STOPP/START interventions to improve medicines management for people aged 65 years and over: a realist synthesis

Jaheeda Gangannagaripalli, Ian Porter, Antoinette Davey, Ignacio Ricci Cabello, Joanne Greenhalgh, Rob Anderson, Simon Briscoe, Carmel Hughes, Rupert Payne, Emma Cockcroft, Jim Harris, Charlotte Bramwell and Jose M Valderas

STOPP/START interventions to improve medicines management for people aged 65 years and over: a realist synthesis

Jaheeda Gangannagaripalli®,[1] Ian Porter,[1]
Antoinette Davey®,[1] Ignacio Ricci Cabello®,[2]
Joanne Greenhalgh®,[3] Rob Anderson®,[1,4]
Simon Briscoe®,[1] Carmel Hughes®,[5] Rupert Payne®,[6]
Emma Cockcroft®,[7] Jim Harris,[1] Charlotte Bramwell®[1]
and Jose M Valderas®[1]*

[1]Health Services and Policy Research Group, University of Exeter Medical School, Exeter, UK
[2]Gerència d'Atenció Primària de Mallorca, Fundació Institut d'Investigació Sanitària Illes Balears – IdISBa, Spain
[3]School of Sociology and Social Policy, University of Leeds, Leeds, UK
[4]Evidence Synthesis & Modelling for Health Improvement (ESMI) Research Group, University of Exeter Medical School, Exeter, UK
[5]School of Pharmacy, Queen's University Belfast, Belfast, UK
[6]Population Health Sciences, University of Bristol, Bristol Medical School, Bristol, UK
[7]National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care for the South West Peninsula, University of Exeter Medical School, Exeter, UK

*Corresponding author

Health Services and Delivery Research

Criteria for inclusion in the *Health Services and Delivery Research* journal
Reports are published in *Health Services and Delivery Research* (HS&DR) if (1) they have resulted from work for the HS&DR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

HS&DR programme

The HS&DR programme funds research to produce evidence to impact on the quality, accessibility and organisation of health and social care services. This includes evaluations of how the NHS and social care might improve delivery of services.

For more information about the HS&DR programme please visit the website at https://www.nihr.ac.uk/explore-nihr/funding-programmes/health-services-and-delivery-research.htm

This report

Abstract

STOPP/START interventions to improve medicines management for people aged 65 years and over: a realist synthesis

Jaheeda Gangannagaripalli[1], Ian Porter,[1] Antoinette Davey[1],
Ignacio Ricci Cabello[2], Joanne Greenhalgh[3], Rob Anderson[1,4],
Simon Briscoe[1], Carmel Hughes[5], Rupert Payne[6],
Emma Cockcroft[7], Jim Harris,[1] Charlotte Bramwell[1]
and Jose M Valderas[1*]

[1]Health Services and Policy Research Group, University of Exeter Medical School, Exeter, UK
[2]Gerència d'Atenció Primària de Mallorca, Fundació Institut d'Investigació Sanitària Illes Balears –
 IdISBa, Spain
[3]School of Sociology and Social Policy, University of Leeds, Leeds, UK
[4]Evidence Synthesis & Modelling for Health Improvement (ESMI) Research Group, University of Exeter
 Medical School, Exeter, UK
[5]School of Pharmacy, Queen's University Belfast, Belfast, UK
[6]Population Health Sciences, University of Bristol, Bristol Medical School, Bristol, UK
[7]National Institute for Health Research Collaboration for Leadership in Applied Health Research and
 Care for the South West Peninsula, University of Exeter Medical School, Exeter, UK

*Corresponding author j.m.valderas@exeter.ac.uk

Background: Drug-related problems and potentially inappropriate prescribing impose a huge burden on patients and the health-care system. The most widely used tools for appropriate prescription in older adults in England and in other European countries are the Screening Tool of Older People's Prescriptions (STOPP)/Screening Tool to Alert to the Right Treatment (START) tools. STOPP/START tools support medicines optimisation for older adults.

Objectives: To identify, test and refine the programme theories underlying how interventions based on the STOPP/START tools are intended to work, for whom, in what circumstances and why, as well as the resource use and cost requirements or impacts.

Design: A realist synthesis.

Setting: Primary care, hospital care and nursing homes.

Patients: Patients aged ≥ 65 years.

Interventions: Any intervention based on the use of the STOPP/START tools.

Review methods: Database and web-searching was carried out to retrieve relevant evidence to identify and test programme theories about how interventions based on the use of the STOPP/START tools work. A project reference group made up of health-care professionals, NHS decision-makers, older people, carers and members of the public was set up. In phase 1 we identified programme theories about STOPP/START interventions on how, for whom, in what contexts and why they are intended to work. We searched the peer-reviewed and grey literature to identify documents

relevant to the research questions. We interviewed experts in the field in our reference group to gain input on our list of candidate context–mechanism–outcome configurations, to identify additional context–mechanism–outcome configurations and to identify additional literature and/or relevant concepts. In phase 2 we reviewed and synthesised relevant published and unpublished empirical evidence and tested the programme theories using evidence from a larger set of empirical studies.

Results: We developed a single logic model structured around three key mechanisms: (1) personalisation, (2) systematisation and (3) evidence implementation. Personalisation: STOPP/START-based interventions are based on shared decision-making, taking into account patient preferences, experiences and expectations (mechanisms), leading to increased patient awareness, adherence, satisfaction, empowerment and quality of life (outcomes). Systematisation: STOPP/START tools provide a standardised/systematic approach for medication reviews (mechanisms), leading to changes in professional and organisational culture and burden/costs (outcomes). Evidence implementation: delivery of STOPP/START-based interventions is based on the implementation of best evidence (mechanisms), reducing adverse outcomes through appropriate prescribing/deprescribing (outcomes). For theory testing, we identified 40 studies of the impact of STOPP/START-based interventions in hospital settings, nursing homes, primary care and community pharmacies. Most of the interventions used multiple mechanisms. We found support for the impact of the personalisation and evidence implementation mechanisms on selected outcome variables, but similar impact was achieved by interventions not relying on these mechanisms. We also observed that the impact of interventions was linked to the proximity of the selected outcomes to the intervention in the logic model, resulting in a clearer benefit for appropriateness of prescribing, adverse drug events and prescription costs.

Limitations: None of the available studies had been explicitly designed for evaluating underlying causal mechanisms, and qualitative information was sparse.

Conclusions: No particular configuration of the interventions is associated with a greater likelihood of improved outcomes in given settings.

Study registration: This study is registered as PROSPERO CRD42018110795.

Funding: This project was funded by the National Institute for Health Research (NIHR) Health Services and Delivery Research programme and will be published in full in *Health Services and Delivery Research*; Vol. 9, No. 23. See the NIHR Journals Library website for further project information.

Contents

CONTENTS

x

List of tables

List of figures

List of boxes

Glossary

Acute care A branch of secondary health care where a patient receives active but short-term treatment for a severe injury or episode of illness or an urgent medical condition or during recovery from surgery.

Adverse drug event An injury resulting from the use of a drug.

Adverse drug reactions Responses to drugs that are noxious and unintended, and which occur at doses normally used in humans for prophylaxis, diagnosis or therapy of disease, or for the modification of physiological function. Includes events due to treatment failures and medication errors.

Age UK A registered charity in the UK formed on 25 February 2009 and launched on 1 April 2009, which combined the operations of the previously separate charities Age Concern England and Help the Aged to form the UK's largest charity for older people.

Beers criteria Standard guidance for clinicians to prevent potentially inappropriate prescribing in patients aged ≥ 65 years.

British Geriatrics Society A professional body of specialists in the health care of older people in the UK.

British National Formulary A pharmaceutical reference in the UK that contains a wide spectrum of information and advice on prescribing and pharmacology, along with specific facts and details about many medicines available on the UK NHS.

Comorbidity The existence of more than one illness or disease in one person at the same time.

Deprescribing Planned and supervised process of dose reduction or stopping of medication that might be causing harm, or no longer be of benefit.

Dosage regimen A schedule of doses of a medicine, including the time between doses, the duration of treatment and the amount to be taken each time.

Drug regimen A prescribed systematic form of treatment for a course of drug(s).

Drug-related problem An event or circumstance involving drug therapy that actually or potentially interferes with desired health outcomes.

Implementation (of evidence) A mechanism by which some Screening Tool of Older People's Prescriptions/Screening Tool to Alert to the Right Treatment-based interventions aim to achieve its impact, which captures the degree to which an intervention oriented to support the medication review focuses on promoting the dissemination and use of the best available evidence.

Inappropriate prescribing Prescription of medications where risk outweighs the benefit; failure to use a safer alternative drug; the misuse of a drug, including incorrect dosage and duration of treatment; use of drugs with significant drug–drug and drug–disease interactions; and, finally, the omission of beneficial drugs.

Incentivisation The practice of building incentives into an arrangement or system to motivate the actors within it.

Life expectancy A statistical measurement of how long a person is expected to live.

Medication error Any preventable event that may cause or lead to inappropriate medication use or patient harm while the medication is in the control of the health-care professional, patient or consumer. Such events may be related to professional practice, health-care products, procedures and systems, including prescribing, order communication, product labelling, packaging, nomenclature, compounding, dispensing, distribution, administration, education, monitoring and use.

Medication review A structured evaluation of a patient's medicines, with the aim of optimising medicines use and improving health outcomes.

Medicines optimisation A person-centred approach to safe and effective medicines use to ensure people obtain the best possible outcomes from their medicines. Medicines optimisation applies to people who may or may not take their medicines effectively.

Morbidity Consequences and complications (other than death) that result from a disease.

Mortality The measure of the number of deaths in a particular population.

Patients and public involvement The involvement of patients and the public in the design, management and conduct of research or the design of health services.

Personalisation (of the medication review) A mechanism by which some Screening Tool of Older People's Prescriptions/Screening Tool to Alert to the Right Treatment-based interventions aim to achieve its impact, which captures the degree to which the medication review process and outcomes are tailored to each individual patient.

Pharmacodynamics A response of the body to a drug.

Pharmacokinetics The study of the bodily absorption, distribution, metabolism and excretion of drugs.

Pharmacotherapy The use of medicine in the treatment of diseases, conditions and symptoms.

Polypharmacy Concurrent use of multiple medication items by one individual.

Potentially inappropriate prescribing Prescriptions that introduce a significant risk of an adverse drug-related event when there is evidence for an equally or more effective alternative medication.

Prescribing errors Errors that may be defined as the incorrect drug selection for a patient. Such errors can include errors of dose, quantity or indication or the prescribing of a contraindicated drug.

Screening tool A tool aimed at identifying who may have a high risk or likelihood of a particular problem, disease or condition.

Systematisation (of the medication review process) A mechanism by which some Screening Tool of Older People's Prescriptions/Screening Tool to Alert to the Right Treatment-based interventions aim to achieve its impact, which captures the degree to which the medication review is formalised for systematic implementation.

List of abbreviations

ADE	adverse drug event
ADR	adverse drug reaction
BGS	British Geriatrics Society
CDSR	Cochrane Database of Systematic Reviews
CENTRAL	Cochrane Central Register of Controlled Trials
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CMO	context–mechanism–outcome
GP	general practitioner
HMIC	Health Management Information Consortium
NHS EED	NHS Economic Evaluation Database
NICE	National Institute for Health and Care Excellence
PPI	patient and public involvement
QALY	quality-adjusted life-year
RAMESES	Realist And Meta-narrative Evidence Syntheses: Evolving Standards
RCT	randomised controlled trial
RPS	Royal Pharmaceutical Society
START	Screening Tool to Alert to the Right Treatment
STOPP	Screening Tool of Older People's Prescriptions

Plain English summary

Background

The Screening Tool of Older People's Prescriptions (STOPP) and Screening Tool to Alert to the Right Treatment (START) are the most widely used approaches for ensuring suitable prescribing of medicine in older adults in Europe. They help support medication reviews and give a set of suggestions in terms of reducing medication burden (STOPP) and adding potentially beneficial therapy (START). Despite wide use, we do not fully understand how the approach works, who it works for and whether or not there are particular situations where this approach works best.

Methods

In the first phase, we identified some ideas about how the STOPP/START approach is intended to work. These ideas were based on reports, existing research and interviews with clinicians who use STOPP/START tools, and other experts. In the second phase, we looked at existing research to support these ideas. A Patient Advisory Group helped to shape the conduct of the research project, analysis and interpretation of our findings.

Results

We developed ideas about how the approach worked, including that the reviews of medicines were personalised, evidence based and had a systematic approach. In the search for existing evidence, we identified 40 relevant studies that showed the following:

- The STOPP/START approach was used differently in different settings (such as hospitals and general practices).
- STOPP/START tools were used as an evidence-based approach to ensure that medicine use was appropriate. Many studies demonstrated a positive impact.
- There was limited evidence of the use of STOPP/START tools leading to personalised medicines reviews.

Conclusion

Medicines reviews using the STOPP/START approach are used to ensure that prescribing is based on available evidence. However, they are not generally personalised and do not ensure that prescribing is tailored to the circumstances of each patient.

Scientific summary

Background

Drug-related problems and potentially inappropriate prescribing can lead to adverse drug events, hospitalisation and death, thereby imposing a huge burden on patients and the health-care system. Discontinuation of inappropriate medications through early detection can trigger a cascade of positive effects from the improvement of care quality and reduction of adverse drug events to the improvement of the cost-effectiveness of pharmacological therapeutic plans and quality of life of older people. The most widely used tools for appropriate prescription in older adults in England and in other European countries are the Screening Tool of Older People's Prescriptions (STOPP)/Screening Tool to Alert to the Right Treatment (START) tools, which have been specifically developed for supporting medicines optimisation for older adults. A deep understanding of how interventions based on the use of these tools work, for whom they work, in what contexts and why is currently lacking. This study aimed to use a realist approach to synthesise the evidence on the interventions based on STOPP/START tools.

Aims and objectives

Our aim was to understand how, when and why interventions based on the STOPP/START tools improve medicines management in older people. Our objectives were as follows:

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

Methods

Searches were conducted in bibliographic databases, websites and the sets of citations of included studies to identify and test programme theories about how interventions based on use of the STOPP/START tools work. In phase 1 we identified programme theories that explain how interventions were supposed to work and in what circumstances to capture in detail the reasoning that underlies these interventions. Subsequently, in phase 2, we reviewed the empirical evidence to determine the extent to which these expectations were met in practice. A project reference group comprising health-care professionals, NHS decision-makers, older people, carers and members of the public was set up to ground the study in real-life experience. In phase 1 we identified programme theories about STOPP/START-based interventions on how, for whom, in what contexts and why they are intended to work, and whether or not patients are being involved in shared decision-making in stopping or starting medicines. We conducted electronic searches of grey literature to identify generic guidance and policy documents, and electronic searches of the peer-reviewed literature to identify position pieces, comments, letters, editorials and critical pieces relevant to the research questions, as well as reports of studies using the tools. Interviews with experts in the field in our reference group contributed to the identification and refinements of programme theories. In phase 2 we tested the programme theories. We reviewed and synthesised relevant published and unpublished empirical evidence.

Results

We developed a single logic model structured around three key mechanisms (embedded within specific sets of context–mechanism–outcome configurations) through which STOP/START tools were thought to work: (1) personalisation (i.e. the degree to which the medication review is tailored to each patient), (2) systematisation (i.e. the degree to which the medication review is systematically implemented) and (3) evidence implementation (i.e. the degree to which the intervention focuses on promoting the dissemination and use of the best available evidence). In addition, we identified mechanism-specific outcomes and prioritised specific theories, as supported by our patient and public involvement advisory team.

Forty studies were identified for testing the programme theory. These studies evaluated the impact of STOPP/START-based interventions, mostly in hospital settings but also in nursing homes and primary care. None of these studies was primarily oriented to study the pathways to impact for the interventions. Most of the interventions were designed to target multiple mechanisms. No distinct pattern of configuration of mechanisms was identified as specific of interventions for each type of setting. In hospitals and nursing homes, the interventions usually involved both a pharmacist and a specialist doctor, whereas in primary care settings, only a general practitioner was involved. Most interventions did not involve patients and carers in the review process.

We found evidence providing some support for the impact of the personalisation mechanism of STOPP/START-based interventions on the specific outcomes of patient satisfaction and adherence. We could not test the prioritised theories for this mechanism because of the lack of studies reporting on relevant aspects.

We did not find any evidence providing support for an impact of the systematisation mechanism of STOPP/START-based interventions on any of the specific outcomes in the logic models. Lack of studies reporting on relevant aspects also prevented testing of the prioritised theories for this mechanism.

We found evidence providing support for an impact of the evidence implementation mechanism of STOPP/START-based interventions on appropriateness of prescribing/deprescribing, adverse drug events and quality of life, some support for a reduction in falls, and consistent evidence on the lack of impact on health-care use and mortality. We also found evidence in support of two of the prioritised theories for this mechanism, and were able to confirm the theory that reduction of adverse outcomes was the result of improvement in medication appropriateness and to refute the theory that administrations by non-geriatricians and non-clinical pharmacists was linked with negative results. We could not test other theories because of the lack of evidence.

We observed that the impact of interventions was linked to the proximity of the selected outcomes to the intervention in the single logic model (i.e. more clear benefit on appropriateness of prescribing, adverse drug events and prescription costs), but that the impact was not related to what mechanisms had been targeted, nor to the setting. We did not find any evidence that studies targeting more mechanisms achieved better outcomes than those targeting fewer, nor was any other configuration linked with better outcomes.

Owing to the nature of the available evidence, it was not possible to conduct a realist economic synthesis. We did, however, identify the following key drivers of costs:

- time
- pay rates and additional training required for health professionals to use and interpret STOPP/START tools
- costs of information technology systems or software
- the complexity and comprehensiveness of the decision-making process

- how many medications per patient are inappropriate/harmful/unnecessary compared with how many medications per patient are indicated/omitted/needed and currently not being taken
- the cost of these medications
- the rates of adherence to prescribed or recommended medicines, and the benefits of those medications (harms/adverse reactions avoided and health improvements)
- the related cost savings/impacts linked to these.

In addition, we determined the potential cost savings and impacts linked to these medication review-based interventions.

One intervention looked particularly promising and relevant to the UK context. In this intervention, patients were invited by the community pharmacy team to hold regular consultations with a pharmacist to review their medication using STOPP/START tools and to discuss risk of falls, pain management, adherence and general health. The design of the intervention targeted all three proposed mechanisms and found positive evidence for improved adherence to medical treatment, improved quality of life and a reduction in falls, with costs per quality-adjusted life-year estimates that ranged up to £32,466.

Conclusions

Current evidence on the impact of these interventions does not support the superiority of any particular configuration of the interventions for any given setting. Future studies aiming to uncover the underlying mechanisms can contribute to designing potentially more effective configurations of STOPP/START interventions.

Study registration

This study is registered as PROSPERO CRD42018110795.

Funding

Chapter 1 Introduction

Background

Up to 30% of hospital admissions of older people are associated with drug-related problems resulting from adverse drug reactions (ADRs).[1–3] The concurrent use of multiple medications (polypharmacy) is associated with both 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 because older patients with multiple morbidities are often excluded from clinical trials.[13]

Potentially inappropriate medication in older adults

The prescription of potentially inappropriate medications to older people is highly prevalent, ranging from 12% in community-dwelling elderly people to 40% in nursing home residents in Europe and the USA.[14] Older people are particularly vulnerable to inappropriate prescribing because of their multiple drug regimens, comorbid conditions and age-associated physiological processes.[15] Drug-related problems and potentially inappropriate prescribing are highly prevalent among older adults and can exert a significant disease burden. In addition, they have been associated with adverse drug events (ADEs) (which includes ADRs and medication errors), leading to hospitalisation and death, and increased health resource utilisation.[16–18] A 2007 national population study in Ireland by Cahir *et al.*[19] estimated a cost of potentially inappropriate prescribing of > €38M.

Medicines in older people are considered appropriate when they have a clear evidence-based indication, are cost-effective and are well tolerated in the majority of the population. Although potentially inappropriate prescribing and ADRs are generally associated with overprescription, potentially inappropriate prescribing may also occur when a patient is not prescribed appropriate medication for the treatment and prevention of a disease or condition.[20] This might occur for many reasons, such as ageism, fear of adverse events, economic concerns and lack of prescribing knowledge.[21–23]

Screening tools for medication optimisation in older adults

Detecting potentially inappropriate prescribing early may prevent ADEs and cost-effectiveness of medicines, which gives an opportunity for improving quality of care in older adults.[10,24] In addition, the quality of life of older people can be improved by discontinuing inappropriate medications.[25] 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 lists of appropriate medications for older people to predominantly avoid harmful prescriptions in the older population.

