A model-driven approach to better identify older people at risk of depression

Chiara Gennaro, Omar Paccagnella and Paola Zaninotto

Abstract

Depression in later life is one of the most common mental disorders. Several instruments have been developed to detect the presence or the absence of certain symptoms or emotional disorders, based on cut-off points. However, the use of a cut-off does not allow to identifying depression subtypes or distinguish between mild and severe depression. As a result, depression may be under or over diagnosed in the older population. This paper aims at applying a model-driven approach to classify individuals into distinct subgroups, based on different combinations of depressive and emotional conditions. This approach is based on two different statistical solutions: a latent class analysis is first applied to the items collected by the depression scale and, according to the final model, the probability of belonging to each class is calculated for every individual. Lastly, a factor analysis on these classes is performed for obtaining a reduced number of clusters easy to interpret. We use data collected through the EURO-D scale in a large sample of older individuals, participants of the sixth wave of the SHARE (Survey of Health, Ageing and Retirement in Europe) survey. We show that by using such model-based approach it is possible to classify individuals in a more accurate way than the simple dichotomisation *depressed* vs *non-depressed*.

Introduction

Depression is one of the most common mental disorders affecting older people, which often coexists with other medical illnesses and physical disabilities (Richard *et al.*, 2004). People with diagnosed clinical depression or depressive symptomatology are at higher risk of mortality (Bruce and Leaf, 1989; Penninx *et al.*, 1999; Pulska *et al.* 1999). The evidence of higher prevalence of depression (minor and major) among women than in men is one of the most widely documented finding, in both population-based and clinical studies (Dennerstein, Astbury and Morse, 1993; Kessler, 2000, 2003). The gender differences in depression have been found to persist also during later life (Beekman, Copeland and Prince, 1999; Cole and Dendukuri, 2003; Pagán-Rodríguez and Pérez, 2012).

Later life depression is characterised by a large heterogeneity in symptoms and significantly decreases quality of life of the older population. Nevertheless, there is not a general agreement in the scientific literature on what constitutes clinically significant depression, nor on how profiling depressive and emotional symptoms (Blazer, 2003).

The identification of substantial and meaningful subgroups of depressive symptom profiles among the old population has important implications for research, public health policies and clinical practice. First, the structure of depressive symptoms in older adults may be different and potentially more heterogeneous with respect to the one in younger adults (Hybels *et al.*, 2009). Furthermore, the public health burden of depressive symptoms in older adults is expected to increase rapidly, given the large proportion of people in this group. Not surprisingly, the theme of the 2017 World Health Day campaign was depression.

Several screening tools for depressive symptoms are used in clinical and non-clinical settings. These can be divided into two broader sets of instruments: the first, designed to collect information on symptoms, conditions or signs (in other words diagnostic criteria) that best reflect diagnoses of mental disorders. Of this type are the Diagnostic Interview Schedule (DIS) (Robins *et al.*, 1981), that assesses disorders using the definitions and criteria of the Diagnostic and Statistical Manual of Mental Disorders, Third Edition (DSM-III) (American Psychiatric Association, 1980), the CIDI (Composite International Diagnostic Interview) (Robins *et al*., 1988), designed for assessing disorders based on the definitions and criteria of the Diagnostic and Statistical Manual of Mental Disorders, Third Edition, revised (DSM-III-R) (American Psychiatric Association, 1987), CIDI extensions like the World Mental Health CIDI (Kessler and Üstün, 2004) or the Structured Clinical Interview for DSM (SCID) (Spitzer *et al*., 1992) and its subsequent revisions. While DIS diagnoses are exclusively based on the definitions and criteria of the DSM, CIDI is also based on the WHO International Classification of Disease (ICD). The second set of instruments are designed to measure more generic factors (i.e. psychological distress), collecting the presence or the absence of certain symptoms or emotional disorders, such as the Centre for Epidemiologic Study Depression Scale (CESD) (Radloff, 1977), the Beck Depression Inventory (BDI) (Beck, Steer and Garbin, 1988) and the EURO-D (Prince *et al*., 1999a) scales. Usually, this kind of instrument provides a mental health score and cut-off points to be used to classify individuals in having (or not) depression or mental health disorders.

The EURO-D scale

The EURO-D is a depression scale that was developed and validated by the EURODEP Concerted Action Programme. The scale includes 12 items: depression, pessimism, wishing death, guilt, sleep, interest, irritability, appetite, fatigue, concentration, enjoyment and tearfulness (Prince *et al*., 1999a). This scale is the result of the harmonization of five depression measures, three interviewer-administered scales that generate clinical diagnoses (Geriatric Mental State-AGECAT, SHORT-CARE and Comprehensive Psychopathological Rating scale) and two self-reported depression symptom scales (Centre for Epidemiological Studies Depression and Zung Self-Rating Depression scale). EURO-D is a symptom-oriented scale, meaning that it identifies the presence (or not) of some depressive or emotional manifestations: its score ranges from 0 (the lowest level of depression) to 12 (the highest level), but it does not provide diagnoses of any mental disorders. The complete list of the EURO-D questions is reported in the online supplementary material (Table A1).

Construct validity of this scale was supported by Larraga *et al*. (2006), Ploubidis and Grundy (2009) and, with some limitations, Brailean *et al*. (2015). Prince *et al*. (1999b) showed that the EURO-D scale could be reduced into two factors, Affective Suffering and Motivation. Brailean *et al*. (2015) and Guerra *et al*. (2015) found evidence on the validity of this scale with the two-dimensional structure across different populations. Overall, the Affective Suffering factor is characterised by very good cross-cultural measurement properties, while the Motivation factor shows heterogeneity in factor loading patterns and item calibrations across countries (Castro-Costa *et al*., 2008; Prince, 2013; Portellano-Ortiz *et al*., 2017). There is strong support for applying in the empirical analysis both the full 12-item EURO-D scale and the Affective Suffering factor score derived from it, while the use of the Motivation factor in cross-cultural studies is suggested with some caution (Prince, 2013).

Dewey and Prince (2005) defined clinically significant depression with a EURO-D score greater than three. This cut-off point was validated in the EURODEP study, across the continent and against a variety of clinically relevant indicators (people with scoring above this level would be likely to be diagnosed as suffering from a depressive disorder, for which therapeutic intervention would be indicated). However, several studies have shown that a higher optimal cut-off point (four, five or even above) should be required in order to detect probable depression individuals (Castro-Costa *et al*., 2007; Jirapramukpitak *et al*., 2009; Guerra *et al*., 2015).

