothesis that adults living in the "rush hour of life" are most exposed to environmental demands, and therefore may feel most impaired by not being able to perform at their best. Indeed, younger and middle-aged adults report encountering minor stressors more frequently, and perceiving them as more severe, than do older adults (Charles et al., 2010; Carstensen et al., 2011; Charles et al., 2016).

A study comparing symptom presentation in MDD with onset before and after age 60 found that two symptoms predict earlier onset: feelings of worthlessness and guilt, and depressed mood (Heun et al., 2000). These findings suggest this symptom may have clinical implications starting at middle age, which is consistent with the effect on high impairment we found. Another study found that subjects who had experienced earlier onset depression were more likely after the age of 65 to report feelings of worthlessness and guilt in the last month (Gallagher et al., 2009). This is a symptom worthy of particular attention as well for being associated with complicated depression, and found to predict concurrent and post-remission suicide attempts in clinical samples (Wakefield & Schmitz, 2016, 2017).

Low mood, on the other hand, is one of the most prevalent symptoms in epidemiological studies of general and clinical samples alike (Tebeka et al., 2018). This is to some extent expected since it is one of the two required diagnostic criteria for MDD. Low mood was the most impairing symptom in a study by Fried and Nesse (2014) which examined the functional impact of individual symptoms in a clinical sample. There were three findings in common with those of our study: first, a large weight of low mood on functional impairment in both community and clinical samples (compared to other symptoms); second, the nonsignificant effect of gender on impairment; and third, the significant effect of age. A second, similar study conducted by Tweed (1993), found that concentration difficulties and low mood, together with other symptoms, statistically predict both concurrent and post-recovery impairment. The comparability of these results to ours is also affected by the age range of the samples. Fried and Nesse's sample was similar to the community sample we used in terms of age distribution. The age range in Tweed's study was not reported, but the mean age was similar.

Thus, our study reinforces some previous findings from observations on clinical samples, and brings novel results suggesting that age moderates the effects of depressive symptoms on functional impairment. Another novel result was that, regardless of the varying size of regression weights, the effect of the symptoms on high impairment was conjoint or common (i.e., symptom-specific contributions were rather small). The outcome we used was a single item querying functional impairment specifically caused by the depressive symptoms, in contrast to other studies which used wider, generic impairment measures. The PHQ item should be a more valid indicator of depression-related functional impairment than are indicators assessing unspecific disability, particularly in community samples which are rather heterogeneous (i.e., general population samples may present important comorbidity with other disorders, which is reduced in clinical samples by exclusion criteria).

Sum-scores of symptom ratings on depression inventories are generally used as a proxy for depression severity. We found that the depression sum-scores in this cross-sectional sample population reached a peak among those just older than age 50 and then decreased, which is generally consistent with findings from earlier studies (Stone et al., 2010; Blanchflower & Oswald, 2016; Cheng et al., 2017; Schaakxs et al., 2017). However, the effects on impairment were most noticeable for the age group of 31 to 50 years; age group 51 to 65 years followed a similar but less pronounced trend. If our findings on self-reported functional impairment generalize to objective disability, it would mean that the age-composition of the population should be taken into account when analyzing the impact of depression. The predictive power of the models was mostly due to the depressive symptoms alone (about 85% in both models). However, epidemiological factors may have small effect sizes and be theoretically important for understanding etiological mechanisms, or may exert a meaningful confounding effect (Kraemer et al., 2001).

The phenomenon that severe depressive symptoms were less strongly related to functional impairment at older ages is in agreement with socioemotional selectivity theory, which accounts for the improvement commonly observed in wellness-related indicators at older ages (Carstensen et al., 2003; Haigh et al., 2018). It postulates that as adults age, they tend to select and evocate more positive stimuli in favor of emotional self-regulation. Accordingly, older adults are supposed to report less depressive symptoms, and probably tend to undervalue their impact as well. The theory complementarily hypothesizes that the so-called negativity bias typically found in younger ages is abandoned in favor of more positive information (Carstensen & DeLiema, 2018). Many of our findings further support these hypotheses: mid-aged adults reported more symptoms and perceived them as more severely impairing than did adults in retirement age, both in terms of sum-scores and single symptoms. However, the result that older adults reported feeling slightly more impaired by mild symptoms is not fully accommodated by the theory. A possible explanation is that the depressive symptoms of older adults

are not well captured by general depression scales like the PHQ-9 (e.g., they lack content validity), or that age-dependent response styles influence symptom reports. However, there was no evidence of the latter according to DIF analyses.

Our findings motivate considering a wider range of symptoms, both in terms of severity (i.e., from mild to severe presentations) and content (i.e., symptomatology in older populations), when assessing the impact of depressive symptoms throughout adulthood. Further research may shed light on whether middle-aged adults are indeed more prone to feeling highly impaired by depressive symptoms, or whether general purpose depression scales fail to properly capture depression in late adulthood. These two interpretations of our findings do not exclude each other. Longitudinal approaches would be useful in future, covering bio-psychological and contextual factors (e.g., taking into account the effects of self-perceived role demands), as well as making use of domain-specific, more comprehensive functional impairment indicators (McKnight & Kashdan, 2009).

Limitations

The NHANES datasets are representative of the U.S. adult population, and thus, appropriate caution should be taken when generalizing our results to other countries, or to individuals aged under 18. Additionally, our findings are based on multiple cross-sectional samples and only regard between-person variability patterns. Regarding the PHQ-9, the amount of information it provides is rather limited, favoring brevity over comprehensiveness. For instance, having thoughts about death in general versus planning suicide most likely imply different levels of burden, even though they were collapsed into a single questionnaire item in the PHQ-9. Similarly, functional impairment was measured with only one item. Measurement could be improved by using multiple indicators of impairment. Additionally, the items in the questionnaire are only a portion of the range of existing depressive symptoms (Fried, 2017). Apart from these nine, which correspond to the diagnostic criteria for MDD in

the DSM-5, many symptoms (e.g., crying spells or physical pain) may play a role in self-reported functional impairment and were not taken into account in this study.

Finally, comorbidity was taken into account only partially in this study. It is common that depressive symptoms co-occur with other mental syndromes (Rush et al., 2005). The symptoms included in the PHQ-9 could be due to or aggravated by other mental syndromes. This was not controlled for in our analyses, since the NHANES data do not include mental syndromes other than depression. While physical medical conditions are examined in more detail in NHANES, these are self-reported in the interview and therefore the data may be less reliable than register-based medical information.

References

(CDC), Centers for Disease Control and Prevention, 2017. National Health and Nutrition Examination Survey Data (NHANES).

(CDC), Centers for Disease Control and Prevention, 2015. NHANES Web Tutorials: When and How to Construct Weights When Combining Survey Cycles [WWW Document]. URL https://www.cdc.gov/nchs/tutorials/nhanes/SurveyDesign/Weighting/Task2.htm

(APA), American Psychiatric Association, 2013. Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub.

