UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTER OF HEALTH SCIENCES
DEPARTMENT OF NEUROPSYCHIATRY

Depression among medical students: an international comparison of prevalence estimates and predictors

Eduardo Karpovisch
Prof. Vicent Balanzá-Martínez
Prof. Beatriz Atienza-Carbonell
Dr. Fabio Porru
Prof. Maurício Scopel Hoffmann

Coordinator:
Prof. João Pedro Gonçalves Pacheco

Santa Maria, Brazil

2024

Abstract

Depression is a debilitating and costly mental health problem. Despite this, its prevalence and associated factors among medical students remain inconclusive. Methodological variations in assessment tools and their cutoff points possibly contributed to inconsistencies in prevalence estimates among previous studies. Stressors and their correlation with depressive symptoms have been identified in some studies; however, the underutilization of standardized assessment tools constitutes another limitation. . To fill this gap, this project aims to provide a refined estimate of the prevalence of depression among medical students, investigate the influence of methodological variations on previous estimates, examine associated factors using a standardized international protocol, and explore how distinct cultures influence academic stressors and their relationship to depressive symptoms. Data will be collected through an online questionnaire administered after an explanation of the project in classrooms at various medical schools in multiple countries. The questionnaire will encompass instruments and inventories designed to assess depressive symptoms and stressors. To mitigate potential confounders, data collection will occur mid-semester, scheduled away from examination periods or just after receiving marks from the previous semester. To address the limitations of previous studies, the measurement invariance of screening instruments will be tested, and prevalence estimates will be adjusted by test characteristics (sensitivity and specificity) using updated Bayesian techniques.

Keywords: depression; mental health; psychiatric epidemiology; medical students; psychometrics

Summary

Introduction

Depression or depressive disorders are characterized by persistent feelings of sadness, loss of interest or pleasure, and a significant impact on daily functioning. It stands as a major public health concern, affecting millions worldwide and significantly impacting their quality of life (Ferrari et al., 2022). Recognized for their demanding academic schedules, sleep deprivation, and anticipation of future professional responsibilities, medical students constitute a population notably vulnerable to depressive symptoms (Tam et al., 2019). However, pinpointing the exact prevalence of depressive disorders and the factors contributing to depressive symptoms in this population remains a challenge. Previous research shows a confusing picture, with estimates varying considerably across studies. This inconsistency possibly stems, among other factors, from methodological differences, including the use of different tools for assessing depressive symptoms and the application of various cutoff points to classify cases of major depression (Rotenstein et al., 2016).

The wide variability in assessment tools is a significant problem. It hinders researchers and policymakers from obtaining a precise estimate of the prevalence of major depression and depressive symptoms and its potential impact on their academic performance, professional development, and overall well-being of medical students. Furthermore, it makes it difficult to develop effective, evidence-based interventions to support this population. This study aims to bridge this knowledge gap by employing a standardized international protocol to assess depression prevalence and its association with academic stressors among medical students across multiple countries. Below we discuss four main problems in the literature that we plan to address with this research.

Problem 1: Heterogeneity of previous estimates

Despite the high prevalence of depressive symptoms among medical students, global review studies reveal a significant degree of variability (Bacchi & Licinio, 2015; Puthran et al., 2016; Rotenstein et al., 2016). This underscores the importance of establishing a consensus-based assessment protocol and conducting thorough investigations into factors that may influence these variations. As highlighted by an seminal meta-analysis by Rotenstein et al., 2016, methodological factors likely play an important role in explaining the variability of depression estimates. This includes the selection of screening tools, time periods, methods for the data collection and minimum necessary response rates.

Generally, depressive symptoms among students are assessed by self-reported questionnaires, instead of clinical interviews, which is the standard method for the diagnosis of a depressive disorder. In Rotentein et al. (2016) meta-analysis, for example, only one of 183 analyzed studies used clinician-administered instruments. When screening tools are used, apparent prevalence estimates need to be adjusted for the uncertainty of the test (sensitivity and specificity) (Gonçalves Pacheco et al., 2023) to derive accurate prevalence estimates.

Problem 2: Measurement invariance of screening instruments

Measurement invariance is a statistical property that ensures the comparability of the same construct across various groups, such as diverse countries, languages, or genders (Vandenberg & Lance, 2000). Measurement variance between groups implies that the comparison between them might not be appropriate. This property is commonly assessed during confirmatory factor analysis. An inherent limitation in prior studies is the absence of previous testing of measurement invariance for the tools that were used, potentially restricting comparisons among various countries. Consequently, the influence of diverse curricular structures and learning environments on the prevalence of depressive symptoms remains unclear.

Problem 3: Gaps in the exploration of sociocultural and psychological determinants

Previous research has identified numerous factors associated with depressive symptoms, including female gender (Pacheco et al., 2019), personality traits (Shi et al., 2015), and recent stressors like the COVID-19 pandemic (Jia et al., 2022). These factors can also influence an individual's stress response, potentially increasing psychological distress, a key determinant of mental well-being (Cohen et al., 2016; Kessler, 1997). For medical students specifically, stressors range from academic workload and performance anxiety to lifestyle factors (Pacheco et al., 2023; Porru et al., 2022; Puthran et al., 2016). However, several limitations exist in current research, such as the need for measures of stressors that are relevant to the context of medical schools. Medical students may experience distinct educational stressors and coping strategies compared to other non-health science students. The demanding nature of the medical school curriculum, coupled with academic and clinical pressures, can significantly impact their mental health in unique ways. In addition, disparities in medical education may exist across various countries. These differences can

manifest in diverse ways, such as variations in practical rotations during the course, disrespectful academic environments (Dahlin et al., 2005; Porru et al., 2022), inadequate practical experience leading to feelings of incapacity, or an overwhelming workload(Tempski et al., 2012).

These potential variations in medical courses underscore the necessity of employing specific tools tailored for this population. Previously developed instruments such as the Medical Student Stress Profile (O'Rourke et al., 2010) and the Medical Students' Stressor Questionnaire (Jayarajah et al., 2020) are not readily accessible across various countries and languages. Additionally, it might be useful to explore the degree of overlap between these instruments and stressors captured by more general, widely validated and culturally adapted inventories related to higher education, such as the Higher Education Stress Inventory (Dahlin et al., 2005). Validating these tools in various languages is crucial to enable an accurate comparison between countries. This approach could allow for a comprehensive understanding of the distinctive stressors encountered within medical education across diverse cultural and linguistic contexts, highlighting the need for international collaboration to enable cross-country comparisons.

Problem 4: Lack of exploration of lifestyle behaviors

Previous studies have linked depressive symptoms among medical students to unhealthy lifestyles (Giner-Murillo et al., 2021; Wong et al., 2024). Notably, changes in diet, influenced by the high academic demands inherent in medical education stressors, warrant attention (Bergeron et al., 2017). Medical education can contribute to increased sedentary behavior, (Dyrbye et al., 2017) and may disrupt sleep patterns due to irregular schedules and demanding academic pressures (Jahrami et al., 2020). Substance abuse might serve as a coping mechanism to handle academic stress, emphasizing the need to comprehend the prevalence and periods during academic life when usage is most intense (Atienza-Carbonell et al., 2022). An additional aspect that remains unclear is the relationship between screen time and outdoor time with overall well-being and depressive symptoms (Balanzá-Martínez & Cervera-Martínez, 2022; Lavados-Romo et al., 2023). However, there are several unanswered questions regarding the potential role of different lifestyle behaviors on the mental health of medical students, including dietary patterns, physical activity, sedentary behavior, sleep patterns, social interaction, stress management, substance use, screen time, and outdoor exposure (Balanzá-Martínez et al., 2021; Hutchesson et al., 2022).

Understanding these environmental factors is essential to inform strategies to prevent depressive symptoms in this population.

Justification

Depressive symptoms are particularly prevalent among medical students. However, global review studies reveal high heterogeneity in reported prevalence of depressive symptoms, limiting a precise identification of positively and negatively associated factors and hindering the development of preventive and management approaches. Employing guiding principles to select depressive symptoms assessment instruments, estimating depression prevalence with updated techniques, testing measurement invariance of multiple instruments, and employing a unified framework for assessing stressors in an international collaboration could significantly advance the development of prevention and detection strategies for depression and depressive symptoms among medical students.

Objectives

Main objective

The main objective of this project is to estimate the prevalence of depression among medical students using a refined methodology considering the previously identified limitations.