Aims and hypotheses

Although widely used, a cut-off to identify people at risk of depression does not discriminate among depression subtypes nor it allows to distinguish between mild and severe depression. As a result, depression may be under- or over-diagnosed (according to the value chosen as cut-off). Nevertheless, older people may differentiate from others for the combinations of symptoms that characterise their mental disorders.

Therefore, the hypothesis underlying this work is that a more accurate classification (according to the number and type of the reported items) of the older individuals than the simple dichotomisation "depressed" vs "non-depressed" may be more helpful to researchers, clinicians and clinical investigators. The main objective of this work is to investigate the usefulness of an alternative approach, based on two *model-driven* solutions, for analysing data collected by the EURO-D scale. We want to identify meaningful subgroups of depressive symptom profiles among a population sample of older individuals.

Data and method

The sample

This study is based on data from the sixth wave of the Survey of Health, Ageing and Retirement in Europe (SHARE), collected in 2015 (Börsch-Supan, 2017) (DOI: https://doi.org/10.6103/SHARE.w6.600). See Börsch-Supan *et al*. (2013) and Malter and Börsch-Supan (2017) for methodological details.

SHARE is a panel survey that collects detailed cross-national information on health, socio-economic status and social and family networks of citizens aged 50 and over from a large set of European countries, ranging from the Scandinavian and Baltic area to Mediterranean nations.

The analysed sample is composed by 64.716 individuals (who answered to all items of the EURO-D scale), living in 18 countries (Austria, Belgium, Croatia, Czech Republic, Denmark, Estonia, France, Germany, Greece, Italy, Israel, Luxembourg, Poland, Portugal, Slovenia, Spain, Sweden and

Switzerland). Sample size ranges from more than 5.000 observations in Belgium, Estonia, Italy and Spain to less than 2.000 in Israel, Luxembourg, Poland and Portugal.

In order to maximise the large sample size, we analyse the pooled dataset rather than by country. The EURO-D scale was specifically developed as a harmonised instrument across European countries (Prince *et al*., 1999a). As a consequence, from its initiation each SHARE country has supported high quality translation procedures (Harkness, 2005), in order to guarantee cross-country homogeneity in understanding the meaning of every question by respondents.

The statistical methods

Our approach is based on two different statistical solutions – Latent Class (LC) analysis and Factor Analysis (FA) – to classify individuals with different patterns of depressive disorders. Specifically, in the LC analysis the probability of belonging to each class is calculated for every individual. However, the LC solution usually results in a large number of classes. In order to create number of clusters easier to interpret, in a second stage we apply a factor analysis to the obtained classes.

The LC approach is one of the most innovative and powerful solutions to classify individuals into distinct subgroups, based on differing combinations of depressive conditions. These subgroups (or classes) set up the categories of a categorical latent variable: units within the same LC are homogeneous according to certain criteria, while observations coming from different LCs are dissimilar from each other in some ways (Vermunt and Magidson, 2004). The LC approach is becoming increasingly popular to determine typologies of depressive symptoms among older people, both in clinic-based and in population-based samples (Bogner *et al*. 2009; Lee *et al.*, 2012; Veltman *et al*., 2017). Overall, findings are consistent with the DSM classification scheme and can provide complementary information to DSM diagnostic groups. The LC approach differs from other traditional latent variable methodologies in that the latent variable is categorical, rather than continuous. Then, differently from traditional analyses for clustering, such as a cluster analysis, the latent class approach is a model-based solution.

The main assumption of LC models is that within each LC, the L manifest variables are assumed to be independent, the so-called *local independence* assumption (Hagenaars and McCutcheon, 2002). Bivariate Residuals (BVRs) are usually calculated for checking violation of this crucial assumption. BVRs are conceptually similar to the Modification Indices in the structural equation modelling approach: low values indicate good model fit, while large values identify correlations between the associated variable pairs (which have not been adequately explained by the model). There is no general agreement about the thresholds for judging the smallness of the BVRs (Oberski, van Kollenburg and Vermunt, 2013), however we adopt the value of 3.84 as the benchmark, as suggested by van Kollenburg, Mulder and Vermunt (2015). Violation of the local independence assumption could lead to a poor fit model (Vermunt and Magidson, 2003). There are different solutions for solving this problem, in particular increasing the number of LCs or allowing for direct relationships between the associated items. In our analysis, particular attention is addressed to check this assumption validity.

The conditional response probabilities of a LC model are modelled through logit specifications and active covariates may (or not) be introduced. In our model specification no active covariate is used.

The posterior membership probability, that is the probability of belonging to a certain LC, can be obtained by the Bayes rule and it is used to assign individuals to latent classes. The most common classification rule is modal assignment. Model parameters can be estimated by maximum likelihood, obtained by an adapted EM algorithm (Vermunt, 2003).

The LatentGOLD and Stata software packages are used for the LC and FA estimations, respectively.

Results

The sample is mainly characterised by female respondents (56 per cent), aged 67.2 (±10.3 s.d.) years, on average, even if a large cross-country heterogeneity is nevertheless present. Mediterranean countries show the largest proportions of low educated people, while Nordic countries are characterised by a large ratio of high educated individuals. Approximately 18 per cent of men and 36 per cent of women are not living with a partner. About 40 per cent of respondents reported being in a fair or bad health, and half of the sample reports at least two chronic diseases. One fourth of the respondents are currently in paid employment, however, this proportion varies from about 15 per cent in Austria, Portugal and Slovenia to over 40 per cent in Denmark.

The EURO-D score

Results in Table 1 show that 27 per cent of respondents (20 per cent of men and 33 per cent of women) report being depressed according to the EURO-D scale. When we closely look at the combinations of the reported symptoms, we find that the majority of the depressed people indicates four or five items, and only a small proportion of them (for women is larger than 7.5 per cent) report nine or more symptoms. The distribution of the reported disorders among respondents not classified as depressed shows a mode of zero for men and one for women: three items are indicated by more than 20 per cent of the old women and less than 15 per cent of men.

Figures 1 and 2 provide some information on the symptoms; the most reported are depression, sleep, irritability, fatigue and (only for women) tearfulness, the least reported are suicidality and guilt.