Blanchflower, D.G., Oswald, A.J., 2016. Antidepressants and age: A new form of evidence for U-shaped well-being through life. J. Econ. Behav. Organ. 127, 46–58. https://doi.org/10.1016/j.jebo.2016.04.010

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach (2nd ed), Ecological Modelling. https://doi.org/10.1016/j.ecolmodel.2003.11.004

Carstensen, L.L., DeLiema, M., 2018. The positivity effect: a negativity bias in youth fades with age. Curr. Opin. Behav. Sci. 19, 7–12. https://doi.org/10.1016/J.COBEHA.2017.07.009

Carstensen, L.L., Fung, H.H., Charles, S.T., 2003. Socioemotional Selectivity Theory and the Regulation of Emotion in the Second Half of Life. Motiv. Emot. 27, 103–123. https://doi.org/10.1023/A:1024569803230

Carstensen, L.L., Turan, B., Scheibe, S., Ram, N., Ersner-Hershfield, H., Samanez-Larkin, G.R., Brooks, K.P., Nesselroade, J.R., 2011. Emotional experience improves with age: Evidence based on over 10 years of experience sampling. Psychol. Aging 26, 21–33.

https://doi.org/10.1037/a0021285

Charles, S.T., Luong, G., Almeida, D.M., Ryff, C., Sturm, M., Love, G., 2010. Fewer ups and downs: daily stressors mediate age differences in negative affect. J. Gerontol. B. Psychol. Sci. Soc. Sci. 65B, 279–86. https://doi.org/10.1093/geronb/gbq002

Charles, S.T., Mogle, J., Urban, E.J., Almeida, D.M., 2016. Daily events are important for age differences in mean and duration for negative affect but not positive affect. Psychol. Aging 31, 661–671. https://doi.org/10.1037/pag0000118

Cheng, T.C., Powdthavee, N., Oswald, A.J., 2017. Longitudinal Evidence for a Midlife Nadir in Human Well-being: Results from Four Data Sets. Econ. J. 127, 126–142. https://doi.org/10.1111/ecoj.12256

Choi, S.W., 2016. lordif: Logistic Ordinal Regression Differential Item Functioning Using IRT.

Cramèr, H., 1946. Mathematical Methods in Statistics. Princeton University Press, Princeton.

Fattori, A., Neri, L., Bellomo, A., Vaggi, M., Mencacci, C., ILDE Study Group, 2017. Depression severity and concentration difficulties are independently associated with HRQOL in patients with unipolar depressive disorders. Qual. Life Res. 26, 2459–2469. https://doi.org/10.1007/s11136-017-1595-4

Foster, S., Mohler-Kuo, M., 2018. New insights into the correlation structure of DSM-IV depression symptoms in the general population v. subsamples of depressed individuals. Epidemiol. Psychiatr. Sci. 27, 288–300. https://doi.org/10.1017/S2045796016001086

Fried, E.I., 2017. The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. J. Affect. Disord. 208, 191–197.

Fried, E.I., 2015. Problematic assumptions have slowed down depression research: why symptoms, not syndromes are the way forward. Front. Psychol. 6, 309.

Fried, E.I., Nesse, R.M., 2015a. Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. BMC Med. 13, 72.

Fried, E.I., Nesse, R.M., 2015b. Depression is not a consistent syndrome: an investigation of unique symptom patterns in the STAR* D study. J. Affect. Disord. 172, 96–102.

Fried, E.I., Nesse, R.M., 2014. The impact of individual depressive symptoms on impairment of psychosocial functioning. PLoS One 9, e90311.

Gallagher, D., Mhaolain, A.N., Greene, E., Walsh, C., Denihan, A., Bruce, I., Golden, J., Conroy, R.M., Kirby, M., Lawlor, B.A., 2009. Late life depression: a comparison of risk factors and symptoms according to age of onset in community dwelling older adults. Int. J. Geriatr. Psychiatry 25, 981–987. https://doi.org/10.1002/gps.2438

Gelin, M.N., Zumbo, B.D., 2003. Differential Item Functioning Results May Change Depending On How An Item Is Scored: An Illustration With The Center For Epidemiologic Studies Depression Scale. Educ. Psychol. Meas. 63, 65–74. https://doi.org/10.1177/0013164402239317

Gonda, X., Pompili, M., Serafini, G., Carvalho, A.F., Rihmer, Z., Dome, P., 2015. The role of cognitive dysfunction in the symptoms and remission from depression. Ann. Gen. Psychiatry 14, 27. https://doi.org/10.1186/s12991-015-0068-9

Haigh, E.A.P., Bogucki, O.E., Sigmon, S.T., Blazer, D.G., 2018. Depression Among Older Adults: A 20-Year Update on Five Common Myths and Misconceptions. Am. J. Geriatr. Psychiatry 26, 107–122. https://doi.org/10.1016/j.jagp.2017.06.011

Haushofer, J., Fehr, E., 2014. On the psychology of poverty. Science 344, 862–7.

https://doi.org/10.1126/science.1232491

Hegeman, J.M., de Waal, M.W.M., Comijs, H.C., Kok, R.M., van der Mast, R.C., 2015. Depression in later life: A more somatic presentation? J. Affect. Disord. 170, 196–202. https://doi.org/10.1016/J.JAD.2014.08.032

Hegeman, J.M., Kok, R.M., van der Mast, R.C., Giltay, E.J., 2012. Phenomenology of depression in older compared with younger adults: Meta-analysis. Br. J. Psychiatry 200, 275–281. https://doi.org/10.1192/bjp.bp.111.095950

Heun, R., Kockler, M., Papassotiropoulos, A., 2000. Distinction of early- and late-onset depression in the elderly by their lifetime symptomatology. Int. J. Geriatr. Psychiatry 15, 1138–1142.

Jokela, M., Virtanen, M., Batty, G.D., Kivimäki, M., 2016. Inflammation and Specific Symptoms of Depression. JAMA Psychiatry 73, 87. https://doi.org/10.1001/jamapsychiatry.2015.1977

Jorm, A.F., 2000. Does old age reduce the risk of anxiety and depression? A review of epidemiological studies across the adult life span. Psychol. Med. 30, S0033291799001452. https://doi.org/10.1017/S0033291799001452

Kessler, R.C., Essex, M., 1982. Marital Status and Depression: The Importance of Coping Resources. Soc. Forces 61, 484–507. https://doi.org/10.1093/sf/61.2.484

Kim, S., Thibodeau, R., Jorgensen, R.S., 2011. Shame, guilt, and depressive symptoms: A meta-analytic review. Psychol. Bull. 137, 68–96. https://doi.org/10.1037/a0021466

Kocalevent, R.-D., Hinz, A., Brähler, E., 2013. Standardization of the depression screener Patient Health Questionnaire (PHQ-9) in the general population. Gen. Hosp. Psychiatry 35, 551–555. https://doi.org/10.1016/J.GENHOSPPSYCH.2013.04.006

Kraemer, H.C., Stice, E., Kazdin, A., Offord, D., Kupfer, D., 2001. How Do Risk Factors Work Together? Mediators, Moderators, and Independent, Overlapping, and Proxy Risk Factors. Am. J. Psychiatry 158, 848–856. https://doi.org/10.1176/appi.ajp.158.6.848

Kroenke, K., Spitzer, R.L., 2002. The PHQ-9: A New Depression Diagnostic and Severity Measure. Psychiatr. Ann. 32, 509–515. https://doi.org/10.3928/0048-5713-20020901-06

Kroenke, K., Spitzer, R.L., Williams, J.B.W., Löwe, B., 2010. The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. Gen. Hosp. Psychiatry 32, 345–359. https://doi.org/10.1016/j.genhosppsych.2010.03.006

Lam, R.W., Kennedy, S.H., McIntyre, R.S., Khullar, A., 2014. Cognitive Dysfunction in Major Depressive Disorder: Effects on Psychosocial Functioning and Implications for Treatment. Can. J. Psychiatry 59, 649. https://doi.org/10.1177/070674371405901206

Le Bon, O., Le Bon, S.D., 2014. Age distribution curve in psychiatric admissions inversely correlates with Life Satisfaction. Psychiatry Res. 219, 232–234. https://doi.org/10.1016/j.psychres.2014.05.001

Lee, R.S.C., Hermens, D.F., Porter, M.A., Redoblado-Hodge, M.A., 2012. A meta-analysis of cognitive deficits in first-episode Major Depressive Disorder. J. Affect. Disord. 140, 113–124. https://doi.org/10.1016/J.JAD.2011.10.023

Lorenzo-Luaces, L., 2015. Heterogeneity in the prognosis of major depression: from the common cold to a highly debilitating and recurrent illness. Epidemiol. Psychiatr. Sci. 24. https://doi.org/10.1017/S2045796015000542

Lumley, T., 2014. survey: analysis of complex survey samples.