Adverse drug events and cost inefficiencies can be reduced by identifying potentially inappropriate prescribing.[10] Screening tools for supporting medicines optimisation of older adults have been developed in an attempt to address this. The most widely used are the Beers criteria for potentially inappropriate medication use in older adults and the combination of the Screening Tool of Older People's Prescriptions (STOPP) and Screening Tool to Alert to the Right Treatment (START). Almost half of the drugs in the Beers criteria are unavailable for prescribers in Europe,[26,27] several of the drugs are not contraindicated in older people, as per the *British National Formulary*,[10] whereas other

contraindicated drugs are omitted.[26] Given these shortcomings, the STOPP/START tools have recently received increased interest in European countries.[28]

Development of STOPP/START tools

Two versions of STOPP/START were developed by a team of researchers at the University of Cork, Ireland. The first iteration. STOPP/START (version 1), was developed in 2008.[28] They are valid, reliable and comprehensive screening tools that enable the prescribing physician to appraise an older patient's prescription drugs in the context of his/her concurrent diagnoses. To establish their content validity, a two-round Delphi survey was conducted with an 18-member expert panel in geriatric pharmacotherapy from academic centres in Ireland and the UK. Sixty-five of the 68 original STOPP criteria and all 22 of the original START criteria received consensus from the expert panel. Inter-rater reliability of the STOPP/START tools was also assessed by determining the kappa statistic and high levels were established between physicians and pharmacists.

Owing to the changing evidence base behind the initial version of STOPP/START tools, the licensing of important new drugs since 2008 and a long list of potentially inappropriate medications included in the first version, an updated version of the tools was deemed necessary. In addition, a number of STOPP/START criteria were no longer considered completely accurate or relevant, a small number (12 in total) of criteria were lacking in clinical importance or prevalence and there were some important criteria that were absent from STOPP/START version 1. The updated version (i.e. version 2) was developed based on an up-to-date literature review and consensus validation among a European panel of experts to reflect Europe-wide prescribing practices in the general population of older people (see *Appendix 1*).

Impact of STOPP/START tools

An updated recent Cochrane systematic review of the evidence base for interventions to improve the appropriate use of polypharmacy for older people concluded that it was unclear whether or not they could improve appropriate polypharmacy.[29] Although previous versions of the review highlighted STOPP/START as the basis for promising interventions,[30] no specific examination was conducted of STOPP/START-based interventions.

A systematic review based on 13 randomised and non-randomised studies that was conducted up to 2012 using STOPP/START tools in community-dwelling, acute care and long-term care older patients found that its use reduced falls, hospital length of stay, care visits and medication cost.[31] A more recent systematic review of randomised clinical trials was conducted on the effectiveness of the STOPP/START tools.[12] This review identified four randomised controlled trials (RCTs). Two studies were judged as having a low risk of bias and two studies were judged as having a moderate to high risk of bias. Meta-analysis of the four RCTs found that the STOPP tool reduced potentially inappropriate medications rates, but study heterogeneity prevented the calculation of a meaningful statistical summary. There was evidence that using STOPP/START tools was associated with a reduction in falls, delirium episodes, hospital length of stay, care visits (primary and emergency) and medication costs, but there was no evidence of improvements in quality of life or mortality. The authors concluded that their use may be effective in improving prescribing quality, and in improving clinical, humanistic and economic outcomes.[12]

Current National Institute for Health and Care Excellence (NICE) guidance on medicines optimisation includes the recommendation to consider using a screening tool for adults such as the STOPP/START to identify potential medicines related patient safety incidents in some groups of people, especially in relation to adults taking multiple medicines (polypharmacy), adults with chronic or long-term conditions and older people.[32] 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 computerised systems. Even this particular application of the decision tools would benefit from knowledge on the mechanisms by which such interventions work more generally to optimise their implementation and to maximise their benefits.

Rationale for realist methodology

To date, and to the best of our knowledge, no attempt has been made to appraise evidence to explain how, when and why intervention based on the STOPP/START tools might improve medicines management in older people.

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. The current evidence base for the STOPP/START tool, as reviewed by conventional methods, is limited for all types of population (especially frail older and community-living patients receiving primary care). Both a Cochrane review[29] and a systematic review by Greenhalgh *et al.*[33] demonstrates the importance of eliciting with sufficient detail the mechanisms that lead to success of interventions and the context in which those interventions are conducted to appraise and eventually maximise their transferability.[33,34] Rather than simply assessing whether interventions based on STOPP/START tools are effective or cost-effective in a binary sense, review methods seek to explain how and why these 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 regimens, is a priority for the NHS.

We can start to investigate the patterns in realist synthesis. 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 is complementary to the two previous reviews,[29,33] which did not have a realist approach. The review provides greater understanding and insight into how, when and why interventions work in practice, and the differences between settings, clinician types, patient groups, patient ages and shared decision-making, all of which are very important in terms of optimising implementation of the STOPP/START tools. We have also broadened our search to include studies from 2014 to 2019 and searched for evidence from non-RCTs and qualitative studies, unlike the other two reviews. In addition, we extracted the information from the relevant reports from the NHS.

In contrast to 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:

- Evidence about the supposed (i.e. programme theory) and actual costs and cost-effectiveness of interventions and comparators should be sought and captured in a wider range of study types, including economic evaluations, costing studies, effectiveness studies and other study types about the interventions or programmes of interest.
- It should be 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 ADEs and 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[35] and more recent clarifications of the realist notion of mechanism[36] clearly acknowledge that intervention mechanisms are fundamentally how people, providers and 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.

Chapter 2 Aims and objectives

Aim

The overall aim of this synthesis was to understand how, when and why the use of the STOPP/START tools improves medicines management in older people.

Objectives

The objectives were as follows:

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

Chapter 3 Methods

A realist synthesis was conducted following Realist And Meta-narrative Evidence Syntheses: Evolving Standards (RAMESES II) guidelines for realist synthesis.[37] The approach was chosen to inform the implementation of initiatives using STOPP/START tools, as it recognises that interventions are not universally successful and that outcomes are context dependent.[35] Realist synthesis is a theory-based approach that seeks to explain how context shapes the mechanisms through which programmes produce intended and unintended outcomes.[35] This is accomplished through the identification and refinement of underlying programme theories. These seek to explain how the generative mechanisms triggered by the resources offered by the intervention or programme are shaped by conditions or contexts, and the pattern of outcomes produced. In this case, context refers to the ways in which the intervention is designed and the circumstances in which the intervention is implemented, which may either support or constrain how participants respond to the intervention, whereas outcome refers to both the intended and unintended effects of the intervention on participants.

The study was registered in the International prospective register of systematic reviews PROSPERO 2018 CRD42018110795: Jaheeda Gangannagaripalli, Joanne Greenhalgh, Rob Anderson, Ian Porter, Simon Briscoe, Emma Cockcroft, Carmel Hughes, Rupert Payne, Jim Harris, Ignacio Ricci-Cabello, Jose M Valderas. How, when and why do STOPP/START tools-based interventions improve medicines management for older people: a realist synthesis. URL: www.crd.york.ac.uk/prospero/display_record. php?ID=CRD42018110795.

Phase 1: identification of programme theories

In this phase we identified programme theories or ideas and assumptions about how the STOPP/ START tools are intended to work, for whom, in what contexts, why (e.g. ideas about what drug groups and/or conditions these tools applied to, and why they do not work and in what patient population) and whether or not (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 did not actively seek theoretical mechanisms that exclusively explain the cost or cost-effectiveness of STOPP/START-based interventions, but rather sought to identify and make explicit which underlying CMOs are resource-consuming or cost-generating, in what ways and to what extent. The aim of this first phase of realist synthesis was to capture in detail the reasoning that underlies the intended outcomes of an intervention.

In summary, we produced a longlist of programme theories with CMO configurations (from peer-reviewed and grey literature, interviews with experts and PPI consultations) that were refined to a single logic model structured around three key mechanisms (embedded within specific sets of CMO configurations) through which STOPP/START tools were expected to work (i.e. personalisation, systematisation and evidence implementation).

Search strategy

We carried out searches of relevant bibliographic databases to identify theories and assumptions about how the STOPP/START tools are intended to work. The search aimed to identify position pieces, commentaries, letters, editorials and critical pieces on the STOPP/START tools wherein theories and assumptions about the tool would be discussed, but not necessarily tested. The search strategy was developed using MEDLINE (via Ovid) by an information specialist (SB) in consultation with the review team. A combination of indexing terms (e.g. medical subject headings in MEDLINE) and free-text terms (i.e. title and abstract terms in the bibliographic records) were used. We included search terms that

described the STOPP/START tools combined with a published publication-type search filter that included search terms such as 'comment', 'letter', 'opinion', 'views' and 'news or newspaper article'.[38] A publication date limit of 2008 to the date of the search was applied. No English-language filter was applied.

The search strategy was finalised in MEDLINE and translated for use in a selection of relevant bibliographic databases, including:

- Cumulative Index to Nursing and Allied Health Literature (CINAHL) (via EBSCO*host*)
- EMBASE (via Ovid)
- Health Management information Consortium (HMIC) database (via Ovid).

The search strategies for all databases and the number of hits retrieved per database are reported in *Appendix 2, Tables 8–11*. Searches of all databases were carried out on 29 October 2018. The results were exported to EndNote X8 (Clarivate Analytics, Philadelphia, PA, USA) and de-duplicated using the automated de-duplication feature and manual checking.

Screening

Documents were selected based on relevance (i.e. whether or not data could contribute to theory building). A random sample of 10% of articles were selected, read, assessed and discussed by three reviewers (JBG, JG and JMV) using a preliminary set of inclusion/exclusion criteria (*Box 1*).

The remaining 90% of articles were completed by the same reviewers independently. However, a number of these necessitated discussion between the reviewers, as they were 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 CMOs.)[39] We therefore refined our studies based on the knowledge we acquired about the theory development of the impact of the STOPP/START tools. At the end of this process, we retained the papers that provided the clearest examples of the ideas underpinning STOPP/START-based interventions.

BOX 1 Phase 1: preliminary set of inclusion/exclusion criteria

Inclusion criteria (any criterion could be met)

- The paper describes how STOPP/START tools and/or STOPP/START tools-based interventions work, which may or may not include a formal theoretical framework.
- The paper reviews ideas and/or provides a critique about how STOPP/START tools and/or STOPP/START tools-based interventions are intended to work.
- The paper provides opinions of how STOPP/START tools and/or STOPP/START tools-based interventions do/do not work.
- The paper outlines, discusses or reviews problems with the implementation of STOPP/START tools and/or STOPP/START tools-based interventions.
- The paper outlines, discusses or reviews potential unintended consequences of STOPP/START tools and/or STOPP/START tools-based interventions.

Exclusion criterion

- Studies including children and adults aged < 65 years.

Inclusion and exclusion criteria for searches in phase 1

Data extraction

The realist review method obtained information by note-taking and annotation rather than standardised data extraction (using a specific software or data extraction form), as used in a traditional systematic review. Documents were examined for data that contributed to theories on how an intervention was supposed to work, which were then highlighted, noted and given an appropriate label. Three researchers (JBG, JG and JMV) initially extracted information from a small set of documents to ensure consistency and then each of them extracted information separately for a subset of papers. The information was subsequently organised by one researcher (JBG). Quality appraisal is integrated into the synthesis narrative, rather than reported separately. Formal quality assessment was not undertaken as part of phase 1.

Synthesis and strategy for prioritisation and finalisation of the candidate programme theories

The final candidate theories were summarised through narrative synthesis, using text, summary tables and a single logic model to summarise individual papers/reports and draw insights across papers/reports.

The initial set of theories on how STOPP/START-based interventions are intended to work were informed by the following steps:

- searching and identifying programme theories from the literature
- qualitative interviews with experts in the field
- ongoing engagement with PPI, including two workshops to ground the study in real-life experience.

The initial list of draft theories from the literature, findings from the interviews and refined list of draft theories from an initial meeting with the Patient Advisory Group were presented to the project team in February 2019. Following productive discussions among the team members, a narrative of each draft theory was developed during the meeting. The group was asked to reflect on the triangulation of findings, to validate theories and to prioritise those appropriate for realist review. From working with the group, we were able to discuss and produce a flow chart with emerging examples of the causal links between CMOs. After working and reworking on the flow chart, we were able to develop the first draft of the single logic model (see *Appendix 3, Figure 16*). The draft model was circulated to the project team, and with the subsequent feedback we produced the candidate programme theories.

Interviews with experts

To prioritise the most important or explanatory programme theories, we consulted topic experts, NHS stakeholders and a patient group. Expert consultations were achieved through ongoing conversations with experts in the field on why the STOPP/START tools work, who they work for, in what circumstances and why. This involved semistructured interviews with experts to triangulate the emerging theories resulting from the literature review. Given the scope and the realist nature of these interviews (i.e. realist hypotheses are confirmed or abandoned not through saturation but rather through relevance and rigour), we expected that approximately 10 interviews would meet our needs, although the final number was confirmed by how interviews uncovered how emerging mechanistic and contextual factors might contribute to the outcome patterns emerging from the empirical literature.[35,40]

We aimed to include and recruited clinicians with a special interest in medicines optimisation, drug safety and pharmacotherapy, experts involved in medicines optimisation at NHS RightCare (Leeds, UK), providers of care in primary and secondary care [e.g. geriatricians, general practitioners (GPs) and other clinicians], care home managers, and academics and those involved in developing education and guidance for older people. In addition, researchers involved in the development of the STOPP/START tools were recruited through Academic Health Science Networks, Clinical Research Networks, the Royal Pharmaceutical Society (RPS) and through co-applicant networks. Following the advice from the board, we considered inviting British Geriatrics Society (BGS) experts. We also considered inviting contributors to the NHS

England *Toolkit for General Practice in Supporting Older People Living with Frailty*[41] and the NICE Multimorbidity Guideline Committee.[42] Ethics approval for the interviews was granted by the University of Exeter Medical School Research Ethics Committee on 19 November 2018 (reference RG/CB/18/09/181).

Interviews were based on a pre-established stakeholder topic guide (see *Appendix 4*) that was developed collaboratively by three of the researchers (JBG, JG and JMV) and based on the emerging findings of the literature review. Interviews were planned to last up to 30 minutes and were conducted either face to face or by telephone at the participant's convenience. Interviews were audio-recorded, transcribed and analysed thematically by two researchers (JG and IP)[43] to expand and refine our programme theories. The purpose of the expert interviews was threefold: (1) to gain input on our list of candidate CMO configurations, (2) to identify additional CMO configurations and (3) to identify additional literature and/or relevant concepts that we may have missed. This was supported by further literature searches in this phase. Interviews were conducted after the patient advisory meeting and before completing candidate theory development.

Phase 2: testing the theory

Phase 2 tested the programme theories obtained from phase 1 (through exploratory search and expert consultation) by synthesising using published and unpublished empirical quantitative and qualitative evidence.

Search strategy

The search to identify studies to test the programme theory included searches of bibliographic databases, checking the reference lists of included studies, forwards citation searching of included studies and searching the websites of relevant organisations. The bibliographic database search was developed using MEDLINE (via Ovid) by an information specialist (SB) in consultation with the review team. We developed a highly sensitive search that identified all studies that mentioned the STOPP/ START tools in the title or abstract fields of the databases searched. No publication-type filter or any other search terms were used. This approach was taken because the number of hits retrieved by searching for STOPP/START tools terminology was feasible to screen in full by two reviewers within the available time and resources. Therefore, the strategies were not designed to target the evidence for prioritised theories, but were very broad and aimed at identifying all empirical studies of interventions based on the use of STOPP/START tools. No indexing (e.g. medical subject headings in MEDLINE) was used, as we did not identify any indexing terms for the STOPP/START tools that were not also used for other prescribing tools. A publication date limit of 2008 to the date of the search was applied. No English-language filter was applied.

The final search strategy was finalised in MEDLINE and translated for use in a selection of relevant bibliographic databases. The full set of bibliographic databases searched included:

1. Cochrane Central Register of Controlled Trials (CENTRAL) (via Cochrane Library)
2. CINAHL (via EBSCO*host*)
3. Cochrane Database of Systematic Reviews (CDSR) (via Cochrane Library)
4. EMBASE (via Ovid)
5. HMIC (via Ovid)
6. MEDLINE (via Ovid)
7. NHS Economic Evaluation Database (NHS EED) (via the Centre for Reviews and Dissemination website)
8. Social Policy & Practice (via Ovid).

The search strategies for all databases and the number of hits retrieved per database are reported in *Appendix 5, Tables 12–19*. Searches of all databases were carried out on 11 June 2019. The results were exported to EndNote X8 and de-duplicated using the automated de-duplication feature and manual checking.

The reference lists of included studies identified by the bibliographic database searches were manually inspected. Forward citation searching of included studies was carried out on 11 October 2019 using, primarily, Web of Science. If a study was not indexed in Web of Science, we used Scopus or Google Scholar (Google Inc., Mountain View, CA, USA) instead, depending on where the study was indexed. The results of forward citation searching were exported to EndNote X8 and de-duplicated before screening.

We searched the repository for grey literature (i.e. OpenGrey). The searches and number of hits are reported in *Appendix 6, Table 20*. URLs that linked to the search results were sent to the review team by e-mail for screening.

We additionally searched the websites of the following organisations on 11 October 2019 (see *Appendix 6*):

- Age UK (URL: www.bgs.org.uk/search/node)
- BGS (URL: www.bgs.org.uk).

Screening

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 CMOs).[39] We therefore assessed studies based on the knowledge we acquired about the theory development of the impact of the STOPP/START tools. Title and abstract screening were carried out by three reviewers (AD, IP and IRC) in Rayyan (Rayyan Software/Qatar Computing Research Institute, Doha, Qatar), with JMV resolving discrepancies (*Box 2*).

BOX 2 Phase 2: preliminary set of inclusion/exclusion criteria

Inclusion criteria (all criteria needed to be met)

- The paper describes the evaluation of an intervention (medication reviews or equivalent) using STOPP/START tools or the paper presents an economic evaluation and/or cost analysis of STOPP/START-based interventions.
- The patient population of the study is adults aged ≥ 65 years.
- The study reported in the paper has any of the following designs: comparative effectiveness study (e.g. a RCT), process evaluation, review of primary research (if method is stated), qualitative research, surveys, history, descriptions of models of care, uncontrolled before and after, cohort and reanalysis of routine data, systematic reviews and narrative reviews.

Exclusion criteria (no criteria should be met)

- Studies including children and adults aged < 65 years.
- Intervention: not a medication review or where review takes place but does not explicitly use STOPP/START tools (although if any doubt include at this stage).
- Study design/publication type: editorials, opinion pieces and advertorials.
- Abstract only (no full text exists).
- Protocols of ongoing reviews.

Inclusion and exclusion criteria for documents for phase 2

Data extraction

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. As such, we created a reference framework to guide extraction, which was intended to be a guide through the process, rather than to be applied in a routine one-size-fits-all manner. The extraction framework had three sections:

1. Standard information collected on all studies (e.g. year, author, title, study type, aims, outcomes, etc.).
2. Evidence of CMOs in studies to support the candidate programme theories.
3. Evidence in studies to refute/refine CMOs in the candidate programme theories.

The data extraction form was piloted by four reviewers (JMV, IP, JBG and AD) and approved by other project team members. The form was then used to extract the above information from all included studies by four reviewers (IP, JBG, AD and IRC). IP was responsible for co-ordinating and merging the data extracts from all the reviewers.

Quality appraisal

Quality assessment used the concept of rigour (i.e. whether or not the methods used to generate the relevant data are credible and trustworthy). Sources were classified as conceptually rich (thick) or weak (thin). 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. Quality appraisal from this perspective is integrated into the synthesis narrative, rather than reported separately.

Appraisal tools may also be used to determine rigour if appropriate.[44] We used the Critical Appraisal Skills Programme[45] to further determine the rigour of the included studies. The importance of transparency in a realist review is similar to that of systematic reviews to ensure the validity, reliability and verifiability of findings and conclusions.[39,44] With the Critical Appraisal Skills Programme, study quality was considered low if half or less than half of the criteria were met, medium if more than half but less than three-quarters of the criteria were met, and high if most were met.[46] The scores were produced, summed and ranked high and medium quality for both randomised and non-randomised studies.

