# BMJ Open

# Translation and adaptation of the STOPP-START screening tool to Portuguese for detecting inappropriate prescriptions in older people: a protocol

Luís Monteiro [iD],[1,2] Matilde Monteiro-Soares,[3,4] Liliane Vélia Mendonça,[4] Inês Ribeiro-Vaz,[4,5] Carlos Martins [iD],[3,4] Andreia Teixeira[1,6,7]

For numbered affiliations see end of article.

**Correspondence to**
Dr Luís Monteiro;
luismonteiro.net@gmail.com

## ABSTRACT

**Introduction** Rational prescribing for older adults is a challenge because they usually exhibit multimorbidity and multimedication. One available and reliable tool to tackle this issue consists of the Screening Tool of Older People's Prescriptions (STOPP) and the Screening Tool to Alert to Right Treatment (START), which has been associated with improvements in clinical outcomes. Our goal here is to translate and validate the STOPP-START screening tool for use with Portuguese general practitioners/family physicians.

**Methods and analysis** The study will be conducted in four phases: phase I—translation of the STOPP-START screening tool to Portuguese; phase II—data collection of patient data; phase III—intrarater reliability and agreement study; and phase IV—inter-rater reliability and agreement study.

**Ethics and dissemination** This study was approved by the Ethics Committee of the Central Health Region of Portugal (where the study will take place). Every participant will sign a written consent form. We intend to publish the full article in a related peer-reviewed journal, conference presentations, reports and in a PhD thesis.

### Strengths and limitations of this study

► This study will develop the first Portuguese version of the Screening Tool of Older People's Prescriptions and the Screening Tool to Alert to Right Treatment criteria.
► This is the first study in a Portuguese primary care setting that aims to develop a useful tool for the appropriate prescription of older patients.
► The main limitation of the study is that it is focused on Portugal and it may not apply to other countries where Portuguese is not the main language.

## INTRODUCTION

In Organisation for Economic Co-operation and Development countries, the number of older adults is increasing[1] as well as their life expectancy.[2][3]

Caring for older adults is a challenge for healthcare systems[4] because older adults are more likely to have more than one chronic disease.[5][6] For example, multimorbidity in the elderly can be higher than 90% in Portugal.[5] Therefore, adults aged ≥65 years are more likely to be prescribed with multiple drugs[7–9] and may be more susceptible to inappropriate medication use.[10–12]

Potentially inappropriate medications (PIMs) can be described as the use of medications that potentially have more risks than benefits even though safer pharmacological and non-pharmacological alternatives

are available.[10] Potentially inappropriate prescription is a different concept than PIM, and includes the overprescription, underprescription and misprescription of medications (eg, inappropriate dose or duration).[13]

There are various tools to help physicians identify PIM such as the Beers Criteria[14] and the Potentially Inappropriate Medications in the Elderly list.[15] The combination of the Screening Tool of Older People's Prescriptions (STOPP) and the Screening Tool to Alert to Right Treatment (START)[16][17] is another widely used tool. One of the advantages of this tool is that it not only considers PIM, but also the indications to start an appropriate medication (START).

Versus other tools, some studies have shown that the STOPP-START tool can identify a significantly higher proportion of patients requiring hospitalisation as a result of PIM-related adverse events,[16] can reduce the highest number of medications and can identify more potential major clinical issues.[18] The criteria for STOPP-START have been associated with improvement in prescribing quality and clinical outcomes.[19] These criteria have been adapted for other languages, such as French.[20] In this adaptation, 50 data sets of patients hospitalised in an academic geriatric

department were analysed independently by one geriatrician and one general practitioner. They considered 87 STOPP-START criteria of the original version. The data sets involved 418 prescribed medications. The proportions of positive and negative inter-rater agreements were 99% and 95%, respectively, for STOPP and 99% and 88% for START; Cohen's κ coefficients were 0.95 for STOPP and 0.92 for START. The results indicated an excellent inter-rater agreement.

