

STOPP/START criteria for potentially inappropriate prescribing in older people: version 2

DENIS O'MAHONY^{1,2}, DAVID O'SULLIVAN³, STEPHEN BYRNE³, MARIE NOELLE O'CONNOR², CRISTIN RYAN⁴, PAUL GALLAGHER²

¹Geriatric Medicine, University College Cork, Cork, Munster, Ireland

²Geriatric Medicine, Cork University Hospital, Cork, Munster, Ireland

³School of Pharmacy, University College Cork, Cork, Munster, Ireland

⁴School of Pharmacy, Queen's University, Belfast, Northern Ireland, UK

Address correspondence to: D. O'Mahony. Tel: (+353) 21 4922396; Fax: (+353) 21 4922829. Email: denis.omahony@ucc.ie

Abstract

Purpose: screening tool of older people's prescriptions (STOPP) and screening tool to alert to right treatment (START) criteria were first published in 2008. Due to an expanding therapeutics evidence base, updating of the criteria was required.

Methods: we reviewed the 2008 STOPP/START criteria to add new evidence-based criteria and remove any obsolete criteria. A thorough literature review was performed to reassess the evidence base of the 2008 criteria and the proposed new criteria. Nineteen experts from 13 European countries reviewed a new draft of STOPP & START criteria including proposed new criteria. These experts were also asked to propose additional criteria they considered important to include in the revised STOPP & START criteria and to highlight any criteria from the 2008 list they considered less important or lacking an evidence base. The revised list of criteria was then validated using the Delphi consensus methodology.

Results: the expert panel agreed a final list of 114 criteria after two Delphi validation rounds, i.e. 80 STOPP criteria and 34 START criteria. This represents an overall 31% increase in STOPP/START criteria compared with version 1. Several new STOPP categories were created in version 2, namely antiplatelet/anticoagulant drugs, drugs affecting, or affected by, renal function and drugs that increase anticholinergic burden; new START categories include urogenital system drugs, analgesics and vaccines.

Conclusion: STOPP/START version 2 criteria have been expanded and updated for the purpose of minimizing inappropriate prescribing in older people. These criteria are based on an up-to-date literature review and consensus validation among a European panel of experts.

Keywords: *inappropriate prescribing, older people, STOPP/START criteria*

Introduction

Adverse drug reactions (ADRs) in older people currently represent a serious and growing public health problem [1]. Polypharmacy and inappropriate prescribing (IP) are well-known risk factors for ADRs, which commonly cause adverse clinical outcomes in older people [2, 3]. IP encompasses potentially inappropriate medications (PIMs) and potential prescribing omissions (PPOs) [4]. STOPP/START criteria for potential IP in older people recognise the dual nature of IP by including a list of PIMs (STOPP criteria) and PPOs (START criteria). Since the first iteration of STOPP/START criteria in 2008 [5], our research group has shown a

number of important properties of STOPP/START criteria, namely:

- STOPP criteria medications are significantly associated with adverse drug events (ADEs), unlike Beers 2003 criteria medications [6].
- STOPP/START criteria as an intervention applied at a single time point during hospitalisation for acute illness in older people significantly improve medication appropriateness [7], an effect that is maintained 6 months post-intervention.
- STOPP/START criteria as an intervention applied within 72 h of admission significantly reduce ADRs (with an

absolute risk reduction of 9.3%; number needed to treat = 11) and average length of stay by 3 days in older people hospitalised with unselected acute illnesses [8, <https://www.clinicaltrials.gov/ct2/show/study/NCT01467050>, 19 June 2014, date last accessed]).

For these reasons, we contend that STOPP/START criteria have practical clinical value. Although these findings are recent, it has become clear that an updated version of STOPP/START criteria is required due to a changing evidence base underpinning the first version of STOPP/START, the licensing of important new drugs since 2008 and the recognition of a more extensive list of PIMs than had been included in version 1. In addition, a number of STOPP/START criteria were no longer considered completely accurate or relevant, e.g. the use of calcium channel blockers (of any kind) in patients with chronic constipation (STOPP criterion) and the use of aspirin for primary prevention of cardiovascular disease in patients with diabetes (START criterion). It was also clear that a small number (12 in total) of criteria in STOPP/START version 1 were lacking in clinical importance or prevalence and were therefore of less relevance compared with other criteria in the list. In addition, there were some important criteria that were absent from STOPP/START version 1. Finally, we considered that STOPP/START criteria would be enhanced by seeking the input of a wider ranging panel of experts from across Europe than the panel of Irish and UK experts involved in the validation of version 1; this was to reflect Europe-wide prescribing practices in the general population of older people.