Maj, M., 2011. Refining the diagnostic criteria for major depression on the basis of empirical evidence.

Acta Psychiatr. Scand. 123, 317–317. https://doi.org/10.1111/j.1600-0447.2011.01680.x

Mander, B.A., Winer, J.R., Walker, M.P., 2017. Sleep and Human Aging. Neuron. https://doi.org/10.1016/j.neuron.2017.02.004

Marc, L.G., Raue, P.J., Bruce, M.L., 2008. Screening performance of the 15-item geriatric depression scale in a diverse elderly home care population. Am. J. Geriatr. psychiatry 16, 914–21. https://doi.org/10.1097/JGP.0b013e318186bd67

McKnight, P.E., Kashdan, T.B., 2009. The importance of functional impairment to mental health outcomes: A case for reassessing our goals in depression treatment research. Clin. Psychol. Rev. 29, 243–259. https://doi.org/10.1016/J.CPR.2009.01.005

Molarius, A., Janson, S., 2002. Self-rated health, chronic diseases, and symptoms among middle-aged and elderly men and women. J. Clin. Epidemiol. 55, 364–370. https://doi.org/10.1016/S0895-4356(01)00491-7

Moussavi, S., Chatterji, S., Verdes, E., Tandon, A., Patel, V., Ustun, B., 2007. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. Lancet 370, 851–858. https://doi.org/10.1016/S0140-6736(07)61415-9

Nagelkerke, N.J.D., 1991. A Note on a General Definition of the Coefficient of Determination. Biometrika 78, 691–692.

Østergaard, S.D., Jensen, S.O.W., Bech, P., 2011. The heterogeneity of the depressive syndrome: when numbers get serious. Acta Psychiatr. Scand. 124, 495–496. https://doi.org/10.1111/j.1600-0447.2011.01744.x

R Core Team, 2017. R: A Language and Environment for Statistical Computing.

Rief, W., Nanke, A., Klaiberg, A., Braehler, E., 2004. Base rates for panic and depression according to the Brief Patient Health Questionnaire: a population-based study. J. Affect. Disord. 82, 271–276. https://doi.org/10.1016/J.JAD.2003.11.006

Rush, A.J., Zimmerman, M., Wisniewski, S.R., Fava, M., Hollon, S.D., Warden, D., Biggs, M.M., Shores-Wilson, K., Shelton, R.C., Luther, J.F., Thomas, B., Trivedi, M.H., 2005. Comorbid psychiatric disorders in depressed outpatients: Demographic and clinical features. J. Affect. Disord. 87, 43–55. https://doi.org/10.1016/J.JAD.2005.03.005

Schaakxs, R., Comijs, H.C., Lamers, F., Beekman, A.T.F., Penninx, B.W.J.H., 2017. Age-related variability in the presentation of symptoms of major depressive disorder. Psychol. Med. 47, 543–552. https://doi.org/10.1017/S0033291716002579

Spitzer, R.L., Kroenke, K., Williams, J.B.W., Group, P.H.Q.P.C.S., 1999. Validation and Utility of a Self-report Version of PRIME-MD.The PHQ Primary Care Study. JAMA 282, 1737. https://doi.org/10.1001/jama.282.18.1737

Spitzer, R.L., Williams, J.B.W., Kroenke, K., Hornyak, R., McMurray, J., 2000. Validity and utility of the PRIME-MD Patient Health Questionnaire in assessment of 3000 obstetric-gynecologic patients: The PRIME-MD Patient Health Questionnaire Obstetrics-Gynecology Study. Am. J. Obstet. Gynecol. 183, 759–769. https://doi.org/10.1067/MOB.2000.106580

Steptoe, A., Deaton, A., Stone, A.A., 2015. Subjective wellbeing, health, and ageing. Lancet 385, 640–648. https://doi.org/10.1016/S0140-6736(13)61489-0

Stone, A.A., Schwartz, J.E., Broderick, J.E., Deaton, A., 2010. A snapshot of the age distribution of psychological well-being in the United States. Proc. Natl. Acad. Sci. U. S. A. 107, 9985–90. https://doi.org/10.1073/pnas.1003744107

Sutin, A.R., Terracciano, A., Milaneschi, Y., An, Y., Ferrucci, L., Zonderman, A.B., 2013. The Trajectory of Depressive Symptoms Across the Adult Life Span. JAMA Psychiatry 70, 803. https://doi.org/10.1001/jamapsychiatry.2013.193

Tebeka, S., Pignon, B., Amad, A., Le Strat, Y., Brichant-Petitjean, C., Thomas, P., Vaiva, G., Roelandt, J.-L., Benradia, I., Etain, B., Rolland, B., Dubertret, C., Geoffroy, P.A., 2018. A study in the general population about sadness to disentangle the continuum from well-being to depressive disorders. J. Affect. Disord. 226, 66–71. https://doi.org/10.1016/j.jad.2017.08.085

Tweed, D.L., 1993. Depression-related impairment: estimating concurrent and lingering effects. Psychol. Med. 23, 373–386.

Wakefield, J.C., Schmitz, M.F., 2017. Symptom quality versus quantity in judging prognosis: Using NESARC predictive validators to locate uncomplicated major depression on the number-of-symptoms severity continuum. J. Affect. Disord. 208, 325–329.

Wakefield, J.C., Schmitz, M.F., 2016. Feelings of worthlessness during a single complicated major depressive episode predict postremission suicide attempt. Acta Psychiatr. Scand. 133, 257–265. https://doi.org/10.1111/acps.12521

Whiteford, H.A., Degenhardt, L., Rehm, J., Baxter, A.J., Ferrari, A.J., Erskine, H.E., Charlson, F.J., Norman, R.E., Flaxman, A.D., Johns, N., others, 2013. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. Lancet 382, 1575–1586.

World Health Organization, 2017. Depression and other common mental disorders. Global health estimates. Geneva.

Table 1. Description of the sample according to age group. NHANES 2005-2016, sampling-weighted.

