

## Supplemental Material 1 – Methods details

### Theoretical rationale for choosing variables of interest to assess their association with the quality of CPGs and their recommendations

- 1) **Handling of conflicts of interest:** In order to have impartial and reliable recommendations, a CPG must not only declare conflicts but also have a handling of conflicts of interest policy. This policy is commonly achieved by editorial independence. Although we have not used in our analyses a variable directly related to the AGREE II domain “Editorial independence”, we have assessed “Handling of conflicts of interest”, which could be considered a proxy of this domain and is a fundamental characteristic for unbiased recommendations.[1]
- 2) **Multiprofessional team:** The development of high-quality CPGs is frequently associated with multidisciplinary teams in order to incorporate different knowledge and apply them in the recommendations. [1,2]
- 3) **Inclusion of patient representative in the team:** This characteristic usually occurs in clinical practice guidelines (CPGs) of higher methodological quality. In addition, interested party’s participation (*i.e.*, patients) is fundamental to the development of acceptable recommendations (Guidelines International Network, 2020). [3]
- 4) **Governmental funding:** Government-funded teams are more likely to have a larger number of professionals involved and more resources allocated allowing a longer and thoroughly review. Developing high quality CPGs usually requires these resources. [1,4]
- 5) **Type of institution:** The type of institution was considered in the model since this variable has been mentioned by previous studies to be associated with higher scores on the AGREE II “rigour of development”. [4]
- 6) **Publication year after 2015:** We hypothesized that newer CPGs and updated or revised versions tend to be associated with higher "rigour of development and highest quality" scores of the recommendations. [4]

### Identifying and selecting CPGs

For the selection of published CPGs, we searched for documents published from January 1, 2011 to December 31, 2021. A period longer than 10 years was considered because some CPGs are not systematically updated. In the occurrence of two or more versions of the same guideline we included only the most recent updated version. The following databases were searched: MEDLINE (via PubMed), Cochrane Library, Embase, PsycINFO, and BVS. Additional searches were performed in the following websites: Australian Clinical Practice Guidelines, National Committee for Health Technology Incorporation (Brazilian Government), Canadian Agency for Drugs and Technologies in Health, Canadian Medical Association, Chilean Ministry of Health, Colombian Ministry of Health and Social Protection, Guidelines International Network, Institute for Clinical Systems Improvement, Portal Guía Salud, Scottish Intercollegiate Guidelines Network, and the National Institute for Health and Care Excellence, Guideline Central, ECRI library. Specific details related to the databases used in this study and the terms used in these repositories are shown below:

#### Medline (via PubMed)

(((((("Guideline" [Publication Type] OR "CPGs as Topic"[Mesh] OR "Practice Guideline" [Publication Type] OR "Health Planning CPGs"[Mesh]) OR "Clinical Protocols"[Mesh])) OR ("Consensus Development Conference, NIH" [Publication Type] OR "Consensus Development

Conference" [Publication Type] OR "Consensus"[Mesh])) OR "Standard of Care"[Mesh])  
 "Guideline" [Publication Type] OR "CPGs as Topic"[Mesh] OR "Practice Guideline"  
 [Publication Type] OR "Health Planning CPGs"[Mesh]) OR "Clinical Protocols"[Mesh])) OR  
 ("Consensus Development Conference, NIH" [Publication Type] OR "Consensus Development  
 Conference" [Publication Type] OR "Consensus"[Mesh])) OR "Standard of Care"[Mesh]))))  
 AND (("Depressive Disorder"[Mesh] OR "Depressive Disorder, Major"[Mesh] OR Depressive  
 Disorders OR Disorder, Depressive OR Disorders, Depressive OR Neurosis, Depressive OR  
 Depressive Neuroses OR Depressive Neurosis OR Neuroses, Depressive OR Depression,  
 Endogenous OR Depressions, Endogenous OR Endogenous Depression OR Endogenous  
 Depressions OR Depressive Syndrome OR Depressive Syndromes OR Syndrome, Depressive  
 OR Syndromes, Depressive OR Depression, Neurotic OR Depressions, Neurotic OR Neurotic  
 Depression OR Neurotic Depressions OR Melancholia OR Melancholias OR Unipolar  
 Depression OR Depression, Unipolar OR Depressions, Unipolar OR Unipolar Depressions))  
 Embase:  
 #1 'practice guideline'/mj OR 'consensus development'/exp/mj OR 'clinical protocol'/mj #2  
 'depression'/exp  
 #3 #1 AND #2

### Cochrane Library

# 1—MeSH descriptor: [Guideline] explode all trees  
 # 1—MeSH descriptor: [Consensus] explode all trees  
 # 1—MeSH descriptor: [Clinical Protocols] explode all trees  
 # 1—#1 OR #2 OR #3  
 # 1—MeSH descriptor: [Depression] explode all trees

### PsychINFO

((Any Field: (depression))) AND ((Any Field: (guideline)) OR (Any Field: (consensus)) OR  
 (Any Field: ("clinical protocol"))) AND Year: 2011 To 9999  
 ((Any Field: (depression))) AND ((Any Field: (guideline)) OR (Any Field: (consensus)) OR (Any  
 Field: ("clinical protocol"))) AND Year: 2011 To 9999