Data synthesis

Synthesis of the diverse sources of evidence was conducted through a structured process that included the following approaches:

1. Juxtaposition of sources of evidence (e.g. where evidence about the STOPP/START tools from one source enables insights into evidence about outcomes from another source).
2. Reconciling of sources of evidence (i.e. where results differ in apparently similar circumstances, further investigation is appropriate to find explanations of why these different results occurred).
3. Adjudication of sources of evidence (on the basis of methodological strengths or weaknesses).[47]
4. Consolidation of sources of evidence (i.e. where outcomes differ in particular contexts, an explanation can be constructed of how and why these outcomes occur differently).

Based on the standardised information extracted for each study, each intervention was assigned to one or more mechanisms by consensus from two of the reviewers (IP and JMV).

For each causal link leading in the logic models, we summarised the number and quality of studies supporting, refuting or qualifying it. We first categorised the link's evidential support in each study as providing:

- no evidence (where the study did not provide any empirical evidence on the link)
- positive evidence (where there were studies providing empirical evidence supporting the link in the direction predicted by the logic model)
- negative evidence (where there were studies providing empirical evidence on the link but not in the direction predicted by the logic model or the associations did not reach statistical significance).

We further categorised the strength of each causal link's evidential support across all studies as:

- positive evidence [i.e. having a substantial majority of positive evidence studies (> 70%)]
- mixed evidence (i.e. having both positive and negative evidence studies and both categories exceeding 30% of studies)
- negative evidence [i.e. having a substantial majority of negative evidence studies (> 70%)]
- no evidence (i.e. all studies classified as providing 'no evidence').

Patient and public involvement

The PPI lead used her extensive contacts to establish and manage a patient group that worked with us throughout the review. A group of five older people with experience of taking multiple medicines assisted in theory development, shaping the focus of the review after screening and synthesis stages in the review process, interpretation and dissemination of findings. The group consisted of five patients (two men and three women) who were all aged $\geq$ 50 years, with multiple health conditions and taking multiple medications. The role on the advisory group for this project was advertised to a larger Patient Advisory Group (i.e. Peninsula Public Involvement Group), which is part of the National Institute for Health Research Applied Research Collaboration South West Peninsula. The group provided regular feedback on documents and met regularly to support and provide direction to the research.

We worked with the Patient Advisory Group at all phases of the project to ensure that the research questions and outcomes of interest remained important and aligned to the needs of patients, and the group members was particularly critical in supporting key stages of the project. Specific Patient Advisory Group meetings were organised to address key issues that arose during the implementation of the review:

- Further review of the application in the light of the feedback received and finalising the protocol, once funding is in place. This meeting was replaced with the development of an initial list of draft theories.
- Discussion with reviewers in relation to the screening criteria. This was not considered important, as the screening criteria were very short and succinct and we did not think that holding a meeting was an efficient use of resources. Instead, we worked further with the Patient Advisory Group to refine the programme theories.
- Interpretations of findings from the review.
- Dissemination of the findings, including the preparation of the final report.

EC provided support and training for patient representatives, which included an introduction to realist synthesis. The patient co-applicant (JH) attended all research project team meetings. This ensured that all discussions throughout the project were informed by patient experience.

Chapter 4 Results: identification of the programme theories

Phase 1 literature search results

Sixty studies were identified after de-duplication in the initial search (*Table 1*). Thirty-one abstracts were selected after sifting. At full-text screening, 17 papers were selected for inclusion, as they provided the clearest examples of the ideas underpinning STOPP/START-based interventions. Forwards and backwards citation tracking of the 17 key articles, and additional iterative searches in Google as key subtheories emerged (Google search first 10 pages) identified a further 14 papers. The full texts of these papers were then read and notes were taken about the key ideas and assumptions regarding how the STOPP/START-based interventions were intended to work. Of the 31 papers identified, eight were purposively selected as 'best exemplars' of the ideas reflected in the papers as a whole. Regular discussion of these ideas among the project group and circulation and feedback of draft working papers outlining the theories ensured that the full range of different theories was represented. *Figure 1* summarises the flow of studies from identification through to inclusion in the final document.

Initial programme theories from the literature

How is the underlying problem conceptualised?

Chronic diseases and multiple medical conditions prevail in older people; therefore, there is a high prevalence of polypharmacy, which has been associated with negative health outcomes.[5] Lack of awareness of pharmacokinetics and pharmacodynamics changes in older people can contribute to inappropriate medicine use, resulting in adverse drug effects. Some of these reactions may be confused with progression of the given pathology or with some typical age-related syndromes.[48]

The risks of potentially inappropriate medications outweigh their benefits, especially when a safer or more effective alternative therapy is available. The use of potentially inappropriate medications is considered a major problem among older people, and may contribute to increased risk of adverse drug effects and to developing drug–drug and drug–disease interactions.[19]

Patients are often treated by different clinicians following either an outpatient visit or an acute admission, and the consultant may not be aware of the existing medication that a patient may be on. Most often, these health conditions are treated in general practice, which results in the attribution to GPs of the responsibility of whether or not prescribing has taken into account the patient's clinical history.

TABLE 1 Bibliographic database search results

Database	Number of documents
MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations	29
EMBASE	37
HMIC	5
CINAHL	5
Total results	76
Unique results	60

```
┌─────────────────────────────────────────────────┐
│ Records identified from MEDLINE, EMBASE, CINAHL and HIMC │
│                    (n = 76)                       │
└─────────────────────────────────────────────────┘
              │
              │          ┌──────────────────────────┐
              ├─────────▶│   Duplicates removed     │
              │          │        (n = 16)          │
              ▼          └──────────────────────────┘
┌──────────────────────────┐
│  Number of records after │
│    duplicates removed    │
│        (n = 60)          │
└──────────────────────────┘
              │          ┌──────────────────────────┐
              ├─────────▶│ Excluded after title/abstracts │
              │          │        screened          │
              ▼          │        (n = 29)          │
┌──────────────────────────┐ └──────────────────────┘
│ Records after titles/abstracts │
│        screening         │
│        (n = 31)          │
└──────────────────────────┘
              │          ┌──────────────────────────┐
              ├─────────▶│ Excluded after full texts │
              │          │        screened          │
              ▼          │        (n = 14)          │
┌──────────────────────────┐ └──────────────────────┘
│     Included studies     │
│        (n = 17)          │
└──────────────────────────┘
```

Articles subjected to citation tracking
(n = 17)

Total records from citation tracking
(n = 511)

Excluded after title and abstract screening
(n = 503)

Grey literatue
(n = 6)

Number of full texts screened
(n = 8)

Included studies
(n = 31)

FIGURE 1 Flow diagram for identification of relevant documents for theory development.

Role of the professionals in prescribing: who does what?

Professional boundaries play a key role in prescribing practices. There is an implicit assumption that clinicians with expertise in older people (e.g. geriatric pharmacotherapists) will be more aware of medications that are inappropriate for older people. There are different perspectives on the onus being on GPs or consultants to ensure that they prescribe drugs that do not lead to drug–drug interactions or ADEs, compared with pharmacists picking these issues up during a review of medications. Furthermore, if GPs identify these problems, it remains to be determined if and how GPs can change/challenge the prescribing of consultants, or reflect on their own prescribing practice or that of their colleagues. Finally, if and when pharmacists pick these up, it is unclear if (and, if so, how) pharmacists can challenge or modify prescribing by GPs or consultants.

These challenges operate within existing sets of expertise, relationships, professional boundaries and remits between GPs, consultants, pharmacists and nurse prescribers. These play out on a 'macro/meso level' in terms of national/professional norms that exist about the remit of different practitioners and the expertise of specialists compared with generalists, but is also subject to local variations in the ways in which these different clinicians work together. They are also likely to be influenced by the existing relationship and knowledge a GP or a consultant has with a patient (and therefore their knowledge about their history, etc.). These sets of relationships form an important background against which the STOPP/START tools are used.

Context of a STOPP/START medication review

To address these problems, current NICE guidance on medicines optimisation includes as a recommendation to consider using a screening tool for adults such as the STOPP/START to identify potential medicines related patient safety incidents in some groups of people, especially in relation to adults taking multiple medicines (polypharmacy) and adults with chronic or long-term conditions and older people.[32] Medication reviews can take several forms and vary in terms of their comprehensiveness, purpose and whether or not the patient is present during the review. Medication reviews could be carried out by the prescribers themselves or by an independent reviewer (usually a pharmacist). However, in the immediate context of the medication review, the process of the review might position them differently, either as someone reflecting on their own prescribing and making changes to their own practice, or as someone who is reviewing another clinician's practice or decisions and therefore attempting to change their practice. Most of these patients are seen by a number of different physicians/prescribers, each of them with a partial knowledge of the patient, the patient's problems and the medications. This makes it unlikely that anybody could use it for reviewing their own prescribing. The only exception (certainly at a stretch) would be the GP. This has a number of implications for thinking about the context in which the STOPP/START tools are used, how we conceptualise any programme theories and the kinds of higher-level theories we might draw on to think about it:

- It changes the nature of the intervention and the underlying assumptions about how it is intended to work (i.e. it changes the programme theories).
- It has implications for the broader theories that might come into play to think about the relevant contextual factors that might shape the intervention. For example, if it is about reviewing and making changes to someone else's prescribing decisions or behaviour, theories about interprofessional boundaries come into play, and if it is about reviewing own practice then audit and feedback theories might come into play.

Some types of review are carried out in the absence of the patient (e.g. prescription reviews), whereas others are carried out with the patient present (e.g. a 'clinical medication review'). Therefore, the type and nature of the medication review will also form an important context that might shape how STOPP/START tools work, especially in relation to some of the programme theories around patient-centred care. Several pieces of guidance advocate the use of the STOPP/START or other tools to assist in these processes. Exactly how they are intended to assist and, therefore, the underlying programme theories about their use vary.

What resources do STOPP/START tools offer?

The STOPP/START tools provide a comprehensive list of medications that clinicians should consider stopping in older people generally or, for some medications, in older people with certain conditions. In addition, they provide a list of medications that clinicians should consider starting if certain conditions are met.

Positive programme theories about how they are expected to work are as follows:

- The STOPP/START tools provide a systematic and structured way of carrying out the medication review process (resource). The assumption is that clinicians will use the tools to structure the way in which the review is carried out [i.e. the tools could either replace the usual way in which the task is carried out or the tools might somehow integrate the structure provided by the tool with usual practice (mechanism)]. The use of a more structured approach is expected to enable the clinician to identify a larger number of potentially inappropriate medications, ADRs or unmet need than if clinical judgement had been used alone (outcome).
- As the STOPP/START tools are evidence based, comprehensive and structured, the assumption is that they can be used with little need for clinical judgement (mechanism) and, therefore, can be used by a range of clinicians who may not be specialists (context) in older people to identify potentially inappropriate medication. This in turn increases the capacity for medication reviews to take place, rather than just to rely on clinicians who are specialists in older people.

- The identification of potentially inappropriate medications using the STOPP/START tools will prompt the clinician to start or guide discussion with the patient about whether they wish to stop or start particular medications and what their priorities are (mechanism), which is expected to result in care that is more patient centred (outcome). However, this is likely to depend on whether or not the patient is actually present at the review (context).

These theories envisage STOPP/START tools as screening tools and a means of identifying problems. However, to work this way, this assumes a number of contextual features being present or that certain mechanisms will trigger these features. These include the following:

- The clinician undertaking the review either knows about, or is able to access information about, the patient's existing conditions and existing medication.
- The clinician undertaking the review agrees with the criteria set out in the tool (i.e. they do not clash with clinical judgement).
- The clinician undertaking the review has the time to go through the entire tool systematically.
- Clinicians will know how to act when potentially inappropriate medications have been identified and that the clinician will discuss this with the patient.

We identified a number of opposing theories about STOPP/START tools that are expected to work:

- The STOPP/START tools offer resources that enable clinicians to identify potential medication problems, it does not provide resources about the context of the individual patient in terms of how these criteria apply to specific patients. Therefore, the clinician needs to integrate their clinical judgement, information gathered from a discussion with the patient and the alerts from the STOPP/START tools (mechanism) to make a decision about whether to stop or start any medicines (outcome). This challenges the idea that the tool can be used without the need for clinical judgement and raises questions about the degree of specialist knowledge that is needed.
- Some authors have argued that one of the reasons why a clinician's use of the tool may not enable them to identify as many potentially inappropriate medicines (outcome) is because clinicians do not use it 'rigorously', implying that clinicians ignore some of the alerts identified through the tool and therefore do not follow through on stopping medication when the tool guidance identifies it as potentially inappropriate (mechanism). This theory implies that the tool is intended to act not just as a means of identifying problems, but as a decision support tool whereby the tool acts as an instruction to decisions (i.e. the guidance should be followed, not just considered). This implies that clinical judgement is not required.

These two different positions highlight the contested nature of the tools and their role and purpose vis-à-vis clinical judgement and patient input. One set of theories position the tools as something that provides advice or alerts to clinicians, whereas the other set of theories position the tools as decision support tools:

1. Paper versions of the tools (resource/context) are time-consuming to use and require familiarity with the content of the tools (context). As many clinicians do not have time or are not familiar with their content (context), they do not use the tools or skim over them (mechanism) and this means that potentially inappropriate medications are missed (outcome).
2. Clinicians may not have all the information to hand (context) and so they are unable to use the tools to carry out the review with the full information (mechanism).

Potential solutions to these two problems are as follows:

- STOPP/START tools should be used electronically (context) and the clinician using the tool can draw on information in a patient's electronic health record to carry out the review.
- It may be possible to design an algorithm based on the tools that carries out the review automatically (context) and simply presents the clinician with an alert about potentially inappropriate medications (resource). The clinician is then expected to accept and trust this alert and follow through on the advice accordingly (mechanism).

We might also consider a further set of theories a bit further down the implementation chain about how, when and why clinicians do actually make prescribing changes based on the use of the tool. This might depend on who uses the tool (e.g. a pharmacist), their remit to instigate changes and who actually makes the prescribing changes (e.g. GP, consultant) and the roles, remits and relationships between these different social actors. For example, if a pharmacist uses the tool to conduct a review, they may only be able to advise what changes might need to be made to either the GP or the consultant, who may choose to follow this advice or may choose to ignore it.

We further discussed the initial theories that arose from the literature and developed the succinct list of theories (*Table 2*).

Expert interviews

Eleven semistructured interviews were conducted with experts in the field (i.e. GPs, geriatricians, pharmacists and academics) either face to face ($n = 1$) or by telephone ($n = 10$) at participants' convenience (*Table 3*) (see *Appendix 4* for interview topic guide). The one participant who attended a face-to-face interview was not reimbursed for travel, as she was an academic based in the University of Exeter. Eight participants were from the UK, one was from the Netherlands, one was from Canada and one was from Ireland.

A total of nine themes emerged from the analysis, which are discussed below.

Familiarity with STOPP/START tools
Some participants were very familiar with the STOPP/START tools and others had become familiar with it only recently. Participants reflected on how familiar different professional groups were with the STOPP/START tools. Some professionals (particularly GPs) seemed not to be familiar with the STOPP/START tools at all.

What are STOPP/START tools and what is their purpose?
Participants expressed mixed views about the role of STOPP/START tools in the decision-making process, such as whether patients should be involved in shared decision-making or if decisions should be solely the decision of the health-care professional. Some participants considered the STOPP/START tools as prompts and others considered the tools as checklists. In addition to supporting decision-making, some participants said that they can be used for improving prescribing appropriateness and some others said that they can be used as an education tool. Participants also discussed the need for using clinical knowledge and clinical judgement when applying the STOPP/START tools.

What are STOPP/START tools?
Some participants considered STOPP/START tools as prompts, some considered them as a good starting point and others perceived them as tools that can provide guidance. Others described the tools as a checklist and some participants considered them to be a structured approach for conducting medication reviews.

TABLE 2 Initial set of theories

Question	Theory
Why these tools?	• STOPP/START tools reduce health-care costs • STOPP/START tools are widely used/recognised • STOPP/START tools are easy to use • STOPP/START tools are comprehensive • STOPP/START tools can be used to give personalised recommendations to patients • STOPP/START tools promote a person-centred approach to shared decision-making • STOPP/START tools facilitate audit and outcomes' monitoring • The application of STOPP/START tools has a direct effect on the health of individuals and appropriate use of limited resources

To what aim should they be used?	• STOPP/START tools help clinicians avoid unnecessary prescribing cascades by encouraging them to consider medications as possible causes of non-specific symptoms, such as confusion, falls or constipation
	• STOPP is a tool that can be used as a screening/prompt/checklist/decision support tool for:
	  ○ identifying medications with potential risks
	  ○ screening for potentially inappropriate medications
	  ○ rapid identification of potentially inappropriate prescribing
	  ○ identifying adverse drug effects in older people
	• STOPP/START tools promote standardisation of clinical judgement
	• To be of use clinically, STOPP/START tools need to demonstrate a positive effect on adverse outcomes (e.g. reduction in ADEs, a reduction in hospitalisation, a reduction in health-care use, a reduction in mortality)

How to use them?	• STOPP/START tools are aimed at supporting clinicians in making decisions about medications, but do not replace clinical judgement • STOPP/START tools leave room for interpretation • STOPP/START tools can facilitate making decisions about medications without specialist knowledge or with little clinical judgement • Non-medical prescribers are more receptive to the use of STOPP/START tools • STOPP/START tools can improve the training of medical students in the management of polypharmacy • Additional training should be optional when using STOPP/START tools • Staff training is important to ensure consistency, and the receptiveness of prescribers, patients and staff in various settings will have an impact on the uptake and effectiveness of STOPP/START tools in older people • STOPP/START tools do not take into account patient preferences or comorbidities
How should they be implemented?	• Paper versions of the STOPP/START tools are time-consuming and require users to be very familiar with the content of the tool to use it, which rarely happens • Computerised use of STOPP/START tools would be more efficient • Computerised use of STOPP/START tools should be as sensitive and specific as possible, as there will not usually be subsequent evaluation by a clinician • STOPP/START tools are difficult to use in everyday clinical practice • Is it sufficient to use the STOPP/START tools during a single episode of care, or should patients be exposed to the tools on a daily, weekly or monthly basis? • Patients may not understand the notion of deprescribing and they may think clinicians are trying to stop them from having medications they need • STOPP/START tools need to be adapted to reflect local policy • Process redesign is needed to fit STOPP/START tools into the clinical workflow

TABLE 3 Interview participant characteristics

Characteristic	Number/range
Gender	
Female	7
Male	4
Age range (years)	30–60
Expertise	
STOPP/START developers	1
Academics	4
GPs	6
Geriatricians	1
Pharmacists	4

What is the purpose of STOPP/START tools?

Participants described the purpose of STOPP/START tools in different ways. Participants said that STOPP/START tools could be used as an educational tool and for improving quality of prescribing, changing culture, personalising care, shared decision-making and supporting evidence-based practice.

Some participants felt that STOPP/START tools are supposed to get prescribers to think about the appropriateness of the drug and others thought that they should be introduced as a part of a routine medication review process. Some participants proposed undertaking quality improvement activities around STOPP/START tools:

> *And I know certainly here we have HIQA [Health Information and Quality Authority], which is our health information and quality authority. I don't know what the equivalent with you would be in the UK. I don't know, but certainly here anyway, anyone who's over the age of 75 and who has more than four medications, they're obliged to have a 6-monthly structured medication review. And we recommend that you use a tool such as STOPP/START in doing that . . .*
>
> *Interviewee 6, geriatrician*

Some participants considered it an effort to enhance pharmacy expertise to bring about a change in culture in general practice:

> *I know that it depends on where exactly you work, but I think there is some funding from NHS England for pharmacy support. So, that's one way of doing it. Seeing it as pharmacy thing and bringing that pharmacy expertise into practice and changing the culture.*
>
> *Interviewee 5, GP*

Participants considered STOPP/START tools to be individualised tools:

> *You start with the criteria but then you really have to apply them in a patient-oriented manner, think about the patient, what does the patient really want? For example, we have done some work in which we used goal attendance scaling of patients where we first asked the patients what do they really want and then, when the patient says, OK I don't like drugs, I have too many drugs and you then start applying STOPP/START criteria . . .*
>
> *Interviewee 3, pharmacist*

Most of the participants felt that patients or carers should be involved in a shared decision-making process, but others did not think that patient enablement was part of their scope.