The same pattern on the most and least reported symptoms is found among depressed respondents and those reporting not being depressed according to the EURO-D scale.

The two-stages model solution

Several LC models over the 12 EURO-D items are estimated, varying the number of LCs. In Table 2 we show indices of model fit and performances. The Information criteria (AIC, BIC, etc.) are the most used indicators, however in this case their values show very strong similarities across all estimated models (for instance, comparing the model with seven LCs with the one having ten more LCs, the differences are even lower than 0.1 per cent, for all indicators). Similar conclusions may be reached according to the classification errors statistics.

BVR values allow to check violation of the local independence assumption. Models with 14 or less LCs present at least one large BVR (i.e. larger than 3.84). We add other classes to the models, first to guarantee that this assumption is met and then because a large number of classes may better capture particular and specific combinations of the reported symptoms.

Therefore, according to all criteria and indicators, the model with 15 LCs appears as the best solution. Moreover, with respect to a model with a larger number of LCs, it is more parsimonious and shows only one class with a very small size (lower than one per cent).

Table 3 shows the main results, that is the size of each cluster and the individual probabilities of reporting the depressive symptoms conditional to belong to the LC. Additional findings are available in the online supplementary material (Tables B1 and B2), in particular some demographic and socio-economic characteristics of the individuals belonging to each LC (since covariates are not specified as active in the model, the reported outcomes are based on the ex-post assignment of each respondent to the LCs).

The largest class (LC1) groups 37 per cent of the respondents and is characterised by a very low probability for all depressive and emotional symptoms. Figure 3 highlights that all individuals who do not mention any symptoms belong to this class, as well as people with at most two reported items. LCs 2, 3 and 8 are similar each other, because they are characterised by people with a low probability in several items. The majority of respondents in these clusters reports from one to three symptoms, but no more than five (LC3).

Also LCs 7, 10 and 14 are comparable, in the sense they are small (a total size of about 5.5 per cent of people) and are characterised by individuals with a low probability for many depressive indicators and a quite large probability (i.e. higher than 0.5) for a couple of symptoms (for instance, depression and tearfulness in LC7 or fatigue and concentration in LC10). LCs 4, 5 and 6 include about 17 per cent of respondents, having some probabilities of reporting a particular pattern of symptoms: depression, sleep, irritability and fatigue. The total number of cited items ranges from 3 to 9. To some extent, LC13 is similar to this group of classes (i.e. the range of the number of reported items), but the pattern of the symptoms having the highest probabilities is not exactly the same with respect to the previous one (for instance, in LC13 a key role is provided by pessimism). The remaining clusters (LCs 9, 11, 12 and 15) are characterised by a large probability for several depressive indicators: the differences each other involve the type and composition of such symptoms within each class (for instance, in the smallest class – LC15 – nine items show a probability higher than 0.7). About 99.5 per cent of respondents belonging to these LCs report a total number of items larger than three.

The individual probabilities of belonging to each class, computed in the LC analysis, are then used to perform a factor analysis, in order to identify a reduced number of clusters. Results of the factor analysis are summarised in Tables 4 (eigenvalues and the proportion of explained variance by each factor), while Table 5 shows the factor loadings matrix of the chosen solution, after a Varimax rotation. According to the size of the eigenvalues (larger than one), five factors should be retained; however, such solution allows to explain only 56.5 per cent of the total variance. Based on the aims of our analysis, we believe that a greater proportion of variance needs to be explained, at least two third of it. For this reason, we opt for the solution of seven factors, which also appears as a not so large number of depression categories.

Table 6 summarises the size of each retained factor, as well as their composition in terms of the original LCs (from one to four of them). According to the features of these LC compositions, we classify the respondents in the following categories: *very low risk of depression*, *low risk of depression*, *medium risk of depression*, *high risk of depression*, *depressed*, *severely depressed* and *extremely depressed*. Some demographic and socio-economic characteristics of the individuals belonging to each category are reported in the online supplementary material (Table C1).

Robustness checks

In this section we aim at comparing the main similarities and differences in the individual classification according to our approach and the EURO-D indicator. Findings are indeed very interesting and strengthens our results.

According to Figure 4, all respondents in the *extremely depressed* category and nearly all in the *depressed* and *severely depressed* categories are classified as *depressed* also by the EURO-D scale. Gender differences do not appear, in contrast with the evidence within each of the other categories, where the proportion of women classified as depressed according to the EURO-D scale is larger than the one of men. From the *high risk of depression* to the *very low risk of depression* categories, the prevalence of rate of depressed people according to the EURO-D instrument is lower and lower, reaching a minimum in the *very low risk of depression* cluster equals to 3.7 per cent and 6.5 per cent for males and females, respectively. The *middle risk* and the *high risk of depression* clusters show the largest differences by gender.

These conclusions are appealing, but do not take into account differences in sizes among the categories obtained through our approach. To this aim, the online supplementary material (Figure A1) shows the distribution of our seven categories within each group of older people classified as depressed and non-depressed by the EURO-D scale (by gender).

About 55 per cent of individuals defined as non-depressed according to the EURO-D scale fall in either the *very low risk* or the *low risk of depression* category. There is, however, a distinction between male and female respondents: 58 per cent of men and 48 per cent of women are in the *very low risk of depression* category. A similar result (but opposite in sign) arises for the *middle risk of depression* cluster. For the remaining categories of the non-depressed people there are no differences by gender and, in particular, less than 0.25 per cent of respondents are classified in any category from *depressed* to *extremely depressed* in both samples. In such case, it is interesting to underline that all individuals (regardless of the gender) identified as non-depressed by EURO-D and *severely depressed* by our approach report the same pattern of symptoms: depression, pessimism and wishing death.

Conditionally to the depressed group of people obtained according to the EURO-D instrument, no relevant gender differences appear: about 15 per cent of men and 17 per cent of women are in the categories from *very low risk of depression* to *middle risk of depression;* about 28 per cent of men and 25 per cent of women lie in the categories from *depressed* to *extremely depressed.* More than half of the individuals identifying as depressed by the EURO-D scale are classified as at *high risk of depression* according to our approach.

To summarise, all of these results highlight that a large proportion of individuals classified as depressed according to the EURO-D instrument falls in any of our clusters that we have labelled at some risk of depression, with men and women who behave similarly. Furthermore, a non-trivial proportion of older people classified as non-depressed by the EURO-D scale is identified at some risk of depression according to our approach, but, in this case, some gender differences arise.