Secondary objectives

1. Develop a consensus-based assessment protocol to measure depressive symptoms among medical students, considering factors such as screening tools, data collection methods, and response rates.
2. Establish crosswalk tables for the three most common depression assessment instruments.
3. Investigate the measurement invariance of instruments used to assess depressive symptoms across languages, countries, and cultures.
4. Investigate sociocultural and psychological determinants of depressive symptoms among medical students, including sex/gender, and socio-economic status.
5. Identification of most common stressors utilizing validated tools, such as the Higher Education Stress Inventory and discipline-specific stressor questionnaires, to identify prevalent stressors and their associations with depressive symptoms.
6. Examine the influence of various lifestyle behaviors, including dietary/nutritional habits, physical activity, sedentary behavior, sleep patterns, social connections, stress management, alcohol/substance use, screen time, and outdoor exposure on the mental health of medical students.

Methods

Design

An online questionnaire will be used, incorporating established tools like Higher Education Stress Inventory (HESI), Patient Health Questionnaire-9 (PHQ-9), University-Short Multidimensional Inventory Lifestyle Evaluation (U-SMILE), Beck's Depression Inventory (BDI), Center for Epidemiological Studies-Depression (CES-D) along with our proposed new tool called Medical School Stressors Inventory (MSSI), drawing inspiration from Medical Student Stress Profile (MSSP) and Medical Student Stress Questionnaire (MSSQ). These questionnaires will explore various aspects of the medical students' lives, encompassing socio-demographic factors, symptoms of depression, stressors and lifestyle factors. The study intends to explore multiple countries with diverse cultures, languages, and education. By doing so, it aims to uncover both commonalities and disparities and the degree to which they associate with depressive symptoms among medical students across different world regions.

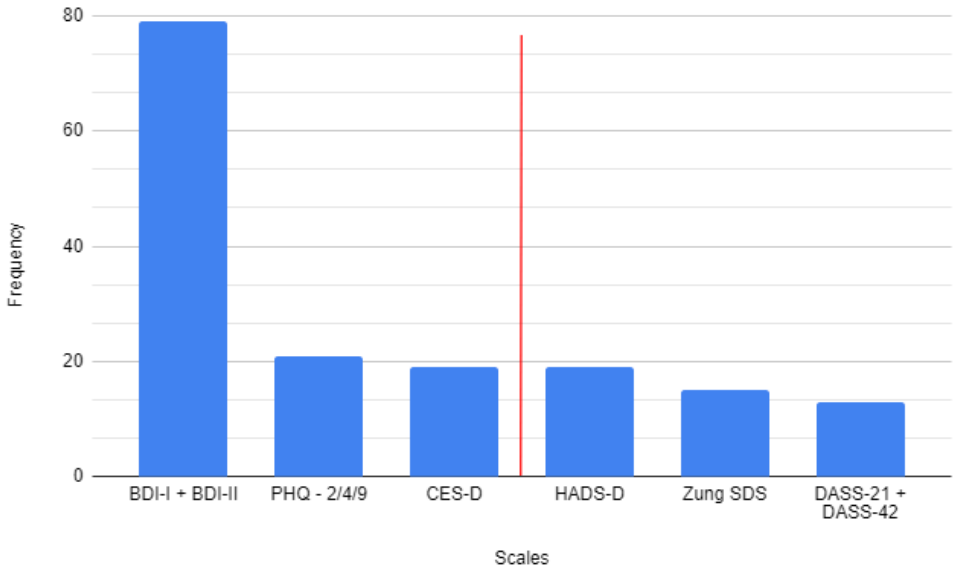
Measures

The selection of depressive symptom tools was based on those instruments commonly utilized in the literature for assessing depressive symptoms, as highlighted in the overview of systematic reviews conducted by Tam et al., in 2019 (Tam et al., 2019) (Figure 1). Employing the identified tools allows the data to be comparable with previous studies and future research in this field.

Additional principles were taken into account for the selection of instruments. The tools were required to (1) be freely accessible, (2) demonstrate established construct validity through exploratory and confirmatory factor analyses (EFAs and CFAs), and (3) have already undergone translation, cultural adaptation, and validation for the initial countries participating in the study (Brazil, Spain and Italy). Furthermore, the administration time for each tool needed to be practical, ensuring that the entire questionnaire, inclusive of all tools, could be completed within approximately 20-30 minutes. The Hospital Anxiety and Depression Scale (HADS-D) and the Center for Epidemiological Studies Depression Scale (CES-D) were utilized in an equal number of studies (Tam et al., 2019), so we opted for the CES-D over the HADS because the HADS was primarily developed to assess depression

among hospitalized patients (Zigmond & Snaith, 1983), whereas the CES-D was designed to explore depression within the broader general population (Radloff, 1977).

Figure 1. Frequency of tools used for assessing depressive symptoms among medical students



Legend: BDI-I and BDI-II: Beck's Depression Inventory; PHQ-2/4/9: Patient Health Questionnaire with 2, 4 or 9 items; CES-D: Center for Epidemiological Studies Depression Scale; HADS-D: Hospital Anxiety and Depression Scale; Zung SDS: Zung Self-Rating Depression Scale; DASS-21: Depression, Anxiety and Stress Scale. The red line separates the tools for inclusion in this project. Figure was generated using data based on Tam et al. (2016).

Socioeconomic status and demographic information

The following information will be collected: sex, age, gender identity, sexual orientation, household income, student's highest level of education, household composition, immigration status, marital status, self-reported previous mental health diagnoses and treatments, and physical conditions. Data on sex and age will be collected to adjust weights so that the sample reflects the demographics of each center.

There is extensive literature on the correlation between socioeconomic status and depressive symptoms (Lorant et al., 2003; Mac-Ginty et al., 2024; Schlax et al., 2019). Therefore, it is important to investigate these variables and their influence on local students at each institution. To address this need, family socioeconomic status/class, financial stress, parental education (Mac-Ginty et al., 2024), and subjective socioeconomic status will be collected. Subjective socioeconomic status will be assessed using the MacArthur Scale of

Subjective Social Status (Adler et al., 2000), available in Appendix 1. Financial stress is contemplated in the Higher Educational Stress Inventory (HESI), discussed below.

We will employ four ultra-brief measures assessing self-reported mental health, general health, general happiness, and well-being, using standardized 5-point Likert-type scales. Participants will be asked: "In general, would you say your health is?" with response options ranging from 1 (Excellent) to 5 (Poor), based on the 36-Item Short Form Survey Instrument (SF-36) (Brazier et al., 1992). For mental health, the question "In general, would you say that your mental health is?" will be used, with the same response options, adapted from the SF-36. Life satisfaction will be measured by asking, "How satisfied are you with your life as a whole?" with response options from 1 (Very Dissatisfied) to 5 (Very Satisfied). Finally, happiness will be assessed with the question "How happy would you say you are overall?" with response options from 1 (Not at all happy) to 5 (Very happy) (Raudenská, 2023).

PHQ-9

The Patient Health Questionnaire-9 (PHQ-9) (Kroenke et al., 2001) comprises nine items evaluated on a Likert-type response scale ranging from 0 to 3. It is utilized to measure the severity of depressive symptoms in individuals and is commonly administered in primary care or medical settings. Regarded as one of the most extensively validated tools for depression screening (El-Den et al., 2018), its nine items are tailored to assess the nine symptom criteria for depression outlined in the Diagnostic and Statistical Manual of Mental Disorders Fourth Edition (DSM-IV), taking around 2 minutes to complete in a touchscreen computer (Fann et al., 2009). In a recent meta-analysis (Levis et al., 2019), varying sensitivity and specificity values were observed for the PHQ-9 when compared to psychiatric interviews. In direct comparison with a fully structured interview, the PHQ-9 with a cut-off of 10 points exhibited a sensitivity of 0.70 and a specificity of 0.84. However, when tested against semi-structured interviews, sensitivity increased to 0.88 and specificity to 0.85. Furthermore, when contrasted with the MINI interview, the PHQ-9 demonstrated a sensitivity of 0.77 along with a specificity of 0.87. Across all these assessment methods, an optimal cut-off score of 10 points was consistently identified. It is also the cut-off score recommended by the developers of the PHQ-9 (Kroenke et al., 2001). The items of the tool are available in Table 1 of the Appendix.