Inter-rater reliability of STOPP and START criteria was also tested between multiple physicians practising independently in Europe.[21] After translation of the criteria into their local language, doctors in Belgium, Czech Republic, Italy, Spain and Switzerland applied the criteria to 20 data sets selected from 200 patients aged ≥65 years of a university teaching hospital in Ireland. The median κ coefficients between raters were 0.93 (0.90 to 0.96) for STOPP criteria and 0.85 (0.82 to 0.91) for START criteria. The results demonstrated good inter-rater reliability of STOPP-START criteria. Therefore, the authors concluded that STOPP and START criteria are generalisable across different European countries and languages.[21]

Reliability and agreement are different concepts but have been used without distinction in many studies.[22] Reliability can be defined as the ratio of variability between scores of the same subjects (by different raters or at different moments) to the total variability of all scores in the sample. Agreement is connected to the question about whether observations are similar or the degree to which they differ.

We aim to make the first translation and validation[23] of the English STOPP-START tool for Portuguese family doctors. In the validation study, we deal with two aspects of reliability and agreement concepts: inter-rater reliability and agreement (different raters using the translated STOPP-START tool assess the same patients), and intra-rater reliability and agreement (the same rater using the translated STOPP-START tool assesses the same subjects at two different moments).

## METHODS AND ANALYSIS
### Study design
This study will be conducted in four phases as illustrated in figure 1 (timeline available in online supplemental appendix I). The first phase (phase I) is the translation to the Portuguese language followed by data collection (phase II).

Phase III consists of an intrarater reliability and agreement study, and phase IV is an inter-rater reliability and agreement study. We made a preregistration on 'Open registries Network' (DOI 10.17605/OSF.IO/SK2RJ).

### Phase I: translation to Portuguese
The translation of the STOPP-START screening tool will follow the Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes Measures.[20] We have already obtained

permission from STOPP-START's authors to translate and validate the tool for Portuguese. We will recruit a key in-country consultant who is a native Portuguese and fluent English speaker and will be the main contact to perform and help with the translation. This consultant will also have a background in health research and experience in translating English documents. We will obtain two independent translations of the STOPP-START. One will be done by the key in-country consultant, and the other will be performed by a forward translator who is also a native Portuguese and fluent English speaker.

The two translations will be reconciled by the research team to obtain a final consensus translation that will be back-translated.

The back-translation (from Portuguese to English) will be done by a professional translator who is a native speaker of English and fluent speaker of Portuguese. This translator will have no prior knowledge of the original English version. Afterwards, the back-translation will be compared with the original to identify any relevant differences.

In the final step, the reconciled Portuguese STOPP-START version will be distributed to a group of 15 general practitioners to verify if there are any interpretation issues. The research team will analyse the results from the application of the STOPP-START tool to prepare the final version.

### Phase II: data collection
#### Design
This will be a cross-sectional, analytical study.

#### Setting
The study will be conducted in a primary care centre in the Centre Region of Portugal.

The health unit is located in Aveiro. Five family doctors follow a total of 8165 patients; 1625 patients aged ≥65 years.

#### Sample size
To calculate the sample size for the validation study, we used the function CIBinary of the kappaSize package of R software.[24] For the intrarater study, we obtained a sample size of 334 subjects considering the following parameters: estimated κ value: 0.68[25]; error margin: 0.1; prevalence of each item of the START criteria: 0.25; number of moments: 2; and significance level: 5%. In the inter-rater study, we obtained a sample size of 205 subjects considering the following parameters: estimated κ value: 0.68[25]; error margin: 0.1; prevalence of each item of the START criteria: 0.25; number of raters: 3; and significance level: 5%. The 205 patients for inter-rater assessment will be randomly selected from the 334 subjects used for the intrarater evaluation.