If we assume as a benchmark the depression status provided by the EURO-D instrument, we might evaluate the goodness of our findings constructing the ROC curve of our classification (Figure A2 in the online supplementary material): the area under the ROC curve is about 0.85, which is a quite good result.

Additionally, in a sensitivity analysis we explore how the results change when choosing a different LC solution, that is, according to the set of indicators, the model with 16 classes. Applying a factor analysis to the probability of belonging to each class, the solution with seven factors explains about 66 per cent of the total variance: in this factor analysis, six eigenvalues are larger than one, while the seventh is equal to 0.984. Exploiting the matrix containing the factor loadings, the retained factors are then constructed as for the model with 15 LCs and the first encouraging result is that we may apply the same labels of the previous classification also to this solution.

Table D1 in the online supplementary material compares the concordance in class belonging of all individuals: more than 95 per cent of people identified as at *low risk* and at *middle risk of depression* in the classification based on 15 LCs are in the same categories also according to the classification based on the 16 LCs model. Very large values may be obtained also for the two extreme categories of the classification. Overall, 79.5 per cent of all individuals are identified in the same category in both solutions and this result is better for males (82.0 per cent) than females (77.7 per cent), as we can see from the online supplementary material (Tables D2 and D3). Some problems arise for people classified as *severely depressed* by the 15 LCs solution: less than ten per cent of them are in the same category also for the 16 LCs model, while the largest proportion of them belongs to the *high risk of depression* cluster according to the 16 LCs solution. Similarly, we may note concordance of classification in the *high risk of depression* category for about 60 per cent of people identified by the 15 LCs estimates, while about 30 per cent are classified in any category from *very low risk* to *middle risk of depression* according to the 16 LCs findings. However, as before, in order to correctly comment this comparison, we have to take into account also the large size differences across these categories.

Looking at the online supplementary material (Tables D2 and D3), 3.2 per cent of male respondents (4.4 per cent of females) classified in any category from *depressed* to *extremely depressed* according to the model with 15 LCs are assigned to any category from *very low risk of depression* to *high risk of* *depression* by the model with 16 LCs. On the other hand, only 2.1 per cent of males (4.9 per cent of females) assigned to the *high risk of depression* or a lower category by the 15 LCs solution are identified as *depressed* (or at higher level) by the other estimation.

Discussion

Using a combination of Latent Class modelling and Factor Analysis we are able to identify seven categories of depressive and emotional problems, from a very low risk of depression cluster to a group of extremely depressed people. These clusters are characterised by different and interesting mixes of symptoms and these differentiations may help epidemiologists, clinicians and health researchers to better observe and understand the presence and the severity of any manifestations of depression among older adults. Moreover, a pithier classification (instead of the classification into a large number of clusters, some of them of small size, as provided by the LC analysis at the beginning) could be more effective, in practice, to be used by any health operators.

It is important to underline that we have labelled the categories according to the main features of the latent classes belonging to each factors. However, clinicians or other experts may rename and reorder these categories in a different way, according to the presence (or not) of some specific characteristics or symptoms they have interest in.

According to our taxonomy, about 40 per cent of the total population is classified *at a very low risk of depression*. All individuals in this cluster present a similar pattern or combination of such symptoms (i.e. no or just one or two reported items). On the other hand, the last three categories are characterised by individuals with a high probability of reporting several depressive and emotional problems. It is likely that in such cases depression corresponds to a clinical diagnosis. The other three categories range from *low risk of depression* to *high risk of depression*: more than half of respondents are in these categories. One third of respondents falls into the *high risk of depression* category: these individuals are not yet classified as depressed, but should be monitored to prevent future development of depression. Gender differences are quite apparent. We observe that nearly half of the older men are in the *very low risk of depression* category, while less than six per cent are classified in the last three categories of depression. Instead, a considerably lower proportion of women than men is in the group *very low risk of depression*. The greatest gender difference is found in the *middle risk of depression* cluster: the number of female respondents who fall in this category is about 1.5 times larger than the one of men. However, the gender difference reduces to only two per cent in the categories from *depressed* to *extremely depressed*.

Lastly, we show interesting comparisons between the approach we adopt and the cut-off point approach of the EURO-D scale. We show that people classified as depressed by the EURO-D scale present different depression characteristics: for example, a large proportion of these individuals (primarily women) are in the *high risk of depression* or a lower category according to our approach. Therefore, our findings can be helpful to further investigate heterogeneity in the manifestation of late-life depression.

Strengths and weaknesses

The main strength of the approach used is the identification of several levels of severity of the depressive disorders, that goes beyond the simple dichotomisation of depression (present or not). Furthermore, this approach allows flexibility in labelling each category that was obtained, according to the nature, type and number of the mental disorders under investigation. Another major strength is that this approach can be easily applied to data collected by any depression instrument, other than the EURO-D scale, designed to detect the presence or the absence of certain symptoms or emotional disorders. Lastly, given that the approach is *strongly* model-driven, subjective evaluation is limited.

Several limitations should be also acknowledged. In factor analysis we could only explain a proportion of the whole variance, nevertheless, it is remarkable the high values of agreement rates between the two solutions we have compared, and the few large discrepancies in the category assignments. Second, in the LC estimation, no active covariates have been specified, but they may be easily introduced to generate the latent classes. Third, we have limited our approach to cross-sectional data for the sake of clarity. However, given the dynamic nature of depression (which often has a chronic course), researchers could be even more interested in modelling developmental trajectories (i.e. the course of a behaviour over age or time) or patterns of change in these outcomes across multiple time points. Latent class models can be easily extended to longitudinal data (dynamic segmentation). In this context, promising approaches involve the development and the estimation of Latent Class Growth Models or Latent Class Markov Models, in order to allow units to change over time the group to which they belong. Lastly, factor analysis seems the most suitable technique for clustering the probabilities obtained after the LC analysis, because its application is straightforward and allows highlighting combinations of patterns of depressive symptoms with some common traits. Other solutions are however suitable, such as building a composite indicator, in particular if some of the 12 EURO-D items might be considered more important than some others.