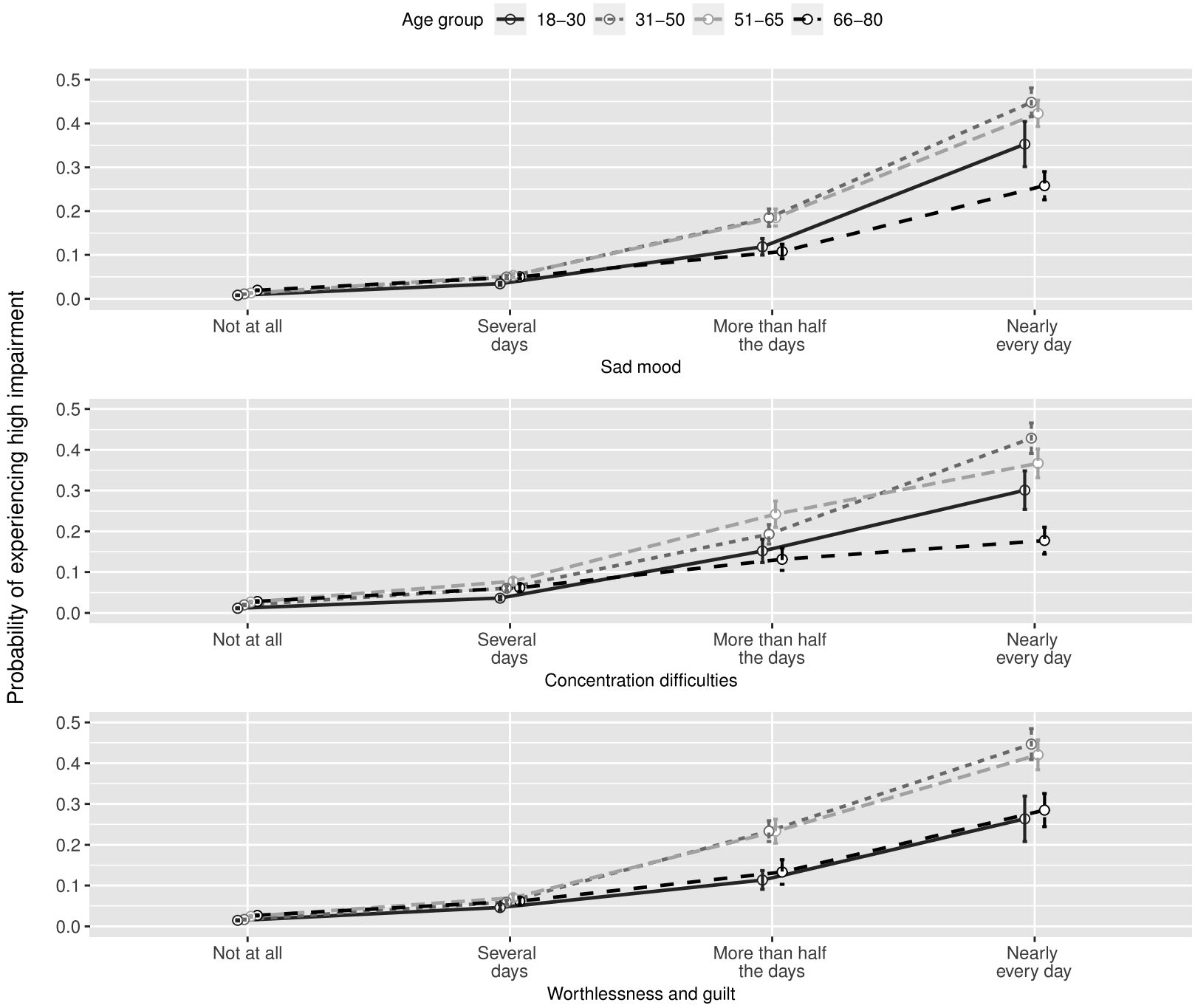
Covariates		**Age groups**			
Categorical covariates		18 to 30	31 to 50	51 to 65	66 to 80
Age group	Total, %	23	36	24	16
Gender	Female, %	50	51	51	57
PHQ-9 Severity score	Minimal (0 to 4)	80	79	78	81
	Mild (5 to 9)	14	14	14	14
	Moderate (10 to 14)	4	5	5	4
	Moderately severe (15 to 19)	1	2	2	1
	Severe (20 to 27)	1	1	1	<1
PHQ item for functional impairment	Not difficult at all	18	26	18	12
	Somewhat difficult	5	8	5	3
	Very difficult	1	1	1	<1
	Extremely difficult	<1	1	<1	<1
Race-ethnicity	Mexican American, %	13	10	5	4
	Other Hispanic, %	7	6	4	3
	Non-Hispanic white, %	58	63	73	80
	Non-Hispanic black, %	14	12	11	8
	Other race, %	8	8	7	5
NHANES sampling year	2005, %	16	17	14	15
	2007, %	16	17	15	15
	2009, %	17	17	16	16
	2011, %	17	16	18	16
	2011, %	17	17	18	18
	2013, %	17	16	18	20
Ratio of family income to poverty	Below poverty threshold, %	24	14	11	11
Marital status	married, %	28	64	65	57
	widowed, %	0	1	5	27
	divorced, %	2	11	16	10
	separated, %	1	3	3	1
	never married, %	52	13	8	3
	living with partner, %	16	9	4	2
Continuous covariates					
PHQ-9 score, mean (SE) [C10, C90]		3.01 (.06) [0, 8]	3.09 (.05) [0, 8]	3.11 (.07) [0, 9]	2.65 (.06) [0, 7]
Number of physical medical conditions, mean (SE)		.080 (.00)	.29 (.01)	.76 (.02)	1.39 (.02)

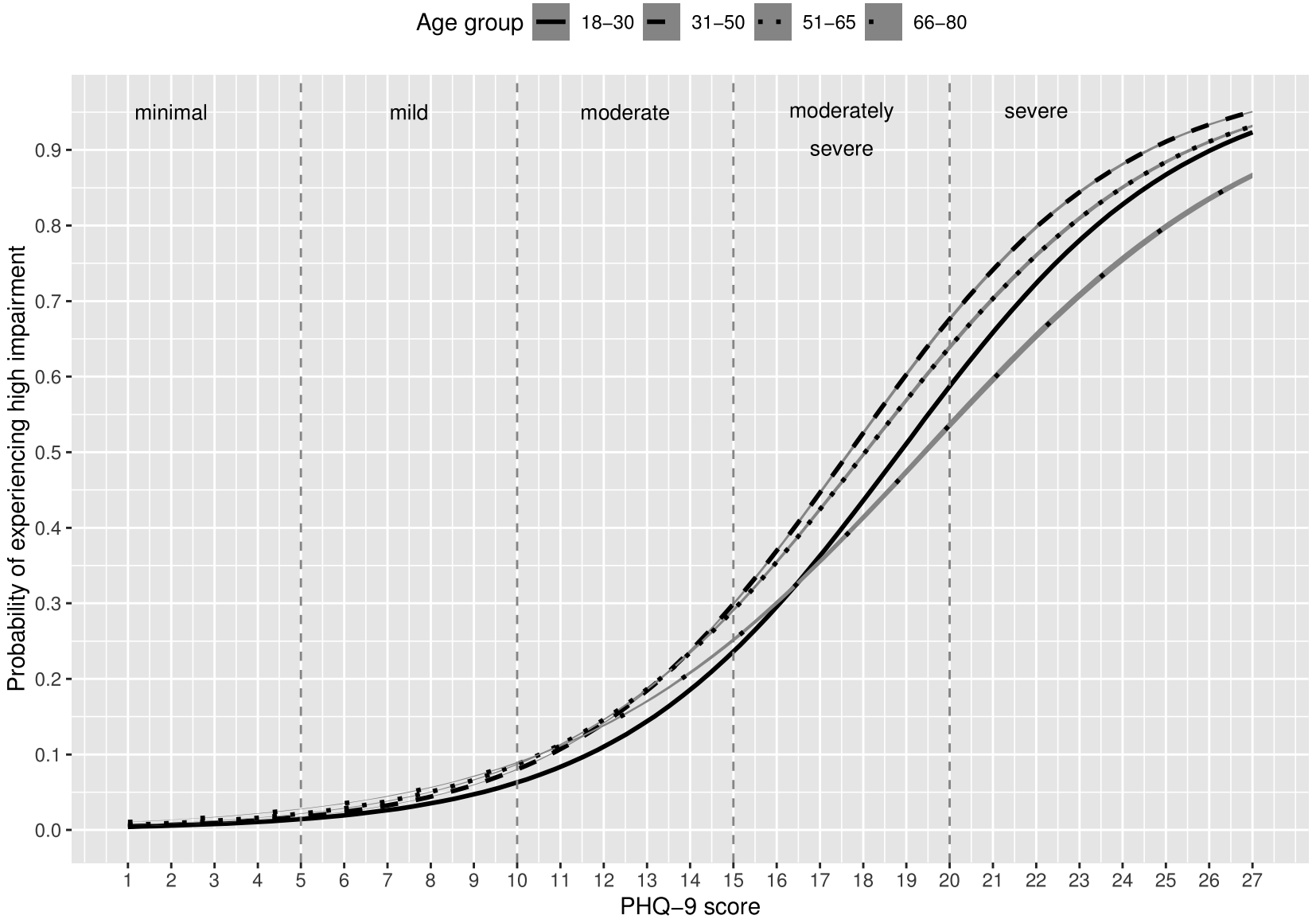
Footnote. The values are percentage units.