### Study procedures
#### Recruitment of patients
Patients will be randomly selected (independent random sampling using computer-generated random digits) from
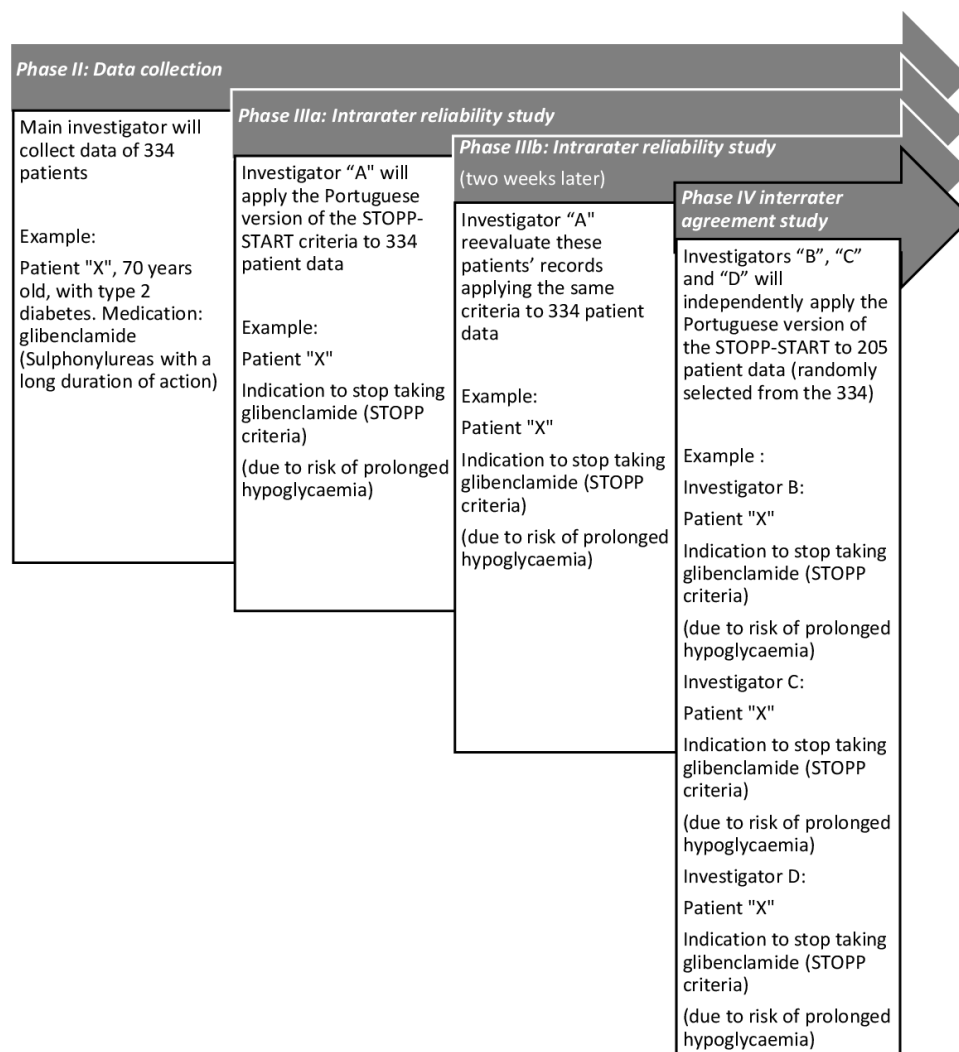
**Figure 1** Flow chart and example. START, Screening Tool to Alert to Right Treatment; STOPP, Screening Tool of Older People's Prescriptions.

a list of patients aged ≥65 years from a primary care centre. They will be invited by telephone to participate in the study. The investigator or a previously trained research associate will then interview the patients in the general practitioner office. Recruitment will continue until 334 patients are enrolled.

Exclusion criteria include incapacity or unwillingness to provide written informed consent, diagnostic of psychotic disorder, institutionalisation and the presence of terminal illness.

At inclusion, the main investigator will collect sociodemographic patient data such as age, gender, educational level, labour status and marital status. Clinical data collection will include identification of total number of medications for chronic diseases, any prescribed drugs, dosage, pharmaceutical dosage form and route of administration, the reason for taking medication, allergies, drug-related conditions and history of adverse drug reactions, and current or past conditions/diseases. A detailed list of current or past conditions/diseases that will be included is given in the online supplemental appendix II.