BDI

The Beck's Depression Inventory (BDI) is one of most used tools for measuring depression severity, developed by Beck et al., in 1961. It consists of 21 items with a 4-point response scale ranging from 0 to 3. Higher values represent more severe symptomatology, with a minimum score of 0 and maximum of 63. In 1996, a new-revised version from the BDI (BDI-II) was published, adapting the tool to DSM-IV-defined symptoms. It takes around 5 to 10 minutes to complete and requires a fifth to sixth grade reading level to complete (Groth-Marnat, 2006). It is available in multiple languages. Currently, several studies employ various thresholds for the Beck Depression Inventory-II (BDI-II). However, a diagnostic meta-analysis indicates that the optimal cut-off for the general population is 13 points, with a sensitivity of 0.86 and a specificity of 0.78 (Von Glischinski et al., 2019). The BDI-II is available as table 2 of the Appendix.

CES-D

The Center for Epidemiological Studies Depression Scale (CES-D) was originally developed by Radloff and colleagues (Radloff, 1977) for use in epidemiological studies of depression in the general population. The tool consists of 20 items, rated in a Likert-type scale ranging from 0 to 3, although multiple shortened versions are available. For this study, we opted for the 20-item version of the questionnaire, for being more frequently used in the literature. The conventional cut-off score is 16, which gives the sensitivity of 0.87 and specificity 0.70, taking around 2 to 5 minutes to complete (Vilagut et al., 2016). The items of the tool are available in Table 3 of the Appendix.

HESI

The Higher Educational Stress Inventory (HESI), conceptualized by Dahlin and colleagues in 2005 (Dahlin et al., 2005), comprises 33 items graded on a Likert-type scale from 1 to 4. This tool serves to assess stressors among university students. These stressors encompass aspects like academic overload, uncertainties about the future, financial constraints, and challenges related to self-esteem and motivation. Notably, the inventory has been applied across various countries, including Sweden, South Korea, Jordan, Uganda, and Brazil, and has received validation for languages such as English (Dahlin et al., 2005), Portuguese (Pacheco et al., 2023), Korean (Shim et al., 2016), Arabic (Masha'al et al., 2021), and Italian (Ferrara et al., 2023). HESI items are available in Table 4 of the Appendix.

U-SMILE

The Short Multidimensional Inventory Lifestyle Evaluation (SMILE; Balanzá-Martínez et al., 2021) was initially designed with 43 items, aiming to evaluate lifestyle behaviors or domains aligned with the contemporary definition of lifestyle by the American College of Lifestyle Medicine: diet/nutrition, substance use, physical activity, stress management, social support , restorative sleep, and environmental exposures. Later, a brief version with 24 items was developed for university students, the "University-SMILE" or U-SMILE (De Boni et al., 2023). Employing a Likert-type scale ranging from 1 to 4, this instrument serves the purpose of assessing overall patterns of lifestyle behaviors among university students. The items of the U-SMILE are available as Table 5 of the Appendix.

MSSI

For the purposes of this project, we introduce a novel tool called the Medical School Stressors Inventory (MSSI, table 8) integrating stress-related questions within the specific context of medical school from two existing tools, the Medical Student Stress Questionnaire(MSSQ) and the Medical Student Stress Profile (MSSP). We opted to not use the MSSQ and the MSSP in this study because they are lengthy, and items significantly overlap with general university life stressors.

The MSSQ, initially developed by Yusoff et al., in 2010 (Yusoff et al., 2010), was designed to explore potential stressors across six distinct domains within the lives of medical students: Academic-related stressors, intrapersonal and interpersonal stressors, teaching and learning-related stressors, social stressors, drive and desire-related stressors, and group activities-related stressors. Comprising 40 items, respondents rate their experiences using a Likert-type scale ranging from 0 to 4, indicating stress levels from "Causing no stress at all" to "Causing severe stress." However, this tool has limited focus on medical-specific contexts and lacks validation in numerous languages, currently being predominantly employed in Asian countries (Yusoff, 2017).

The MSSP, developed by O'Rourke et al. in 2010 (O'Rourke et al., 2010), serves as a tool specifically tailored for educational contexts within the medical field. It identifies facets such as course organization, supervision, time pressures, social pressures, and medical-specific or patient-related issues. This tool involves 52 items where students report the frequency and intensity of each item on a Likert-type scale ranging from 0 to 5. Subsequently, a geometric

mean is calculated for both the frequency and severity responses to determine the score for each item. Due to its complex application process, extensive length (52 items answered twice), both MSSQ and MSSP were revised, items were compared and harmonized for item content. Harmonization was facilitated by the use of Harmony (Moltrecht et al., 2023). A side-by-side comparative display of both tools is available in the Appendix, Table 6. Questions pertaining to the medical context are highlighted in blue for easy identification and reference.

The new proposed inventory was derived by specifically selecting questions related to the medical practice context. Items 22, 39, 46, 47, 48, 49, 50, 51, and 52 were extracted from MSSP, while questions 2, 17, 21, and 24 were chosen from MSSQ. Given the similarity among some of these questions assessing similar aspects, items 46 and 47 from MSSP were excluded as they overlap with item 24 from MSSQ. Additionally, item 51 from MSSP was omitted due to its similarity to item 2 from MSSQ. A table detailing the primary selected items, along with their corresponding original questionnaire numbers and highlighting the excluded items, is provided as Table 7 in the Appendix of the project. The final version of the MSSI has 10 items, and each question in the inventory is rated using a Likert-type scale from 1 to 4, with response options based on the MSSQ levels, as well as minor modifications to some items. The finalized version of the inventory can be accessed as Appendix Table 8.

As this is a newly developed tool, the inventory will be translated into the languages of the participating countries before being incorporated into the project questionnaire. The initial step involves creating a machine translation (Google Translator and GPT-4o) that will be evaluated by two independent researchers, both fluent in English (original version) and the target language. The result of this step will be a translated version by consensus, which will be evaluated by a third researcher proficient in English and the target language. This revised version will then be presented to a focus group of medical students, aiming to identify cultural nuances. In addition, factor analyses, including EFA, CFA, IRT, and measurement invariance assessments will be conducted to ensure the robustness of the statistical properties of the MSSI.

Procedure

Each center will administer the survey midway through the semester to avoid high-stress periods such as final exams or the release of grades from the previous semester. Participants

will complete an online questionnaire on the KoboToolbox platform, designed to be completed in approximately 30 minutes to avoid fatigue.

To ensure comprehensive participation, 5 minutes of a mandatory class (e.g., seminars and clinical rotations in some countries) or lectures with high expected attendance will be used to explain the project's importance and distribute the questionnaire for home completion. A follow-up reminder will be given two weeks later during the same class to minimize class time disruption. Additionally, class representatives will be encouraged to use social media platforms, such as WhatsApp groups, to remind absentees to participate. via social media platforms such as WhatsApp groups that involve their respective classes.

Stratified sampling

In this study, stratified sampling will be employed to select groups encompassing various stages of medical education within participating centers. Given the diversity of medical education systems across different countries, our approach will be adapted to each context to ensure internal consistency. For countries with uniform structures, such as a continuous six-year program, participants will be stratified into three segments representing different stages of the curriculum (e.g., years 1-2, 3-4, 5-6). In cases where medical education follows a different structure, stratification will be adjusted accordingly. We will stratify based on the most relevant educational milestones within each system. This approach allows for a consistent comparison within each country while acknowledging the inherent variability across different settings. Raking will be used to adjust sample cases to match known population distributions of age and gender (as described in the Statistical Analysis section).

Statistical analysis

The statistical analysis of the data will be conducted using the latest versions of R and RStudio software.

Raking

A common method for weighting is iterative proportional fitting, or raking (Battaglia et al., 2009; Yap et al., 2022). This technique adjusts the weights of survey cases so that the sample distribution aligns with known population distributions for selected variables. We will use age and gender for raking, adjusting the weights until all target distributions are

met. Raking requires only the marginal proportions of each variable, without needing the population proportion for every combination of characteristics. This method ensures that the weighted survey sample accurately reflects the demographics of the medical student population at each center. Therefore, each center will be required to provide marginal proportions of age and gender for their participating population as a prerequisite for participation in this study.

Prevalence estimation

The selected tools will undergo dichotomization based on identified cut-offs. Subsequently, the prevalence of depression will be calculated, considering both apparent prevalence and adjusted values accounting for sensitivity and specificity using Bayesian adjustment techniques within the R package TruePrev. For each center, a chi-squared test will be utilized to evaluate the prevalence differences across various instruments. As this prevalence refers to the true prevalence estimate, no differences are expected.

