NIHR HS&DR programme

HS&DR protocol – Project ref: 17/100/13

Version: 2

Date: 19/11/2018

How, when and why do STOPP/START criteria based interventions improve medicines management for older people: a realist synthesis

Chief investigator	**Prof Jose M Valderas**
Sponsor	**University of Exeter**
Funder	**NIHR HS&DR**
NIHR Portfolio number	
Ethics Reference	**University of Exeter Medical School REC Reference: Nov18/D/181**

Protocol version control table

Protocol version	Date	Changes in relation to previous version	Comments
V1	01/10/2018	Not applicable	Final application version
V2	19/11/2018	No changes	As approved by UoE REC

How, when and why does the use of the STOPP/START tool improve medicines management in older people: a realist synthesis

1. Aims and Objectives

***Aim*:** To understand how, when and why the use of the STOPP/START criteria improves medicines management in older people.

Objectives:

1. To identify the ideas and assumptions (programme theories) underlying how interventions based on STOPP/START criteria are intended to work, for whom, in what circumstances and why. Test and refine these programme theories to explain how contextual factors shape the mechanisms through which the STOPP/START criteria produces better outcomes for patients.
2. Identify and describe the resource use and cost requirements or impacts of the different context mechanism outcome (CMO) configurations.

2. Background and Rationale

What is the problem being addressed?

Up to 30% of hospital admissions of older people are associated with drug-related problems related to noxious and unintended responses to drugs (Adverse Drug Reactions (ADRs)) (1-3). The concurrent use of multiple medications (polypharmacy) is associated with ADRs and prescribing errors (4, 5). Patients taking seven or more concurrent medications simultaneously have an 82% risk of an ADR (6). Evidence suggests that between 30% and 55% of admissions due to ADRs could be prevented by more appropriate prescribing (7-9), by appraising age-related changes in pharmacodynamics and pharmacokinetics, balancing risks and benefits (including cost-efficiency and life expectancy), and listening to patients' and carers' concerns (10-12). The task is further complicated by the fact that older patients with multiple morbidities are often excluded from clinical trials (13).

To address this, screening tools for supporting medicines optimization of older adults have been developed. Those most widely used are the Beers criteria for Potentially Inappropriate Medication Use in Older Adults and the combination of the Screening Tools of Older Persons' Prescriptions (STOPP) and to Alert to Right Treatment (START). Almost half of the drugs in the Beers criteria are unavailable for prescribers in Europe (14, 15), several of the drugs are not contra-indicated in older people as per the British National Formulary (BNF) (10) whereas other contra-indicated drugs are omitted (14). Given these shortcomings, the STOPP/START criteria has recently been developed and validated in European countries for use (16). A systematic review identified 13 studies conducted up to 2012 using these criteria in community dwelling, acute care and long-term care older patients found that the use of

these criteria reduces falls, delirium episodes, hospital length-of-stay, care visits (primary and emergency) and medication costs (17).

To date no attempt has been made to understand how, when and why the use of the STOPP/START criteria might improve medicines management in older people by different ways:

a) To understand how, when and why the STOPP/START criteria works differently in different situations for different population
b) Variation in how the STOPP/START criteria are implemented (e.g. by different professionals)
c) Need to understand how to further improve the use of the STOPP/START criteria

Why is the research important in terms of improving the health of the public and/or to patients and the NHS?

The prescription of potentially inappropriate medications to older people is highly prevalent, ranging from 12% in community-dwelling elderly to 40% in nursing home residents in Europe and the United States (18). Older people are particularly vulnerable to inappropriate prescribing because of their multiple drug regimens, co-morbid conditions and age associated physiological processes (19). Drug-related problems and potentially inappropriate prescribing (PIP) are highly prevalent among older adults and exert a significant disease burden, and have been associated with adverse drug events (ADEs (which includes ADRs and medication errors)) leading to hospitalisation and death, and increased health resource utilisation (20-22). A national population study in Ireland by Cahir et.al in 2007 alone, estimated a cost of PIP of over €38 million (23).

Medicines in older people are considered appropriate when they have a clear evidence-based indication, are cost-effective and well tolerated in the majority of the population. While PIP and ADRs are generally associated with over prescription, PIP may also occur when a patient is NOT prescribed appropriate medication for the treatment and prevention of a disease or condition (24). This might occur due to many factors such as ageism, fear of adverse events, economic concerns and lack of prescribing knowledge (25-27).

Detecting PIP early may prevent ADEs and cost-effectiveness of medicines which gives an opportunity for improving quality of care in older adults (10, 28), and quality of life of older people can be improved by discontinuing inappropriate medications (29). Currently, prescribing and reviewing medications are largely based on clinicians' clinical judgement for older patients. However, guidelines and systematically developed evidence-based tools have emerged in recent years to facilitate a comprehensive clinical approach to medication review. These typically consist of appropriate medications for older people to predominantly avoid harmful prescriptions in the older population. Current NICE Guidance on Medicines Optimisation includes as a recommendation: "Consider using a screening tool – for example, the STOPP/START tool in older people – to identify potential medicines-related patient safety incidents in some groups.", especially in relation to adults taking multiple medicines (polypharmacy) and adults with chronic or long-term conditions and older people (30). However, the evidence for this recommendation

is limited. One of the research recommendations in the guideline is for research on the impact of using clinical decision support systems. The focus of the recommendation is specific for computerized systems, even this particular application of the decision tools would be benefit from knowledge on the mechanisms by such an interventions work more generally in order to optimize their implementation and to maximize their benefits.