Conclusion

We have shown that is possible to use a model-based approach for classifying individuals in a more accurate way than the simple dichotomisation *depressed* vs *non-depressed*. Homogeneous groups of people with different levels of depressive or emotional symptoms could be highlighted, in particular those who are at (lower or higher) risk at developing depression. This refined classification of older individuals may provide the basis for improving current protocols for detecting different levels of depressive disorders and help developing customised intervention and treatment programmes.

References

American Psychiatric Association (1980). *Diagnostic and Statistical Manual of Mental Disorders*, Third edition, Washington DC: American Psychiatric Association.

American Psychiatric Association (1987). *Diagnostic and Statistical Manual of Mental Disorders*, Third edition, Revised. Washington DC: American Psychiatric Press.

Beck, A.T., Steer, R.A. and Garbin, M.G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8*, 77-100.

Beekman, A.T., Copeland, J.R. and Prince, M.J. (1999). Review of community prevalence of depression in later life. *British Journal of Psychiatry, 174*, 307-11.

Blazer, D.G. (2003). Depression in Late Life: Review and Commentary. *Journals of Gerontology Series A, 58*(3), 249-65.

Bogner, H., Richie, M., de Vries, H.F. and Morales, K.H. (2009). Depression, Cognition, Apolipoprotein E Genotype: Latent Class Approach to Identifying Subtype. *American Journal of Geriatric Psychiatric, 17*(4), 344-52.

Börsch-Supan, A. (2017). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6*. Release version: 6.0.0. SHARE-ERIC. Data set. DOI: 10.6103/SHARE.w6.600

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., Schaan, B., Stuck, S. and Zuber, S. (2013). Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). *International Journal of Epidemiology, 42*(4), 992-1001.

Brailean, A., Guerra, M., Chua, K.C., Prince, M. and Prina, M.A. (2015). A multiple indicators multiple causes model of late-life depression in Latin American countries. *Journal of Affective Disorders, 184*, 129-36.

Bruce, M.L. and Leaf, P.J. (1989). Psychiatric disorders and 15-month mortality in a community sample of older adults. *American Journal of Public Health, 79*(6), 727-30.

Castro-Costa, E., Dewey, M., Stewart, R., Banerjee, S., Huppert, F., Mendonca-Lima, C., Bula, C., Reisches, F., Wancata, J., Ritchie, K., Tsolaki, M., Mateos, R. and Prince, M. (2007). Prevalence of depressive symptoms and syndromes in later life in ten European countries. The SHARE study. *The British Journal of Psychiatry, 191*, 393-401.

Castro-Costa, E., Dewey, M., Stewart, R., Banerjee, S., Huppert, F., Mendonca-Lima, C., Bula, C., Reisches, F., Wancata, J., Ritchie, K., Tsolaki, M., Mateos, R. and Prince, M. (2008). Ascertaining late-life depressive symptoms in Europe: An evaluation of the EURO-D scale in 10 nations. The SHARE project. *International Journal of Methods in Psychiatric Research, 17*(1), 12-29.

Cole, M.G. and Dendukuri, N. (2003). Risk factors for depression among elderly community subjects: a systematic review and meta-analysis. *American Journal of Psychiatry, 160*(6), 1147-1156.

Dennerstein, L., Astbury, J. and Morse, C. (1993). Psychosocial and mental health aspects of women's health. *World Health Statistics Quarterly, 46*(4), 234- 36.

Dewey, M. and Prince, M. (2005). Mental Health. In A. Börsch-Supan, A. Brugiavini, H. Jürges, J. Mackenbach, J. Siegrist and G. Weber (eds) *Health, ageing and retirement in Europe First results from the Survey of Health, Ageing and Retirement in Europe*, Mannheim: Mannheim Research Institute for the Economics of Aging (MEA), 108-17.

Guerra, M., Ferri, C.P., Llibre, J., Prina, A.M. and Prince, M. (2015). Psychometric properties of EURO-D, a geriatric depression scale: a cross-cultural validation study. *BMC Psychiatry* 15:12.

Hagenaars, J.A and McCutcheon, A.L. (2002). *Applied latent class analysis models*. Cambridge University Press.

Harkness, J. (2005). SHARE Translation Procedures and Translation Assessment. In A. Börsch-Supan and H. Jürges (eds) *The Survey of Health, Ageing and Retirement in Europe – Methodology*, Mannheim: Mannheim Research Institute for the Economics of Aging (MEA), 24-27.

Hybels, C.F., Blazer, D.G., Pieper, C.F., Landerman, L.R. and Steffens, D.C. (2009). Profiles of depressive symptoms in older adults diagnosed with major depression: latent cluster analysis. *American Journal of Geriatric Psychiatric, 17*(5), 387-96.

Jirapramukpitak, T., Darawuttimaprakorn, N., Punpuing, S. and Abas, M. (2009). Validation and factor structure of the Thai version of the EURO-D scale for depression among older psychiatric patients. *Aging & Mental Health, 13*, 899-904.

Kessler, R.C. (2000). Gender differences in major depression: Epidemiological findings. In E. Frank (ed.) *American Psychopathological Association series. Gender and its effects on psychopathology*, Arlington, VA: American Psychiatric Publishing, Inc., 61-84.

Kessler, R.C. (2003). Epidemiology of women and depression. *Journal of Affective Disorders, 74*(1), 5-13.

Kessler, R.C. and Üstün, T.B. (2004). The World Mental Health (WMH) Survey initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research, 13*(2), 93-121.

Larraga, L., Saz, P., Dewey, M.E., Marcos, G., Lobo, A. and the ZARADEMP Workgroup (2006). Validation of the Spanish version of the EURO-D Scale: an instrument for detecting depression in older people. *International Journal of Geriatric Psychiatry, 21*, 1199-1205.

Lee, C.T., Leoutsakos, J.M., Lyketsos, C.G., Steffens, D.C., Breitner, J.C.S. and Norton, M.C. (2012). Latent Class-Derived Subgroups of Depressive Symptoms in a Community Sample of Older Adults: The Cache County Study. *International Journal of Geriatric Psychiatry, 27*, 1061-1069.

Malter, F. and Börsch-Supan, A. (Eds) (2017). *SHARE Wave 6: Panel innovations and collecting Dried Blood Spots*. Munich: Munich Center for the Economics of Aging (MEA).

Oberski, D.L., van Kollenburg, G.H. and Vermunt, J.K. (2013). A Monte Carlo evaluation of three methods to detect local dependence in binary data latent class models. *Advances in Data Analysis and Classification, 7*(3), 267-79.