Participants mentioned that STOPP/START tools are aligned with NICE guidelines and provide the best evidence.

Participants described STOPP/START tools as educational tools for both professionals and patients, and particularly for professionals who may be unfamiliar with geriatric pharmacotherapy.

Target patients

Participants described which patients should be targeted, and proposed that those with multimorbidity and multiple medications might benefit most from the use of STOPP/START tools:

> . . . but often, because of the cognitive problems with many older patients, they themselves may not be able to engage positively or negatively because of their cognitive problems. But that's why we need to have a look at their medications to see if those medications are potentially causing cognitive problems or worsening the cognition, or if they're going to survive long enough to derive any benefit from these drugs.
>
> *Interviewee 6, geriatrician*

Appropriate setting

Participants expressed different views on the specific settings that STOPP/START tools would be applicable to, and considered primary care, hospitals, care homes and nursing homes as potential settings where they could be of value:

> I very much feel STOPP/START feels like a hospital tool. It feels like the sort of tool you can use if you've got a fair bit of time with the patient, you are specifically seeing them for the review, and you're going through the body system.
>
> *Interviewee 5, GP*

> In care homes, in – yeah in the elderly and the [inaudible 17.55] morbidity really because those are the ones that are most of risk of interactions and unnecessary drug side effects.
>
> *Interviewee 8, GP*

> One showing that in a nursing home environment that you can reduce the risk of falls the risk of hospitalisation over a 6-month period by intervening with STOPP and START criteria through multidisciplinary discussion and review of medications.
>
> *Interviewee 6, geriatrician*

Professionals

Participants had mixed views about their role in the use of STOPP/START tools. Some participants said that they can be used by doctors (e.g. geriatricians or GPs), nurses [e.g. clinical nurse specialists, nurse consultants in geriatric medicine or practice nurses (diabetics/respiratory clinics)] or pharmacists (e.g. general practice pharmacists, hospital pharmacists or pharmacy technicians). Most of the participants proposed that a pharmacist as part of a multidisciplinary team and with access to the patient's clinical history is the ideal professional to use STOPP/START tools.

It appears that the purpose of the tools is perceived by some as core to the scope of the practice of geriatricians. Some participants also believed that less experienced professionals could benefit more

from using them than more experienced professionals, who may not benefit from using them at all. This may particularly be the case for geriatricians:

> *I think it can have a big impact. I mean the first example here is what we do so you know independent medication reviews – we use the patient notes and we have the patient there in a face-to-face encounter and we can make recommendations and follow them up so you know the STOPP/START interventions can have a huge impact on that as sort of guidance and I think, useful for people who are less experienced in medication review.*
>
> *Interviewee 9, pharmacist*

> *I think hospitals and geriatricians and, say, clinical nurse specialists, or nurse consultants in geriatric medicine, I think they have a role to play here as well. I think often GPs are very overburdened and they're not geriatricians. GPs have to know about paediatrics, obstetrics, psychiatry, everything. I actually think that this kind of complex problem, perhaps geriatricians should have much more of a role and there should be much more of a demand for them.*
>
> *Interviewee 6, geriatrician*

> *I think the context of where you use it, I think contributes to the success of the tool because medication patient records are not available for example, at say the community pharmacy or in an emergency department or anything like that. I think there are issues around how the best places it's got to be used or it can be used successfully.*
>
> *Interviewee 10, pharmacist*

Resources

Participants discussed various resources that would be needed for ensuring that STOPP/START tools would have the expected impact, including training, time, staff availability and costs.

Training and education

Some participants said that they would need training in using STOPP/START tools, whereas others were using it as a training tool. Some participants mentioned ongoing education as the most effective intervention.

Time

Participants expressed concerns about the time required to do medication reviews using STOPP/START tools. Some participants felt that they were very time-consuming processes and would not be ideal tools to squeeze into shorter consultations in primary care. Other participants agreed that using STOPP/START tools is time-consuming, but considered this to be within the scope of geriatric medicine.

Staff

Participants expressed concerns about lack of staff and staff availability, as medication reviews using STOPP/START tools need additional staff to carry out the review.

Costs

Participants said that implementing and performing medication reviews using STOPP/START tools would involve additional costs, if they are to be widely accessible.

Impact of the STOPP/START tools

Several participants expressed views about the degree of potential impact of using STOPP/START tools. Some participants were positive about their potential, whereas others reflected on several negative impacts.

Uncertainty about degree of potential impact and how to capture it
Participants were unsure about the degree of potential impact and expressed their views on how to capture it:

> It's difficult to say isn't it, because I don't think it seems to quantify the impact of the effects that you had necessarily, it's not easy to say well definitely stopping that Zopiclone, will stop that fall, it's hard to say . . .
>
> *Interviewee 4, pharmacist*

Positive impact on quality of life and patient satisfaction
Participants perceived that STOPP/START tools could reduce hospital admissions, morbidity and mortality, and improve quality of life and patient satisfaction:

> I have seen different things with patients around quality of life, so slowly reducing antipsychotics and sedating drugs to actually people then are no longer bed bound, sit in a chair or lounge, so you can see lots of really good impact when you actually start to take away reducing these medicines. I think it's hard to quantify.
>
> *Interviewee 4, pharmacist*

Challenges, barriers and facilitators for using the tools
Several challenges, barriers and facilitators were identified, including effective patient engagement, the user-friendly nature of the tools (or lack thereof), availability of electronic formats, accessibility and integration of the tools into existing systems.

Challenges
Participants recognised the difficulties in ensuring effective patient engagement and considered approaches for overcoming them:

> I think it's hard because it depends how it's presented and I think it depends on the patient because there are some patients, you know of the people that I can think, there are some patients who absolutely want to be involved, they want to see, you know they want to know why, they want to know the evidence, they'd quite like to see a STOPP/START and this is why and be given the rationale for why you're making suggestions and then equally there are other patients who don't want to know.
>
> *Interviewee 9, pharmacist*

Facilitators
Some participants described the STOPP/START tools as user-friendly. Some participants also believed that availability of electronic formats was crucial for uptake. Participants believed that one of the requirements for the tools to be used in daily practice is to make them widely accessible. Participants proposed that the tools be embedded in everyday practice for conducting medication reviews and integrated into existing systems. Leadership and incentivisation through inclusion and as part of quality improvement initiatives were also mentioned as relevant facilitators:

> I suppose it would mean maybe somebody being a bit of a champion for it. I don't know whether that would be just one lead GP who could disseminate the information and give a bit of training at a clinical meeting or something like that. That's what I would probably do but whoever the lead GP is, they'd need a bit of sort of training on it.
>
> *Interviewee 2, GP*

Barriers
Some other participants found the original version of the STOPP/START tools to be non-user-friendly, hefty and cumbersome. Lack of electronic versions was also considered a barrier to wider uptake.

Suggestions for implementation

Most of the participants were very keen to have automated, computerised STOPP/START tools, but they also raised concerns about the problems with implementation. Some participants felt that there needs to be a more sophisticated system to prioritise actions resulting from using the tools:

> *I think there is always a risk of information or alert overload, so I think perhaps prioritising intervention, to say OK which are the most dangerous or top 10 most risky, instances of potentially inappropriate prescribing and in fact that's what we are doing with this study that we are just in the process of completing, we have identified the top 10 most common, well not just top 10, we have ranked the STOPP/START criteria in order of prevalence in our population . . .*
>
> *Interviewee 7, GP*

Patient and public involvement in theory development

The Patient Advisory Group provided support throughout the process of theory development, including helping to revise drafts and providing feedback. In addition, two meetings were convened with the Patient Advisory Group to discuss specific issues, provide feedback on the initial set of theories (see *Appendix 7*) and provide feedback on a subsequent subset of specific and key theories. There was also a discussion on other concepts and ideas considered as potential theories by the advisory team.

The first meeting took place before conducting the expert interviews. The aims of the project and the concept of a realist synthesis were presented. Instead of using the term 'theory' (which is complex), we instead used the term 'idea'. In total, we had 29 'ideas statements' that were focused on why, what, how and implementation. These simplified statements were derived from initial literature searches. Patients initially worked in two groups, each facilitated by one of the researchers. The groups worked through and discussed each statement. The groups were then asked to rank the ideas based on how important they thought it was to pursue that idea. Notes were made to capture discussions around this decision. In the final section of the workshop, a whole-group discussion was held to discuss the most important ideas (see *Appendix 8*).

In a subsequent meeting, 13 specific theories were shared with the advisory members who discussed if the statements were relevant or not important to pursue. Notes from discussions were captured by the PPI facilitator. The meetings finished with agreement among the group on which three theories should be taken forward.

Integration of the ideas and discussion from the Patient Advisory Group meetings into the wider project was ensured by (1) having researchers present at meetings to hear discussions first hand and (2) sharing notes around the wider research team and discussing these in project group meetings where decisions were finalised.

Key theories that emerged during the meetings are listed below:

- Using STOPP/START tools enables patient preferences, concerns and experience to be taken into account, resulting in increased adherence, satisfaction and empowerment.
- STOPP/START tools provide a structured approach that allows a greater reduction in unnecessary or harmful medication combinations, burden, adverse outcomes and resource use, which is more effective than relying on clinical judgement alone.
- For STOPP/START tools to be effective, the clinician undertaking the review must have time to go through the entire tool systematically. Guidance should be followed, not just considered.

During the second meeting, there was also discussion around specific concepts or ideas that the group considered to be important as potential programme theories. The discussion focused on reluctance of pharmacists to take on the role and concerns about pharmacist limitations in knowledge authority and access to patient notes. Carers' presence during the medication review was also highlighted and communication was suggested as a potential theory. One of the important issues considered in the meeting was the reason triggering the review. The advisory members provided the following examples of potential triggers:

- the patient having problems
- part of Quality and Outcomes Framework ('tick-boxing')
- financial incentives.

Candidate theories

A refined list of theories produced from the literature and results from the interviews were discussed during project team meetings, which resulted in a comprehensive single logic model of how STOPP/START-based interventions operate (see *Appendix 3*). For convenience, this single model is presented in relation to three different mechanisms.

Three main mechanisms were identified: (1) personalisation of the medication review, (2) systematisation of medication review processes and (3) implementation of evidence. These mechanisms are used as the organising principle of programme theories. It was hypothesised that each of these mechanisms has a strong association with particular contexts and outcomes, which we outline below. That is not to say that these are unique or the only possible combinations; however, these are the mechanisms that we consider at this stage to be the strongest. They are not mutually exclusive, but rather are expected to work synergistically. However, the focus was on testing each of the programme theories separately, which, depending on the evidence, would be rejected or refined.

A refined list of theories produced from the literature and from the interviews was developed by the project team. The list was extensively discussed during a project face-to-face team meeting and follow-up consultation, which resulted in a comprehensive single logic model of how STOPP/START interventions operate.

We presented the initial list of theories to the PPI advisory team. Following productive discussions, the PPI advisory group rated the theories based on their perceived importance (see *Appendix 8*) and provided additional feedback. This information was used to prioritise theories for the subsequent (testing) phase.

Personalisation of the medication review
Personalisation of the medication review (*Figure 2*) refers to the degree to which the medication review processes and outcomes are tailored to each individual patient. STOPP/START tools are being promoted to support shared decision-making (taking into account patient preferences, experiences and expectations) (mechanism). For STOPP/START-based interventions to achieve the hypothesised improvement in patient awareness of the medication review process, adherence to the therapeutic plan, satisfaction with care, patient empowerment and quality of life, and to decrease patient burden and support a patient safety culture (outcomes), patients and/or carers need to be present and engaged in the process, and there should be sufficient time to carry out the review [i.e. variation in intervention delivery (context)].

The full set of related key variables (context and outcomes) was included in a logic model (see *Figure 2*). Note that, in *Figures 2–4*, the distance along the horizontal axis between any variable and the corresponding mechanism represents the expected degree of the association along the proximal–distal continuum.
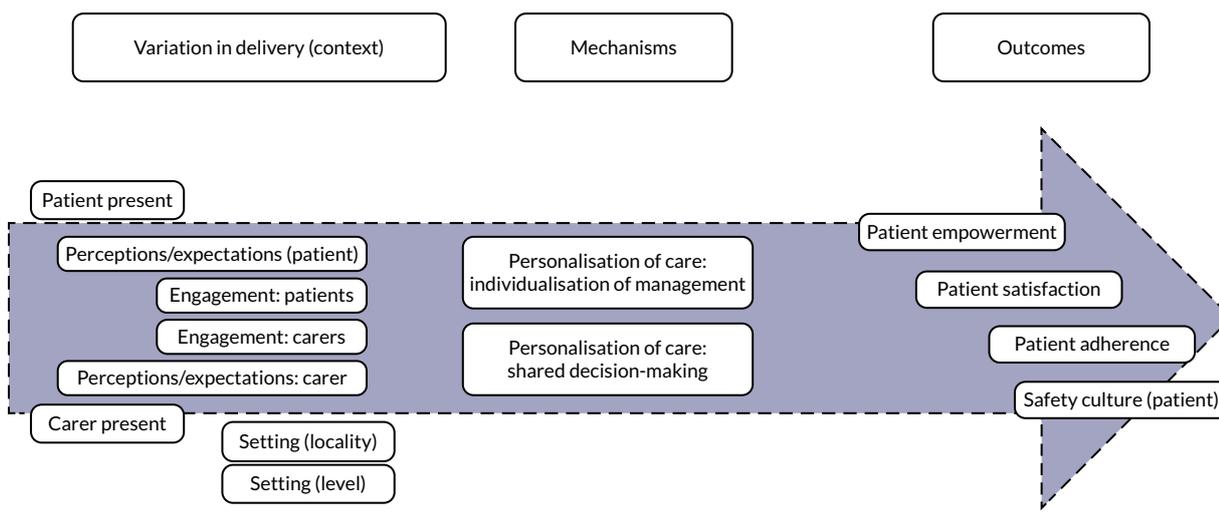
FIGURE 2 Logic model for personalisation.

For example, individualisation of management is expected to trigger a cascade of effects that may result in improved patient empowerment, which can trigger an improvement in patient satisfaction, which can result in increased adherence and eventually modify the safety culture of patients. The link between individualisation of management and the latter outcome is expected to be more tenuous than between individualisation of management and patient empowerment.

Based on discussions and feedback from the PPI advisory group, we prioritised the following theories for phase 2:

- STOPP/START tools lead to greater adherence to pharmacological treatment when patients are present at their medication review.
- STOPP/START tools increase patient empowerment when patient perceptions and expectations are taken into account (elicited and acted on).

Systematisation (of medication review processes)
Systematisation (of medication review processes) (*Figure 3*) refers to the degree to which the medication review is formalised and systematically implemented. STOPP/START tools support a standardised/systematic approach for medication reviews (mechanism). For STOPP/START tools to change professional and organisational culture, burden and costs, and for patients to be aware that a review has taken place (outcomes), there needs to be a delivery system (paper or computer based) that is linked to a patient's medical records and information from previous reviews using STOPP/START tools. Information must also be shared between relevant professionals, teams and organisations [i.e. variation in intervention delivery (context)].

Based on discussions and feedback from the PPI advisory group, we prioritised the following theories for phase 2:

- STOPP/START tools change professional and organisational cultures when there are computerised systems linked to medical records, access to previous STOPP/START tools data, and data-sharing takes place (with relevant professionals, teams and organisations).
- STOPP/START tools reduce professional and organisational burden in medication management when there is sufficient familiarity with the tool.
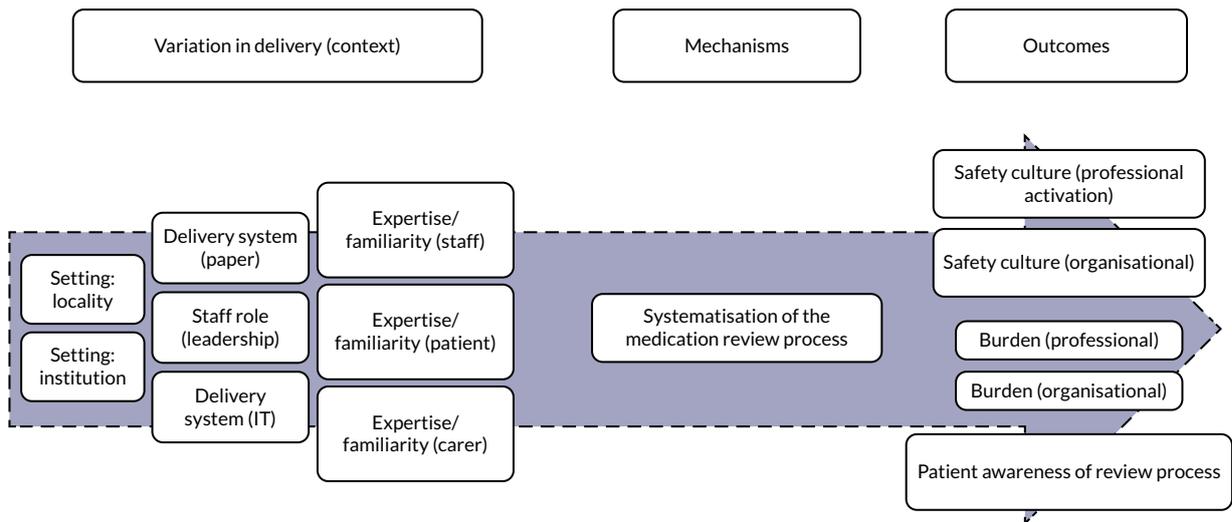
FIGURE 3 Logic model for systematisation. IT, information technology.

Implementation of evidence

Implementation of evidence (*Figure 4*) captures the degree to which an intervention oriented to support the medication review focuses on promoting the dissemination and use of the best-available evidence. Delivery of STOPP/START-based interventions is based on implementation of evidence (mechanism). For STOPP/START-based interventions to result in greater appropriateness of prescribing/deprescribing, reduction in adverse outcomes and hospitalisations, improvements in clinical outcomes, mortality and quality of life, along with having an impact on costs, health-care use, professional burden and organisational culture (outcomes), the supporting evidence base needs to be actively promoted by people in leadership roles and requires training of relevant clinicians (e.g. GPs, pharmacists, geriatricians or nurses) [i.e. variation in delivery of the intervention (context)].
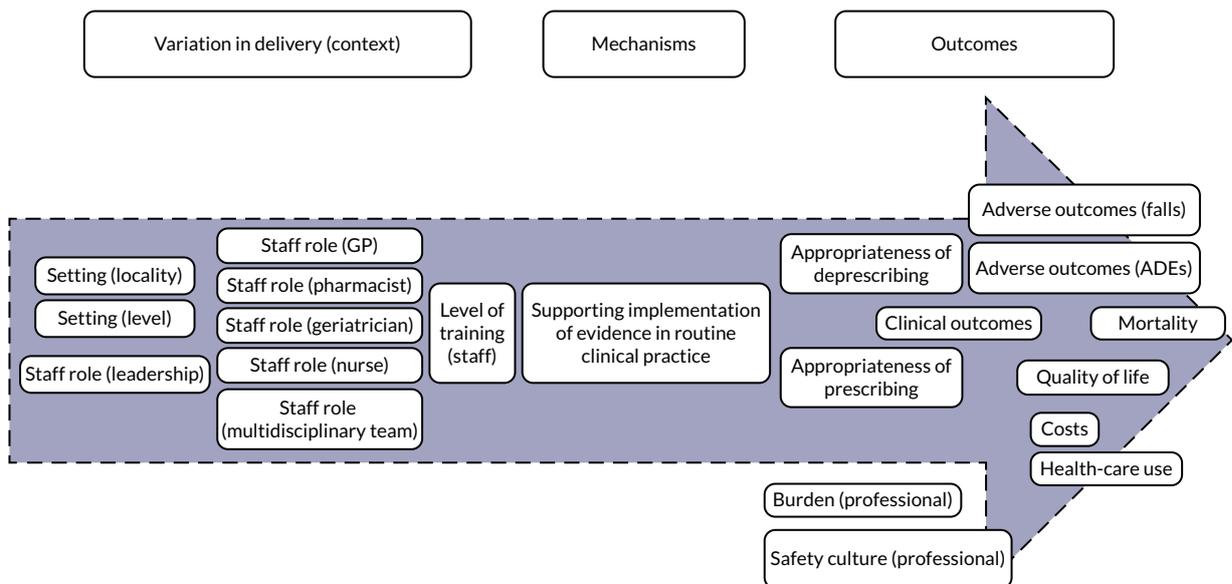
FIGURE 4 Logic model for the implementation of evidence.