The aim of this study was to prepare and validate a new version of STOPP/START criteria so as to reflect more complete and up-to-date sets of PIMs and PPOs that may have serious negative effects on the health and well-being of older people in most clinical settings.

Methods

We proposed new criteria to be added to the 2008 list of STOPP/START criteria on the basis of an expanded evidence base since 2008. We evaluated these proposed new criteria in terms of their clinical importance, accuracy and evidence base. If the proposed new criteria met these requirements, we included them in the first draft of the STOPP/START version 2 criteria for further validation.

We recruited a panel of 19 experts from 13 countries in Europe, who had recognised expertise in Geriatric Medicine and pharmacotherapy in older people. We asked each expert to comment on the 2008 STOPP/START criteria, in particular their opinions on their current validity and relevance. We asked the expert panel members to propose additional STOPP and START criteria and ways to improve the structure and content of the existing criteria. We evaluated all new criteria proposed by the expert panel in terms of their clinical importance, accuracy and evidence base.

We then undertook a process of establishment of the evidence base to support all proposed criteria, including both

the criteria to be retained from STOPP/START version 1 and also the suggested additional criteria. The STOPP/START version 1 criteria were reviewed in terms of the current evidence base to support them. Some of the version 1 criteria were found to lack a firm evidence base, e.g. statin therapy for primary prevention of cardiovascular disease in diabetes mellitus. The authors proposed some new criteria, as did the expert panel members. We then searched PubMed, Embase and Cochrane Library databases for recent published evidence to underpin each version 1 criterion and the proposed new criteria. The key search words relating to each proposed criterion were entered in the three databases and relevant articles identified in the categories: systematic reviews, randomised controlled trials (RCTs) and reviews. In addition, we examined other sources, such as recently published textbooks, the British National Formulary and NICE (<http://www.nice.org.uk/>) and SIGN (<http://www.sign.ac.uk/guidelines/published/numlist.html>, 1 March 2014, date last accessed) treatment guidelines for sources of references. Where we did not find systematic reviews to support a particular criterion, we searched for reviews or RCTs that indicated clearly that the criterion was evidence based and therefore appropriate to include in STOPP/START version 2. Three members of the research team (the principal author and two postgraduate students under the principal author's supervision) read the selected articles to ensure their suitability as support evidence. From the initial list of proposed criteria, we removed criteria that did not have a clear evidence base. The remaining proposed criteria were organised according to physiological systems for further consensus assessment by the expert panel.

We then made available all of the relevant reference articles that constituted the support reference bank for the STOPP/START version 2 draft criteria to the members of the expert panel; we provided the latter in an online reference paper repository, using DropBox[®] software. The Delphi panel members were offered the abstracts and full publication versions of all of the selected articles relating to each proposed criterion. Each of the selected articles provided an evidence base to support each proposed criterion, and it was left to the discretion/need of each Delphi panel member to assess the evidence presented from the selected articles provided by the literature search. We did not ask the Delphi panel members to read all articles in a systematic manner and to provide a standardised rating of each article offered as evidence to support individual criteria. Rather, the articles were provided as a reference repository to be accessed whenever the expert panel members needed to check the supporting reference papers relating to particular proposed STOPP/START criteria.

As with STOPP/START version 1, topics were chosen for inclusion according to their considered importance within each physiological system, provided they had a sound evidence base, following literature search.

When the review of the first draft of version 2 criteria was complete, we sent the draft criteria to each member of the expert panel for review and feedback. We used SurveyMonkey[®]

software to facilitate an online Delphi validation, an established method of achieving consensus [9]. Using the Delphi validation method, we presented each criterion to the expert panel members in the form of a statement, e.g. Antipsychotics (i.e. other than quetiapine or clozapine) in older patients with Parkinsonism or Lewy Body Disease should be avoided due to the risk of severe extra-pyramidal symptoms. Each expert panel member then chose his/her level of agreement with each statement, ranging from 'strongly agree' to 'strongly disagree'. Panellists were also given a 'don't know' option and had the opportunity to comment on each suggested criterion using free text feedback before moving to the next proposed criterion for evaluation.