3. Evidence explaining why this research is needed now

ADEs and cost-inefficiencies can be reduced by identifying PIPs and thus improving the health and quality of life of older people (10). Explicit screening tools such as the STOPP/START tool help in detecting PIPs systematically. A number of systematic reviews have been conducted on interventions to improve the appropriateness of polypharmacy and the STOPP/START tool: a 2012 Cochrane systematic review of the evidence base for interventions to improve the appropriate use of polypharmacy for older people has highlighted STOPP/START criteria as the basis for promising interventions(31). The Cochrane review also highlighted the STOPP/START as validated instruments and demonstrated a reduction in inappropriate prescribing. Another systematic review specifically of the STOPP/START criteria looked at the prevalence of potentially inappropriate prescribing in older adults, and evidence of clinical, humanistic and economic impact (17). Another recent systematic review of randomised clinical trials was conducted on the effectiveness of the STOPP/START criteria(12). However, the current evidence base for the STOPP/START criteria, as reviewed by conventional methods for the systematic review of effectiveness is limited for all types of population (e.g. especially in frail older and community-living patients receiving primary care). The Cochrane review (31) and another systematic review by Greenhalgh and colleagues (32) also indicated that it is important that sufficient detail about mechanisms that lead to success in the intervention and the context in which complex interventions are conducted are reported and understood so the transferability of complex interventions can be assessed(32, 33). Rather than seeking to assess whether the STOPP/START criteria is effective or cost-effective, in any binary sense, it is better to use review methods that seek to explain how and why these tool tools are more or less effective in different circumstances or for different patient groups, thereby facilitating the development of optimal implementation strategies. This is particularly relevant at a time when improving care for older people with complex care needs, who frequently require complex medication regimes, is a priority for the NHS.

Systematic reviews can sometimes provide a picture of 'outcome patterns' and give indications of where the intervention was successful and where it was not. In systematic reviews this variation is ignored and what matters is the mean effect across trials. In realist synthesis we can start to investigate the patterns. Were there any contextual features that were common to the positive trials? How might we explain this? Does this give any hints as to possible mechanisms? This review will be complementary to the two previous reviews, which did not have a realist approach. The proposed review will provide greater understanding and insight into how, when and why interventions work in practice, difference between settings, clinician types, patient groups, patient ages and shared decision making which is very important in terms of optimising implementation of STOPP/START. We will also broaden our

search to include studies from 2014 to till date and examine non-RCTs and qualitative studies, unlike the other two reviews. We will also extract the information from the relevant reports from NHS.

Unlike with a conventional systematic review of cost-effectiveness, in which it is typically presumed that empirical evidence about the costs and cost-effectiveness of an intervention solely exists in economic evaluations or comparative cost analyses, a realist review with a cost-effectiveness review question should take an integrated approach, in which:

- Evidence about the supposed (i.e. programme theory) and actual costs and cost-effectiveness of interventions and comparators is sought and captured in a wider range of study types; including economic evaluations and costing studies, but also in effectiveness studies and other study types about the interventions or programmes of interest.
- It is acknowledged that 'effectiveness outcomes' and 'economic outcomes' are not mutually exclusive categories; most effectiveness outcomes, whether clinical or patient-reported, are also resource-committing and therefore cost-affecting in some way (e.g. the treatment of Adverse Drug Events; the reduction in medications taken).
- The underlying mechanisms of an intervention are the consequences of providing a particular combination of resources.

Both Pawson's original conception of programme theory (34) and more recent clarifications of the realist notion of mechanism (35) clearly acknowledge that intervention mechanisms are fundamentally how people/providers/clients respond to particular (usually new) resources. As well as identifying the resource consumption required for intervention mechanisms to exist, realist review must also then identify the resource changes or commitments implicit in outcomes, and any resource implications of salient contexts.

4. Research Plan/Methods

A realist evaluation will be conducted following Realist And Meta-narrative Evidence Syntheses: Evolving Standards (RAMESES) guidelines for realist synthesis (36). The approach was chosen to inform the implementation of future of the STOPP/START criteria initiatives as it recognises that interventions are not universally successful and that outcomes are context dependent (37). Realist synthesis is a theory-based approach that seeks to explain how context shapes the mechanisms through which programmes produce intended and unintended outcomes (37). This is accomplished through the identification and refinement of underlying programme theories. These seek to explain how the generative mechanisms (M) triggered by the resources offered by the intervention or programme are shaped by conditions or contexts (C) and the pattern of outcomes (O) produced. In this case, context refers to the ways in which the intervention is designed and the circumstances into which the intervention is implemented that may either support or constrain how participants respond to the intervention, while outcome refers to both the intended and unintended effects of the intervention on participants.