The investigator will also collect the following information: presence or absence of ankle oedema, bone mineral density T-scores, history of influenza and pneumococcal vaccination, heart rate (beats per minute), and systolic blood and diastolic blood pressure (mm Hg).

The data are summarised in table 1.

### Data source

We will collect data using electronic health record consultations and clinical patient interviews.

### Database

The information collected will not include information that might identify the patients. Each patient will be numbered from 1 to 334 to protect their identity.

To evaluate data obtained throughout the study, a data safety monitoring board will be set up that will be composed of two external investigators with board expertise in this clinical field and academic and scientific activities.

| Table 1 Patients' data (phase II) | |
|---|---|
| **Patients' data** | |
| Sociodemographic data | Age<br>Gender<br>Educational level<br>Labour status<br>Marital status |
| Clinical data | Number of medications for chronic diseases, prescribed drugs<br>Pharmaceutical dosage form and route of administration, reason for taking medication<br>Allergies<br>Drug-related conditions<br>History of adverse drug reactions<br>Current or past conditions/diseases*<br>Presence or absence of ankle oedema<br>Bone mineral density T-scores<br>History of influenza and pneumococcal vaccination heart rate (bpm)<br>Systolic blood and diastolic blood pressure (mm Hg)<br>Estimated glomerular filtration rate<br>Serum K+ (mmol/L)<br>Serum Na+ (mmol/L) |

*Available at online supplemental appendix II.
bpm, beats per minute.

Following the Portuguese Clinical Research Law, all data recorded during the study will be stored for 5 years in a safe and proper place in the primary investigator's health centre after the closure of the investigation. All data containing participant codes will be destroyed after this period.

### Phase III: intrarater reliability and agreement study
An independent researcher/family doctor (named investigator 'A') will apply the Portuguese version of the STOPP-START criteria to all the patients using the information collected in phase II.

Investigator 'A' is an independent researcher with more than 10 years of experience of clinical practice.

To ensure intrarater reliability and agreement, the same doctor will re-evaluate these patients' records applying the same criteria 2 weeks later to avoid recall bias.[26 27]

### Phase IV: inter-rater reliability and agreement study
Three independent investigators/family doctors (named investigators 'B', 'C' and 'D') will independently apply the Portuguese version of the STOPP-START using the data, collected in phase II, of 205 randomly selected participants.[28] These three physicians are based in different health units and they will only have contact with the corresponding author who will give them the comprised data. Investigators 'B', 'C' and 'D' will independently assess the STOPP and START criteria in each of the 205 data sets and will be invited to give written comments if necessary.

Inter-rater agreement will be assessed by comparing the results of the three raters.

### Statistical analysis
Data will be stored with Microsoft Excel software. Data analyses will be made with SPSS Statistics V.27.0 and the software R.

Categorical variables will be described by absolute and relative frequencies.

Continuous variables will be described by mean and SD if normally distributed or by median and IQR if not normally distributed. Normality will be assessed by observation of histograms and implementation of the Kolmogorov-Smirnov test.

Intrarater/inter-rater reliability will be measured using Cohen's $\kappa$ coefficient and the respective 95% CI.[22] The Cohen's $\kappa$ coefficient will be interpreted as poor ($\kappa \leq 0.2$), fair ($0.21 \leq \kappa \leq 0.40$), moderate ($0.51 \leq \kappa \leq 0.6$), substantial ($0.61 \leq \kappa \leq 0.8$) and good ($0.81 \leq \kappa \leq 1.00$).[29] Intrarater/inter-rater agreement will be assessed using agreement proportions and specific (positive and negative) agreement proportions and the respective 95% CI.[22]

A p value less than or equal to 0.05 will be considered statistically significant.