Based on discussions and feedback from the PPI advisory group, we prioritised the following theories for phase 2:

- STOPP/START-based interventions are associated with increased appropriateness of prescribing/ deprescribing when there are lower levels of familiarity with the evidence base (i.e. non-geriatricians, non-pharmacists and more junior professionals).
- STOPP/START-based interventions reduce adverse outcomes through appropriate prescribing/ deprescribing.
- STOPP/START-based interventions can change professional culture in settings less familiar with the supporting evidence base.

Summary of findings

Programme theories were articulated around three main mechanisms: (1) personalisation of the medication review, (2) standardisation of the review process and (3) implementation of evidence. Each mechanism was linked to mechanism-specific contextual variables and to outcomes. A set of seven theories was prioritised for testing in the subsequent phase, as originally planned. However, given the nature and scope of the available evidence, we were able to test all configurations, not just our prioritised ones.

Chapter 5　Results: testing the programme theories

Study screening and identification

Out of a total of 1359 records, 817 were identified from all databases after deduplication (*Table 4*). These were subjected to screening and full-text assessment for eligibility, and 34 studies were initially included in the review. Following previously forwards and backwards tracking of citations, an additional six studies were included, giving a total of 40 studies (*Figure 5*). We also screened 16 records obtained through grey literature searching; however, none of the studies met our eligibility criteria and, therefore, they were discarded.

Descriptive summary of included studies

All 40 studies had been published in the period 2011–19, and the majority (*n* = 31)[49–81] had been published in the 5 years immediately prior to the conduct of the synthesis (*Table 5*). The studies had been conducted in Europe [mainly Spain (*n* = 13),[51,53,56,58,60,61,73,79,84–88] Belgium (*n* = 5)[49,55,57,74,78] and Ireland (*n* = 3)], North America [Canada (*n* = 2)[50,54] and the USA (*n* = 2)[67,80]] and, in a small number of cases, other regions. Twelve studies had a randomised design.[50,51,53–55,70,71,73,78,82–84] Of note is that only two of the studies[49,50] used qualitative methodology.

The majority of studies had been conducted in hospitals, most frequently across multiple departments (*n* = 25).[49,52,54–58,60,64,66–74,77,78,80,82,83,85,87] Other studies were conducted in nursing homes (*n* = 9),[59,61,62,65,75,79,81,84,88] primary care (*n* = 5)[50,51,53,63,86] or a community pharmacy (*n* = 1)[76]. The majority of studies (*n* = 29)[49,51–53,56–59,62,65,68–71,73–77,79,81–88] used both STOPP and START tools. The 11 studies[50,54,55,60,61,63,64,66,67,72,78] using only the STOPP tool were conducted most frequently in hospitals, in a similar proportion to those including both tools.

Studies carried out in hospitals and nursing homes usually involved more than one professional group. Typically, a pharmacist would administer the STOPP/START tools and a specialist doctor would decide

TABLE 4 Retrieval of documents in bibliographic databases

Database	Number of documents
MEDLINE (all)	338
EMBASE	749
HMIC	5
Social Policy & Practice	3
CINAHL	179
CENTRAL	85
CDSR	0
NHS EED	0
Total	1359
Duplicate results	542
Unique results	817

```
┌────────────────────────────────────┐
│ Records identified from the databases │
│            (n = 1359)                 │
└────────────────────────────────────┘
                │
                │        ┌──────────────────────┐
                ├───────▶│  Duplicates removed    │
                │        │      (n = 542)         │
                ▼        └──────────────────────┘
┌────────────────────────────────────┐
│ Number of records after duplicates   │
│            removed                    │
│            (n = 817)                  │
└────────────────────────────────────┘
                │
                │        ┌──────────────────────┐
                ├───────▶│ Excluded after title/  │
                │        │  abstracts screened    │
                │        │      (n = 720)         │
                ▼        └──────────────────────┘
┌────────────────────────────────────┐
│ Records after titles/abstracts        │
│            screening                  │
│            (n = 97)                   │
└────────────────────────────────────┘
                │
                │        ┌──────────────────────┐
                ├───────▶│ Excluded after full texts│
                │        │       screened          │
                │        │      (n = 63)           │
                ▼        └──────────────────────┘
┌────────────────────────────────────┐
│        Included studies               │
│            (n = 34)                   │
└────────────────────────────────────┘
```
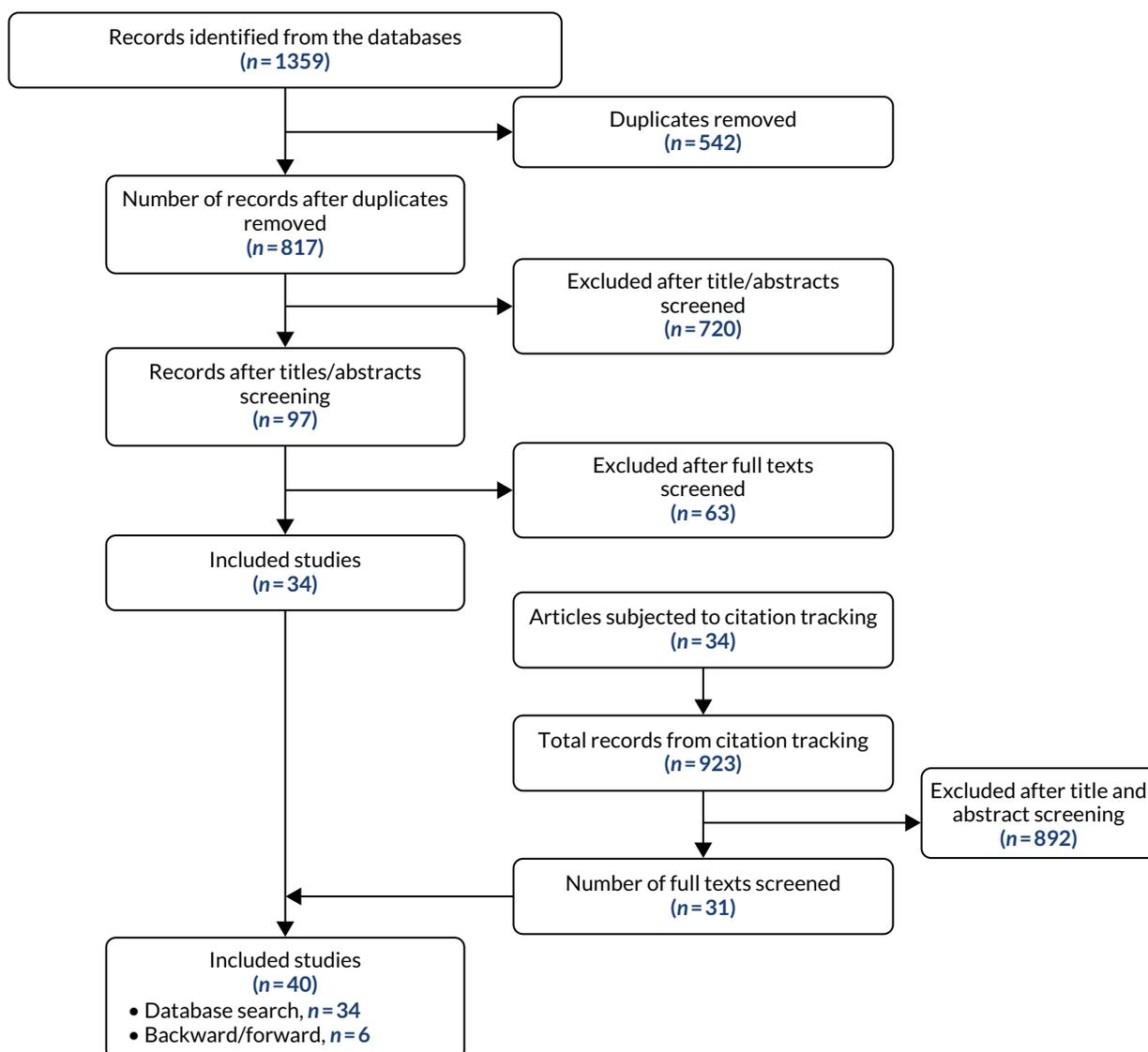

FIGURE 5 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram for documents in phase 2.

whether or not to implement the recommendations. Studies carried out in primary care usually involved one professional, typically a GP, who both administered the STOPP/START tools and made decisions about the implementation of the recommendations.

Critical appraisal

Critical appraisal of the 40 studies was conducted by four independent reviewers (JBG, AD, IRC and Christina Lugones). The quality appraisal of the studies has been presented in separate tables. For non-randomised studies, out of 28 studies, 18 studies[49,56–65,68,69,74,75,77,79,85] were scored as being of high quality and 10 studies[52,66,67,72,76,80,81,86–88] scored as being of medium quality (see *Appendix 9, Table 21*). For RCTs, out of 12 studies, 11 studies[50,51,53–55,70,71,73,82–84] scored as being of high importance and only one study[78] was scored as being of medium importance (see *Appendix 10, Table 22*). Methodological limitations were found in all the studies, including uncontrolled groups, short follow-ups, small sample sizes, potential selection bias and insufficient information on confounding factors. Furthermore, generalisability was limited, as 25[49,52–55,57,58,60–62,64–72,74,77,78,80,82,83,85,87] of the 40 studies were conducted in a single institution or hospital or nursing home or clinical units.

TABLE 5 Study characteristics

Reference	Year	Country	Setting	Tool	Study design	Objective(s)	Quality
Campins et al.[51]	2019	Spain	Primary care	STOPP/START	A multicentre RCT comparing an intervention (medication review) with usual care	To assess the impact of the intervention on the appropriateness of prescribed drugs and associated costs	High
Chandrasekhar et al.[52]	2019	India	Hospital (multiple departments)	STOPP/START	A single-site prospective non-controlled interventional study (medication review and recommendations to managing physician)	To evaluate PIMs/PPOs among hospitalised geriatric patients using STOPP/START, to study morbidity and drug use pattern and to collect feedback from the physicians to improve rational drug therapy.	Medium
Coronado-Vázquez et al.[53]	2019	Spain	Primary care	STOPP/START	A multicentre RCT comparing an intervention (decision support tool) with usual care	To evaluate the effectiveness of the intervention on medication appropriateness	High
Cossette et al.[54]	2017	Canada	Hospital (multiple departments)	STOPP	A single-site RCT comparing an intervention (CAS-based pharmacist–physician intervention) with usual care	To assess the impact of the intervention on PIMs	High
Dalleur et al.[55]	2014	Belgium	Hospital (multiple departments)	STOPP	A single-site RCT comparing an intervention (medication review and recommendations to attending physicians) with usual care	To assess the impact of the intervention on reductions in PIMs for patients at discharge from hospital	High
De Bock et al.[49]	2018	Belgium	Hospital (geriatrics)	STOPP/START	A single-site interventional non-randomised controlled study comparing two interventions based on medication reconciliation, using either GheOPS or STOPP/START, with usual care and collection of participants' experiences	To identify successes of and barriers to the implementation of the intervention	High
Delgado Silveira et al.[56]	2015	Spain	Hospital[a] (multiple departments)	STOPP/START	A multicentre prospective non-controlled interventional study (medication review and recommendations to attending physicians)	To assess the impact of the intervention on drug-related problems and negative outcomes associated with medications. To estimate the prevalence of drug-related problems, negative outcomes associated with medications and PIPs, and the drugs involved	High
Deliens et al.[57]	2016	Belgium	Hospital (geriatric oncology)	STOPP/START	A single-site prospective non-controlled interventional study (medication review and discussion with clinical team)	To assess the impact of the intervention on PIMs and drug–drug interactions	High

TABLE 5 Study characteristics (continued)

Reference	Year	Country	Setting	Tool	Study design	Objective(s)	Quality
Fernández Regueiro et al.[58]	2019	Spain	Hospital (internal medicine)	STOPP/START	A single-site prospective non-controlled interventional study (medication review and recommendations to managing physician)	To assess the impact of the intervention on prescription rates	High
Fog et al.[59]	2017	Norway	Nursing home	STOPP/START	A multicentre prospective non-controlled interventional study (medication review and recommendations to managing physician)	To describe drug-related problems and associated changes in clinical management	High
Frankenthal et al.[82]	2014	Israel	Hospital (geriatrics)	STOPP/START	A single-site RCT comparing an intervention (medication review and recommendations to managing clinicians) with usual care	To assess the impact of the interventions on PIPs, PPOs and other clinical and economic outcomes	High
Gallagher et al.[83]	2011	Ireland	Hospital (multiple departments)	STOPP/START	A single-site RCT comparing an intervention (STOPP/START screening at admission and recommendations to the attending medical team) with usual care	To assess the impact of the intervention on prescribing appropriateness	High
Garay-Bravo et al.[60]	2018	Spain	Hospital (multiple departments)	STOPP	A single-site prospective non-controlled interventional study (review of psychotropic drugs and communication with primary care physician)	To assess the impact of a multidisciplinary intervention on the detection and optimisation of prescriptions of psychotropic drugs	High
García-Caballero et al.[61]	2018	Spain	Nursing home	STOPP	A single-site prospective non-controlled interventional study (computerised standardised reports to relevant physicians)	To determine the types of drugs that originated the largest number of alerts, to evaluate the acceptability of alerts and to estimate the impact of the intervention on resources	High
García-Gollarte et al.[84]	2014	Spain	Nursing home	STOPP/START	A multicentre RCT comparing an education intervention with no intervention	To assess the impact of the intervention on inappropriate prescriptions, health outcomes and resource utilisation	High
Gaubert-Dahan et al.[62]	2019	France	Nursing home	STOPP/START	A single-site prospective non-controlled interventional study (medication review)	To assess the impact of the intervention on inappropriate prescriptions	High
Gibert et al.[63]	2018	France	Primary care	STOPP	A multicentre prospective non-controlled interventional study (educational intervention)	To assess the impact of the intervention on inappropriate prescriptions	High
Gramage Caro et al.[85]	2014	Spain	Hospital (multiple departments)	STOPP/START	A single-site prospective non-controlled interventional study (computer alerts)	To assess the impact of the intervention on PIPs	High

Reference	Year	Country	Setting	Tool	Study design	Objective(s)	Quality
Grion et al.[64]	2016	Italy	Hospital (geriatrics)	STOPP	A single-site prospective non-controlled interventional study (computerised reports for attending physicians)	To assess the impact of the intervention on PIPs	High
Ilić et al.[65]	2015	Serbia	Nursing home	STOPP/START	A multicentre prospective non-controlled interventional study (educational intervention)	To assess the impact of the intervention on appropriateness of prescribing	High
Kimura et al.[66]	2017	Japan	Hospital (multiple departments)	STOPP	A single-site prospective non-controlled interventional study (medication review and discussion with attending physicians)	To assess the impact of the intervention on PIMs	Medium
McNicholl et al.[67]	2017	USA	Hospital (HIV clinic)	STOPP	A single-site prospective observational study (medication review)	To assess PIMs	Medium
Mekdad and Alsayed[68]	2019	Saudi Arabia	Hospital (geriatric cardiology)	STOPP/START	A single site prospective observational study	To estimate PIMs and drug-related problems and to identify risk factors associated with PIMs and drug-related problems	High
Mulvogue et al.[69]	2017	Australia	Hospital (geriatrics)	STOPP/START	A single-site prospective interventional non-controlled study (pharmacist)	To assess the impact of the intervention on PIMs and PPOs	High
Naveiro-Rilo et al.[86]	2014	Spain	Primary care	STOPP/START	Multicentre prospective interventional non-controlled study (medication review)	To assess the impact of the intervention on prescribing and health-related quality of life	Medium
O'Connor et al.[70]	2016	Ireland	Hospital (multiple departments)	STOPP/START	A single-site cluster RCT comparing an intervention (medication review) with usual care	To assess the impact of the intervention on ADRs	High
O'Sullivan et al.[71]	2016	Ireland	Hospital (multiple departments)	STOPP/START	A single-site cluster RCT comparing an intervention (computer-supported medication review) with usual care	To design the computer-supported medication review and to test the impact on its implementation on ADRs	High
Price et al.[50]	2017	Canada	Primary care	STOPP	A multicentre cluster RCT comparing an intervention (CAS) with usual care and interviews with participants	To assess the impact of the intervention on PIMs, and to explain findings and to understand the experiences of the professionals	High
Rossi et al.[72]	2017	Italy	Hospital (geriatrics)	STOPP	A single-site prospective interventional non-controlled study (medication review and discussion with attending physicians)	To assess the impact of the intervention on prescriptions at risk of ADRs	Medium
Santolaya-Perrín et al.[73]	2016	Spain	Hospital (A&E)	STOPP/START	A multicentre RCT comparing an intervention (medication review with recommendations for GPs) with usual care	To determine the prevalence of PIPs	High

TABLE 5 Study characteristics (continued)

Reference	Year	Country	Setting	Tool	Study design	Objective(s)	Quality
Senin Loreto et al.[87]	2013	Spain	Hospital (orthopaedics)	STOPP/START	A single-site prospective interventional non-controlled study (medication review and recommendations for GPs)	To assess the impact of the intervention on PIMs	Medium
Sennesael et al.[74]	2018	Belgium	Hospital (geriatrics)	STOPP/START	A single-site retrospective interventional non-controlled study (screening and discussion with attending physicians)	To assess the impact of the intervention on PIMs and PPOs	High
Silva et al.[75]	2015	Portugal	Nursing home	STOPP/START	A multicentre observational cross-sectional study	To evaluate the need for implementation of pharmaceutical care services	High
Momblona et al.[88]	2011	Spain	Nursing home	STOPP/START	A multicentre prospective interventional non-controlled study (medication review and reports for attending physicians)	To determine the prevalence PIMs and PPOs, and to estimate impact of the intervention on prescribing	Medium
Twigg et al.[76]	2015	UK	Community pharmacies	STOPP/START	A multicentre prospective interventional non-controlled study (medication review and feedback to GPs)	To evaluate the impact of the intervention on falls, adherence, quality of life and costs	Medium
Unutmaz et al.[77]	2018	Turkey	Hospital (geriatrics)	STOPP/START	A single-site retrospective interventional non-controlled study (comprehensive geriatric assessment)	To assess the impact of the intervention on polypharmacy, PIMs, PPOs and costs	High
Van der Linden et al.[78]	2017	Belgium	Hospital (geriatrics)	STOPP	A single-site RCT comparing an intervention (medication review) with usual care	To assess the impact of the intervention on polypharmacy, PIMs, emergency admissions and quality of life	Medium
Weeks et al.[79]	2019	Spain	Nursing home	STOPP/START	A multicentre retrospective interventional controlled study (medication review)	To assess the impact of the intervention on the use of psychotropic medication	High
Whitman et al.[80]	2018	USA	Hospital (geriatric oncology)	STOPP/START	A single-site observational study	To compare medication screening tools for PIMs; and to determine feasibility of interventions	Medium
Zaal et al.[81]	2016	Netherlands	Nursing home	STOPP/START	A multicentre prospective non-controlled interventional study (medication review)	To test the feasibility of the intervention	Medium

A&E, accident and emergency; CAS, computerised alert systems; GheOPS, Ghent Older People's Prescription community Pharmacy Screening; HIV, human immunodeficiency virus; PIM, potentially inappropriate medications; PIP, potentially inappropriate prescribing; PPO, potential prescribing omission.
a Includes a small number of nursing homes.

Overview of mechanisms

We found evidence for each of the three mechanisms (i.e. personalisation of the medication review, systematisation of the medication review process and implementation of evidence) across all the studies (*Figure 6*). The main mechanisms present were standardisation and implementation of evidence, often in combination. Personalisation occurred far less frequently, to the extent that patients were directly involved in the process of using STOPP/START tools, for example through shared decision-making, in a minority of studies. Personalisation was present in 16 studies[1,3,4,6,8–10,12,19,21,31,33,35–37,39] (never on its own and always in combination with other mechanisms). Systematisation was present in 28 studies[1,3–5,8,9,11,13–16,18,19,23–29,31–39] (on its own in three studies and combined with other mechanisms in 26 studies). Finally, implementation of evidence was present in 33 studies[1–7,10–17,20,21,23,24,26,27,29–33,35–39] (on its own in five studies and combined with other mechanisms in 29 studies).