Phase 1: Identifying programme theories

Phase 1 will identify programme theories or ideas and assumptions about how the STOPP/START criteria is intended to work, for whom and in what contexts, why (e.g. ideas about what drug groups and/or conditions are these tools applied to, and why it doesn't work and in what patient population) and whether (and how) patients are being involved in shared decision making in stopping or starting medicines. STOPP/START-based interventions are clearly primarily intended to improve health and avoid harms. Therefore, in Phase 1 of the review, we will not actively seek theoretical mechanisms that exclusively explain the cost or cost-effectiveness of STOPP/START-based interventions; but rather seek to identify and make explicit which underlying mechanisms, contexts or outcomes are resource-consuming or cost-generating, in what ways and to what extent. We will conduct electronic searches of the grey literature to identify guidance documentation and policy documents, and electronic searches of the peer reviewed literature to identify position pieces, comments, letters, editorials and critical pieces which explain how the STOPP/START criteria is intended to work in addition to information provided in reports of studies using the tool. The aim of this first phase of realist synthesis is to capture in detail the reasoning that underlies the intended outcomes of an intervention. To prioritise the most important or explanatorily powerful programme theories, we will consult topic experts and NHS stakeholders and with a patient group. EC will use her extensive contacts to establish and manage a patient group who will work with us throughout the review. Our first meeting with them will focus on identifying how the STOPP/START criteria was intended to work. Expert consultations will be achieved through ongoing conversations with experts in the field on why the STOPP/START criteria works, who it works for, in what circumstances and why. This involves semi-structured interviews with experts to triangulate the emerging theories as resulting from the literature review. Given the scope and the realist nature of these reviews (realist hypotheses are not confirmed or abandoned through saturation but rather through relevance and rigour) we expect that about ten interviews will meet our needs, although the final number will be confirmed if the interviews uncover how emerging mechanistic and contextual factors may contribute to the outcome patterns emerging from the empirical literature (37, 38). We aim to include clinicians with a special interest in medicines optimisation, drug safety and pharmacotherapy, experts involved in medicines optimisation at NHS RightCare, providers of care in primary and secondary care (e.g. Geriatricians, GPs and other clinicians), care home managers, academics and those involved in developing education and guidance for older people, and researchers involved in the development of the STOPP/START criteria, and will be recruited through Academic Health Science Network, Clinical Research Network, Royal Pharmaceutical Society, and through co-applicant networks. In addition we will consider inviting British Geriatrics Society experts (following the advice of the Board). We will also consider inviting contributors to the NHS England "Toolkit for general practice in supporting older people with frailty" (39), and the NICE Multimorbidity Guideline Committee (40). Interviews will be based on a pre-established stakeholder topic guide. Interviews will last up to 30 minutes and will be conducted either face-to-face or telephone at participant's convenience. Participants will be reimbursed for their travel. Interviews will be audio-recorded and analysed analysed retroductively (41) to expand and refine our programme theories. The purpose of the expert interviews is three-fold: (1) to gain input on our list

of candidate CMO configurations, (2) to identify additional CMO configurations, and lastly, (3) to identify additional literature and/or relevant concepts that we may have missed. This will be used as a device to guide searching in the literature.

For this first phase of the review, that primarily concerns the comprehensive development of programme theory, any identified source or study design that provides relevant and rich insights will be used.

Phase 2: testing programme theories

Phase 2 will 'test' the programme theories obtained from phase 1 (through exploratory search and expert consultation), by synthesising using published and unpublished empirical quantitative and qualitative evidence. The programme theories or hypotheses that provide the backbone of the review and determine the search strategy and decisions about study inclusion into the review in order to test and refine these theories. We anticipate that the majority of evidence will be found in the published research literature via searching bibliographic databases. Searches of bibliographic databases will be supplemented by forward- and backward-citation chasing of relevant studies, and searches of grey literature resources and contacting authors, as required.

Phase 2 of the review, the second (cost-effectiveness) objective of the realist review will have dedicated searches to identify any relevant economic evaluations and cost analyses of the use STOPP/START based interventions in older people, and we will assess the quality of these studies using conventional criteria (42). However, their data extraction, quality assessment and synthesis will be integrated with the main body of empirical studies in the realist review. Further, we expect that as well as the types and levels of resource use being context-dependent, we anticipate that the opportunity cost of particular resources will be different in different contexts. The main principles and steps of this approach will be explored and illustrated (43, 44).

The project team will develop a list of relevant search terms to use in the main electronic databases. The information specialist will have a major role in this process, which is further discussed under search strategy section below. A starting point for our evidence searches will be existing systematic reviews of the STOPP/START tools and/or of Interventions aiming at medicines optimisation for improving the appropriateness of prescribing.

Additional studies will be identified through searching electronic databases. We will search the databases relevant to subject area (described in more detail under search strategy) limited to review publication type. Our scoping searches have identified approximately 164 references in MEDLINE (see Appendix 1 for example) which we estimate would retrieve 500-600 references when conducted on all databases.

For this phase of the review, that specifically aims to test/refine the programme theories in relation to quantitative empirical evidence about patient outcomes, the review will include the following types of comparative effectiveness studies of the STOPP/START criteria on the medication profiles of persons 65 years of age or older: Randomised controlled trials (individual- or cluster-randomised), nonrandomised controlled trials, controlled before-and-after studies, uncontrolled before-and-after

studies, observational and cross-sectional studies. The theory testing/refinement phase of the review will also include primary qualitative and mixed methods studies.