### BVS

((guideline) OR (consensus) OR (clinical protocol)) AND (depression) AND (db:("LILACS" OR  
 "IBECIS" OR "WHOLIS" OR "BDENF" OR "BINACIS" OR "INDEXPSI" OR "BIGG" OR  
 "BBO" OR "CUMED")) AND (year\_cluster:[2011 TO 2021])

We included documents between January 1, 2011, to December 31, 2021; containing recommendations for the pharmacological treatment of depression in adults at outpatient care setting, regardless of whether it met the U.S. National Academy of Medicine. If a CPG included recommendations for both children and adults, they were considered. No language restriction was applied; CPGs that were in languages other than Portuguese, English, or Spanish were translated into Portuguese by a professional translation service. We excluded CPGs for specific populations (e.g., treatment of pregnant women); local CPGs developed by hospitals or originations intended to be applied only at a local level and treatment of depression in comorbidity with specific diseases (e.g., depression in patients with diabetes).

Retrieved references were exported to the online platform Rayyan® reference manager. [5] After removal of duplicates, references were screened by two independent researchers. Then searched for the full texts, and these were reviewed in duplicate. Discrepancies between researchers were resolved by consensus. When no consensus was reached, a third evaluator was involved.

### **Extracting the characteristics of the CPGs**

A Google Form was used for the extraction of general CPG data. The process of data extraction by two independent researchers was validated in a previous project [4] focused on osteoporosis CPGs conducted by our team. [6] We extracted the following independent variables: year of publication of the most recent version of the CPGs (2011 to 2014/2015 to 2021); type of institution (Governmental or University; Professional society), inclusion of a patient representative in the team (yes/no), multiprofessional team (yes - different professions/no - only one profession), governmental funding (yes/no), and policy for handling of conflicts of interest (yes/no). The type of institution classification and the funding variables were defined as governmental, even if the governmental institution had partnership with other types of institutions. These factors were selected because they are included in the AGREE II [7], AGREE-REX [8] and IOM. [9] Additionally, they are commonly found in literature articles [4,10] related to guideline's quality.

### **Appraisal of the CPGs quality**

AGREE II is a reliable and validated tool composed by 23 items, clustered in six domains. The appraisal team comprised multidisciplinary researchers, including pharmacists, nurses, and public health professionals previously trained according to protocol. [11] The training consisted of reading the AGREE II manual. [7] Subsequently, the evaluators appraised the CPG quality on chronic pain [12], Gaucher disease [13], and for the treatment of obesity. [14] A discussion was held with previously trained evaluators on the evaluations made. In the next stage of training, the team appraised the most recent two CPGs for hyperthyroidism [15] and urinary tract infection. [16]

The appraisers assigned a score from 1 to 7 for each AGREE II item (following a 7-point Likert scale). Each guideline was appraised by three appraisers as suggested by the AGREE Next Steps Consortium [7] using the AGREE-PLUS Platform. [17] A difference of two or more points in the individual items' [4,11,18] scores was considered discrepant and was resolved by consensus between the appraisers to obtain the final score. A final consolidated score was obtained from 0-100% per domain as suggested by the AGREE II Manual.

### **Appraisal of the quality of the recommendations of the CPGs**

We used the AGREE-REX instrument to appraise the quality of the CPG recommendations. All recommendations were grouped and analysed for each CPG, according to one of the options recommended in the AGREE-REX manual. [8] AGREE-REX consists of nine items clustered into three domains: clinical applicability, values and preferences, and implementability. The same team of appraisers that assessed the quality of the CPGs using AGREE II also assessed the quality of the recommendations. The team was also trained for this evaluation. The three appraisers assigned a score from 1 to 7 for each item (following a 7-point Likert scale). When there was a discrepancy of two points or more, the evaluators discussed it until they reached consensus. A final consolidated score was obtained from 0-100% per AGREE-REX domain as suggested by the AGREE-REX Manual. [8] The scores of each evaluator were inserted in a Google Form®.

### **Statistical data analysis**

Quantitative data were described using mean, standard deviation, median, and interquartile range. Categorical variables were presented as frequencies and percentages.

The CPGs were considered high-quality if they scored  $\geq 60\%$  in the AGREE II Rigour of Development domain, as it has been previously recommended. [4,10,19-22] The CPGs' recommendations were considered high-quality if they scored  $\geq 60\%$  in the AGREE-REX Clinical Applicability domain which considers the appropriateness of recommendations for clinical practice, patient needs, and the intended impact of guideline implementation.

Since the outcomes of high-quality CPGs and high-quality recommendations showed only 17 and 7 events, respectively, instead of analysing the binary characteristics we decided to model the scores directly. We used linear regression analysis and with this approach we were able to estimate the impact that each studied factor had on the AGREE II domain 3 and AGREE-REX domain 1 scores. The regression coefficients were obtained with a simple unadjusted model (univariate) and with an adjusted model (multivariable). Results with p values below 0.05 were deemed statistically significant. Data were processed and analyzed with IBM-SPSS version 25.0.