A Likert scale was used to measure the responses: 0 = don't know; 1 = strongly agree; 2 = agree; 3 = neutral; 4 = disagree; 5 = strongly disagree. The median and inter-quartile range values were calculated for each response in each iteration of the Delphi process. Criteria with a median value of 1 or 2 and a 75th centile value of not >2 were retained. Criteria with a higher median value were excluded.

Following the first validation round, we removed any proposed criteria that did not meet the retention requirements. We then drew up the second draft of the new criteria and proceeded to a second round of Delphi validation, once again using an online method, and inviting free text feedback from panel members on each criterion. As in the first validation round, we excluded those criteria that did not meet the retention requirements detailed previously. We planned to continue this process of repeated Delphi validation rounds until agreement to retain or reject was reached on all proposed criteria before declaring that the validation process was complete.

The construction and validation process of STOPP/START version 2 is summarised as follows:

- Phase 1: Call for review of STOPP/START version 1 criteria and proposal of a new evidence-based criteria/removal of obsolete criteria.
- Phase 2: Draft 1 of STOPP/START version 2 criteria.
- Phase 3: Search of PubMed, Embase and Cochrane databases for systematic reviews, reviews and other references to support STOPP/START version 2 criteria.
- Phase 4: Draft 2 of STOPP/START version 2 criteria, with support literature.
- Phase 5: Delphi validation Round 1 (19 experts).
- Phase 6: Draft 3 of STOPP/START version 2 criteria.
- Phase 7: Delphi validation Round 2 (19 experts).
- Phase 8: Draft 4 of STOPP/START version 2 criteria (final draft).

Results

For the first expert panel consultation, there was a list of 138 proposed STOPP/START criteria for evaluation. These included the original 87 STOPP/START version 1 criteria plus 37 possible additional criteria proposed by the Irish STOPP/START criteria group; we also received suggestions

Table 1. STOPP/START version 1 criteria removed from the proposed version 2 because of weak or equivocal supporting evidence

STOPP criteria
Aspirin with no history of coronary, cerebral or peripheral arterial occlusive symptoms
Calcium channel blockers with chronic constipation
Non-cardioselective beta-blocker with chronic obstructive pulmonary disease
Use of aspirin and warfarin in combination without histamine H2 receptor antagonist (except cimetidine because of interaction with warfarin) or proton pump inhibitor
Dipyridamole as monotherapy for cardiovascular secondary prevention
Aspirin to treat dizziness not clearly attributable to cerebrovascular disease
Phenothiazines in patients with epilepsy
Diphenoxylate, loperamide or codeine phosphate for treatment of severe gastroenteritis
Selective alpha-blockers in males with frequent urinary incontinence, i.e. one or more episodes of incontinence daily
First-generation antihistamines in patients with falls
Long-term opioids in patients with falls
Long-term opioids in those with dementia unless indicated for palliative care or management of moderate/severe chronic pain syndrome
START criteria
Metformin with type 2 diabetes mellitus +/- metabolic syndrome (in the absence of renal impairment, i.e. serum creatinine > 150 µmol/l, or estimated GFR < 50 ml/min/1.73 m ²)
Aspirin for primary prevention of cardiovascular disease in diabetes mellitus
Statin therapy for primary prevention of cardiovascular disease in diabetes mellitus

GFR, glomerular filtration rate.

for a further 14 criteria from the international expert review panel. Of the 138 proposed criteria, 127 criteria had supporting published evidence sufficient to warrant their inclusion in Round 1 of the Delphi validation which yielded 124 criteria with median Likert scores of 1 or 2. One hundred and seven of these 124 criteria had median Likert scores with 75th centile values of 1 or 2 and were retained as validated criteria. The remaining 17 criteria (7 STOPP and 10 START criteria) formed the basis for Delphi consensus Round 2 which achieved consensus on 7 criteria; a third consensus validation round was not required. The full list of references which supports the new STOPP criteria and START criteria is given in Supplementary data, Appendices 1 and 2 and the final list of new STOPP criteria and the new START criteria is given in Supplementary data, Appendices 3 and 4 available in *Age and Ageing* online, respectively.