Justification and explanation why realist synthesis methods are appropriate for this study and how they will be used to synthesise all of the evidence

The goal of realist synthesis is to refine our understanding of how a programme works and the conditions and caveats that influence its success, rather than offering a descriptive summary or mean effect calculation across a family of programmes. Specifically, our synthesis is concerned with understanding the conditions in which blockages or unintended consequences occur (which may prevent or limit the implementation of STOPP/START) and those in which these blockages can be overcome. Synthesis takes several forms. At its most basic realist synthesis is a form of 'triangulation', bringing together information from different primary studies and different study types to explain why a pattern of outcomes may occur. For example, a systematic review on the effectiveness of the STOPP/START criteria reported that all interventions produced improvements in prescribing appropriateness. However, the impact of the intervention differed between studies, due perhaps to the variation in the implementation of the STOPP/START criteria, and its integration into the clinical workflow (12). This suggests there is a 'blockage' or 'obstacle' in the implementation of the criteria. We can then explore potential explanations for these findings. For example, qualitative studies suggest that the GPs imagined key elements for the implementation in daily practice as computerized clinical decision support system (CCDS), education, and multidisciplinary collaborations, especially at care transitions and in nursing homes(45) and utility of the criteria(46). One possible option to address this blockage it to provide training to doctors on appropriateness of prescribing in older patients(47), how to facilitate STOPP/START using CCDS, on multidisciplinary usage of the tool(45) and how the interface works. In this way, different studies can be brought together to understand progress (or otherwise) along the implementation chain. In general we will explore proximal, intermediate and distal in the use of STOPP/START and explain why disruption may occur and identify the circumstances in which it may be overcome.

Another form of synthesis, particularly useful when there is disagreement on the merits of an intervention is to 'adjudicate' between the contending positions. This is not a matter of providing evidence to declare a certain standpoint correct and another one invalid. Rather adjudication assists in understanding the respects in which a particular programme theory holds and those where it does not. For example, in previous studies, despite the use of STOPP/START, patients' conservativeness was mentioned as a barrier to optimising prescribing(48). A recent review highlighted that patients could be both barriers and enablers to de-prescribe drugs(49), and patients should be involved in shared decision making and optimising the treatment. Our review would seek to identify explanations for these contrasting findings to identify the circumstances in which the intended mechanism (motivation to improve) and intended outcomes occur (improved care) and those in which unintended mechanisms (lowered morale, tunnel vision) occur. Thus, it seeks to provide an explanation for the whole pattern of outcomes across studies rather than seek out an average effect. This will enable clinicians to target

the criteria to local conditions more effectively, to identify the ways in which the intervention could be implemented to maximise its impact on patient care and the resources need to support this.

Finally, the main form of synthesis is known as 'contingency building'. All interventions based on STOPP/START criteria make assumptions that they will work under implementation conditions A, B, C and applied in contexts P, Q, R. The purpose of the review is to refine many such hypotheses, enabling us to say that, more probably, A, C, D, E and P, Q, S are the vital ingredients. For example, at an aggregate level there is debate about whether a collaborative approach of multidisciplinary team that include pharmacists reduces the use of potentially inappropriate medications(50). Others have argued that whether a particular intervention in some specific therapeutic points would be beneficial (51, 52). Some others have argued that potentially high clinical value to be obtained from routine deployment of STOPP/START criteria(53). Our review would seek to identify the necessary conditions under which the intervention or interventions results in improvements to the quality of care.

BeHEMoTh criteria

Behaviour: The behaviour of interest in this study will be 'prescribing'

Health condition: older adults taking multiple medicines (polypharmacy) and adults with chronic or long-term conditions

Exclusions: Children and adults <65 years

Models or Theories: We will identify models and theories as the review develops.

Search Strategy

Both the search to identify relevant programme theories and the search to identify evidence to test the programme theories will make use of bibliographic databases, including MEDLINE (via Ovid). Search strategies developed and carried out using MEDLINE will use relevant MeSH terms wherever appropriate; and MeSH terms will be translated for use in other bibliographic databases as required. We will take a lead on relevant MeSH terminology from the recently completed Cochrane systematic review of interventions to improve the appropriate use of polypharmacy in older people (54). This systematic review used the following MeSH heading in the MEDLINE (Ovid) search strategy: STOPP, START, potentially inappropriate medication list, polypharmacy, pharmacotherapy, multimorbidity, comorbidity, polymorbidity, poly-drugs, multiple drugs, multiple medicines, inappropriate drugs/medicines, medication errors, inappropriate prescribing, age, elder, older adults, outpatients, inpatients and general practice.

The search strategy to identify evidence to test the programme theories will be designed and conducted by an experienced information specialist (SB) in close collaboration with the whole research team. SB has extensive experience of conducting searches for complex evidence syntheses, particularly realist reviews.

We anticipate our search strategy to include a combination of the following:

(1) Searching bibliographic databases (using keywords based on the theories identified) including: Embase, MEDLINE, MEDLINE In-Process and Social Policy and Practice (all via Ovid); CINAHL (Cumulative Index to Nursing & Allied Health Literature) (via EBSCO); the Cochrane library, (including the Cochrane Database of Systematic Reviews, CENTRAL and the NHS Economic Evaluation Database [NHS EED]) and other relevant databases identified by the Information Specialist if necessary. NHS EED database is searchable as an archive which might include studies not indexed in other bibliographic databases (http://eprints.whiterose.ac.uk/593/1/nixonj1.pdf). For economic studies, we would also search other, up-to-date, health and social care databases, using a health economics search filter (to be developed as and when needed, adapted from pre-existing health economics filters: https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/filters-to-find-i).

(2) "Cited by" articles search (backward citation chasing) and citations contained in the reference lists of included papers (forward citation chasing) using Web of Science (Clarivate Analytics).

(3) Contacting authors of included papers (4) Grey literature searching via Social Policy and Practice (listed above under 'Searching bibliographic databases') and the websites of relevant organisations.

(4) Grey literature searching via Social Policy and Practice (listed above under 'Searching bibliographic databases') and the websites of relevant organisations.