There was a complete absence of information across all of the included studies of either direct measures of success in implementing the mechanisms or of process evaluation of the implementation of the intervention. Therefore, it was not possible to establish if any of the reported mechanisms for any given intervention had been actually successfully implemented. The associations are therefore established with the interventions as intended, rather than as actually implemented.

Only two studies[49,50] explored experiences of participants, but with very limited reporting of findings.[50] In addition, both studies had a very limited scope (i.e. feedback on a specific computer interface) and insufficient reporting.[49]

Evaluation of causal links

Personalisation of the medication review
All the available evidence used in testing this mechanism relied on configurations including at least another mechanism, and in half of the studies both of the other two mechanisms. None of the included studies assessed interventions targeting this mechanism in isolation.
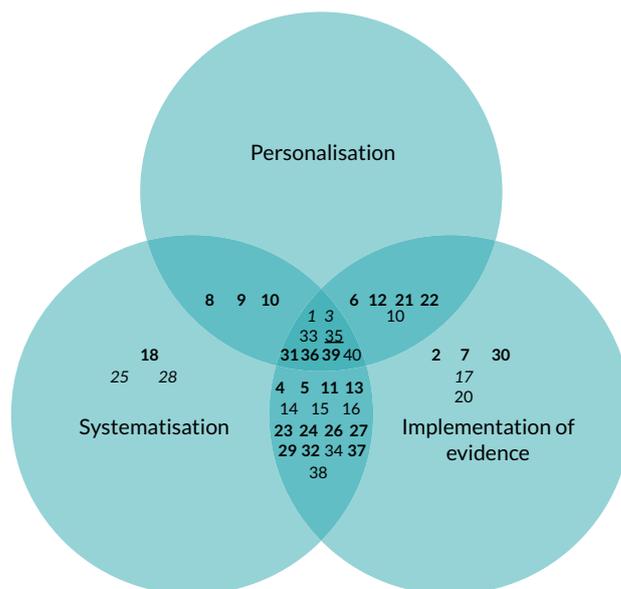
FIGURE 6 Overview of mechanisms across all studies. Bold text, hospital; italicised text, primary care centre; regular text, nursing home; and underlined text, community pharmacy.

The single logic model for first candidate programme theory linked mechanism personalisation of care with the following specific outcomes: patient awareness, empowerment, satisfaction and adherence. Although as many as 16 studies[1,3,4,6,8–10,12,19,21,31,33,35–37,39] reported an intervention relying on this mechanism, only two studies[49,50] considered a relevant outcome in the personalisation logic model. Both of these studies evaluated interventions focusing on the use of both STOPP and START by pharmacists and were designed around both the personalisation and implementation of evidence mechanisms in non-randomised designs. The study by De Bock et al.[49] found positive evidence for improved patient satisfaction in a geriatrics hospital department, whereas the study by Twigg et al.[76] found positive evidence for improved adherence to medical treatment only in the intervention tested in community pharmacies (Figure 7).

The study by Coronado-Vázquez et al.[53] also provided evidence for an association between adherence and appropriateness. Although this was not part of the personalisation logic model, it would support the synergistic effect of interventions using both the mechanisms of personalisation and implementation of evidence.

The two prioritised theories for this mechanism were as follows:

1. STOPP/START tools lead to greater adherence to pharmacological treatment when patients are present at their medication review.
2. STOPP/START tools increase patient empowerment when patient perceptions and expectations are taken into account (elicited and acted on).

It was not possible to test either of these theories, as there was only one study assessing adherence and there was no study assessing patient empowerment.

Systematisation of the medication review process
The single logic model for first candidate programme theory linked mechanism systematisation of care with the following specific outcomes: professional and organisational culture, burden on patient and on professionals, and patient awareness of the medication review process. None of the studies of interventions using this mechanism examined the impact on these outcomes and we therefore could not find any evidence in support of the link between the proposed mechanism and any of the outcomes (Figure 8).
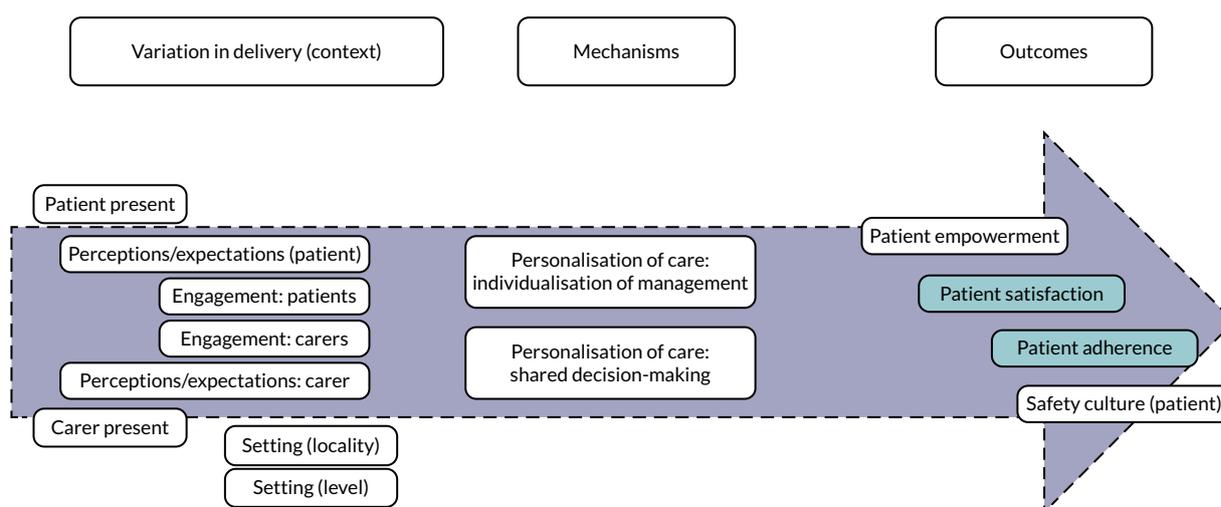
FIGURE 7 Outcomes linked to personalisation across all studies. Light blue shading indicates positive evidence.
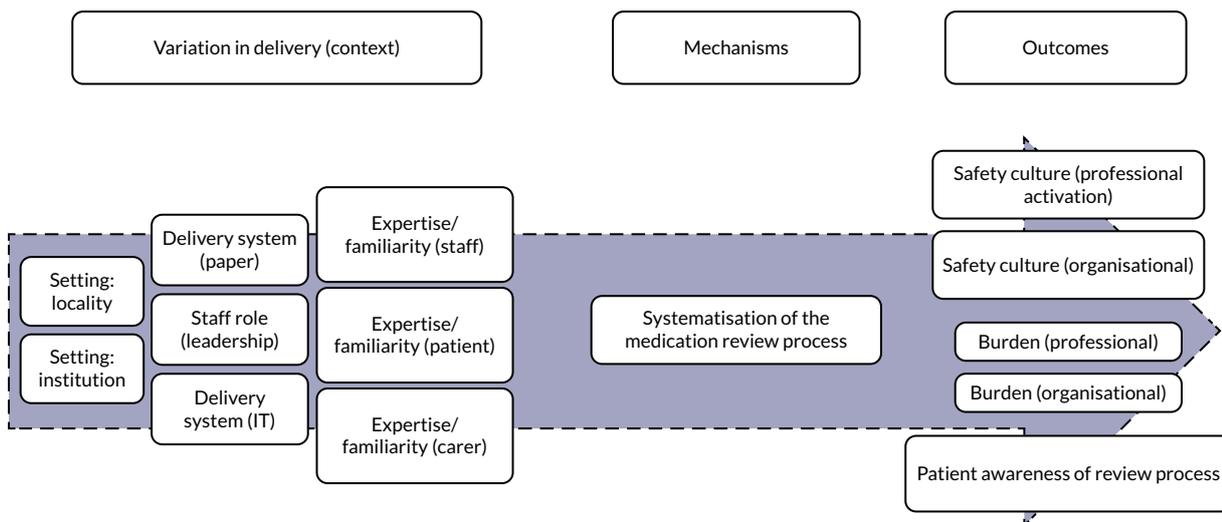
FIGURE 8 Outcomes linked to standardisation across all studies. IT, information technology.

The two prioritised theories for this mechanism were as follows:

1. STOPP/START tools change professional and organisational cultures when there are computerised systems linked to medical records, access to previous STOPP/START tools data, and data-sharing takes place (with relevant professionals, teams and organisations).
2. STOPP/START tools reduce professional and organisational burden in medication management when there is sufficient familiarity with the tool.

It was not possible to test these theories, as there was no study assessing professional and organisational cultures nor professional and organisational burden in medication management.

In the study by Price *et al.*,[50] five physicians provided feedback on their experience with an intervention that focused on the delivery system (i.e. information technology, namely computer alerts). The feedback was oriented mostly at the evaluation of the proposed software solution. More general comments referred to the location on screen and the workflow, which were perceived to be barriers. STOPP alerts had been displayed in a separate location to simple drug alerts, which often meant that the user would need to tab between screens and refresh screens. Participants preferred a single location for all medication-related alerts, regardless of the logic behind those alerts.

Implementation of evidence
The programme theory related to implementation of evidence had a wider set of outcomes than the other mechanisms. More proximal outcomes included appropriateness of prescribing/deprescribing, burden on professionals and safety culture, which would mediate impact on intermediate outcomes [i.e. clinical outcomes, adverse outcomes (ADRs, falls and hospitalisations)], and more distal outcomes (i.e. health-care use, costs, quality of life and mortality).

Nineteen studies[2,3,5–8,11–13,15,18–20,24,28,29,31,32,37] looked at using STOPP/START tools to improve appropriateness of prescribing/deprescribing (all of them, except for four, with mechanism configurations consistent with the logic model for implementation of evidence outcomes). Fourteen studies[3,5,7,8,11–13,15,18–20,24,29,37] found positive evidence (14/19, 74%), whereas the five other studies[4,11,12,26,27] provided negative evidence (*Figure 9*).
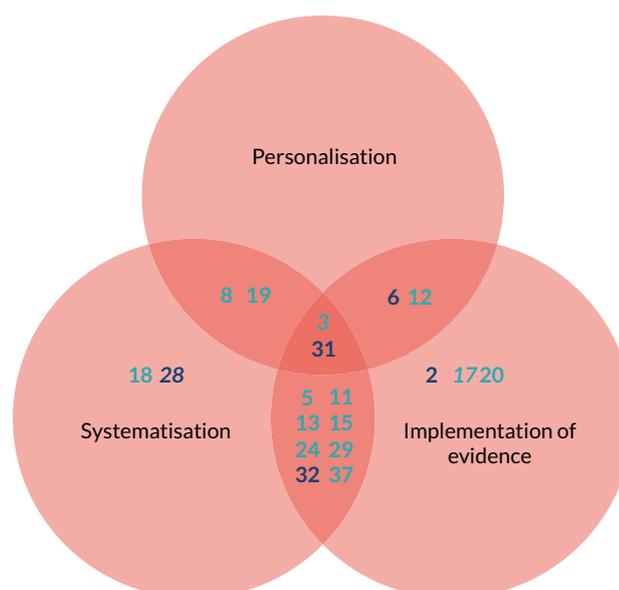
FIGURE 9 Studies focusing on appropriateness of prescribing/deprescribing. Bold text, hospital; dark blue text, negative evidence; italicised text, primary care centre; light blue text, positive evidence; and regular text, nursing home.

Studies providing positive evidence had explored interventions using STOPP/START or only STOPP with a range of different configurations of mechanisms [the majority (14/19)[3,5,6,8,11–13,15,19,24,29,31,32,37] relied on at least two mechanisms]. The studies had all been conducted in hospitals, except for two studies conducted in primary care[17,28] and one study conducted in a nursing home,[20] had used a range of research designs and were mostly considered as being of high quality.

No differences were found in these studies between impact on either appropriateness of prescribing or appropriateness of deprescribing. The five studies[2,6,28,31,32] that found negative evidence had all been conducted in hospitals, used either STOPP/START or only STOPP, had implemented at least two mechanisms and also used both STOPP and START, had used both non-randomised and randomised designs, had medication appropriateness as the primary outcome and had similar quality ratings. The three studies[2,17,20] that studied interventions targeting this mechanism in isolation provided mixed evidence in support of the link.

Three studies[11,26,27] examined the impact of STOPP/START-based interventions on ADEs (*Figure 10*). None of them studied the implementation mechanism in isolation. They all relied on both the implementation of evidence and the systematisation mechanisms, but not the personalisation mechanism, and they all used a randomised design and were of high quality. All three studies found positive evidence of reduction in ADEs.

Five studies[11,15,35,37,38] looked at using either STOPP/START tools or STOPP to reduce falls (*Figure 11*). None of them studied the implementation mechanism in isolation. Two studies took place in nursing homes, two studies took place in hospitals and one study took place in a community pharmacy. The studies relied on both the implementation of evidence and the systematisation mechanism, but not the personalisation mechanism, and used a range of research designs. Three of the studies[11,15,35] provided positive evidence (60%), whereas the other two studies[37,38] provided negative evidence for this link. There were no differences in research design and quality between studies providing positive evidence and those providing negative evidence, except for all negative studies focusing on falls as a secondary rather than a primary outcome.
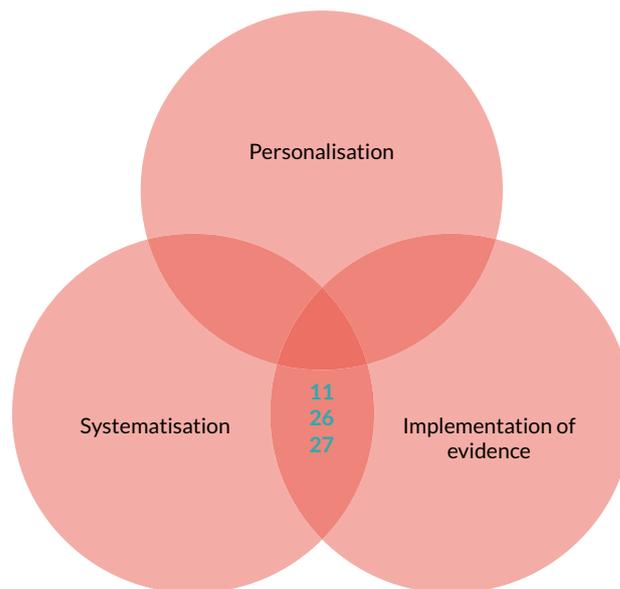
FIGURE 10 Studies focusing on adverse outcomes: ADEs. Light blue text, positive evidence.
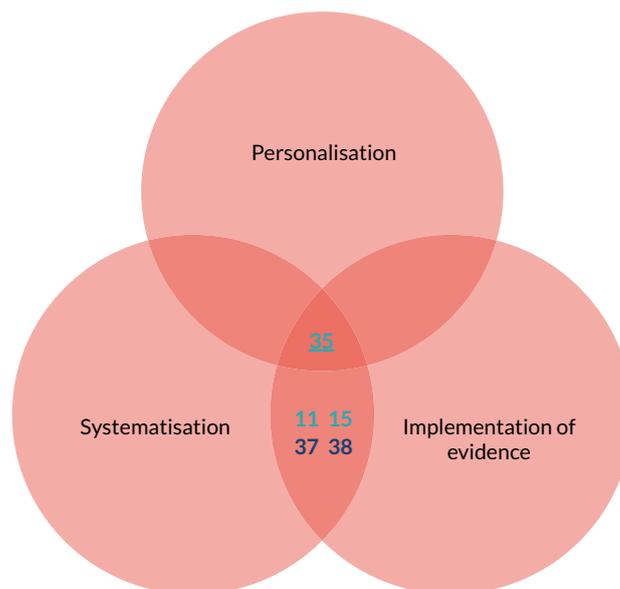
FIGURE 11 Studies focusing on adverse outcomes: falls. Dark blue text, negative evidence; light blue text, positive evidence; and underlined text, community pharmacy.

Eight studies[4,9,11,12,15,26,27,37] examined the impact of using either STOPP/START tools or only STOPP on health-care utilisation (all eight studies used mechanism configurations consistent with the logic model for this outcome, except for one study, and none studied the implementation mechanism in isolation) (*Figure 12*). The studies used a range of randomised and non-randomised research designs. The studies had all been conducted in hospitals [with a single exception (nursing home)] and all of them were considered to be of high quality (with a single exception). The majority of studies[4,9,11,12,26,27] provided negative evidence (6/8, 75%). One of the studies that provided positive evidence was conducted in a nursing home[15] and another study[37] was considered to be of medium quality. Both of these studies used the same configuration of mechanisms, whereas the four other studies[4,11,26,27] provided negative evidence.
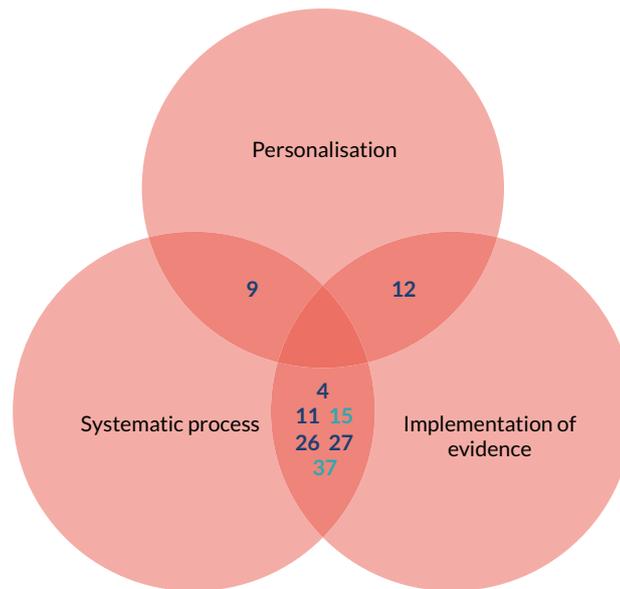
FIGURE 12 Studies focusing on health services utilisation. Dark blue text, negative evidence; light blue text, positive evidence; bold text, hospital; regular text, nursing home.

Four studies[11,25,35,37] examined the impact of STOPP/START-based interventions on quality of life (all of them using a mechanism configuration consistent with the logic model for the outcome, except for one study) (*Figure 13*). Two of the studies[11,37] took place in hospitals, one study took place in a primary care centre[25] and one study took place in a community pharmacy.[35] Three studies[25,35,37] provided positive evidence for this link and one study[11] provided negative evidence. Two studies[35,37] provided positive evidence for STOPP/START-based interventions using non-randomised designs, whereas the third study[25] was a RCT that had evaluated only STOPP. All four studies[11,25,35,37] were considered to be of medium quality. The only study[11] providing negative evidence was a single-site RCT considered to be of high quality.
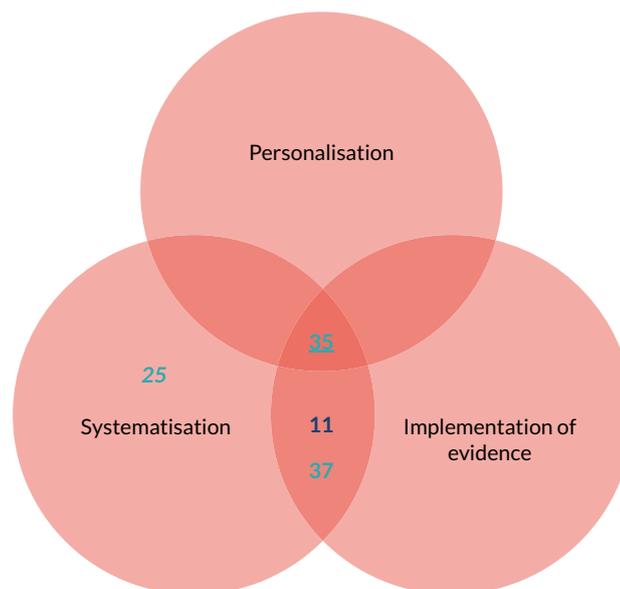
FIGURE 13 Studies focusing on quality of life. Dark blue text, negative evidence; italicised text, primary care centre; light blue text, positive evidence; and underlined text, community pharmacy.