Finally, results of the depression screening instrument with highest measurement invariance will be meta-analyzed using a random-effects model.

Examination of associated factors

Linear regression will be employed to explore the relationships between stressors, lifestyle, and socioeconomic factors.

Crosswalks between depression instruments

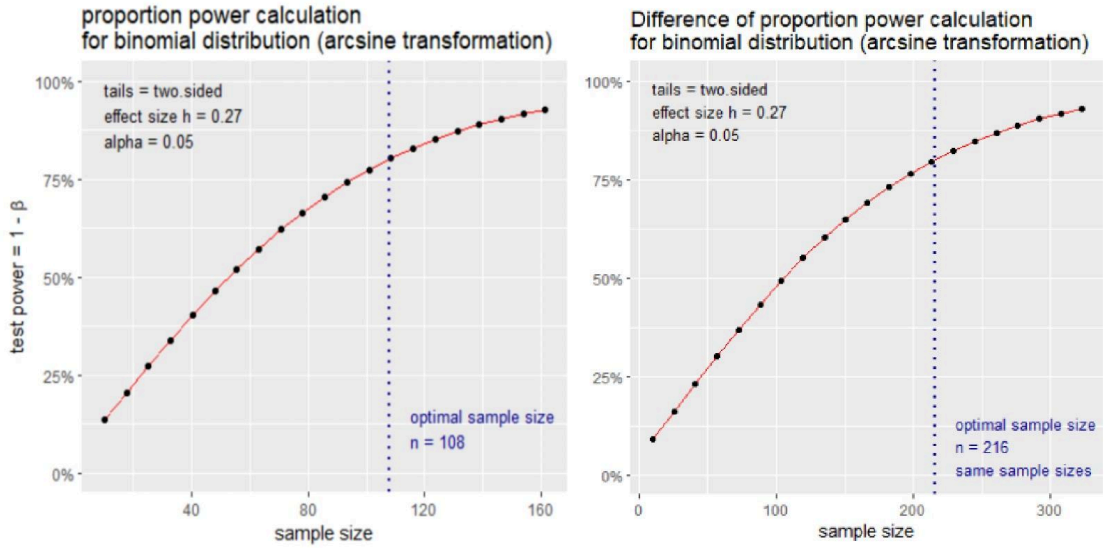
A single-group design (Choi et al., 2014) will be utilized to create crosswalk tables connecting the three selected depression assessment tools: PHQ-9, CES-D, and BDI. All participants will be administered items from each instrument, and scores will be collected for each individual on every measure to create linkage tables. Equipercentile linking, a non-IRT approach, will be employed to estimate these tables. This method establishes a nonlinear linking relationship by aligning scores with corresponding percentile ranks on the score distributions of the instruments. Score smoothing will be conducted due to the susceptibility of this linking procedure to random sampling error.

Measurement invariance allows for discerning whether differences observed in measurements between groups due from actual differences in the intensity of the same construct or if these variations are influenced by other factors, such as gender, language, cultural aspects etc. A multigroup Confirmatory Factor Analysis (CFA) approach will be implemented following the Wu and Estabrook method. This involves a step-by-step application of constraints, assessing the fit indices for each constrained model. The initial constraint involves establishing equal configural equivalence, thereby constraining the model structure to be identical across groups. Subsequently, the same item thresholds across groups are set, followed by equalizing the factor loadings between groups. Finally, the latent intercepts are constrained to achieve scalar equivalence. To ascertain measurement equivalence between these groups, differences in fit indices such as CFI (Comparative Fit Index), RMSEA (Root Mean Square Error of Approximation), and SRMR (Standardized Root Mean Square Residual) will be examined. A discrepancy of less than 0.01 for CFI, 0.015 for RMSEA, and 0.010 for SRMR indicates measurement equivalence. This outcome implies that the same tool effectively measures the same underlying construct across the groups. The measurement invariance of the instruments used in this project (PHQ-9, BDI, CES-D, HESI, U-SMILE and MSSI) will be calculated for groups of gender, age, course period, socioeconomic status and country of data collection.

Sample size

Sample size calculation was implemented using the R "pwr" package. This package incorporates the "pwr.p.test" function, which determines the minimum sample size required to estimate prevalence. The correlation between sample size and test power can be visualized in (Figure 2, left). Based on data from Rotenstein et al.'s meta-analysis, estimating a prevalence of 27% and using a significance level of 0.05 along with a test power of 0.8, the analysis determined a necessary sample size of 108 participants per center for calculating prevalence. To evaluate the minimum sample size required for comparing prevalence between two groups, the function "pwr.2p.test" within R's "pwr" package was utilized. With the same previously provided inputs, this function yielded an output indicating a requirement of 216 participants (Figure 2, right).

Figure 2. Representation of an optimal sample size for a test power of 0.8, in a binomial distribution with an effect size of 27% and a significance level of 0.05' for a single population (left) or two populations (right).



The calculation of sample size for measurement invariance lacks a consensus formula. Typically, recommendations span from 3 to 20 participants per item in the questionnaire for factorial analysis (Mundfrom et al., 2005). This wide range is attributed to variations in measures such as RMSEA, which tends to excessively reject accurate models in smaller sample sizes. Conversely, the χ2 statistic may result in over-rejection of measurement invariance tests when the sample size is larger. Consequently, the sample size for measurement invariance is considered less pivotal compared to other study characteristics such as the number of groups and model size (Putnick & Bornstein, 2016). In line with these considerations, this project adopts a sample size of 300 participants. This decision is based on the HESI tool, which contains 33 items to be validated, ensuring approximately nine participants per item. Additionally, the number 300 is a multiple of both 50 and 60, corresponding to typical medical class sizes. Furthermore, a sample size of 300 aligns with the minimum recommended for factor analysis (Frost et al., 2007).

Ethics approval and consent to participate

The study adheres to the principles outlined in the Declaration of Helsinki. Prior to participating in the survey, all participants will be required to provide informed consent. To safeguard the privacy of participants, collected data will be anonymized. Handling of data will be conducted in accordance with data privacy regulations such as the *Lei Geral de Proteção de Dados* (LGPD) since the data will be stored in Brazil. Participants have the right to withdraw from the study at any point. In accordance with ethical guidelines at certain study sites, students exhibiting symptoms of mental health issues may be directed to specialized mental health services offered by the institutions. Ethical approval will be sought from the Brazilian ethics committee. Centers located outside of Brazil must obtain approvals from their respective ethics committees.

Participants inclusion

Participant inclusion criteria

Participants are eligible for inclusion in the study if they meet the following criteria: fluency in the language used at the participating centers and current enrollment in a Medicine course at the participating center.

Current participating centers

The confirmed participant countries and their associated centers in the project are as follows:

- Spain: Universitat de València
- Brasil: Universidade Federal de Santa Maria
- Italy: awaiting contact feedback.

Inclusion of other participating centers

Project dissemination plans

Dissemination plans for this project prioritize engaging a diverse range of medical schools, necessitating the spread of information across various universities. To achieve this objective, multiple strategies will be employed. Firstly, the project will be shared on the OSF platform, facilitating dissemination among fellow researchers. Utilizing Twitter (X) will target students and researchers potentially interested in participation. Furthermore, a viewpoint article has been submitted for publication, offering an overview of the project's core concepts and inviting collaboration. Additionally, direct outreach via email will be conducted to esteemed experts within universities, soliciting contributions. These concerted efforts aim to enhance awareness about the project and foster engagement from a broad spectrum of medical schools.

Inclusion requisites

Centers wishing to participate need to meet certain requirements, including:

1. Offering a medical course at the university;
2. Appointing a teacher to oversee the project's implementation at the center;
3. Having students from all semesters available for data collection;
4. Provide the marginal proportions of age and gender of the participating center population .

Activities and actions schedule

The project comprises several distinct phases, such as: the recruitment phase, aimed at onboarding new participant centers; the data collection phase; the data analysis phase; and the final phase involving drafting the article for publication. Monthly meetings are scheduled to review project progress. Additionally, there's flexibility to attend conferences and other events at any point during the year to present the project and promote collaborations with other centers, seeking further support for its advancement.

A synthesis of the schedule is available on Table 1 below.