Pagán-Rodríguez, R. and Pérez, S. (2012). Depression and self-reported disability among older people in Western Europe. *Journal of Aging and Health, 24*(7), 1131-1156.

Penninx, B.W., Geerlings, S.W., Deeg, D.J., van Eijk J.T., van Tilburg, W. and Beekman, A.T. (1999). Minor and major depression and the risk of death in older persons. *Archives of General Psychiatry, 56*(10), 889-95.

Ploubidis, G.B. and Grundy, E. (2009). Later-life mental health in Europe: A country-level comparison. *The Journals of Gerontology – Series B, 64*(5), 666-76.

Portellano-Ortiz, C., Garre-Olmo, J., Calvó-Perxas, L. and Conde-Sala, J.L. (2017). Factor structure of depressive symptoms using the EURO-D scale in the over-50s in Europe. Findings from the SHARE project. *Aging & Mental Health*, forthcoming, available online.

Prince, M. (2013). Cross-Cultural Research Methods and Practice, in V. Patel, H. Minas, A. Cohen and M. Prince (eds.) *Global Mental Health: Principles and Practice*, Oxford University Press, 63-80.

Prince, M., Reischies, F., Beekman, A.T., Fuhrer, R., Jonker, C., Kivelä, S.L., Lawlor, B., Lobo, A., Magnusson, H., Fichter, I., van Oyen, H., Roelands, M., Skoog, I., Turrina, C. and Copeland, J.R. (1999a). Development of the EURO-D scale - a European Union initiative to compare symptoms of depression in 14 European centres. *British Journal of Psychiatry, 174*, 330-38.

Prince, M., Beekman, A.T., Deeg, D.J., Fuhrer, R., Kivelä, S.L., Lawlor, B.A., Lobo, A., Magnusson, H., Meller, I., van Oyen, H., Reischies, F., Roelands, M., Skoog, I., Turrina, C. and Copeland, J.R. (1999b). Depression symptoms in late life assessed using the EURO-D scale. Effect of age, gender and marital status in 14 European centres. *British Journal of Psychiatry, 174*, 339-45.

Pulska, T., Pahkala, K., Laippala, P. and Kivelä, S.L. (1999). Follow up study of longstanding depression as predictor of mortality in elderly people living in the community. *British Medical Journal, 318*(7181), 432-33.

Radloff, L.S. (1977). The CES-D scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*, 385- 401.

Richard, B., Birrer, M.D., Sathya, P. and Vemuri, M.D. (2004). Depression in later life: A diagnostic and therapeutic challenge. *American Family Physicians, 69*(10), 2375-2382.

Robins, L.N., Helzer, J.E., Croughan, J.L. and Ratcliff, K.S. (1981). National Institute of Mental Health Diagnostic Interview Schedule: its history, characteristics and validity. *Archives of General Psychiatry, 38*(4), 381-89.

Robins, L.N., Wing, J., Wittchen, H.U., Helzer, J.E., Babor, T.F., Burke, J.D., Farmer, A., Jablenski, A., Pickens, R., Regier, D.A., Sartorius, N. and Towle, L.H. (1988) The Composite International Diagnostic Interview: an epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry, 45*(12), 1069-1077.

Spitzer, R.L., Williams, J.B., Gibbon, M. and First, M.B. (1992). The Structured Clinical Interview for DSM-III-R (SCID). I: History, rationale, and description. *Archives of General Psychiatry, 49*(8), 624-29.

van Kollenburg, G.H., Mulder, J. and Vermunt, J.K. (2015) Assessing model fit in latent class analysis when asymptotics do not hold. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 11*(2), 65-79.

Veltman, E.M., Lamers, F., Comijs, H.C., de Waal, M.W.M., Stek, M.L., van der Mast, R.C. and Rhebergen, D. (2017) Depressive subtypes in an elderly cohort identified using latent class analysis. *Journal of Affective Disorders, 218*, 123- 30.

Vermunt, J.K. (2003). Multilevel latent class models. *Sociological Methodology, 33*, 213-39.

Vermunt, J.K. and Magidson J. (2003) Latent class models for classification. *Computational Statistics and Data Analysis, 41*, 531-37.

Vermunt, J.K. and Magidson, J. (2004) Latent class analysis. In M.S. Lewis-Beck, A. Bryman and Liao T.F. (eds.) *The Sage Encyclopedia of Social Sciences Research Methods*, Thousand Oaks, CA: Sage Publications, 549-53.

TABLES

Table 1. Classification and total number of reported depressive symptoms according to the EURO-D scale, by gender

EURO-D classification	# of reported EURO-D items	Proportion (per cent)		
		Men	Women	All
NON DEPRESSED	0	28.0	17.6	22.1
	1	23.5	18.5	20.7
	2	17.2	16.5	16.7
	3	11.8	14.1	13.1
	Total	80.5	66.6	72.6
DEPRESSED	4	7.7	11.1	9.6
	5	4.9	8.0	6.7
	6	3.0	5.7	4.5
	7	1.8	3.6	2.9
	8	1.1	2.4	1.8
	9	0.6	1.5	1.1
	10	0.3	0.7	0.5
	11	0.1	0.3	0.2
	12	<0.1	0.1	0.1
	Total	19.5	33.4	27.4

Table 2: Model fit indices of the Latent Class analysis over the depressive symptoms

# LCs	AIC	BIC	CAIC	AIC3	# large BVRs	Classification errors	# small size LCs (<1%)
9	648665	649718	649834	648781	14	0.3298	0
10	648549	649720	649849	648678	7	0.3867	0
11	648473	649762	649904	648615	7	0.4014	0
12	648371	649778	649933	648526	4	0.3812	0
13	648316	649841	650009	648484	1	0.3954	1
14	648252	649895	650076	648433	1	0.4237	0
15	648193	649954	650148	648387	0	0.4176	1
16	648171	650050	650257	648378	0	0.4212	2
17	648167	650164	650384	648387	0	0.4359	2

Table 3: LC analysis results: individual clusters and conditional probabilities of reporting depressive symptoms