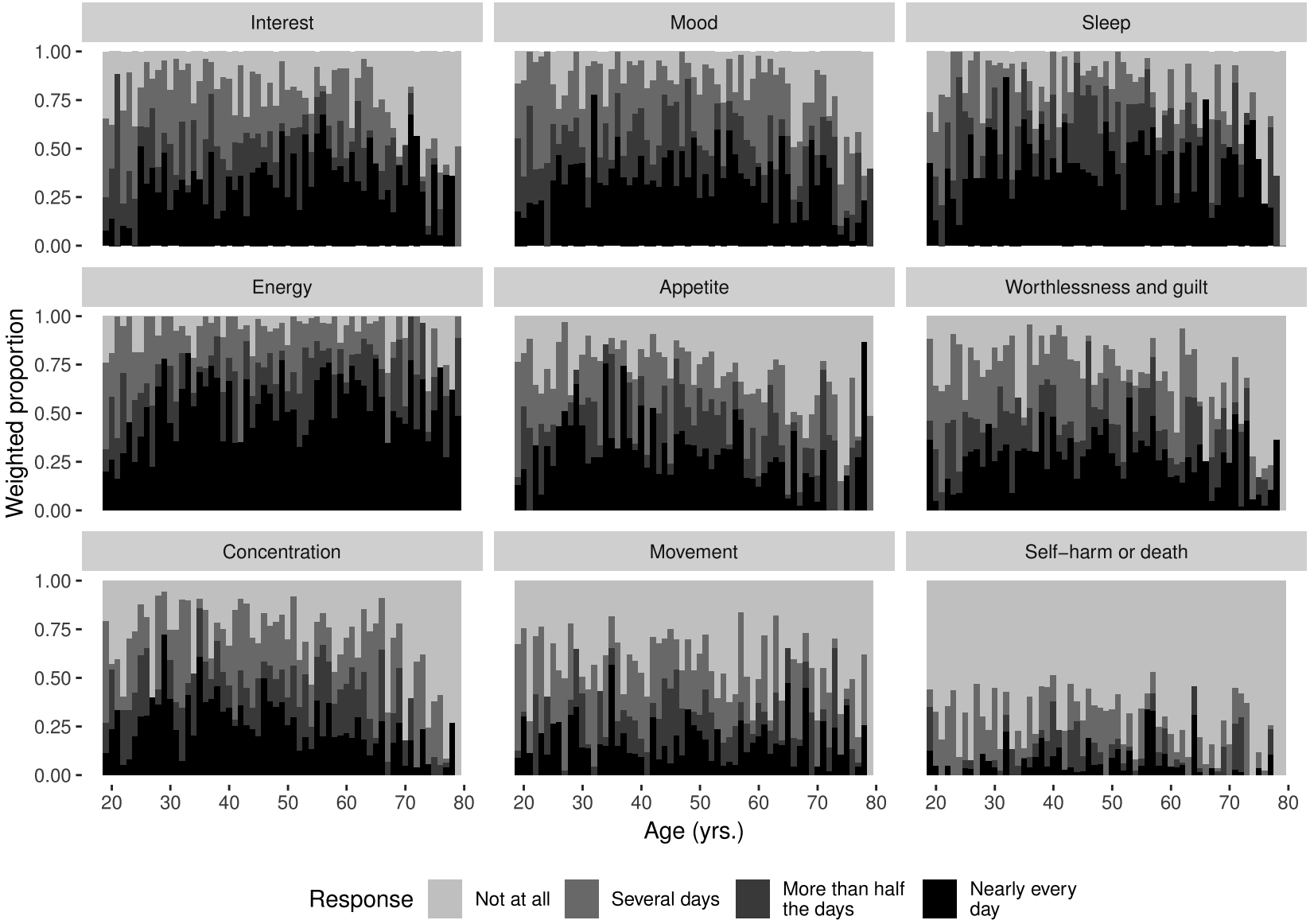
Table 2. Contribution of the symptom to the full model in terms of Nagerlkerke's R^2.

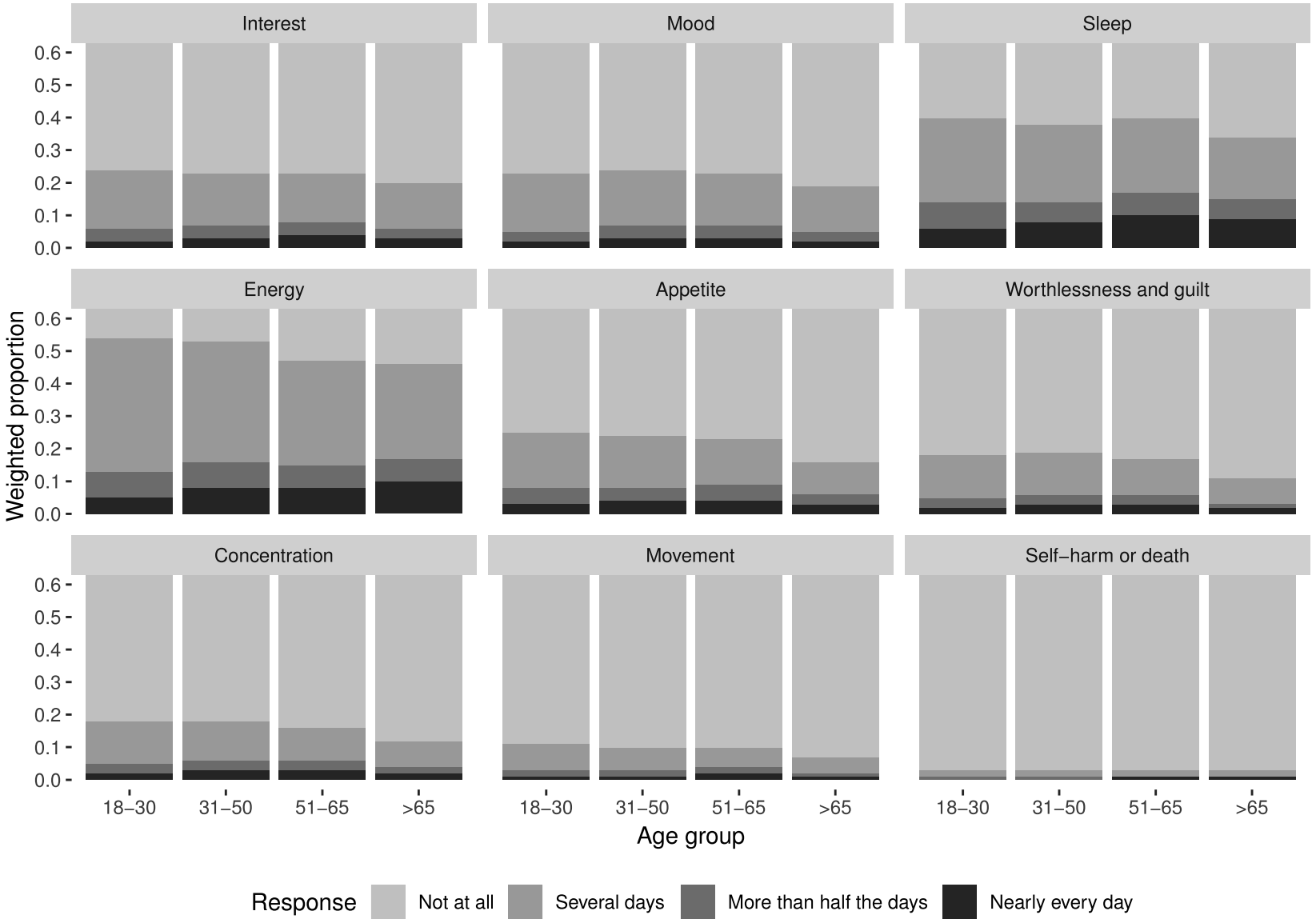
PHQ-9 Symptom	Nagelkerke's R^2 for Step 3 model without the specific symptom	Percentage of variance with respect to full Step 3 model[a]
Anhedonia	.318	3.98
Low mood[b]	.318	4.06
Sleep	.327	1.13
Fatigue	.317	4.31
Appetite	.325	1.76
Feelings of worthlessness and guilt[b]	.322	2.90
Concentration[b]	.321	3.04
Motor	.318	3.85
Self-harm or death thoughts	.326	1.65