Table 1. Annual schedule of activities and actions

Planned steps	2nd Semester 2024	1st Semester 2025	2nd semester 2025	1st semester 2026
Recruiting new members	x	x		
Literature review	x	x	x	x
Monthly meetings	x	x	x	x
Collection of data		x	x	
Data analysis			x	x
Article drafting and publication procedures				x
Dissemination in academic events	x	x	x	x

Expected results

Based on the outlined objectives , the expected results of the project are:

- A research framework applicable across countries and medical schools to assess depressive symptoms and identify potential stressors among medical students.
- Validation of existing assessment tools and a new scale, the MSSI.
- Evaluation of the estimated prevalence of depressive disorders among medical students.
- Identification potential risk and protective factors relevant to depressive symptoms in this population.
- Analyzing the impact of the different contexts of each center on the measures derived from the project instruments.

Limitations

The project is subject to two primary limitations. Firstly, being an observational study, it lacks the capacity to establish causation. Secondly, it does not incorporate a diagnostic interview, potentially leading to discrepancies between the data obtained from the tools and actual caseness. However, the primary aim of this study is to perform population screening employing state-of-the-art techniques, thus providing precise prevalence estimates at the population level rather than the individual level.

References

Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy, White women. *Health Psychology*, *19*(6), 586–592. https://doi.org/10.1037/0278-6133.19.6.586

Atienza-Carbonell, B., Guillén, V., Irigoyen-Otiñano, M., & Balanzá-Martínez, V. (2022). Screening of substance use and mental health problems among Spanish medical students: A multicenter study. *Journal of Affective Disorders*, *311*, 391–398. https://doi.org/10.1016/j.jad.2022.05.090

Bacchi, S., & Licinio, J. (2015). Qualitative Literature Review of the Prevalence of Depression in Medical Students Compared to Students in Non-medical Degrees. *Academic Psychiatry*, *39*(3), 293–299. https://doi.org/10.1007/s40596-014-0241-5

Balanzá-Martínez, V., & Cervera-Martínez, J. (2022). Lifestyle Prescription for Depression with a Focus on Nature Exposure and Screen Time: A Narrative Review. *International Journal of Environmental Research and Public Health*, *19*(9), 5094. https://doi.org/10.3390/ijerph19095094

Balanzá-Martínez, V., Kapczinski, F., De Azevedo Cardoso, T., Atienza-Carbonell, B., Rosa, A. R., Mota, J. C., & De Boni, R. B. (2021). The assessment of lifestyle changes during the COVID-19 pandemic using a multidimensional scale. *Revista de Psiquiatría y Salud Mental*, *14*(1), 16–26. https://doi.org/10.1016/j.rpsm.2020.07.003

Battaglia, M. P., Hoaglin, D. C., & Frankel, M. R. (2009). Practical Considerations in Raking Survey Data. *Survey Practice*, *2*(5), 1–10. https://doi.org/10.29115/SP-2009-0019

Bergeron, N., Al-Saiegh, S., & Ip, E. J. (2017). An Analysis of California Pharmacy and Medical Students' Dietary and Lifestyle Practices. *American Journal of Pharmaceutical Education*, *81*(8), 5956. https://doi.org/10.5688/ajpe5956

Brazier, J. E., Harper, R., Jones, N. M., O'Cathain, A., Thomas, K. J., Usherwood, T., & Westlake,

L. (1992). Validating the SF-36 health survey questionnaire: New outcome measure for primary care. *BMJ*, *305*(6846), 160–164. https://doi.org/10.1136/bmj.305.6846.160

Choi, S. W., Schalet, B., Cook, K. F., & Cella, D. (2014). Establishing a common metric for depressive symptoms: Linking the BDI-II, CES-D, and PHQ-9 to PROMIS Depression. *Psychological Assessment*, *26*(2), 513–527. https://doi.org/10.1037/a0035768

Cohen, S., Gianaros, P. J., & Manuck, S. B. (2016). A Stage Model of Stress and Disease. *Perspectives on Psychological Science*, *11*(4), 456–463. https://doi.org/10.1177/1745691616646305

Dahlin, M., Joneborg, N., & Runeson, B. (2005). Stress and depression among medical students: A cross-sectional study. *Medical Education*, *39*(6), 594–604. https://doi.org/10.1111/j.1365-2929.2005.02176.x

De Boni, R. B., Mota, J. C., Schuch, F. B., Pires, D. A., Matias, T. S., Monteiro-Junior, R. S., Deslandes, A. C., Silva, D. R., Moura, H. F., Galvão-Coelho, N., Jesus-Moraleida, F. R. de, Kapciznki, F., & Balanzá-Martinez, V. (2023). U-SMILE: a brief version of the Short Multidimensional Inventory on Lifestyle Evaluation. *Trends in Psychiatry and Psychotherapy*. http://dx.doi.org/10.47626/2237-6089-2023-0722

Dyrbye, L. N., Satele, D., & Shanafelt, T. D. (2017). Healthy Exercise Habits Are Associated With Lower Risk of Burnout and Higher Quality of Life Among U.S. Medical Students. *Academic Medicine*, *92*(7), 1006–1011. https://doi.org/10.1097/ACM.0000000000001540

El-Den, S., Chen, T. F., Wong, E., & O'Reilly, C. L. (2018). The psychometric properties of depression screening tools in primary healthcare settings: A systematic review. *Journal of Affective Disorders*, *225*, 503–522. https://doi.org/10.1016/j.jad.2017.08.060

Fann, J. R., Berry, D. L., Wolpin, S., Austin-Seymour, M., Bush, N., Halpenny, B., Lober, W. B., & McCorkle, R. (2009). Depression screening using the Patient Health Questionnaire-9 administered on a touch screen computer. *Psycho-Oncology*, *18*(1), 14–22. https://doi.org/10.1002/pon.1368

Ferrara, P., Delli Poggi, A., Ruta, F., Porru, F., Destrebecq, A., & Terzoni, S. (2023). Misurazione dello stress nella formazione infermieristica: Validazione della Versione italiana della Higher Education Stress Inventory (HESI) in un campione di studenti del Corso di Laurea in Infermieristica. *PROFESSIONI INFERMIERISTICHE*, *76*(1), 21–27.

Ferrari, A. J., Santomauro, D. F., Herrera, A. M. M., & Whiteford, H. A. (2022). Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019. *The Lancet Psychiatry*, *9*(2), 137–150. https://doi.org/10.1016/S2215-0366(21)00395-3

Frost, M. H., Reeve, B. B., Liepa, A. M., Stauffer, J. W., & Hays, R. D. (2007). What Is Sufficient Evidence for the Reliability and Validity of Patient-Reported Outcome Measures? *Value in Health*, *10*, S94–S105. https://doi.org/10.1111/j.1524-4733.2007.00272.x

Giner-Murillo, M., Atienza-Carbonell, B., Cervera-Martínez, J., Bobes-Bascarán, T., Crespo-Facorro, B., De Boni, R. B., Esteban, C., García-Portilla, M. P., Gomes-da-Costa, S., González-Pinto, A., Jaén-Moreno, M. J., Kapczinski, F., Ponce-Mora, A., Sarramea, F., Tabarés-Seisdedos, R., Vieta, E., Zorrilla, I., & Balanzá-Martínez, V. (2021). Lifestyle in Undergraduate Students and Demographically Matched Controls during the COVID-19 Pandemic in Spain. *International Journal of Environmental Research and Public Health*, *18*(15), 8133. https://doi.org/10.3390/ijerph18158133

Gonçalves Pacheco, J. P., Salum, G., & Hoffmann, M. S. (2023). Estimated Prevalence of Perinatal Depression in Low- and Middle-Income Countries Considering

Uncertainty of Test Properties. *JAMA Psychiatry*, *80*(8), 856. https://doi.org/10.1001/jamapsychiatry.2023.2016

Groth-Marnat, G. (2006). *Handbook of Psychological Assessment* (Fourth). John Wiley & Sons, Inc.