### Patient and public involvement
No patient or member of the public will be involved in the design of this protocol or the establishment of intervention and the outcome measures.

### DISCUSSION
Appropriate prescriptions for older patients are a quality standard for healthcare. General practitioners are the main prescribers and they struggle to identify PIM as well as potential prescribing omissions. The STOPP-START tool is an easy way to manage the care of older patients. It is easier for daily use when adapted for the language of the prescriber.

This study is innovative because it is the first development of a Portuguese version of the STOPP-START criteria. Our research will not be merely a translation but also an adaptation done by independent general practitioners that will potentially increase the use of this version in the primary care setting.

Our research has some limitations such as the fact that even though it will be Portuguese language adaption of the STOPP-START criteria, it is only focused on Portugal and may not apply to other countries where Portuguese is used. This adapted version of STOPP-START is exclusively focused towards primary healthcare centres.

### ETHICS AND DISSEMINATION
Every participant will sign a written consent form (online supplemental appendix III). The identity of all participants will be protected throughout the study. The documents used to collect the data of the participants will

contain only an identification code of each participant using a number from 1 to 334.

This protocol was approved on 30 July 2020 by the Ethics Committee of the Central Health Region of Portugal with the reference number 034-2020.

We intend to publish the full article in a related peer-reviewed journal, and results will also be disseminated in conference presentations, reports and in a PhD thesis.

**Author affiliations**
[1]Center for Health Technology and Services Research; Faculty of Medicine, University of Porto, CINTESIS, Porto, Portugal
[2]Unidade de Saúde Familiar, USF Esgueira +, Aveiro, Portugal
[3]MEDCIDS—Department of Community Medicine, Information and Health Decision Sciences, FMUP, Porto, Portugal
[4]Center for Health Technology and Services Research, Oporto University Faculty of Medicine, CINTESIS, Porto, Portugal
[5]Porto Pharmacovigilance Centre, FMUP, Porto, Portugal
[6]MEDCIDS—Department of Community Medicine, Information and Decision in Health; Faculty of Medicine, University of Porto, FMUP, Porto, Portugal
[7]IPVC – Instituto Politécnico de Viana do Castelo, Viana do Castelo, Portugal