Five studies[9,12,27,37,38] looked at using either STOPP/START tools or STOPP to reduce mortality, all of them as a secondary outcome and all of them using configurations aligned with the logic model (except for one study). None of the studies considered the implementation mechanism in isolation (*Figure 14*). The studies had been conducted in hospitals, except for one study[38] that was conducted in a nursing home. The studies used both randomised and non-randomised designs and were considered to be of mostly high quality. All five studies[9,12,27,37,38] provided negative evidence for the impact of STOPP/START-based interventions on mortality.

The evidence on the economic impact of STOPP/START-based interventions is reviewed separately in *Chapter 6*.

In summary, we were able to confirm pathways from the proposed mechanisms to appropriateness of prescribing/deprescribing, ADEs and quality of life. We found mixed evidence for the impact on falls and we found consistent evidence of lack of impact on health-care use and mortality (*Figure 15*).

The three prioritised theories corresponding to this mechanism were subsequently tested.

STOPP/START-based interventions are associated with increased appropriateness of prescribing/deprescribing when there are lower levels of familiarity with the evidence base (e.g. non-geriatricians, non-pharmacists and more junior professionals)

There was only one study[56] that compared different professional groups, as defined by training/ expertise, applying STOPP/START tools or receiving recommendations from a STOPP/START-based medication review. The study took place primarily in hospitals but also included nursing homes. The acceptance of the intervention was significantly higher for physicians who had a higher degree of academic specialisation in elderly patients. There was, however, no study linking appropriateness of prescribing/deprescribing with training/expertise of health professionals. It was therefore not possible to directly test the theory. Instead, we compared categories of professionals across studies providing positive and negative evidence and either using STOPP/START to make recommendations or receiving such recommendations and acting on them.
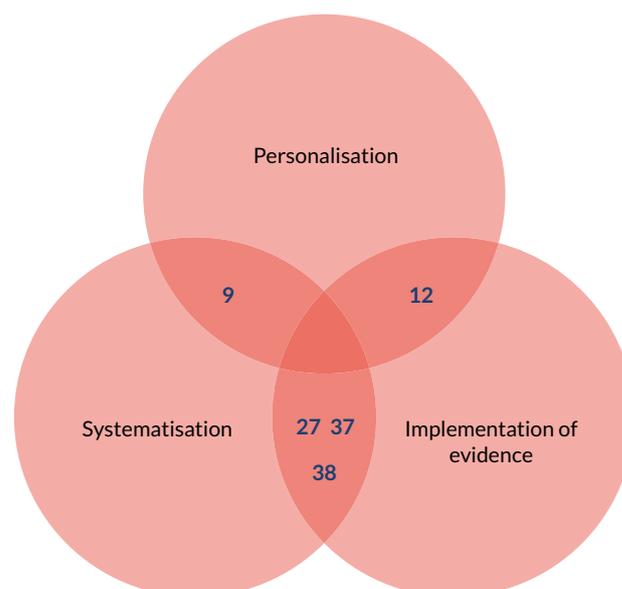
FIGURE 14 Studies focusing on mortality. Dark blue text, negative evidence; bold text, hospital; and regular text, nursing home.
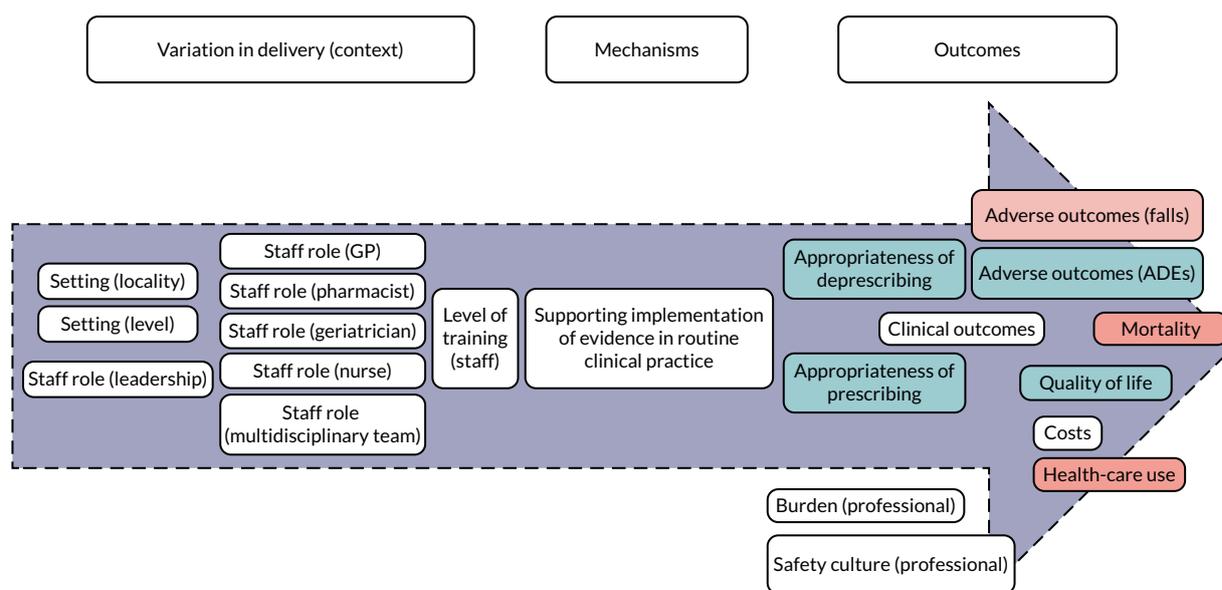
FIGURE 15 Outcomes linked to implementation of evidence across all studies. Light blue shading indicates positive evidence; orange shading indicates negative evidence; light orange shading indicates mixed evidence; and white indicates no evidence available.

In 14 studies[3,5,8,11–13,15,17–20,24,29,37] providing positive evidence, the responsibility in the administration of STOPP/START tools fell on clinical pharmacists ($n = 4$), geriatricians ($n = 3$), GPs ($n = 2$), psychiatrists ($n = 1$) and the research team ($n = 1$), with computerised reports and alerts being used in another three studies (in the additional study, geriatricians were trained in the use of STOPP/START tools but not in a structured way). In five studies[2,6,28,31,32] providing negative evidence, the administration was completed by the research team ($n = 2$), clinical pharmacists ($n = 1$) and computer alerts ($n = 1$). There was therefore no information on the role of junior staff and no evidence that administration by non-geriatricians and non-clinical pharmacists was linked with negative results.

The 14 studies[3,5,8,11–13,15,17–20,24,29,37] providing positive evidence fed back information to geriatricians ($n = 9$), consultants of multiple hospital departments ($n = 3$) and GPs ($n = 3$). The five studies[2,6,28,31,32] providing negative evidence included GPs ($n = 2$), consultants of multiple hospital departments ($n = 1$) and geriatricians ($n = 1$). Therefore, there was again no information on the role of junior staff. Although in this case there was some indication that receipt of the information by non-pharmacists and non-geriatricians was more frequently associated with negative results, it was still the case that a larger number of studies suggested benefits of the provision of STOPP/START information to these groups. This theory was therefore refuted.

Of note, one study[60] showed a greater probability of maintaining recommendations at 3 months in patients whose intervention report was sent to their GP and for those whose electronic prescription was changed, than the remaining patients. This also needs to be put in context with the findings of another study,[49] in which a pharmaceutical hospital discharge letter was sent to the general physician and community pharmacist. Although 17 of 18 pharmacists found the information useful and saved the information in the pharmacy software or patient file, only 7 of 16 GPs were positive. The main reasons to not support this were (1) information was already available in the standard discharge letter, (2) it was unnecessary due to information being available in an upcoming shared electronic patient file and (3) additional administrative burden (another letter to read and file).

STOPP/START-based interventions reduce adverse outcomes through appropriate prescribing/deprescribing

All of the three studies[70,71,82] that assessed the impact of the STOPP/START-based interventions on ADRs provided positive evidence for this link. One study[82] also reported an improvement in the appropriateness of prescribing/deprescribing, and the two other studies[70,71] failed to report any information on this. Of three studies[76,82,84] that demonstrated a reduction on falls, two[82,84] had also demonstrated a positive impact of the intervention in appropriateness of prescribing/deprescribing, but the third[76] failed to report on that particular proximal outcome. Of the other two studies[76,78] that examined the impact of the interventions on falls, one study[78] also provided information on appropriateness of prescribing/deprescribing, and it was for positive evidence of improvement. This study,[78] which had been conducted in geriatric wards in Belgium, was not different from those observing improvements in both appropriateness and reduction in falls in terms of type of mechanism configuration of the intervention, setting or research design, but was of lesser quality and the only study using only STOPP. With the exception of one study,[76] all other studies[78,82,84] reported on both improvement in medication appropriateness and reduction of adverse outcomes. This theory was therefore confirmed.

STOPP/START tools can change professional culture in settings less familiar with the supporting evidence base

It was not possible to test the third theory because no study assessed professional culture.

Variation in care delivery (context)

With the few exceptions presented earlier, setting was the only variable, other than specialisation of health professionals involved, that allowed us to explore variation in care delivery. Studies used a range of configurations of mechanisms across settings. We did not find any clear patterns of association between setting, mechanisms and outcomes.

Hospitals

Twenty-five studies[49,52,54–58,60,64,66–70,72–74,77,78,80,82,83,85,87] were conducted in hospitals. The most common type of intervention was for a pharmacist to administer the STOPP/START tools and for a physician to implement recommendations.

None of these studies assessed outcomes linked to the personalisation mechanism (i.e. patient empowerment, satisfaction, adherence and safety culture), neither did they assess outcomes linked to the systematisation mechanism (i.e. professional and organisational safety culture, professional and organisational burden and patient awareness of the medication review). All the evidence provided in these studies focused on outcomes linked to the implementation of evidence mechanism. They provided positive evidence of improvement in medication appropriateness (10/14)[5,8,9,11–13,19,24,29,37] and reduction in ADEs (3/3),[11,26,27] mixed evidence on improvements in falls (1/2)[11] and quality of life (1/2),[37] and negative evidence of any positive impact on the use of health services (6/7),[4,9,11,12,26,27] and mortality (4/4).[9,27,37,38]

Nursing homes

Nine studies[59,61,62,65,75,79,81,84,88] were conducted in nursing homes. The studies took place in Spain (n = 4),[61,79,84,88] Norway (n = 1),[59] France (n = 1),[62] Serbia (n = 1),[65] Portugal (n = 1)[75] and the Netherlands (n = 1).[81] There was a fairly even split in the administration of STOPP/START tools by pharmacists and physicians, with one study[61] utilising automatic administration via software. As with studies carried out in hospital settings, there was a trend for recommendations to be implemented by physicians, although it was more common for multidisciplinary teams to be involved.

Studies in nursing homes provided positive evidence of the impact on prescribing/deprescribing appropriateness (2/2)[15,20] and service utilisation (1/1),[15] mixed evidence of positive impact on falls (1/2) and quality of life (1/2), and negative evidence of any positive impact on mortality (1/1).[38]

Primary care

Five studies[50,51,53,63,86] were conducted in primary care centres. The studies took place in Spain ($n = 3$),[51,53,86] Canada ($n = 1$)[50] and France ($n = 1$).[63] Typically, GPs both administered STOPP/START tools and implemented the information/recommendations. A study by Campins et al.[51] was the exception, in that an experienced clinical pharmacist evaluated all drugs prescribed using the STOPP/START tools, then discussed recommendations with the physician, which were then discussed with the patient. The studies by Gibert et al.[63] and Naveiro-Rilo et al.[86] were notable in that GPs were given specific training in using the STOPP/START or STOPP tools. In the Naveiro-Rilo et al.[86] study, doctors were trained on the use of the tools over two sessions.

Studies in primary care provided positive evidence of the impact on quality of life (1/1)[25] and mixed evidence of positive impact on prescribing/deprescribing appropriateness (1/2).

Community pharmacies

Only one study[76] was conducted in a community pharmacy, which was notable for being the only one carried out in the UK. The STOPP/START tools were administered by a pharmacist in an intervention that targeted all three mechanisms, and a physician implemented the recommendations. The study found positive evidence for improved adherence to medical treatment and quality of life, and a reduction in falls.[76]

Comparative studies

In a single study[56] that included multiple settings (primarily hospitals but also nursing homes) and evaluated a pharmacist-led intervention, acceptance of the interventions was higher in nursing homes than in hospitals.

Public and patient involvement in theory testing

The PPI advisory team provided support throughout the process of theory testing, revising drafts and providing feedback. In addition, two meetings were convened with the PPI advisory team in phase 2 to provide feedback on various aspects, including rewording of programme theories, the concept of personalisation (see *Appendix 8*, *Box 3*), summary of the results and the *Plain English summary*.

In this meeting, the advisory team also provided advice on dissemination, which is covered in *Chapter 8*.

Summary of findings

The vast majority of studies evaluated interventions targeting multiple mechanisms, and the majority of those that did examined the effect on outcomes other than those identified as mechanism-specific outcomes in the logic model. It was therefore not possible to test the mechanisms in isolation, with the exception of the link between implementation of evidence and appropriateness of prescribing/deprescribing, for which we found mixed evidence (two studies[17,20] providing positive evidence and one providing negative evidence[2]).

When all the studies targeting a mechanism, either in isolation or most frequently in combination with others, we found the following.

Personalisation

- Confirmation of a positive impact on patient satisfaction (hospital) and patient adherence (community pharmacies).
- Neither confirmation nor refutation that STOPP/START-based interventions lead to greater adherence to pharmacological treatment when patients are present at their medication review (no evidence).
- Neither confirmation nor refutation of a positive impact on patient empowerment and safety culture of patients, or, accordingly, that STOPP/START-based interventions increase patient empowerment when patient perceptions and expectations are taken into account (elicited and acted on) (no evidence).

Systematisation

- Neither confirmation nor refutation of a positive impact on professional and organisational culture, burden on patient/professional, or patient awareness of the medication review process (no evidence).
- Accordingly, neither confirmation nor refutation that STOPP/START-based interventions change professional and organisational cultures when there are computerised systems linked to medical records, access to previous STOPP/START data and data-sharing takes place (with relevant professionals, teams and organisations). Likewise, neither confirmation nor refutation that STOPP/START-based interventions reduce professional and organisational burden in medication management when there is sufficient familiarity with the tool (no evidence).

Implementation of evidence

- Confirmation of a positive impact on appropriateness of prescribing/deprescribing (hospitals, nursing homes, primary care), ADEs (hospitals) and quality of life (hospitals, nursing homes, community pharmacy).
- Confirmation that STOPP/START-based interventions reduce adverse outcomes through appropriate prescribing/deprescribing (hospitals, nursing homes).
- Neither confirmation nor refutation for a positive impact on falls (nursing homes).
- Refutation of a positive impact on health-care use (hospitals) and mortality (hospitals, nursing homes).
- Refutation that STOPP/START-based interventions are associated with increased appropriateness of prescribing/deprescribing when there are lower levels of familiarity with the evidence base (non-geriatricians, non-pharmacists and more junior professionals) (hospitals, nursing homes, primary care).

We did not find any evidence that studies targeting more mechanisms achieved better outcomes than those targeting fewer, nor was any other configuration linked with better outcomes.

Chapter 6 Economic evaluation

Overview of types of studies and economic evidence

Our screening processes identified 13 studies[28,49,51,61,62,70,75–77,80–82,89] that explicitly aimed to collect and analyse costs, reported the estimated cost or resource requirements of the intervention, or reported the estimated cost or resource impacts of the changes in medication or monitoring. On closer inspection, one study[49] was found not to contain such information and was deemed to not be an economic study. Another study, a RCT by Gallagher *et al.*,[28] among older hospitalised patients at a single centre in Ireland, also did not report cost differences between the intervention and control groups. However, the Gallagher *et al.*[28] study did report the outcomes of the Medication Appropriateness Index at the level of each Medication Appropriateness Index criterion, including one that captured the proportion of drugs that were not the 'least expensive available'. This showed that, at 6 months' follow-up, 72.4% of drugs taken by patients in the intervention group were the least expensive available, compared with 66% of those in the control group (difference $p < 0.001$). However, the cost savings due to this difference were not estimated. Both of these studies were excluded.[28,49]

Table 6 shows the main study characteristics of the remaining eleven included studies.[51,61,62,70,75–77,80–82,89] They were conducted in nine different countries (Ireland,[70,89] Spain,[51,61] Israel,[82] France,[62] Portugal,[75] Turkey,[77] the UK,[76] the USA[80] and the Netherlands[81]) and based on study data collected between 2013 and 2017 (where reported). Even among these 11 studies,[51,61,62,70,75–77,80–82,89] there was considerable variation between the studies in the interventions evaluated, the implementation of medication recommendations, the care setting and patient group, and the age and degree of polypharmacy among patients.

There was no standard way in which STOPP/START tools were used within the medication review strategies evaluated. In the economic studies, the STOPP/START tools were used as follows:

- Sometimes alone and sometimes in conjunction with other tools for assessing complex needs or the appropriateness of medication.[76,77,81]
- Sometimes using only a subselection of the criteria in the STOPP/START tools.[76]
- The tools were usually administered by a clinical pharmacist and with the patient. Although in some studies, the tools were administered by a geriatrician or other medical professional[62,90] and sometimes without initially involving the patient, using computer software to screen medication lists.[61]

There was also considerable variation in what then happened to the information, how changes in medication were categorised, how medications were considered, by whom, how any changes in medication were agreed/implemented, and how medications were reviewed/monitored. In some studies, it was a one-off passing on of the information to the medical team (e.g. O'Connor *et al.*[90]), whereas in others there was a more comprehensive, ongoing, multidisciplinary process of implementing and monitoring changes in medication.

The care setting and patient group also varied considerably, and included hospital inpatients,[83,90] hospital outpatients,[77,80] older people in residential/nursing homes,[61,62,75,82] community-dwelling adults[51,76] and adults in centres for people with intellectual disabilities.[81]

Finally, there was very wide variation in the age range and mean age of patients in different studies, and the extent of polypharmacy in the recruited study samples. Most studies recruited patients aged either $\geq$ 65 years or $\geq$ 70 years. The mean age of patients in these six studies[51,70,75,80,82,83] ranged from 79 to 84 years. In contrast, one study[81] was not conducted in older people. The study by Zaal *et al.*[81] in

TABLE 6 Main characteristics of included economic studies

Study and country	Intervention	Population and setting	Patients (n)	Drugs (n)	Age (years)
Gaubert-Dahan et al.[62] France	Medication review was conducted by an expert geriatrician with attending physician and pharmacist. A nurse was interviewed prior to the review about medication administration (regarding help with taking medication, crushing medication and resident comments about their medication). The medication review was performed from the electronic medical record with STOPP/START	Nursing home residents	52	NR	84 (SE 9)
Campins et al.[51] Spain	A clinical pharmacist evaluated all prescribed drugs using the GP–GP algorithm and basing their decision about appropriateness on STOPP/START. The pharmacist discussed recommendations with a physician face to face to make final recommendations, which were then discussed with the patient. A final decision was agreed by the physicians with patients in a face-to-face visit	Community-dwelling polymedicated (i.e. receiving eight or more drugs) elderly people	490[a]	≥ 8	≥ 70 (mean 79.1, SD 5.4)
Frankenthal et al.[82] Israel	A medication review by a pharmacist using STOPP/START tools. Recommendations were discussed with a chief physician, who decided whether or not to accept these recommendations and implement prescribing changes	Residents in a chronic care geriatric facility	358	≥ 1	≥ 65 (mean 82.7, SE 8.7)
García-Caballero et al.[61] Spain	A pilot computer tool that automatically processes medication lists, providing the corresponding STOPP alerts and facilitating the elaboration of standardised reports	Entire nursing home population	115	≥ 1 (mean 6.77, SD 2.92)	No age limit (range 46–102, mean 79, SD 11.44)
Gallagher et al.[89] Ireland	A pharmacist applied a structured pharmacist review of medication supported by computerised decision support software	All patients admitted to hospital (for medical or surgical care) via the emergency department	361	0	≥ 65 (median 77, IQR 71–84)
O'Connor et al.[70] Ireland	A single-time point presentation to attending physicians of potentially inappropriate medications according to STOPP/START	Individuals consecutively admitted to hospital	732	8/9	≥ 65 (mean 78 or 80 in each group, IQR 72–84 and 73–85 in each group)

TABLE 6 Main characteristics of included economic studies (*continued*)

Study and country	Intervention	Population and setting	Patients (n)	Drugs (n)	Age (years)
Silva et al.[75] Portugal	A pharmacist conducted structured interviews with each patient and consulted patient medical records to gather demographic data and information on health problems and medications used, with support for drug-related problems identification using PCNE classification, official drug information sources and STOPP/START	Older people in nursing homes	31	≥ 5	≥ 65 (mean 81.65, SE 6.86)
Twigg et al.[76] UK	A pharmacist conducted the medication review, discussed it with the patient and agreed what action should be taken in relation to contacting a GP	Patients taking four or more medicines and using community pharmacies	620	≥ 4	≥ 65
Unutmaz et al.[77] Turkey	A pharmacist conducted the medication review using STOPP/START and CGA (battery of eight scales)	All patients attending the geriatric outpatient clinic at a hospital	1579	Mean 5.3, SE 3.4	No limit (mean 74 and 77, two groups)
Whitman et al.[80] USA	A pharmacist conducted CGA followed by polypharmacy assessment using Beers criteria, STOPP/START and the Medication Appropriateness Index. Deprescribing occurred after discussion with the pharmacist, geriatric oncologist, patient and caregiver	Adult patients with cancer in a university hospital geriatric oncology clinic	26	Mean 12, range 5–24	≥ 65 (mean 81, range 65–92)
Zaal et al.[81] Netherlands	Medication review supported by STRIP, including STOPP/START	Adults with an intellectual disability living in a centralised or dependent setting	27	Mean 10, range 5–29	No limit median 45 (range 18–80)

CGA, comprehensive geriatric assessment; IQR, interquartile range; NR, not reported; PCNE, Pharmaceutical Care Network Europe; SD, standard deviation; SE, standard error; STRIP, Systematic Tool to Reduce Inappropriate Prescribing.
a 245 in the intervention.

the Netherlands used STOPP/START tools in adults with an intellectual disability (mean age 45 years, range 18–80 years). The extent of polypharmacy in study patients/residents varied, including people taking at least one medication,[82] people taking four or more medications[75,76,80,81] and people taking eight or more medications.[51,71] Where reported, the median or mean number of medicines being taken by patients/residents before receiving the intervention varied from 5.3[77] to 12.[80]

Resource use and cost of providing STOPP/START-based interventions

Of the 11 studies[51,61,62,70,75–77,80–82,89] included as economic studies or reporting cost information, only five[51,76,80,81,83] estimated intervention costs (e.g. estimates of the time taken by pharmacists or other

professionals to administer the screening tools, review medications and implementation of any recommended changes in medication). These studies reveal wide variation in estimated staff time requirements for delivering the interventions, in part reflecting important differences in the specific combination of intervention components delivered, what type(s) of clinical staff deliver them and the care setting (e.g. hospital inpatient or community-dwelling older adults). The per patient intervention costs (with country and data year) for these studies was €27 in Campins *et al.*[51] (Spain, 2016), £99 in Twigg *et al.*[76] (UK, 2012), €55 in Gallagher *et al.*[83] (Ireland, 2011), €184 in Zaal *et al.*[81] (the Netherlands, 2016) and US$25 in Whitman *et al.*[80] (USA, 2018). Where stated, pharmacist time accounted for the majority of these costs.