Fifteen of the criteria from STOPP/START version 1 were not included in STOPP/START version 2 (Table 1) on the basis of a lack of sufficiently robust or consistent evidence in the published literature [10]. Table 2 details those criteria rejected during the Delphi consensus validation of STOPP/START version 2.

Discussion

STOPP/START criteria are important for several reasons. Since the first iteration of STOPP/START in 2008, there have been 74 published articles describing the use of

Table 2. Proposed criteria rejected by the expert panel for inclusion in STOPP/START version 2 using Delphi consensus

Rejected new STOPP criteria
Diuretic for treatment of hypertension with concurrent urinary incontinence (may exacerbate incontinence)
SSRIs with concurrent bleeding diathesis, prescription of anticoagulants or antiplatelet agents (increased risk of bleeding in general), active peptic ulcer disease or concurrent NSAID prescription (risk of gastrointestinal bleeding)
SSRIs in patients with previous history of major non-traumatic bleeding or in combination with drugs that may promote peptic ulceration, e.g. NSAIDs (increased risk of recurrent major bleeding)
Aspirin, clopidogrel, dipyridamol, vitamin K antagonists, direct thrombin inhibitors or factor Xa inhibitors with concurrent high bleeding risk, i.e. HAS-BLED score ≥ 3 ; HAS-BLED (hypertension, abnormal renal/liver function, stroke, bleeding history, labile INRs, elderly (age > 65 years), drugs that promote bleeding/alcohol)
Antidepressants of any kind in patients with recurrent falls
Rejected new START criteria
Memantine for moderate–severe Alzheimer's disease
Dopamine agonist (e.g. ropinirole or pramipexole) for Restless Legs Syndrome once iron deficiency has been excluded
Statin therapy in diabetes mellitus, unless the patient is at end of life or more appropriate for palliation
Phosphodiesterase type-5 inhibitor with persistent erectile dysfunction

SSRI, selective serotonin reuptake inhibitor; NSAID, non-steroidal anti-inflammatory drug.

While these criteria have a significant supportive evidence, the expert panel did not judge them to be of such high importance as to be considered potentially inappropriate in every case where they are encountered.

STOPP/START criteria in the PubMed database [<http://www.ncbi.nlm.nih.gov/pubmed/?term=stopp+criteria> (14 November 2013, date last accessed)], including 5 review articles and 45 original research articles involving STOPP/START criteria in various clinical scenarios. These publications originate from 24 countries. A recent Australian study comparing Beers criteria, STOPP/START criteria and prescribing indicators in Elderly Australians criteria concluded that the number and scope of drug-related problems identified by pharmacists was best represented by STOPP/START criteria [11].

The fact that STOPP/START criteria have been successfully applied for both research and practical clinical purposes in several countries in Europe, Asia, Australia, North America and South America indicates that the criteria probably have true global relevance. The relevance of STOPP/START criteria is further supported by the tangible clinical benefits demonstrated in the studies completed by our group, alluded to in the Introduction.

Version 2 of STOPP/START, with 114 criteria, represents a 31% increase in the total number of criteria included in version 1. This number of criteria may be considered unwieldy by some users, particularly those in busy clinical practice. Development of STOPP/START software applications has opened up the real possibility of applying the criteria in routine clinical practice globally.

Undoubtedly, some criteria have greater clinical importance than others, and there may be an argument in favour of hierarchical prioritisation within the overall set of STOPP/

START criteria. However, we considered that such hierarchical prioritisation might introduce unnecessary complexity to using STOPP/START, particularly when the criteria refer to *potential* rather than absolute medication inappropriateness.

Inevitably, there will be comparisons between STOPP/START version 2 and Beers criteria version 4 published in 2012 [12]. Although we have included a new section in STOPP criteria containing three implicit prescribing rules, STOPP/START, like Beers criteria, are essentially explicit criteria for PIMs. However, some important essential differences between STOPP/START and Beers criteria remain, principally the list of PPO's (START criteria) and the avoidance of mention of some Beers criteria drugs that are now absent from most European drug formularies, e.g. guanabenz, reserpine, mesoridazine, estazolam, trimethobenzamide and metaxalone.