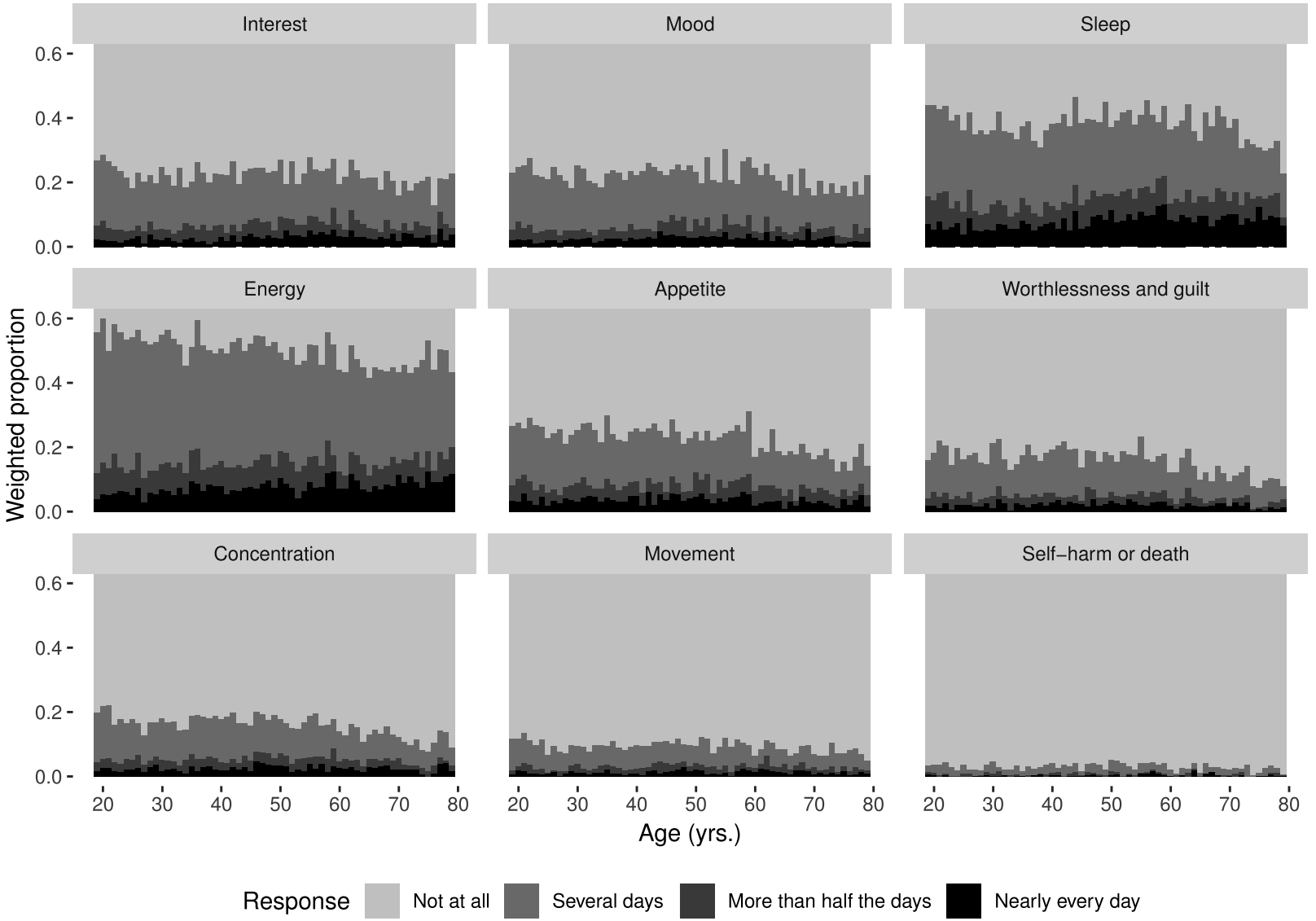
Footnote: Values for Nagelkerke's Pseudo R^2 registered when excluding individual PHQ-9 symptoms from the model in Step 3. [a]These values are with respect to the full model R^2=.331 (e.g. 1 - .318/.331 = 3.98). [b]The effect of this symptom on high functional impairment was moderated by age group.

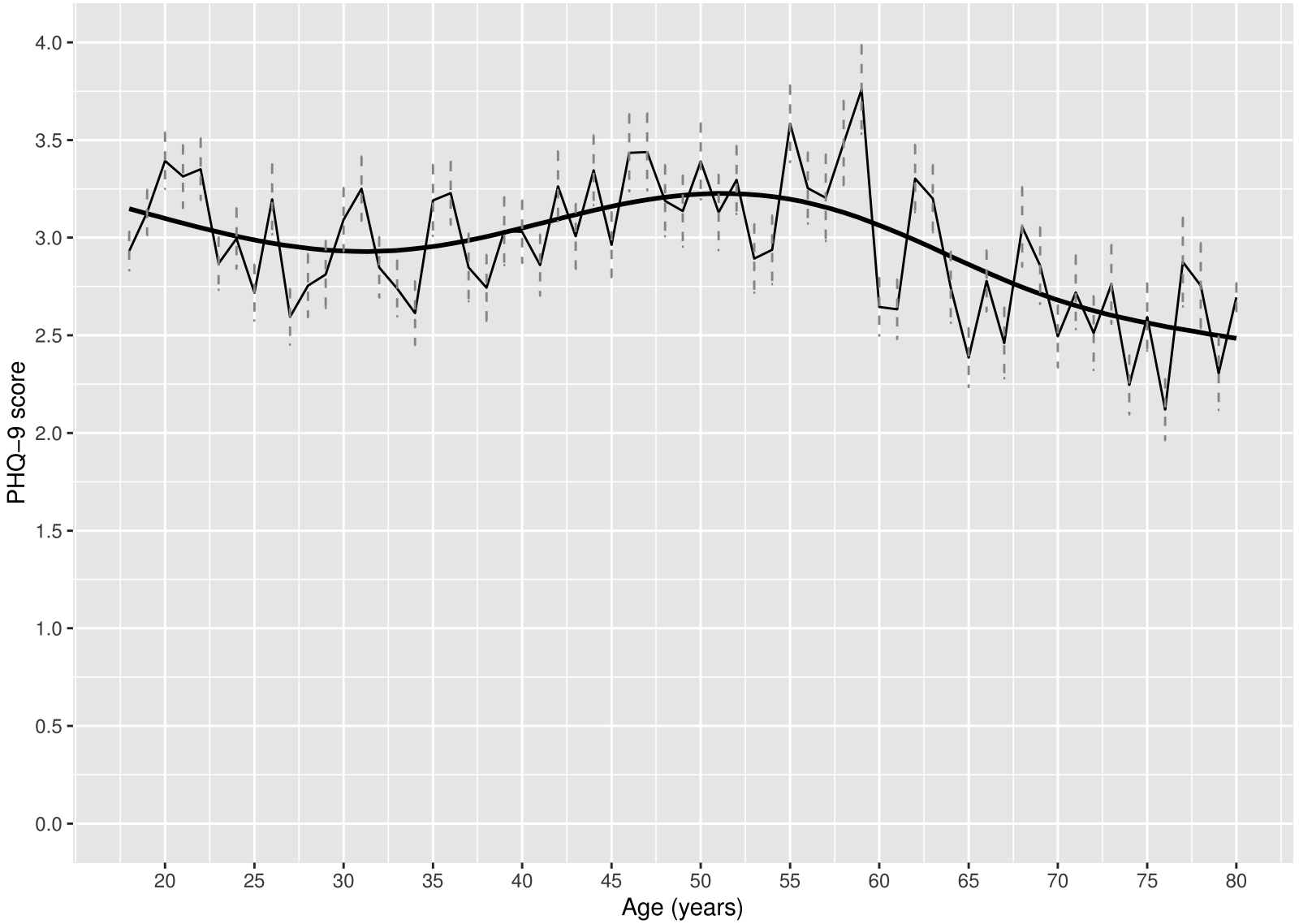

Age group
18–30
31–50
51–65
66–80
Probability of experiencing high impairment
0.0
0.1
0.2
0.3
0.4
0.5
Not at all
Several days
More than half the days
Nearly every day
Sad mood
Concentration difficulties
Worthlessness and guilt