Hutchesson, M. J., Whatnall, M. C., Yazin, N., Fenton, S., Duncan, M. J., Kay-Lambkin, F. J., & Burrows, T. L. (2022). Health behavior interventions for university students measuring mental health outcomes: A scoping review. *Frontiers in Public Health*, *10*, 1063429. https://doi.org/10.3389/fpubh.2022.1063429

Jahrami, H., Dewald-Kaufmann, J., Faris, M. A.-I., AlAnsari, A. M. S., Taha, M., & AlAnsari, N. (2020). Prevalence of sleep problems among medical students: A systematic review and meta-analysis. *Journal of Public Health*, *28*(5), 605–622. https://doi.org/10.1007/s10389-019-01064-6

Jayarajah, U., Lakmal, K., Athapathu, A., Jayawardena, A. J., & De Silva, V. (2020). Validating the Medical Students' Stressor Questionnaire (MSSQ) from a Sri Lankan medical faculty. *Journal of Taibah University Medical Sciences*, *15*(5), 344–350. https://doi.org/10.1016/j.jtumed.2020.08.003

Jia, Q., Qu, Y., Sun, H., Huo, H., Yin, H., & You, D. (2022). Mental Health Among Medical Students During COVID-19: A Systematic Review and Meta-Analysis. *Frontiers in Psychology*, *13*, 846789. https://doi.org/10.3389/fpsyg.2022.846789

Kessler, R. C. (1997). THE EFFECTS OF STRESSFUL LIFE EVENTS ON DEPRESSION. *Annual Review of Psychology*, *48*(1), 191–214. https://doi.org/10.1146/annurev.psych.48.1.191

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9 Validity of a Brief Depression Severity Measure. *Journal of General Internal Medicine*, *16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Lavados-Romo, P., Andrade-Mayorga, O., Morales, G., Muñoz, S., & Balboa-Castillo, T. (2023).

Association of screen time and physical activity with health-related quality of life in college students. *Journal of American College Health*, *71*(5), 1504–1509. https://doi.org/10.1080/07448481.2021.1942006

Levis, B., Benedetti, A., & Thombs, B. D. (2019). Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: Individual participant data meta-analysis. *BMJ*, l1476. https://doi.org/10.1136/bmj.l1476

Lorant, V., Deliège, D., Eaton, W., Robert, A., Philippot, P., & Ansseau, M. (2003). Socioeconomic Inequalities in Depression: A Meta-Analysis. *American Journal of Epidemiology*, *157*(2), 98–112. https://doi.org/10.1093/aje/kwf182

Mac-Ginty, S., Lira, D., Lillo, I., Moraga, E., Cáceres, C., Araya, R., & Prina, M. (2024). Association Between Socioeconomic Position and Depression, Anxiety and Eating Disorders in University Students: A Systematic Review. *Adolescent Research Review*. https://doi.org/10.1007/s40894-023-00230-y

Masha'al, D., Hayajneh, A. A., & Tawalbeh, L. I. (2021). Psychometric evaluation of the Arabic version of the higher education inventory scale for nursing students. *BMC Psychiatry*, *21*(1), 77. https://doi.org/10.1186/s12888-021-03082-9

Moltrecht, B., McElroy, E., & Scopel Hoffmann, M. (2023). *Software Profile: Harmony: a web-tool for retrospective, multilingual harmonisation of questionnaire items using natural language processing*. https://doi.org/10.31219/osf.io/9rp3s

Mundfrom, D. J., Shaw, D. G., & Ke, T. L. (2005). Minimum Sample Size Recommendations for Conducting Factor Analyses. *International Journal of Testing*, *5*(2), 159–168. https://doi.org/10.1207/s15327574ijt0502_4

O'Rourke, M., Hammond, S., O'Flynn, S., & Boylan, G. (2010). The Medical Student Stress Profile: A tool for stress audit in medical training: Stress audit of medical students. *Medical Education*, *44*(10), 1027–1037. https://doi.org/10.1111/j.1365-2923.2010.03734.x

Pacheco, J. P. G., Hoffmann, M. S., Braun, L. E., Medeiros, I. P., Casarotto, D., Hauck, S., Porru, F., Herlo, M., & Calegaro, V. C. (2023). Translation, cultural adaptation, and validation of the Brazilian Portuguese version of the Higher Education Stress Inventory (HESI-Br). *Trends in Psychiatry and Psychotherapy*. https://doi.org/10.47626/2237-6089-2021-0445

Pacheco, J. P. G., Silveira, J. B., Ferreira, R. P. C., Lo, K., Schineider, J. R., Giacomin, H. T. A., & Tam, W. W. S. (2019). Gender inequality and depression among medical students: A global meta-regression analysis. *Journal of Psychiatric Research*, *111*, 36–43. https://doi.org/10.1016/j.jpsychires.2019.01.013

Porru, F., Schuring, M., Bültmann, U., Portoghese, I., Burdorf, A., & Robroek, S. J. W. (2022). Associations of university student life challenges with mental health and self-rated health: A longitudinal study with 6 months follow-up. *Journal of Affective Disorders*, *296*, 250–257. https://doi.org/10.1016/j.jad.2021.09.057

Puthran, R., Zhang, M. W. B., Tam, W. W., & Ho, R. C. (2016). Prevalence of depression amongst medical students: A meta-analysis. *Medical Education*, *50*(4), 456–468. https://doi.org/10.1111/medu.12962

Putnick, D. L., & Bornstein, M. H. (2016). Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research. *Developmental Review*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement*, *1*(3), 385–401. https://doi.org/10.1177/014662167700100306

Raudenská, P. (2023). Single-item measures of happiness and life satisfaction: The issue of cross-country invariance of popular general well-being measures. *Humanities and Social Sciences Communications*, *10*(1), 861. https://doi.org/10.1057/s41599-023-02299-1

Rotenstein, L. S., Ramos, M. A., Torre, M., Segal, J. B., Peluso, M. J., Guille, C., Sen, S., & Mata, D. A. (2016). Prevalence of Depression, Depressive Symptoms, and Suicidal Ideation Among Medical Students: A Systematic Review and Meta-Analysis. *JAMA*, *316*(21), 2214. https://doi.org/10.1001/jama.2016.17324

Schlax, J., Jünger, C., Beutel, M. E., Münzel, T., Pfeiffer, N., Wild, P., Blettner, M., Kerahrodi, J. G., Wiltink, J., & Michal, M. (2019). Income and education predict elevated depressive symptoms in the general population: Results from the Gutenberg health study. *BMC Public Health*, *19*(1), 430. https://doi.org/10.1186/s12889-019-6730-4

Shi, M., Liu, L., Yang, Y.-L., & Wang, L. (2015). The mediating role of self-esteem in the relationship between big five personality traits and depressive symptoms among Chinese undergraduate medical students. *Personality and Individual Differences*, *83*, 55–59. https://doi.org/10.1016/j.paid.2015.03.050

Shim, E.-J., Jeon, H. J., Kim, H., Lee, K.-M., Jung, D., Noh, H.-L., Roh, M.-S., & Hahm, B.-J. (2016). Measuring stress in medical education: Validation of the Korean version of the higher education stress inventory with medical students. *BMC Medical Education*, *16*(1), 302. https://doi.org/10.1186/s12909-016-0824-9

Tam, W., Lo, K., & Pacheco, J. (2019). Prevalence of depressive symptoms among medical students: Overview of systematic reviews. *Medical Education*, *53*(4), 345–354. https://doi.org/10.1111/medu.13770

Tempski, P., Bellodi, P. L., Paro, H. B., Enns, S. C., Martins, M. A., & Schraiber, L. B. (2012). What do medical students think about their quality of life? A qualitative study. *BMC Medical Education*, *12*(1), 106. https://doi.org/10.1186/1472-6920-12-106

Vandenberg, R. J., & Lance, C. E. (2000). A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organizational Research Methods*, *3*(1), 4–70. https://doi.org/10.1177/109442810031002

Vilagut, G., Forero, C. G., Barbaglia, G., & Alonso, J. (2016). Screening for Depression in the General Population with the Center for Epidemiologic Studies Depression (CES-D): A Systematic Review with Meta-Analysis. *PLOS ONE*, *11*(5), e0155431. https://doi.org/10.1371/journal.pone.0155431

Von Glischinski, M., Von Brachel, R., & Hirschfeld, G. (2019). How depressed is "depressed"? A systematic review and diagnostic meta-analysis of optimal cut points for the Beck Depression Inventory revised (BDI-II). *Quality of Life Research*, *28*(5), 1111–1118. https://doi.org/10.1007/s11136-018-2050-x

Wong, V. W.-H., Yiu, E. K.-L., Ng, C. H., Sarris, J., & Ho, F. Y.-Y. (2024). Unraveling the associations between unhealthy lifestyle behaviors and mental health in the general adult Chinese population: A cross-sectional study. *Journal of Affective Disorders*, *349*, 583–595. https://doi.org/10.1016/j.jad.2023.12.079

Yap, S., Luo, Q., Wade, S., Weber, M., Banks, E., Canfell, K., O'Connell, D. L., & Steinberg, J. (2022). Raking of data from a large Australian cohort study improves generalisability of estimates of prevalence of health and behaviour characteristics and cancer incidence. *BMC Medical Research Methodology*, *22*(1), 140. https://doi.org/10.1186/s12874-022-01626-5

Yusoff, M. S. B. (2017). A Systematic Review on Validity Evidence of Medical Student Stressor Questionnaire. *Education in Medicine Journal*, *9*(1), 1–16. https://doi.org/10.21315/eimj2017.9.1.1

Yusoff, M. S. B., Rahim, A. F. A., & Yaacob, M. J. (2010). The Development and Validity of the Medical Student Stressor Questionnaire (MSSQ). *ASEAN Journal of Psychiatry*, *11*(1).