In STOPP/START criteria, we decided not to indicate the comparative clinical relevance/severity of each criterion, since we considered almost all of the potential instances of IP in STOPP and START lists to be non-trivial, i.e. potentially serious.

There are several PIM criteria sets in the published literature [4]. However, only five published studies describe the application of PIM criteria *as an intervention tool* for improving medication appropriateness [13–16]. Three of these five studies describe the use of Beers criteria or variations of Beers criteria [13–15], one study deals with inappropriate prescribing in the elderly tool (IPET) criteria [16] and one study with STOPP/START criteria [7]. In our opinion, only those sets of PIM criteria that have tangible clinical benefit when applied as an intervention merit serious attention. Three of these five intervention studies used either Beers criteria or an adaptation of Beers criteria or IPET criteria as both the intervention and the primary outcome measure [13–15]. Gill *et al.* [16] reported that 37.9% of PIMs identified by IPET were discontinued by the prescribing physician, i.e. the intervention and the outcome measure were essentially the same. Although all five studies report benefit, only the study by Gallagher *et al.* [7] described an intervention (STOPP/START criteria) that was distinct from the outcome measure (Medication Appropriateness Index).

All explicit IP criteria essentially aim to improve medication appropriateness and/or avoid potentially serious ADRs and ADEs. IP criteria are clinically relevant if they significantly reduce the rate of ADRs or ADEs when applied prospectively to an unselected population of older patients in a particular clinical setting. STOPP/START criteria meet this essential requirement for clinical relevance on the basis of a highly significant reduction in ADR incidence in older hospitalised patients whose medication has been adjusted according to STOPP/START criteria compared with similar older patients receiving standard pharmaceutical care [8].

The present validation study shows the need to update and revise explicit IP criteria on a regular basis. The total number of STOPP/START criteria has increased by 31% from version 1 to version 2 between 2008 and 2013. Most of the extra criteria do not pertain to new drugs with new

indications arriving on the market during that 5-year time interval. Rather, they arise from new trial information, new systematic reviews and expert panel suggestions for additional criteria.

Although there are research data to indicate that the first iteration of STOPP/START as an intervention has clinical relevance in the acute hospital setting, it remains to be seen whether STOPP/START version 2 offers further ADR/ADE prevention benefits to older patients in various clinical settings. A recently funded European Commission Seventh Framework Programme project, called SENATOR [<http://www.senator-project.eu>] (19 June 2014, date last accessed), will examine the efficacy of a new pharmacotherapy optimisation software intervention based largely on STOPP/START criteria. The primary outcome measure of the international multi-centre RCT designed to test SENATOR software's efficacy will be ADR incidence in older people hospitalised with acute illness.

Key points

- PIMs and PPOs are commonly encountered in older people. PIMs and PPOs are closely related to ADEs and ADRs, but they are preventable.
 - STOPP/START criteria have been shown to be significantly associated with ADEs in acutely ill older people, unlike Beers criteria.
 - In single-centre RCTs, STOPP/START criteria used as an intervention significantly improve medication appropriateness and reduce the incidence of ADRs in older people in hospital, compared with standard pharmaceutical care.
 - Since the first publication of STOPP/START criteria in 2008, the therapeutics evidence base as it applies to older people has expanded significantly, indicating the need for updating and revision of STOPP/START criteria. The present study describes this process, resulting in a 31% increase of the number of STOPP/START criteria compared with the 2008 version, i.e. 114 criteria.
-

Acknowledgements

We are deeply grateful to the members of the expert advisory panel for their individual contributions to the updation and validation of STOPP/START version 2. The panel included the following persons: Prof. Stephen Jackson (UK), Dr Una Martin (UK), Prof. Sylvie Legrain (France), Dr Marie Laroche (France), Dr Alfonso Cruz-Jentoft (Spain), Prof. Antonio Cherubini (Italy), Prof. Graziano Onder (Italy), Prof. Tischa van der Cammen (The Netherlands), Dr Paul Jansen (Netherlands), Dr Jens-Ulrik Rosholm (Denmark), Prof. Eva Topinkova (Czech Republic), Prof. Tomasz Grodzicki (Poland), Prof. Adalsteinn Gudmundsson (Iceland), Prof. Jean-Pierre Baeyens (Belgium), Prof. Mirko Petrovic (Belgium), Prof. Anne Spinewine (Belgium), Prof. Pierre Olivier Lang

(Switzerland), Dr Thomas Fruehwald (Austria) and Dr Wolfgang von Renteln-Kruse (Germany).