	LC														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
LC size (per cent)	37.5	15.3	15.1	7.3	6.7	2.9	2.6	2.5	1.8	1.6	1.6	1.6	1.4	1.3	0.8
Depression	0.030	0.705	0.081	0.969	0.790	0.964	0.930	0.148	0.962	0.239	0.913	0.944	0.245	0.583	0.962
Pessimism	0.070	0.058	0.093	0.091	0.283	0.299	0.156	0.473	0.458	0.005	0.985	0.797	0.982	0.109	0.777
Suicidality	0.005	0.014	0.014	0.190	0.007	0.155	0.076	0.027	0.626	0.078	0.163	0.683	0.137	0.044	0.562
Guilt	0.011	0.083	0.059	0.280	0.139	0.084	0.031	0.014	0.573	0.092	0.045	0.263	0.061	0.014	0.301
Sleep	0.104	0.341	0.387	0.698	0.531	0.769	0.362	0.128	0.823	0.388	0.507	0.585	0.546	0.615	0.957
Interest	0.010	0.011	0.030	0.076	0.142	0.457	0.071	0.079	0.538	0.281	0.622	0.223	0.315	0.038	0.942
Irritability	0.073	0.315	0.261	0.698	0.645	0.570	0.153	0.111	0.815	0.292	0.473	0.509	0.293	0.318	0.634
Appetite	0.013	0.037	0.046	0.144	0.035	0.475	0.070	0.035	0.360	0.314	0.276	0.234	0.242	0.340	0.724
Fatigue	0.070	0.214	0.403	0.693	0.612	0.894	0.308	0.179	0.938	0.651	0.867	0.646	0.691	0.993	0.927
Concentration	0.031	0.026	0.170	0.211	0.291	0.582	0.178	0.166	0.590	0.514	0.664	0.343	0.472	0.161	0.898
Enjoyment	0.033	0.009	0.063	0.033	0.193	0.276	0.142	0.356	0.384	0.300	0.992	0.246	0.408	0.011	0.878
Tearfulness	0.045	0.369	0.073	0.605	0.269	0.700	0.587	0.052	0.795	0.186	0.569	0.551	0.094	0.229	0.819

Table 4: Eigenvalues and proportion of explained variance of the Factor Analysis for solution with 15 LCs

Factor	Eigenvalue	Proportion of explained variance	Cumulative of explained variance
1	2.600	0.173	0.173
2	1.901	0.127	0.300
3	1.536	0.102	0.403
4	1.305	0.087	0.490
5	1.133	0.076	0.565
6	0.982	0.066	0.631
7	0.924	0.062	0.692
8	0.834	0.056	0.748

Table 5: Factor loadings of the Factor Analysis for the solution with 15 LCs (after Varimax rotation, only loadings larger than 0.5 are reported)

Cluster	Factor							Uniqueness
	1	2	3	4	5	6	7	
1			-0.7252					0.1091
2		0.8564						0.2140
3	-0.6656							0.3136
4	0.7552							0.3285
5	0.5693							0.4259
6	0.5000							0.4581
7		0.8072						0.2926
8							0.8727	0.2058
9					0.5634			0.3538
10			0.7142					0.3436
11						0.7264		0.2971
12					0.9004			0.1666
13						0.6754		0.3226
14			0.5570					0.4652
15				0.8035				0.3229

Table 6: Categories obtained according to our approach based on 15 LCs

Category	LCs				Name	Size (per cent)		
						Men	Women	All
1	1	10	14		Very low risk of depression	48.3	34.4	40.4
2	8				Low risk of depression	3.2	1.9	2.5
3	2	7			Middle risk of depression	13.8	21.0	17.9
4	3	4	5	6	High risk of depression	29.0	34.2	32.0
5	11	13			Depressed	2.9	3.2	3.0
6	9	12			Severely depressed	2.3	4.2	3.4
7	15				Extremely depressed	0.5	1.1	0.8