Appendix

Appendix 1. MacArthur Scale of Subjective Social Status

Instructions: Think of this ladder as representing where people stand in our society. At the top of the ladder are the people who are the best off, those who have the most money, most education, and best jobs. At the bottom are the people who are the worst off, those who have least money, least education, and worst jobs or no job. Mark a "X" on the rung that best represents where you think you stand on the ladder.

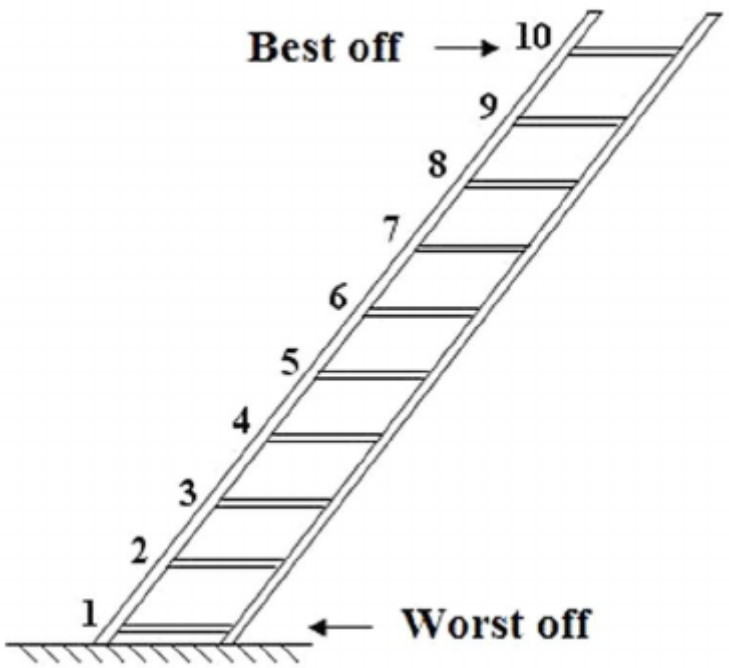


Table 1. Patient Health Questionnaire-9 (PHQ-9)

Instructions: Over the last 2 weeks, how often have you been bothered by the following problems?

	Not at all	Several days	More than half the days	Nearly every day
Little interest or pleasure in doing things				
Feeling down, depressed or hopeless				
Trouble falling or staying asleep, or sleeping too much				
Feeling tired or having little energy				
Poor appetite or overeating				
Feeling bad about yourself - or that you are a failure or have let yourself or your family down				
Trouble concentrating on things, such as reading the newspaper or watching television				
Moving or speaking so slowly that other people could have noticed? Or the opposite - being so fidgety or restless that you have been moving around a lot more than usual				
Thoughts that you would be better off dead or of hurting yourself in some way				

Table 2. Beck Depression Inventory (BDI)

Instructions: Please read each group of statements carefully. And then pick out the one statement in each group that best describes the way you have been feeling during the past two weeks, including today. Check the statement you have picked.

1.
( ) I do not feel sad.
( ) I feel sad
( ) I am sad all the time and I can't snap out of it.
( ) I am so sad and unhappy that I can't stand it.

2.
( ) I am not particularly discouraged about the future.
( ) I feel discouraged about the future.
( ) I feel I have nothing to look forward to.
( ) I feel the future is hopeless and that things cannot improve.

3.
( ) I do not feel like a failure.
( ) I feel I have failed more than the average person.
( ) As I look back on my life, all I can see is a lot of failures.
( ) I feel I am a complete failure as a person.

4.
( ) I get as much satisfaction out of things as I used to.
( ) I don't enjoy things the way I used to.
( ) I don't get real satisfaction out of anything anymore.
( ) I am dissatisfied or bored with everything.

5.
( ) I don't feel particularly guilty
( ) I feel guilty a good part of the time.
( ) I feel quite guilty most of the time.
( ) I feel guilty all of the time.

6.
( ) I don't feel I am being punished.
( ) I feel I may be punished.
( ) I expect to be punished.
( ) I feel I am being punished.

7.
( ) I don't feel disappointed in myself.
( ) I am disappointed in myself.
( ) I am disgusted with myself.
( ) I hate myself.

8.

( ) I don't feel I am any worse than anybody else.
( ) I am critical of myself for my weaknesses or mistakes.
( ) I blame myself all the time for my faults.
( ) I blame myself for everything bad that happens.

9.
( ) I don't have any thoughts of killing myself.
( ) I have thoughts of killing myself, but I would not carry them out.
( ) I would like to kill myself.
( ) I would kill myself if I had the chance

10.
( ) I don't cry any more than usual.
( ) I cry more now than I used to.
( ) I cry all the time now.
( ) I used to be able to cry, but now I can't cry even though I want to.

11.
( ) I am no more irritated by things than I ever was.
( ) I am slightly more irritated now than usual.
( ) I am quite annoyed or irritated a good deal of the time.
( ) I feel irritated all the time.

12.
( ) I have not lost interest in other people.
( ) I am less interested in other people than I used to be.
( ) I have lost most of my interest in other people.
( ) I have lost all of my interest in other people.

13.
( ) I make decisions about as well as I ever could.
( ) I put off making decisions more than I used to.
( ) I have greater difficulty in making decisions more than I used to.
( ) I can't make decisions at all anymore.

14.
( ) I don't feel that I look any worse than I used to.
( ) I am worried that I am looking old or unattractive.
( ) I feel there are permanent changes in my appearance that make me look unattractive
( ) I believe that I look ugly.

15.
( ) I can work about as well as before.
( ) It takes an extra effort to get started at doing something.
( ) I have to push myself very hard to do anything.
( ) I can't do any work at all.

16.
( ) I can sleep as well as usual.
( ) I don't sleep as well as I used to.
( ) I wake up 1-2 hours earlier than usual and find it hard to get back to sleep.

( ) I wake up several hours earlier than I used to and cannot get back to sleep.

17.
( ) I don't get more tired than usual.
( ) I get tired more easily than I used to.
( ) I get tired from doing almost anything.
( ) I am too tired to do anything.

18.
( ) My appetite is no worse than usual.
( ) My appetite is not as good as it used to be.
( ) My appetite is much worse now.
( ) I have no appetite at all anymore.

19.
( ) I haven't lost much weight, if any, lately.
( ) I have lost more than five pounds.
( ) I have lost more than ten pounds.
( ) I have lost more than fifteen pounds.

20.
( ) I am no more worried about my health than usual.
( ) I am worried about physical problems like aches, pains, upset stomach, or constipation.
( ) I am very worried about physical problems and it's hard to think of much else.
( ) I am so worried about my physical problems that I cannot think of anything else.

21.
( ) I have not noticed any recent change in my interest in sex.
( ) I am less interested in sex than I used to be.
( ) I have almost no interest in sex.
( ) I have lost interest in sex completely.

Table 3. Center for Epidemiological Studies Depression Scale (CES-D)

Instructions: Below is a list of some ways you may have felt or behaved. Please indicate how often you have felt this way during the last week by checking the appropriate space.