The search for evidence will be progressively extended and re-focused (based on the identified sources) as the review progresses and greater insight is attained into the issues concerned.

Review Strategy

Documents will be selected based on relevance i.e. whether data can contribute to theory building and/or testing. A random sample of 10% of articles will be selected, read, assessed and discussed by three reviewers using a preliminary set of inclusion/exclusion criteria. The remaining 90% will be completed by the same reviewers independently. However a number of these may require discussion/joint reading between the readers as they may be pivotal papers or difficult to understand/integrate). In realist reviews, the study itself is rarely used as the unit of analysis; instead realist reviews may consider small sections of the primary study e.g. the introduction or discussion sections, to test a very specific hypothesis about the relationships between context, mechanism, and outcomes (55). We will thus select and review studies based on what new knowledge they bring to our thinking about the theory of the impact of the STOPP/START criteria.

The realist review method obtains information by note-taking and annotation rather than standardised data extraction as used in a traditional systematic review. Documents are examined for data that contribute to theories on how an intervention is supposed to work which are then highlighted, noted and given an approximate label. The reviewer may make use of forms to assist the sifting, sorting and annotation of primary source materials but do not take the form of a single, standard list of questions as used in a traditional systematic review. Three researchers will independently extract data from all potentially relevant full-text articles.

Synthesis of the diverse sources of evidence included in a realist review is conducted through a process of reasoning that is structured around the following activities (56):

(1) Juxtaposition of sources of evidence – for example, where evidence about the STOPP/START criteria in one source enables insights into evidence about outcomes in another source.

(2) Reconciling of sources of evidence – where results differ in apparently similar circumstances, further investigation is appropriate in order to find explanations for why these different results occurred.

(3) Adjudication of sources of evidence – on the basis of methodological strengths or weaknesses.

(4) Consolidation of sources of evidence – where outcomes differ in particular contexts, an explanation can be constructed of how and why these outcomes occur differently.

Quality appraisal

Formal quality assessment will not be undertaken as part of Phase 1. Instead, the lead reviewer along with other team members will make judgements about the relevance and value of each source in contributing to theory development on a source by source basis.

For Phase 2 of the review, which aims to test/refine the programme theories in relation to evidence about patient outcomes from empirical effectiveness and qualitative studies, the reviewers will formally assess study quality.

Quality assessment will use the concept of rigour -whether the methods used to generate the relevant data are credible and trustworthy. Sources will be classified conceptually rich (thick) or weaker (thin) (57). This strategy has been found to be practical and useful in theory-driven synthesis as it allows the reviewer to focus on the stronger sources of programme theories without excluding weaker sources that may make an important contribution (58). There will not be the use of a single standard list of questions used on all data, as used in a traditional systematic review.

Review synthesis

The final realist programme theory will be summarised through narrative synthesis, using text, summary tables, a logic model and where appropriate graphics to summarise individual papers/reports and draw insights across papers/reports. In terms of the outputs for this review objective, we will produce both a (conventional) cost-effectiveness 'matrix' to map the variation in quantitative results of economic evaluations of STOPP/START, and a narrative synthesis which uses a wider set of our included studies to try and explain the heterogeneity of cost and cost-effectiveness findings in terms of relevant mechanisms, related contexts and outcomes. Rather than seeking to produce a reliable 'average' cost-effectiveness finding, the realist review will aim to identify those circumstances, and with which configurations and levels of different resources, STOPP/START interventions are most effective and cost-effective. This approach is also consistent with the revised goals and methods advocated by other health economists when evaluating the cost-effectiveness of complex interventions (59, 60). The results of the synthesis will be written up according to RAMESES guidelines for reporting realist reviews (34). The results will produce useful knowledge for academics, patients, professionals, and NHS and policy makers.

5. Project management

JV will be responsible for the co-ordination of the project and the supervision of the research associates. The review will be carried out by research associates (JBG, IRC, MA) under the supervision of and with assistance from JV and with regular methodological support from JG and RA. Quality assurance will be carried out by JBG. Our NHS stakeholder and topic experts (CH, RP) will provide topic oriented support throughout the review but specifically in relation to (1)informing the focus of the review;(2)interpreting the findings of the review and (3)advising and planning the dissemination of the review.

SB will be responsible for designing, testing, refining and running the searches and setting up and managing the Endnote database. The project team (JV, CH,RP, SB, JG, RA, IRC, MA, EC, JH, JBG) will meet at five points during the review (face-to-face where possible and otherwise by teleconference) and circulate discussion points and feedback as appropriate. EC will be responsible for recruiting members to our patient advisory group and chairing the patient group meetings and JH, JV, JBG and MA will be part of the meetings. The patient/public advisory group will meet four times, to provide input on scope, methods, theory development, synthesis, interim findings and dissemination plans.

To ensure the emerging synthesis is recorded, shared among and shaped by all team members, our management team (JV, JBG, IRC, MA, JG, EC) will produce a series of Working Papers to summarise key processes within the review that will be shared for comment, edit and refinement. Co-ordination of the comments will be provided by JV. These series of working papers will constitute the basis of briefing papers for our patient advisory group and will also provide material for the final report to HS&DR.

As discussed under research timetable, we have 4 milestones each to complete phase 1, phase 2 of the review, synthesising the information and submitting final report to HS&DR.