FIGURES

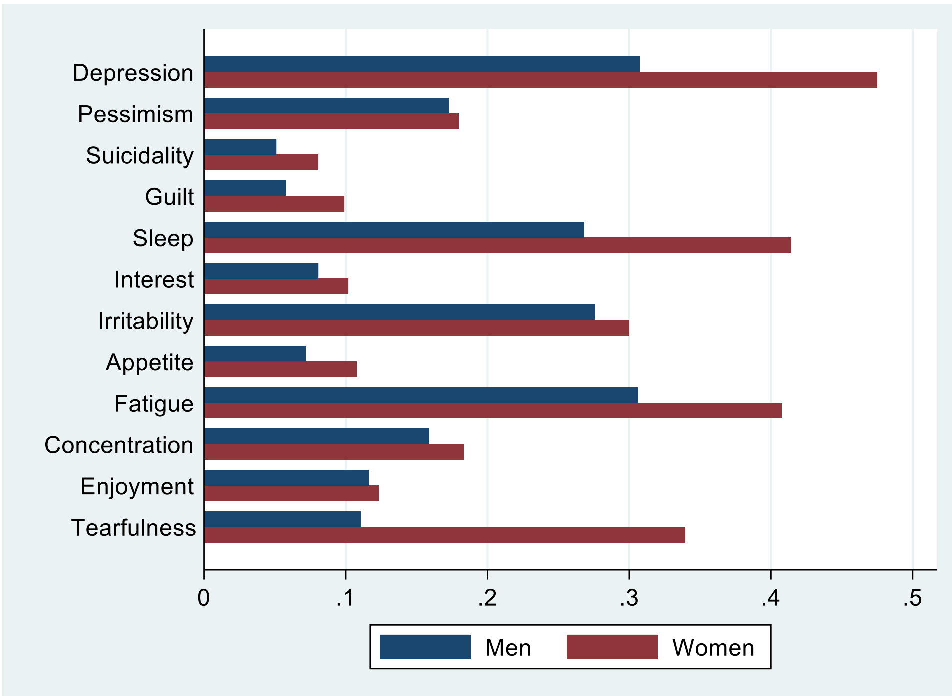

Figure 1: Distribution of the reported depression symptoms, by gender

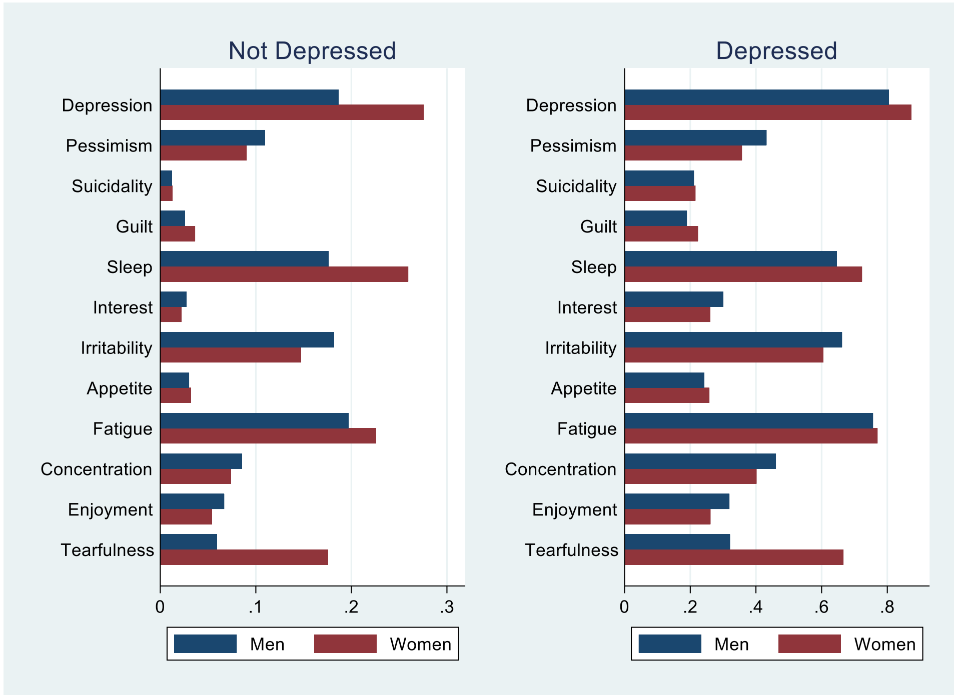

Figure 2: Distribution of the reported depression symptoms, by gender and EURO-D classification

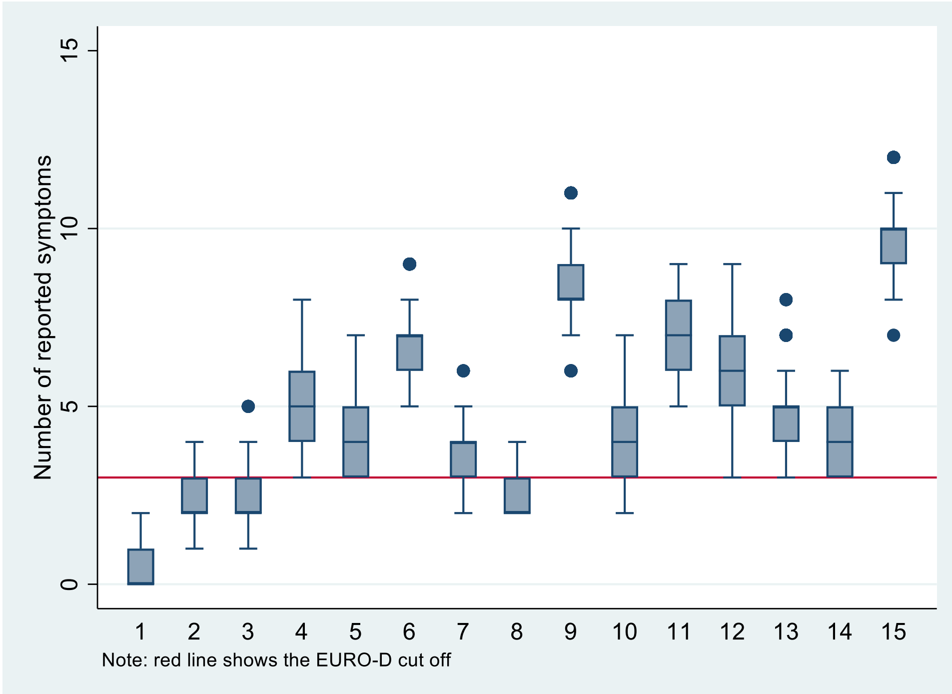

Figure 3: Distribution of the reported depression symptoms, by latent class

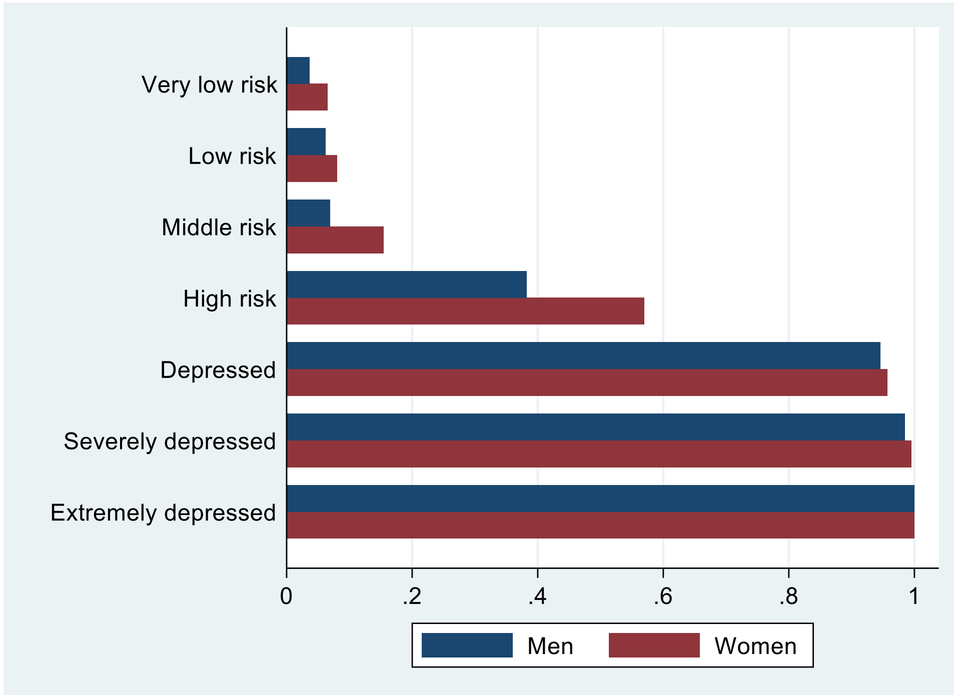

Figure 4: Proportion of the older people classified as depressed according to the EURO-D scale within each category identified according to our approach, by gender