Age group
18–30
31–50
51–65
66–80
minimal
mild
moderate
moderately severe
severe
Probability of experiencing high impairment
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
PHQ–9 score


Interest
Mood
Sleep
Energy
Appetite
Worthlessness and guilt
Concentration
Movement
Self−harm or death
Weighted proportion
1.00
0.75
0.50
0.25
0.00
20
30
40
50
60
70
80
Age (yrs.)
Response
Not at all
Several days
More than half the days
Nearly every day


Interest
Mood
Sleep
Energy
Appetite
Worthlessness and guilt
Concentration
Movement
Self−harm or death
Weighted proportion
0.0
0.1
0.2
0.3
0.4
0.5
0.6
18−30
31−50
51−65
>65
Age group
Response
Not at all
Several days
More than half the days
Nearly every day


Interest
Mood
Sleep
Energy
Appetite
Worthlessness and guilt
Concentration
Movement
Self−harm or death
Weighted proportion
0.0
0.2
0.4
0.6
20
30
40
50
60
70
80
Age (yrs.)
Response
Not at all
Several days
More than half the days
Nearly every day


PHQ–9 score
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
20
25
30
35
40
45
50
55
60
65
70
75
80
Age (years)

Conflict of Interests. The authors declare none.

Contributors. RG-V performed the statistical analyses and wrote the first draft of the manuscript. All three authors edited and reviewed the manuscript.

Role of the Funding Source. RG-V was supported by Emil Aaltonen Foundation. MJ was supported by the Academy of Finland (311578).

Acknowledgements. None.

Supplementary Table S1. Differential Item Functioning analyses for the PHQ-9 according to age groups 18-30, 31-50, 51-65, and 66-80. NHANES 2005-2016.

PHQ-9 item	Uniform DIF, Nagelkerke's R^2 change	Non-uniform DIF, Nagelkerke's R^2 change	Total DIF effect, Nagelkerke's R^2 change
Anhedonia	.001	.001	.002
Low mood	.000	.000	.000
Sleep	.000	.000	.001
Fatigue	.001	.003	.004
Appetite	.002	.000	.002
Worthlessness and guilt	.002	.000	.002
Concentration	.001	.000	.001
Motor	.000	.000	.000
Self-harm or death	.003	.000	.004
Difficulties to carry out normal life activities	.000	.001	.001

Note. In order to be flagged for DIF, a change in Nagelkerke's R^2 ought to be ≥.02.

Supplementary Table S2. Differential Item Functioning analyses for the PHQ-9 according to gender. NHANES 2005-2016.

PHQ-9 item	Uniform DIF, Nagelkerke's R^2 change	Non-uniform DIF, Nagelkerke's R^2 change	Total DIF effect, Nagelkerke's R^2 change
Anhedonia	.000	.000	.000
Mood	.000	.000	.000
Sleep	.000	.000	.001
Fatigue	.002	.000	.002
Appetite	.004	.000	.004
Worthlessness and guilt	.000	.000	.000
Concentration	.000	.001	.001
Motor	.002	.000	.002
Self-harm or death	.002	.000	.002
Difficulties to carry out normal life activities	.001	.000	.001

Note. In order to be flagged for DIF, a change in Nagelkerke's R^2 ought to be ≥.02.

Supplementary Table S3. Logistic regression models predicting functional impairment derived from the PQ-9 symptoms. Sample-weighted model estimates corresponding to the Steps 1 to 4. NHANES 2005-2016.

Step	Predictors	b	s.e.	Z	p-value	AIC	R^2
1	Sum-score	.31	.02	14.43	<.001***	4574.57	.316
	Age_1	.23	.31	.75	.45		
	Age_2	.54	.33	1.74	.10		
	Age_3	.96	.34	2.80	<.01**		
	Sum-score*Age_1	.01	.02	.23	0.82		
	Sum-score*Age_2	-.02	.03	-.84	.40		
	Sum-score*Age_3	-.06	.03	-2.15	.03*		
2a	Interest	1.19	.10	12.28	<.001***	5526.90	.210
	Age_1	.22	.33	.66	.51		
	Age_2	.33	.37	.89	.38		
	Age_3	1.04	.39	2.68	.01**		
	Interest*Age_1	.07	.12	.56	.58		
	Interest *Age_2	.00	.13	-.01	.99-.		
	Interest *Age_3	-.43	.14	-3.05	<.001***		
2b	Mood	1.40	.11	13.25	<.001***	5315.18	.233
	Age_1	.36	.36	.99	.32		
	Age_2	.50	.38	1.132	.19		
	Age_3	1.19	.40	2.99	<.001***		
	Mood *Age_1	-.05	.13	-.37	.71		
	Mood *Age_2	-.08	.13	-.61	.54		
	Mood *Age_3	-.48	.15	-3.26	<.001***		
2c	Sleep	.84	.11	7.56	<.001***	6123.41	.149
	Age_1	.32	.39	.80	.42		
	Age_2	.77	.41	1.87	.06		
	Age_3	1.17	.45	2.57	.01**		
	Sleep *Age_1	.03	.13	.22	.82		
	Sleep *Age_2	-.13	.14	-.93	.35		
	Sleep *Age_3	-.40	.15	-2.67	.01**		
2d	Energy	1.18	.13	9.24	<.001***	5713.72	.191
	Age_1	.51	.48	1.06	.29		
	Age_2	.20	.49	.41	.68		
	Age_3	.96	.53	1.81	.07		
	Energy *Age_1	-.04	.16	-.26	.80		
	Energy *Age_2	.13	.16	.81	.42		
	Energy *Age_3	-.25	.17	-1.51	.13		
2f	Appetite	1.01	.10	9.76	<.001***	5983.35	.163
	Age_1	.49	.34	1.47	.14		
	Age_2	1.09	.36	3.04	<.001***		
	Age_3	1.28	.38	3.37	<.001***		
	Appetite*Age_1	-.03	.12	-.26	.79		