Our criteria of success of the project as it progresses will be measured against completion of the tasks described above and production of the following knowledge: (1) answering the aim of the project; (2) developing a comprehensive taxonomy of the 'programme theories' underlying the STOPP/START criteria;(3)Identify and explain the most important mechanisms by which the STOPP/START criteria is thought to improve outcomes for patients;(4)identification of the potential unintended consequences of using the STOPP/START criteria;(5)Variation in how the STOPP/START criteria are implemented (e.g.by different professionals);(6)an understanding of how to further improve use of the criteria (e.g.in different population);(7)dissemination of our findings to a range of audiences;(8) engaging the attention of NHS managers, policy-makers and other stakeholders, influencing the future implementation of the STOPP/START criteria.

A potential barrier to the success of the project will be the volume and diversity of literature that could be included in the review. At the outset, we recognise that it will be difficult to review all possible programme theories associated with the STOPP/START criteria and we will be selective in the programme theories that are put to review. To ensure the rigour and the relevance of this selection process to the NHS, we have included in our project plan explicit steps to work with our project reference group to inform this selection process. This will enable us to be confident that the programme theories

that we do review are those with the most potential to facilitate the future implementation of the STOPP/START criteria within the NHS.

6. Ethical approval

We have received ethical approval for this study from the University of Exeter Medical School Research Ethics Committee with approval reference Nov18/D/181.

7. Patient and public involvement

We engaged early on the Peninsula Patient and Public Involvement (PenPIG) team in the development of the proposal, and a researcher and PPI facilitator joined the core team (EC). She was involved in drafting the proposal and co-ordinated and facilitated the participation of several lay members of PenPIG (5), who read and commented on early drafts of the project, helped further develop the research idea and shaped how patients would be involved throughout the course of the project. PenPIG consists of 15 members, a number of which are older adults living with multiple health conditions and taking multiple medications. Other members of the group have experience of caring for elderly relatives. The group have vast experience of being involved in research, including projects on polypharmacy and deprescribing. One of them (JH) had a particular interest in the topic and expressed an interest to be further involved. He more extensively participated in the drafting of the proposal and joined the team as a co-applicant.

Patient and public involvement (PPI) will be facilitated by EC through PenPIG and further supported by JH. PPI will be fully integrated into all phases of the project to ensure the research questions and outcomes of interest remain important and aligned to the needs of patients, and will be particularly critical in supporting four project key stages: 1) Further review of the application in light of the feedback received and finalising the protocol, once funding is in place; 2) Discussion with reviewers in relation to the screening criteria; 3) Interpretations of findings from the review; and 4) Dissemination of the findings, including the preparation of the final report. EC will provide support and training for patient representatives, as needed. The PPI co-applicant (JH) will attend all research project team meetings. This will ensure that all discussions throughout the project will be informed by a patient voice.

Reference List

1. Franceschi M, Scarcelli C, Niro V, Seripa D, Pazienza AM, Pepe G, et al. Prevalence, clinical features and avoidability of adverse drug reactions as cause of admission to a geriatric unit: a prospective study of 1756 patients. Drug safety. 2008;31(6):545-56.
2. Kongkaew C, Noyce PR, Ashcroft DM. Hospital admissions associated with adverse drug reactions: a systematic review of prospective observational studies. The Annals of pharmacotherapy. 2008;42(7):1017-25.
3. Buajordet I, Ebbesen J, Erikssen J, Brors O, Hilberg T. Fatal adverse drug events: the paradox of drug treatment. Journal of internal medicine. 2001;250(4):327-41.
4. Alhawassi TM, Krass I, Bajorek BV, Pont LG. A systematic review of the prevalence and risk factors for adverse drug reactions in the elderly in the acute care setting. Clinical interventions in aging. 2014;9:2079-86.
5. Hajjar ER, Cafiero AC, Hanlon JT. Polypharmacy in elderly patients. The American journal of geriatric pharmacotherapy. 2007;5(4):345-51.
6. Goldberg RM, Mabee J, Chan L, Wong S. Drug-drug and drug-disease interactions in the ED: analysis of a high-risk population. The American journal of emergency medicine. 1996;14(5):447-50.
7. Hanlon JT, Pieper CF, Hajjar ER, Sloane RJ, Lindblad CI, Ruby CM, et al. Incidence and predictors of all and preventable adverse drug reactions in frail elderly persons after hospital stay. The journals of gerontology Series A, Biological sciences and medical sciences. 2006;61(5):511-5.
8. Doucet J, Jego A, Noel D, Geffroy CE, Capet C, Coquard A, et al. Preventable and Non-Preventable Risk Factors for Adverse Drug Events Related to Hospital Admissions in the Elderly. Clinical Drug Investigation. 2002;22(6):385-92.
9. Chan M, Nicklason F, Vial JH. Adverse drug events as a cause of hospital admission in the elderly. Internal Medicine Journal. 2001;31(4):199-205.
10. O'Mahony D, Gallagher PF. Inappropriate prescribing in the older population: need for new criteria. Age and ageing. 2008;37(2):138-41.
11. Turner G, Clegg A. Best practice guidelines for the management of frailty: a British Geriatrics Society, Age UK and Royal College of General Practitioners report. Age and ageing. 2014;43(6):744-7.
12. Hill-Taylor B, Walsh KA, Stewart S, Hayden J, Byrne S, Sketris IS. 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. Journal of clinical pharmacy and therapeutics. 2016;41(2):158-69.
13. Orwig D, Rickles NM, Martin LG. Methodological issues in pharmacotherapy research in older adults. The American journal of geriatric pharmacotherapy. 2011;9(3):173-89.
14. Fialova D, Topinkova E, Gambassi G, Finne-Soveri H, Jonsson PV, Carpenter I, et al. Potentially inappropriate medication use among elderly home care patients in Europe. Jama. 2005;293(11):1348-58.
15. van der Hooft CS, Jong GW, Dieleman JP, Verhamme KM, van der Cammen TJ, Stricker BH, et al. Inappropriate drug prescribing in older adults: the updated 2002 Beers criteria--a population-based cohort study. British journal of clinical pharmacology. 2005;60(2):137-44.
16. 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. International journal of clinical pharmacology and therapeutics. 2008;46(2):72-83.
17. Hill-Taylor B, Sketris I, Hayden J, Byrne S, O'Sullivan D, Christie R. Application of the STOPP/START criteria: a systematic review of the prevalence of potentially inappropriate prescribing in older adults, and evidence of clinical, humanistic and economic impact. Journal of clinical pharmacy and therapeutics. 2013;38(5):360-72.
18. Gallagher P, Barry P, O'Mahony D. Inappropriate prescribing in the elderly. Journal of clinical pharmacy and therapeutics. 2007;32(2):113-21.