**Twitter** Luís Monteiro @luismonteiro140 and Carlos Martins @mgfamiliarnet

**ORCID iDs**
Luís Monteiro http://orcid.org/0000-0003-0784-5770
Carlos Martins http://orcid.org/0000-0001-8561-5167

## REFERENCES

1 OECD. Elderly population (indicator). Available: https://data.oecd.org/pop/elderly-population.htm#indicator-chart [Accessed 12 May 2019].
2 Organization WH, 2019. Available: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health
3 Organization WH. *World report on ageing and health*, 2015.
4 Organization WH. *The global strategy and action plan on ageing and health*, 2019.
5 Prazeres F, Santiago L. Prevalence of multimorbidity in the adult population attending primary care in Portugal: a cross-sectional study. *BMJ Open* 2015;5:e009287.
6 Formiga F, Ferrer A, Sanz H, *et al*. Patterns of comorbidity and multimorbidity in the oldest old: the Octabaix study. *Eur J Intern Med* 2013;24:40–4.
7 Lu W-H, Wen Y-W, Chen L-K, *et al*. Effect of polypharmacy, potentially inappropriate medications and anticholinergic burden on clinical outcomes: a retrospective cohort study. *CMAJ* 2015;187:E130–7.
8 Gnjidic D, Hilmer SN, Blyth FM, *et al*. Polypharmacy cutoff and outcomes: five or more medicines were used to identify community-dwelling older men at risk of different adverse outcomes. *J Clin Epidemiol* 2012;65:989–95.
9 Payne RA, Avery AJ, Duerden M, *et al*. Prevalence of polypharmacy in a Scottish primary care population. *Eur J Clin Pharmacol* 2014;70:575–81.
10 Motter FR, Fritzen JS, Hilmer SN, *et al*. Potentially inappropriate medication in the elderly: a systematic review of validated explicit criteria. *Eur J Clin Pharmacol* 2018;74:679–700.
11 Marengoni A, Angleman S, Melis R, *et al*. Aging with multimorbidity: a systematic review of the literature. *Ageing Res Rev* 2011;10:430–9.
12 Weng M-C, Tsai C-F, Sheu K-L, *et al*. The impact of number of drugs prescribed on the risk of potentially inappropriate medication among outpatient older adults with chronic diseases. *QJM* 2013;106:1009–15.
13 Tommelein E, Mehuys E, Petrovic M, *et al*. Potentially inappropriate prescribing in community-dwelling older people across Europe: a systematic literature review. *Eur J Clin Pharmacol* 2015;71:1415–27.
14 By the 2019 American Geriatrics Society Beers Criteria® Update Expert Panel. American geriatrics Society 2019 updated AGS beers Criteria® for potentially inappropriate medication use in older adults. *J Am Geriatr Soc* 2019;67:674–94.
15 Holt S, Schmiedl S, Thürmann PA. Potentially inappropriate medications in the elderly: the PRISCUS list. *Dtsch Arztebl Int* 2010;107:543–51.
16 Gallagher P, O'Mahony D. STOPP (screening tool of older persons' potentially inappropriate prescriptions): application to acutely ill elderly patients and comparison with beers' criteria. *Age Ageing* 2008;37:673–9.
17 O'Mahony D, O'Sullivan D, Byrne S, *et al*. STOPP/START criteria for potentially inappropriate prescribing in older people: version 2. *Age Ageing* 2015;44:213–8.
18 Boland B, Guignard B, Dalleur O, *et al*. Application of STOPP/START and beers criteria: compared analysis on identification and relevance of potentially inappropriate prescriptions. *Eur Geriatr Med* 2016;7:416–23.10.1016/j.eurger.2016.03.010
19 Hill-Taylor B, Walsh KA, Stewart S, *et al*. Effectiveness of the STOPP/START (screening tool of older persons' potentially inappropriate Prescriptions/Screening tool to alert doctors to the right treatment) criteria: systematic review and meta-analysis of randomized controlled studies. *J Clin Pharm Ther* 2016;41:158–69.
20 Lang P-O, Hasso Y, Belmin J, *et al*. [STOPP-START: adaptation of a French language screening tool for detecting inappropriate prescriptions in older people]. *Can J Public Health* 2009;100:426–31.
21 Gallagher P, Baeyens J-P, Topinkova E, *et al*. Inter-Rater reliability of STOPP (screening tool of older persons' prescriptions) and start (screening tool to alert doctors to right treatment) criteria amongst physicians in six European countries. *Age Ageing* 2009;38:603–6.
22 Kottner J, Audigé L, Brorson S, *et al*. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *J Clin Epidemiol* 2011;64:96–106.
23 Wild D, Grove A, Martin M, *et al*. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (pro) measures: report of the ISPOR Task force for translation and cultural adaptation. *Value Health* 2005;8:94–104.
24 Rotondi MA. kappaSize: sample size estimation functions for studies of interobserver agreement. R package version 1.1, 2013. Available: https://CRAN.R-project.org/package=kappaSize
25 Gallagher P, Ryan C, Byrne S, *et al*. STOPP (screening tool of older person's prescriptions) and start (screening tool to alert doctors to right treatment). consensus validation. *Int J Clin Pharmacol Ther* 2008;46:72–83.
26 Streiner DL NG. *Health measurement scales: a practical guide to their development and use*. 3rd ed. Oxford, England: Oxford University Press, 20032003.
27 Beemster TT, van Velzen JM, van Bennekom CAM, *et al*. Test-Retest reliability, agreement and responsiveness of productivity loss (iPCQ-VR) and healthcare utilization (TiCP-VR) questionnaires for sick workers with chronic musculoskeletal pain. *J Occup Rehabil* 2019;29:91–103.
28 Shoukri MM. *Measures of interobserver agreement*, 2004.
29 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.