Conflicts of interest

D.O'M. and S.B. hold a patent, with others, called A Prescription Decision Support System (based on STOPP & START prescribing guidelines), lodged with the European Patent Office (Munich); patent No. 11757950.8–1952. They also possess a shareholding in Clinical Support Information Systems[®], a software company established for the purpose of developing and marketing STOPP & START criteria as a set of software products.

Funding

This research was supported by grant funding provided by the Health Research Board of Ireland (HRB; grant number HRA_HSR/2010/14).

Supplementary data

Supplementary data mentioned in the text are available to subscribers in *Age and Ageing* online.

References

1. Scott I, Jayathissa S. Quality of drug prescribing in older patients: is there a problem and can we improve it? *Intern Med J* 2010; 40: 7–18.
2. Atkin PA, Veitch PC, Veitch EM, Ogle SJ. The epidemiology of serious adverse drug reactions among the elderly. *Drugs Aging* 1999; 14: 141–52.
3. Onder G, van der Cammen TJ, Petrovic M, Somers A, Rajkumar C. Strategies to reduce the risk of iatrogenic illness in complex older adults. *Age Ageing* 2013; 42: 284–91.
4. O'Connor MN, Gallagher P, O'Mahony D. Inappropriate prescribing: criteria, detection and prevention. *Drugs Aging* 2012; 29: 437–52.
5. Gallagher P, Ryan C, Byrne S, Kennedy J, O'Mahony D. 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.
6. Hamilton H, Gallagher P, Ryan C, Byrne S, O'Mahony D. Potentially inappropriate medications defined by STOPP criteria and the risk of adverse drug events in older hospitalized patients. *Arch Intern Med* 2011; 171: 1013–9.
7. Gallagher PF, O'Connor MN, O'Mahony D. Prevention of potentially inappropriate prescribing for elderly patients: a randomized controlled trial using STOPP/START criteria. *Clin Pharmacol Ther* 2011; 89: 845–54.
8. O'Connor MN. Adverse Drug Reactions in Older People During Hospitalisation: Prevalence, Risk factors and Recognition. Unpublished MD thesis. University College Cork, 2013.
9. Dalkey NC. Delphi P-3704 RAND. Santa Monica, CA: RAND Corporation, 1967.

D. O'Mahony *et al.*

10. Sackett DL, Straus SE, Richardson WS *et al.* Evidence-Based Medicine: How to Practice and Teach EBM, 2nd edition. Edinburgh: Churchill Livingstone, 2000.
11. Curtain CM, Bindoff IK, Westbury JL, Peterson GM. A comparison of prescribing criteria when applied to older community-based patients. *Drugs Aging* 2013; 30: 935–43.
12. American Geriatrics Society 2012 Beers Criteria Update Expert Panel. American Geriatrics Society updated Beers Criteria for potentially inappropriate medication use in older adults. *J Am Geriatr Soc* 2012; 60: 616–31.
13. Spiker EC, Emptage RE, Giannamore MR, Pedersen CA. Potential adverse drug events in an indigent and homeless geriatric population. *Ann Pharmacother* 2001; 35: 1166–72.
14. Blozik E, Born AM, Stuck AE, Benninger U, Gillmann G, Clough-Gorr KM. Reduction of inappropriate medications among older nursing-home residents: a nurse-led, pre/post-design, intervention study. *Drugs Aging* 2010; 27: 1009–17.
15. Dunn RL, Harrison D, Ripley TL. The beers criteria as an out-patient screening tool for potentially inappropriate medications. *Consult Pharm* 2011; 26: 754–63.
16. Gill SS, Misiaszek BC, Brymer C. Improving prescribing in the elderly: a study in the long term care setting. *Can J Clin Pharmacol* 2001; 8: 78–83.

Received 7 March 2014; accepted in revised form 4 July 2014