19. Spinewine A, Schmader KE, Barber N, Hughes C, Lapane KL, Swine C, et al. Appropriate prescribing in elderly people: how well can it be measured and optimised? Lancet (London, England). 2007;370(9582):173-84.
20. Hanley GE, Morgan S, Reid RJ. Explaining prescription drug use and expenditures using the adjusted clinical groups case-mix system in the population of British Columbia, Canada. Medical care. 2010;48(5):402-8.
21. van Mil JW, Westerlund LO, Hersberger KE, Schaefer MA. Drug-related problem classification systems. The Annals of pharmacotherapy. 2004;38(5):859-67.
22. Wahab MS, Nyfort-Hansen K, Kowalski SR. Inappropriate prescribing in hospitalised Australian elderly as determined by the STOPP criteria. Int J Clin Pharm. 2012;34(6):855-62.
23. Cahir C, Fahey T, Teeling M, Teljeur C, Feely J, Bennett K. Potentially inappropriate prescribing and cost outcomes for older people: a national population study. British journal of clinical pharmacology. 2010;69(5):543-52.
24. Lipton HL, Bero LA. Undermedication among geriatric outpatients: results of. Annual review of gerontology and geriatrics. 1992;12:95.
25. Cherubini A, Corsonello A, Lattanzio F. Underprescription of beneficial medicines in older people: causes, consequences and prevention. Drugs & aging. 2012;29(6):463-75.
26. Byrne S, O'Mahony D, Hughes C, Parsons C, Patterson S, McCormack B. An evaluation of the inappropriate prescribing in older residents in long term care facilities in the greater Cork and Northern Ireland regions using the STOPP and Beers' criteria. Centre for Aging Research and Development in Ireland. 2011.
27. Hanlon JT, Schmader KE, Ruby CM, Weinberger M. Suboptimal prescribing in older inpatients and outpatients. Journal of the American Geriatrics Society. 2001;49(2):200-9.
28. Tangiisuran B, Wright J, Van der Cammen T, Rajkumar C. Adverse drug reactions in elderly: challenges in identification and improving preventative strategies. Age and ageing. 2009;38(4):358-9.
29. Garfinkel D, Mangin D. Feasibility study of a systematic approach for discontinuation of multiple medications in older adults: addressing polypharmacy. Archives of internal medicine. 2010;170(18):1648-54.
30. NICE. Medicines optimisation: the safe and effective use of medicines to enable the best possible outcomes 2015 2015 [Available from: https://www.nice.org.uk/guidance/ng5/chapter/recommendations#terms-used-in-this-guideline.
31. Patterson SM, Cadogan CA, Kerse N, Cardwell CR, Bradley MC, Ryan C, et al. Interventions to improve the appropriate use of polypharmacy for older people. The Cochrane database of systematic reviews. 2014(10):Cd008165.
32. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. The Milbank quarterly. 2004;82(4):581-629.
33. Wells M, Williams B, Treweek S, Coyle J, Taylor J. Intervention description is not enough: evidence from an in-depth multiple case study on the untold role and impact of context in randomised controlled trials of seven complex interventions. Trials. 2012;13(1):95.
34. Pawson R. Nothing as practical as a good theory. Evaluation. 2003;9(4):471-90.
35. Dalkin SM, Greenhalgh J, Jones D, Cunningham B, Lhussier M. What's in a mechanism? Development of a key concept in realist evaluation. Implementation Science. 2015;10(1):49.
36. Wong G, Westhorp G, Manzano A, Greenhalgh J, Jagosh J, Greenhalgh T. RAMESES II reporting standards for realist evaluations. BMC medicine. 2016;14(1):96.
37. Pawson R. The science of evaluation: a realist manifesto: Sage; 2013.
38. Manzano A. The craft of interviewing in realist evaluation. Evaluation. 2016;22(3):342-60.
39. England NHS. Toolkit for general practice in supporting older people living with frailty. NHS England, London[Google Scholar]. 2017.
40. NICE. Membership of Multimorbidity Guideline Committee 2016 2016 [Available from: https://www.nice.org.uk/guidance/ng56/documents/committee-member-list-2. .