	Rarely or none of the time	Some or a little of the time	Occasionally or a moderate amount of the time	Most or all of the time
I was bothered by things that usually don't bother me				
I did not feel like eating; my appetite was poor				
I felt that I could not shake off the blues even with help from my family and friends				
I felt that I was just as good as other people				
I had trouble keeping my mind on what I was doing				
I felt depressed				
I felt that everything I did was an effort				
I felt hopeful about the future				
I thought my life had been a failure				
I felt fearful				
My sleep was restless				
I was happy				
I talked less than usual				
I felt lonely				
People were unfriendly				
I enjoyed life				
I had crying spells				
I felt sad				
I felt that people disliked me				
I could not get "going"				

Table 4. Higher Educational Stress Inventory (HESI)

Instructions: Below you will find several statements about your relationship and personal experience with the academic environment. Please, for each of them, indicate your level of agreement.

	Totally disagree	Somewhat disagree	Somewhat agree	Totally agree
Studies control my life and I have little time for other activities				
I feel that my teachers treat me with respect				
I am worried that I will not acquire all the knowledge needed for my future profession				
The studies have created anonymity and isolation among students				
The teachers often fail to clarify the aims of the studies				
The studies stimulate my personal development				
The professional role presented in the training conflicts with my personal views				
The teachers give encouragement and personal attention				
There is a competitive attitude among students				
I am satisfied with my choice of career				
I feel that the studies have played a role in creating a cold and impersonal attitude among students				
As a student, my financial situation is a worry				
My fellow students support me				
I worry about long working hours and responsibilities in my future career				
The training is characterized by an atmosphere where weakness and personal shortcomings are not accepted				
As a student you are often expected to participate in situations where your role and function is unclear				
I am proud of my future profession				
I feel that I am less well treated because of my sex				
I am able to influence the studies				
The insight I have had into my future profession has made me worried about the stressful workload				

There is too much focus on passive learning of fact and too little on active seeking of knowledge and time for reflection				
Expectations from my family have influenced my choice of career too much				
I am worried about accommodation				
I feel that I am less well treated because of my ethnic background				
I meet many future colleagues that seem dejected or dissatisfied in their profession				
I feel that the training is preparing me well for my future profession				
Student union activities promote a sense of community and contribute to a better working environment for students				
I am worried about my future economy and my ability to repay student loans				
The education is highly characterized by group activities with unclear goals and with too much responsibility placed on the student				
The literature is too difficult and extensive				
The pace of studies is too high				
The training demands that I join in situations that I find unethical				
The teachers often give feedback on the students' knowledge and skills				

Table 5. University - Short Multidimensional Inventory on Lifestyle Evaluation (U-SMILE)

Instructions: In the last month, how often in your daily routine…

	Always	Often	Seldom	Never
When shopping for food, do you check labels for ingredients such as the quantity of salt?				
Do you eat processed food (frozen food such as pizza, French fries, puff pastries, deep-fried foods and canned foods)?				
Do you eat healthy foods such as fresh fruits, fresh vegetables, wholegrain, legumes or nuts?				
Do you drink 5 or more doses (men) or 4 or more doses (women) of alcoholic beverages on a single occasion, which means within 2 hours?				
Do you smoke tobacco (cigarette, electronic cigarette, cigar, pipe, smokeless tobacco)?				
Do you use marijuana or hashish?				
Do you use other drugs (cocaine, crack, amphetamines, ecstasy, opioids without medical prescription, and others)?				
Do you exercise for at least 30 minutes daily (or 150 minutes a week)?				
Do you play at least 2 hours of team sports (like soccer, volleyball, basketball, rugby, etc.) a week?				
Do you feel good after performing physical activity?				
Do you use any strategy or psychological support to deal with stress (for instance meditation, mindfulness or psychotherapy)?				
Do you practice a faith or religion?				
Do you manage to sleep between 7 and 9 hours per night?				
Do you feel rested with the number of hours you sleep?				
Do you maintain a regular sleep schedule?				
Do you interact with your friends and/or relatives?				
Do you feel that you are part of a group of friends, the community or the society?				
Do you have someone you trust who listens to your problems or concerns?				

Do you have someone in your life to go out or have fun with when you feel like it?				
Do you make yourself available to support your significant ones?				
Do you spend more than 2 hours a day watching TV, playing computer games, video games or on the internet?				
Do you spend time on a computer / smartphone within one hour of going to sleep?				
Are you in touch with nature (for instance parks, beach, countryside, mountains)?				
Do you feel your relationship to nature, that is all living things, is an important part of who you are?				

Table 6. Questions from MSSP and MSSQ (medical education questions highlighted in gray)

MSSP		MSSQ	
Item number	**Item content**	**Item number**	**Item content**
1	Not getting on with peers	1	Tests/Examinations
2	Unsupportive supervisors	2	Talking to patients about personal problems
3	Unsupportive department	3	Conflicts with other students
4	Unsupportive lecturers	4	Merit ranking system in finals
5	Loneliness	5	Verbal or physical abuse by other student(s)
6	Fear of failure	6	Parental wish for you to study medicine
7	Strain on relationship	7	Need to do well (self-expectation)
8	No time for extracurricular activities	8	Not enough study material
9	Examinations	9	Conflict with faculty staff
10	Peer competition	10	Heavy workload
11	Financial pressure	11	Participation in class discussion
12	Rote learning	12	Falling behind in reading schedule
13	No time for exercise	13	Participation in class presentation
14	No time for rest	14	Lack of guidance from teacher(s)
15	Dependent on financial support	15	Feeling of incompetence
16	Not enough time	16	Uncertainty of what is expected of me
17	Boredom	17	Not enough medical skill practice
18	Inadequate student facilities	18	Lack of time for family and friends
19	Course interferes with home life	19	Highly competitive learning context
20	Peers do not pull weight	20	Teacher- lack of teaching skills
21	Hard to talk to supervisor	21	Unable to answer questions from patients
22	Inadequate patient facilities	22	Inappropriate assignments
23	Partner resents time spent	23	Having difficulty understanding the context
24	Not getting enough credit	24	Facing illness or death of the patients
25	Too much responsibility	25	Getting poor marks
26	No say in course organization	26	Poor motivation to learn

27	Unsupportive people outside the course	27	Lack of time to review what has been learnt
28	Poor lecturers	28	Verbal or physical abuse by teacher(s)
29	Unsocial hours	29	Frequent interruption of my work by others
30	Unsupportive peers	30	Inability to answer the question from the teachers
31	Lecturers do not listen	31	Conflict with teacher(s)
32	Discrimination	32	Unwillingness to study medicine
33	Less able students getting credit	33	Large amount of content to be learnt
34	Complaining among peers	34	Need to do well (imposed by others)
35	Workload	35	Not enough feedback from teacher(s)
36	Work / family conflict	36	Unjustified grading/ranking process
37	Isolated on course	37	Lack of recognition for work done
38	Inadequate guidance	38	Working with computers (teaching and learning)
39	Unpleasant patients	39	Working with computers (Staff and Students)
40	Insufficient support staff	40	Family responsibility
41	Course versus social life		
42	Going nowhere		
43	Placed in physical danger		
44	Course unchangeable		
45	Consequences of decisions		
46	Death of patients		
47	Terminally ill patients		
48	Chronic patients		
49	Fear of infection		
50	Physical examinations		
51	Discussing personally sensitive matters		
52	Discussing sexual matters		

Table 7. First version of Medical Student Stress Inventory.

Item number	Tool	Item content
2	MSSQ	Talking to patients about personal problems
17	MSSQ	Not enough medical skill practice
21	MSSQ	Unable to answer questions from patients
24	MSSQ	Facing illness or death of the patients
22	MSSP	Inadequate patient facilities
39	MSSP	Unpleasant patients
46	MSSP	Death of patients
47	MSSP	Terminally ill patients
48	MSSP	Chronic patients
49	MSSP	Fear of infection
50	MSSP	Physical examinations
51	MSSP	Discussing personally sensitive matters
52	MSSP	Discussing sexual matters

Note: numbers represent the respective item number from the original tools; excluded questions in the final version are highlighted in gray.

Table 8. Final version of Medical Student Stressor Inventory (MSSI)

Instructions: Please specify the level of stress caused by each aspect of your daily routine as a medical student over the last month.

	Not at all	A little	Somewhat	A lot	Extremely	Not part of my routine
Talking to patients about their personal problems						
Not enough practical experience or medical skills						
Unable to answer questions from patients						
Facing patient's illness or death						
Inadequate patient facilities						
Unpleasant patients						
Chronic patients						
Fear of infection						
Physical examinations						
Discussing sexual matters						