	Appetite*Age_2	-.23	.13	-1.79	.07		
	Appetite*Age_3	-.44	.14	-3.23	<.001***		
2g	W&G[a]	1.09	.10	11.26	<.001***	5608.50	.202
	Age_1	-.02	.31	-.07	.94		
	Age_2	.39	.33	1.18	.24		
	Age_3	.76	.36	2.11	.04*		
	W&G[a]*Age_1	.15	.12	1.29	.20		
	W&G[a]*Age_2	.04	.12	.30	.77		
	W&G[a]*Age_3	-.26	.12	-1.84	.07		
2h	Concentration	1.22	.10	12.16	<.001***	5697.97	.192
	Age_1	.65	.32	2.02	.04*		
	Age_2	1.09	.33	3.28	<.001***		
	Age_3	1.42	.36	3.91	<.001***		
	Concentration*Age_1	-1.0	.12	-.83	.41		
	Concentration*Age_2	-.22	.12	-1.76	.08		
	Concentration*Age_3	-.55	.13	-4.31	<.001***		
2i	Movement	1.24	.11	11.42	<.001***	5898.14	.172
	Age_1	.71	.28	2.54	.01**		
	Age_2	.99	.31	3.22	<.001***		
	Age_3	.59	.34	1.72	.09		
	Movement*Age_1	-.16	.13	-1.19	.23		
	Movement*Age_2	-.26	.14	-1.87	.06		
	Movement*Age_3	-.21	.14	-1.47	.14		
2j	SH&D[b]	1.19	.16	7.27	<.001***	6176.12	.144
	Age_1	.01	.30	.05	.96		
	Age_2	.42	.32	1.30	.19		
	Age_3	.53	.34	1.54	.12		
	SH&D[b] *Age_1	.31	.21	1.47	.14		
	SH&D[b] *Age_2	.05	.21	.22	.83		
	SH&D[b] *Age_3	-.32	.23	-1.41	.16		
3	Age_1	.17	.56	.30	.76	4563.42	.331
	Age_2	.64	.57	1.12	.26		
	Age_3	1.35	.64	2.10	.04*		
	Interest	.33	.16	2.05	.04*		
	Mood	.77	.17	4.53	<.001***		
	Sleep	-.07	.15	-.49	.63		
	Energy	.44	.19	2.32	.02*		
	Appetite	.20	.13	1.54	.12		
	W&G[a]	-.14	.16	-.93	.35		
	Concentration	.60	.15	3.99	<.001***		
	Movement	.58	.14	3.98	<.001***		
	SH&D[b]	-.01	.27	-.05	.96		
	Interest*Age_1	.07	.20	.33	.74		
	Interest *Age_2	.17	.20	.85	.39		
	Interest *Age_3	-.17	.21	-.81	.42		

	Mood *Age_1	-.42	.21	-1.99	.05*		
	Mood *Age_2	-.32	.21	-1.55	.12		
	Mood *Age_3	-.34	.22	-1.54	.12		
	Sleep *Age_1	.16	.18	.90	.37		
	Sleep *Age_2	.02	.19	.13	.90		
	Sleep *Age_3	.16	.19	.84	.40		
	Energy *Age_1	-.06	.23	-.28	.78		
	Energy *Age_2	.21	.23	.90	.37		
	Energy *Age_3	.06	.23	.27	.78		
	Appetite*Age_1	.01	.16	.06	.96		
	Appetite*Age_2	-.17	.17	-1.01	.31		
	Appetite*Age_3	-.07	.19	-.38	.70		
	W&G[a]*Age_1	.52	.19	2.67	.01**		
	W&G[a]*Age_2	.42	.19	2.18	.03*		
	W&G[a]*Age_3	.24	.21	1.14	.25		
	Concentration*Age_1	-.30	.18	-1.73	.08		
	Concentration*Age_2	-.44	.18	-2.33	.02*		
	Concentration*Age_3	-.54	.19	-2.76	.01**		
	Movement*Age_1	-.23	.18	-1.27	.20		
	Movement*Age_2	-.30	.18	-1.63	.10		
	Movement*Age_3	.07	.20	.33	.74		
	SH&D[b] *Age_1	.45	.31	1.43	.15		
	SH&D[b] *Age_2	.24	.31	.79	.43		
	SH&D[b] *Age_3	.03	.35	.07	.94		

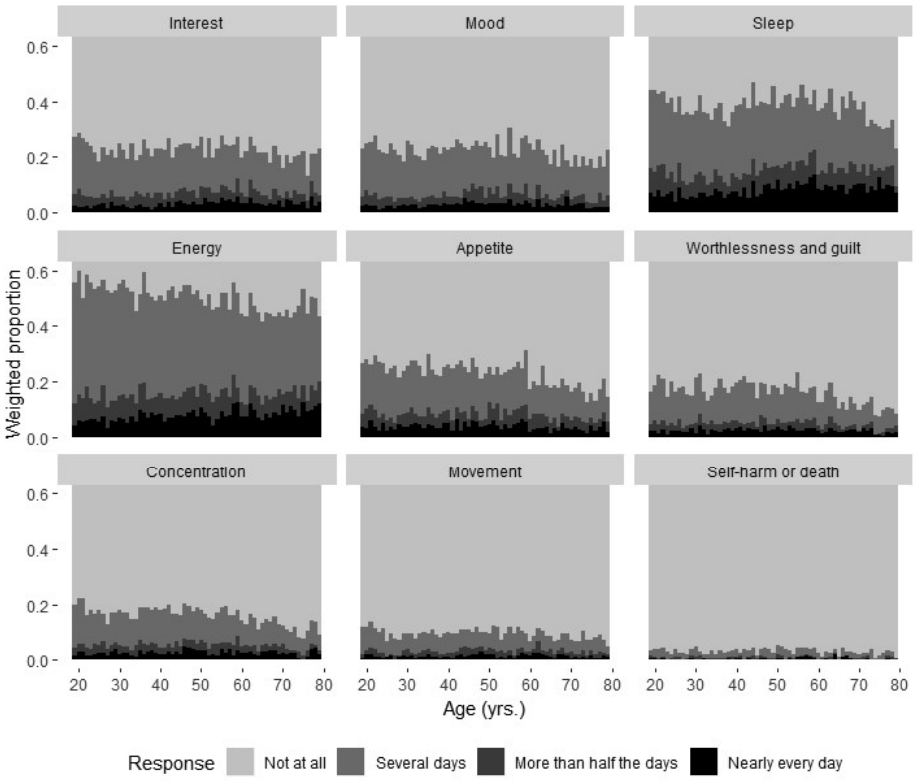
Footnote. Coefficients of covariates excluded for readability. The reference category of age group is 18-30, de other groups are 31-50, 51-65, and ≥66. b: unstandardized logistic regression coefficient, s.e.: standard error of estimation, *Z*: Z statistic, AIC: Akaike's Information Criterion, lower values indicate better model fit. R^2: Nagelkerke's R^2, obtained by comparing to the null model (i.e. intercept only). [a] Feelings of worthlessness and guilt, [b] Self-harm and death thoughts.

Supplementary Table S4. Sensitivity analyses of Step 1 excluding three items from the sum-score. NHANES 2005-2016.

Step	Predictors	b	s.e.	Z	p-value	AIC	R^2
1	Sum-score[a]	.36	.03	13.71	<.001***	4837.17	.285
	Age_1	.23	.32	.72	.47		
	Age_2	.48	.34	1.44	.15		
	Age_3	.80	.36	2.24	.02*		
	Sum-score[a]*Age_1	.02	.04	.41	.69		
	Sum-score[a]*Age_2	-.03	.04	-.64	.52		
	Sum-score[a]*Age_3	-.08	.04	-1.75	.08		

Footnote. Coefficients of covariates excluded for readability. The reference category of age group is 18-30, de other groups are 31-50, 51-65, and ≥66. b: unstandardized logistic regression coefficient, s.e.: standard error of estimation, Z: Z statistic, AIC: Akaike's Information Criterion, lower values indicate better model fit. R^2: Nagelkerke's R^2, obtained by comparing to the null model (i.e. intercept only). [a]The sum-score was calculated using the responses to the six symptoms only (anhedonia, appetite, sleep, fatigue, movement, self-harm or death), therefore leaving out the effects of the symptoms showing interaction with age group in the Step 3 of Table S3 (low mood, feelings of worthlessness and guilt, and concentration problems).

Figure S1. Sample-weighted proportions of endorsement of PHQ-9 symptoms according to age. NHANES sampling years 2005-2016.



Footnote. The panels are truncated in the Y-axis at height .60 to aid visual inspection.

Figure S2. Sample-weighted proportions of endorsement of PHQ-9 symptoms according to age for participants reporting that depressive symptoms make it very or extremely difficult to function. NHANES sampling years 2005-2016 (n= 1,032).