41. Danermark B, Ekstrom M, Jakobsen L. Explaining society: An introduction to critical realism in the social sciences: Routledge; 2005.
42. Evers S, Goossens M, De Vet H, Van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. International journal of technology assessment in health care. 2005;21(2):240-5.
43. Anderson R, Hardwick R. Realism and resources: Towards more explanatory economic evaluation. Evaluation. 2016;22(3):323-41.
44. Anderson R, hardwick R, Pearson M, Byng. Using realist approaches to explain the costs and cost-effectiveness of programmes. Chapter 8 in: Emmel N, Greenhalgh J, Manzano A, Monaghan M, and Dalkin S, eds. Doing Realist Research. London: Sage 2018.
45. Dalleur O, Feron JM, Spinewine A. Views of general practitioners on the use of STOPP&START in primary care: a qualitative study. Acta Clinica Belgica. 2014;69(4):251-61.
46. Clyne B, Cooper JA, Hughes CM, Fahey T, Smith SM. 'Potentially inappropriate or specifically appropriate?' Qualitative evaluation of general practitioners views on prescribing, polypharmacy and potentially inappropriate prescribing in older people. BMC Family Practice. 2016;17(1):109.
47. Topinková E, Baeyens JP, Michel J-P, Lang P-O. Evidence-based strategies for the optimization of pharmacotherapy in older people. Drugs & aging. 2012;29(6):477-94.
48. Spinewine A, Swine C, Dhillon S, Franklin BD, Tulkens PM, Wilmotte L, et al. Appropriateness of use of medicines in elderly inpatients: qualitative study. BMJ. 2005;331(7522):935.
49. Reeve E, To J, Hendrix I, Shakib S, Roberts MS, Wiese MD. Patient Barriers to and Enablers of Deprescribing: a Systematic Review. Drugs & Aging. 2013;30(10):793-807.
50. Shade MY, Berger AM, Chaperon C. Potentially inappropriate medications in community-dwelling older adults. Research in gerontological nursing. 2014;7(4):178-92.
51. Barry PJ, Gallagher P, Ryan C, O'Mahony D. START (screening tool to alert doctors to the right treatment)—an evidence-based screening tool to detect prescribing omissions in elderly patients. Age and ageing. 2007;36(6):632-8.
52. Salgueiro E, Elizarde BC, Elola AI, Garcia-Pulido B, Nicieza-Garcia ML, Manso G. [The most common STOPP/START criteria in Spain. A review of the literature]. Revista espanola de geriatria y gerontologia. 2018.
53. Mori ALPM, Carvalho RC, Aguiar PM, de Lima MGF, Rossi MdSPN, Carrillo JFS, et al. Potentially inappropriate prescribing and associated factors in elderly patients at hospital discharge in Brazil: a cross-sectional study. International Journal of Clinical Pharmacy. 2017;39(2):386-93.
54. Cooper JA, Cadogan CA, Patterson SM, Kerse N, Bradley MC, Ryan C, et al. Interventions to improve the appropriate use of polypharmacy in older people: a Cochrane systematic review. BMJ Open. 2015;5(12).
55. Kastner M, Estey E, Perrier L, Graham ID, Grimshaw J, Straus SE, et al. Understanding the relationship between the perceived characteristics of clinical practice guidelines and their uptake: protocol for a realist review. Implementation science : IS. 2011;6:69.
56. Pearson M, Chilton R, Woods HB, Wyatt K, Ford T, Abraham C, et al. Implementing health promotion in schools: protocol for a realist systematic review of research and experience in the United Kingdom (UK). Systematic Reviews. 2012;1:48-.
57. Pearson M, Hunt H, Cooper C, et al. Intermediate care: a realist review and conceptula framework. Final Report. NIHR Service Delivery and Organisation Programme, 2012.
58. Pawson R: Evidence Based Policy: A Realist Perspective. London: Sage; 2006.
59. Drummond M. Evidence-based decisions and economics: an agenda for research. Evidence-Based Decisions and Economics: Health Care, Social Welfare, Education and Criminal Justice. 2010:181-2.
60. Payne K, Thompson A. Economic evaluations of complex interventions. Complex interventions in health: An overview of research methods Abingdon, Oxfordshire: Routledge. 2015.

Appendix 1: Example 'Theory' Search strategy for first phase of the review

Database: Ovid MEDLINE(R) <2007 to 2017> Search Strategy:

#1. Polypharmacy.tw.

#2. exp polypharmacy/

#3. ((inappropriate* or suboptim* or sub-optim* or unnecessary or incorrect* or excess* or multip* or concurrent* or inadvert*) adj2 (medici* or medicat* or prescrib* or prescription* or drug*)).ti,ab.

#4. ((over adj1 (prescrib* or prescript*)) or (over-prescrib* or overprescrib*) or ("or more" adj (medication* or prescrib* or prescript*))).ti,ab.

#5. ((under adj1 prescrib*) or underprescrib* or under-prescrib*).ti,ab.

#6. ((multi-drug* or multidrug*) adj2 (prescrib* or prescription* or regimen? or therap* or treatment?)).ti,ab.

#7. 1 or 2 or 3 or 4 or 5 or 6

#8. (elder* or geriatric*).ti,ab.

#9. exp geriatrics/

#10. 8 or 9

#11. STOPP?.tw.

#12. (stopp* adj2 criter*).tw.

#13. (start* adj2 criter*).tw.

#14. ((stop* and start*) adj2 criteri*).tw.

#15. "Screening Tool of Older Persons' Potentially Inappropriate Prescriptions".tw.

#16. "Screening Tool to Alert Doctors to Right Treatment".tw.

#17. 11 or 12 or 13 or 14 or 15 or 16

#18. 7 and 10 and 17