Full economic evaluations

Only two studies[76,89] analysed the costs and effectiveness of two or more alternatives together:

1. Gallagher *et al.*:[89] a cost-effectiveness analysis based on a cluster RCT of pharmacists in Ireland, applying a structured review of medication (based on the STOPP/START tools) among older people admitted to hospital via the emergency department. Incremental costs were calculated and compared with the number of ADRs.
2. Twigg *et al.*:[76] a cost–utility analysis [i.e. estimating the cost per quality-adjusted life-year (QALY)] was conducted of a new service in 25 community pharmacies in England, with older patients who receive four or more medicines receiving a medication review (which used a subset of the STOPP/START tools alongside assessment of falls risk, pain management and medication adherence).[76]

Both studies[76,89] were judged to be of high quality overall (as assessed using the Consensus Health Economic Criteria; see *Appendix 11, Table 22*),[91] although the study by Gallagher *et al.*[89] would have more valid estimates of effectiveness, having been conducted alongside a RCT. Although both studies[76,89] made detailed costings of the intervention (e.g. pharmacist time) and professional time/ appointment costs after medication review, neither study estimated the cost savings/impacts from subsequent altered medication. Exploration of uncertainty was also limited in both studies.

In the Gallagher *et al.* study,[89] in a hospital in Ireland, the STOPP/START-based intervention was both lower in cost and more effective (fewer ADRs) than usual care.[89] The mean total cost of care was €13,250 (standard deviation €15,530) in the intervention group and €15,465 (standard deviation €19,310) in the usual-care group, albeit over the short time horizon of 10 days or time to discharge. There were 50 ADRs in the intervention group and 78 ADRs in the usual-care group (14% and 21% of patients, respectively). The stochastic analysis showed that even with a health system being unwilling to pay anything to prevent ADRs, the intervention would be > 70% likely to be cost-effective.

In the Twigg *et al.* study,[76] in community pharmacies in England, the intervention and cost of health-care use was both more expensive (NHS costs £398.76 vs. £179.41) and more effective than in the 6 months before the intervention, generating 0.007 additional QALYs. The 6-month incremental cost-effectiveness ratio was therefore £32,466 per QALY, which is slightly less cost-effective than would normally be recommended for the NHS (using the typical thresholds for cost-effectiveness used by NICE). The study[76] also presents 12-month results, but purely based on linear extrapolation of the 6-month study results, and so we do not present them here. The analysis cautiously and reasonably concludes that the intervention 'has the potential to be cost-effective if these increases [in QALYs] can be maintained at no further cost'.[76] However, the authors also note that, being based on a before-and-after study, their findings should be treated with caution.

Costs or savings from changing medications

All nine cost analyses and cost comparison studies[51,61,62,70,75,77,80–82] estimated the difference in medication costs due to conducting medication reviews with STOPP/START tools and optimising medication. This was typically calculated as the cost savings due to stopping potentially inappropriate medications less the additional costs of dispensing or prescribing potentially indicated/omitted medications. Given the very wide variety of countries/currencies, data years, interventions, patient groups and methods for estimating these costs, we make no attempt to discuss them or synthesise these findings in any way, but instead have listed them in *Table 7*.

All studies estimated a cost saving to health-care providers based on changes in prescribed or dispensed medications (in most studies, before vs. after the medication review intervention). These estimates, however, exclude the cost of providing the intervention, and this was estimated in three of the studies that also estimated the cost of changing medications.[51,80,81]

TABLE 7 Estimated cost savings due to reviewed and changed medications

Study and country	Year (data)	Design	Saving (increase) in drug costs?	Per patient/resident
Campins *et al.*[51] Spain	2016[a]	Cost comparison	Yes	Intervention: €1062 (95% CI €962 to 1163) Control: €1211 (95% CI €1108 to 1315) Difference: €149 (per year; $p = 0.010$ MWU)
García-Caballero *et al.*[61] Spain	2012	Cost analysis (one arm)	Yes	€32.77 per patient per year
Gaubert-Dahan *et al.*[62] France	2015	Cost analysis (one arm)	Yes	€20.21 ± €31.34 per resident (at 3 months)
O'Connor *et al.*[70] Ireland	NR	Cost comparison	Yes	Intervention: €73.16 (IQR €38.68–121.72) Control: €90.62 (IQR €49.38–162.53) ($p < 0.001$ WRT)
Silva *et al.*[75] Portugal	2014	Cost analysis (one arm)	Yes	Whole study group: €423 Estimated €13.64 per patient
Unutmaz *et al.*[77] Turkey	2016	Cost analysis (one arm)	Yes	US$7.2 per month
Zaal *et al.*[81] Netherlands	2014	Cost analysis (one arm)	Unclear[a]	Assuming a 67% (probably unrealistic) implementation rate of recommended changes: savings of €25,617. Estimated €949 per patient
Frankenthal *et al.*[82] Israel	NR	Cost comparison	Yes	US$29 per patient per month ($p < 0.001$ RMA)
Whitman *et al.*[80] USA	2017[a]	Cost analysis (one arm)	Yes	US$4282

CI, confidence interval; IQR, interquartile range; MWU, Mann–Whitney *U*-test; NR, not reported; RMA, repeated-measure analysis of variance; WRT, Wilcoxon rank-sum test.
a   Data year was not reported but has been inferred from other information in the paper.

Realist consideration of economic studies

The data extraction form for each economic study included data fields for identified causal mechanisms and related contexts to capture any explanations of the variations in costs or cost-effectiveness described in studies (e.g. between subgroups in a study). Unfortunately, all studies were primarily descriptive rather than explanatory, and none examined (or therefore attempted to explain) differences in costs or cost-effectiveness for different subgroups that received the intervention. Furthermore, most of the studies were primarily effectiveness studies that happened to collect and report some cost data, and so cost-related methods, results and interpretation were described in minimal detail. These factors meant that we were unable to perform a realist synthesis of economic studies.

Summary of findings

Although we were unable to perform a realist synthesis of economic studies, it is possible to say, in a less overtly theoretically informed way, what seems to drive the costs and potential cost savings of these medication review-based interventions. The following factors seem most important:

- The time, pay rate and additional training required by (typically) pharmacists to apply the screening tools.
- The time, pay rate and additional training required by physicians/doctors to interpret the information from screening tools.
- The cost of any information technology systems or software that are used to apply the tools or support the decisions made using the information from them (and the degree to which they save time or lead to more effective reviewing and monitoring decisions).
- The complexity and comprehensiveness of the decision-making process for implementing the recommendations from the tools and turning them into plans for stopping inappropriate medication and starting indicated/omitted medications.
- The extent to which suboptimal medication is due to how many medications per patient are inappropriate/harmful/unnecessary medication compared with how many medications per patient are indicated/omitted/needed medications currently not being taken. The higher the level of polypharmacy, the greater the chance that some medicines are clinically inappropriate.
- The cost of these medications (both inappropriate and omitted), for example the mixture of generic compared with new/patented drugs being stopped or started.
- The rates of adherence to prescribed or recommended medicines.
- The benefits of those medications (harms/adverse reactions avoided and health improvements) and the related cost savings/impacts linked to these.

Chapter 7 Discussion

Summary of findings

In this realist synthesis we have developed a programme theory for how interventions based on the use of STOPP/START tools achieve their impact. We have developed a single logic model structured around three key mechanisms (personalisation, systematisation and evidence implementation), have identified mechanism-specific outcomes and have prioritised specific theories that are supported by our Patient Advisory Team.

For testing the programme theories, we have identified 40 studies that have evaluated the impact of STOPP/START-based interventions, mostly in hospital settings but also in nursing homes and primary care. None of these studies was oriented at studying the pathways to impact for the interventions and only two studies[49,50] explored the perspectives of participants in relation to the experience of implementing the intervention.

- We have found evidence providing some support for an impact of the personalisation mechanism of STOPP/START-based interventions on the specific outcomes of patient satisfaction and adherence. We could not test the prioritised theories for this mechanism because of the lack of studies reporting on relevant aspects.
- We have not found any evidence providing support for an impact of the systematisation mechanism of STOPP/START-based interventions on any of the specific outcomes in the logic models, nor have we been able to test the prioritised theories for this mechanism because of a lack of studies reporting on relevant aspects.
- We have found evidence providing support for an impact of the evidence implementation mechanism of STOPP/START-based interventions on the specific outcomes of appropriateness of prescribing/deprescribing, ADEs and quality of life, some support for reduction in falls and consistent evidence on lack of impact on health-care use and mortality. We also found evidence in support of two of the prioritised theories for this mechanism. We were able to confirm the theory that reduction of adverse outcomes was the result of improvement in medication appropriateness. In addition, we were able to refute the theory that administrations by non-geriatricians and non-clinical pharmacists is linked with negative results. We could not test other theories because of the lack of evidence.

We found that most of the interventions were designed to use multiple mechanisms, regardless of the setting, but that the impact was not related to what mechanisms had been targeted or to the setting. We did not find any evidence that studies of interventions targeting more mechanisms achieved better outcomes than those targeting fewer mechanisms, nor was any other configuration linked with better outcomes. We observed that the impact of interventions was linked to the proximity of the selected outcomes to the intervention in the logic model.

Owing to the nature of the available evidence, it was not possible to conduct a realist economic synthesis. We did, however, identify the following key drivers of costs:

- time
- pay rates and additional training required for health professionals to use and interpret STOPP/START tools
- costs of information technology systems or software
- the complexity and comprehensiveness of the decision-making process
- how many medications per patient are inappropriate/harmful/unnecessary compared with how many medications per patient are indicated/omitted/needed and currently not being taken

- the cost of these medications
- the rates of adherence to prescribed or recommended medicines, and the benefits of those medications (harms/adverse reactions avoided and health improvements)
- the related cost savings/impacts linked to these.

In addition, we determined the potential cost savings and impacts linked to these medication review-based interventions.

One intervention looked particularly promising and relevant to the UK context. In this intervention, patients were invited by the community pharmacy team to attend regular consultations with a pharmacist to review their medication using STOPP/START tools, and discuss risk of falls, pain management, adherence and general health. The design of the intervention targeted all three proposed mechanisms and found positive evidence for improved adherence to medical treatment and quality of life, and a reduction in falls. The cost per QALY estimates ranged from £11,885 to £32,466, depending on the assumptions made.

Findings in context

The most recent systematic review, specifically focusing on interventions based on STOPP/START tools, identified four RCTs and concluded that their use may be effective in improving prescribing quality and in improving clinical, humanistic and economic outcomes.[12] We identified a substantially larger number of RCTs. This is not surprising, as most of the studies included had been published after the publication of this review. Within the limits of this previous review, our findings are broadly consistent with it.

Of note, only one study[76] had been conducted in the UK. However, the vast majority of studies had been conducted in either western European countries[49,51,53,55–58,60,61,70,71,73,74,78,79,83–88] or in countries that can be considered to have comparable health systems[50,54,69] (notably Canada and Australia) and population needs[50,54,67,69,80] (the latter two, as well as the USA). In particular, one in three of all studies and three in five of those conducted in primary care were conducted in Spain, a country with a national health service and a similarly strong primary care subsystem.

This review of economic studies of STOPP/START-based interventions strongly echoes the conclusions of a 2014 systematic review of economic evaluations of clinical pharmacist interventions on hospital inpatients (i.e. that the quality of studies is poor, but nevertheless positive cost savings are usually demonstrated).[92]

We found, appraised and summarised the findings of 11 studies that reported the cost of STOPP/START-based interventions. The studies contained considerable variation in the interventions evaluated, the extent of implementation of medication recommendations, the countries, care setting and patient group, and the age and degree of polypharmacy among patients. Therefore, no aggregative synthesis, to produce an average empirical result across studies, is feasible.

Only two studies[76,89] were economic evaluations: (1) a cost-effectiveness study,[89] which showed that the STOPP/START-based intervention reduced both the number of ADRs and reduced costs, and (2) a cost–utility study,[76] which concluded that the intervention had the potential to be cost-effective if the attributed gains in QALYs achieved after 6 months were real, and could be maintained in the longer term.

Nine cost analyses or cost comparison studies[51,61,62,70,75,77,80–82] were also included. Despite the variation in methods, health systems, medication review interventions, patient populations and care settings, they all demonstrated a saving in medication costs due to the STOPP/START-based interventions

(although most studies were based on before vs. after data). However, most did not estimate the cost per patient of receiving the intervention and so whether or not a net cost saving would be achieved is unclear. In contrast, strangely, neither of the full economic evaluations assessed cost savings/changes due to changes in medication.

Strengths

We have conducted and reported our review following the RAMESES II quality and reporting standards[37] (see *Appendix 13*, *Table 24*). We have provided a detailed account of how we applied a realist approach, our search methods and the process of theory elicitation and selection, so that the processes for identifying, elaborating and refining programme theories can be traced. We feel that these strategies enhance transparency and therefore improve the validity of our review. The specific focus of this review has also ensured that we have been able to draw on the broad international literature and triangulate the comprehensiveness in the study identification with experts both within and outside the research team.

The strength of the review of economic evidence lies in the comprehensive methods of searching, data extraction, quality assessment and narrative synthesis conducted by an experienced health economist. A key strength of the evidence is that most of the studies have been conducted in the last 5 years. The two full economic evaluations were well conducted and well reported (the quality assessment did not suffer from unreported information).

Limitations

Some limitations of this work are due to the sources of evidence on which it relies. The lack of any process evaluation or other research on the implementation of the tools beyond the evaluation of the effectiveness has limited our ability to test our programme theory. In particular, the limitations of available qualitative research has made it necessary to rely on the interpretation of researchers conducting the evaluations. In addition, information on the degree to which each of the mechanisms actually operated (e.g. measurement of the degree to which patients and/or professionals are satisfied that the patients were involved in the decision-making process) was notably absent for all mechanisms. The review has largely focused on testing links between mechanism and outcome. The links between context and mechanism (or context and outcome) have not been explored in detail because of the nature of the available evidence; however, context has been explored in the corresponding sections. Our observation that the nature and/or number of mechanisms used by each intervention is not associated with the impact of the intervention needs to be therefore taken with caution, and this would warrant replication in studies using appropriate measures.

The main weaknesses of the review of economic studies are the major heterogeneity in the included studies and the often minimal level of description of the costing methods, resource use and cost results. It would be fair to say that there are not even a pair of studies that are sufficiently similar in type of intervention, patient group/care setting, extent of polypharmacy, outcomes and follow-up, and study design for their results to be meaningfully compared. In all studies that measured the cost impacts of changes in medication, there was always a cost saving. The uncertainty comes from not rigorously and fully assessing other cost impacts (e.g. adverse drug effects avoided, falls averted) and other potential health benefits (e.g. reduced pain, increased quality of life, deaths prevented) in the longer term, and from not using rigorous study designs (e.g. cluster randomised trials).

There are limitations of this work that can be attributed to the methodology used. Realist synthesis is flexible and iterative. We acknowledge that if another team conducted a realist review in this area, they might have identified a different programme theory and therefore tested different ideas and

assumptions about how and in which circumstances STOPP/START-based interventions work. We also acknowledge that we cannot rule out the possibility that an important study was missed from our review. However, for the reasons outlined above, we do not feel that this is likely and the potential implications are likely to be small, as there was consistency in our main observations across the included studies. Finally, we acknowledge that the approach to conducting expert interviews was pragmatic and that time may have been a factor limiting our ability to engage in an in-depth examination, critique and prioritisation of the programme theories.

Implications for future research

The following implications are presented in order of priority.

Interventions
It would be worth replicating the clinical effectiveness and cost-effectiveness study[76] of the intervention delivered through community pharmacies in more settings. It was the only study[76] tested in the UK, one of the few studies that simultaneously considered the three proposed mechanisms of action and one of the studies that demonstrated better results in the use of STOPP/START tools. The nature, scope and target population of the intervention are all particularly relevant to primary care UK economies and offer promising potential in addressing the needs of frail and potentially pre-frail people.

Study design
Studies of the effectiveness of STOPP/START tools have mainly selected outcomes related to appropriateness of prescribing. There has been little exploration of other relevant and longer-term outcomes, such as patient-reported activation, experiences and outcomes, medication concordance or patient safety culture. We have also observed a mismatch in the vast majority of studies between the mechanisms targeted by the interventions and outcomes selection. Studies should be conducted that explore the impact of the tools on outcomes, such as the ones outlined above, and should carefully select outcomes so as to maximise alignment with the hypothesised mechanisms. This would provide more valuable evidence both for testing the single logic model and for informing policy on the potential benefits of routine use of the tools.

We found very limited research using qualitative research to support the evaluation of STOPP/START-based interventions. Qualitative information on experiences of implementing, delivering or being involved in such interventions (e.g. as a health professional, patient or carer) would be a significant addition to the literature, and would provide key information for further testing and refining the proposed logic model.

Economic evaluation
Although savings may have been demonstrated, these were often achieved in the context of uncosted or unreliably costed interventions and short-term capturing of a relatively limited set of patient-relevant outcomes. More high-quality economic evaluations and more comprehensive and fully reported costing studies are required alongside any future clinical effectiveness evaluations. As well as more fully reporting the resources required, and the outcomes and resource impacts produced, studies should endeavour to explain what drives the particular mix and levels of resource use and cost differences compared with usual care. Only when more studies do this better will more valuable realist, explanatory syntheses of cost-effectiveness research be possible.[93]

Primary care
Studies of the effectiveness of STOPP/START tools in PRIMARY CARE consistently evaluated interventions in which the primary care physician both used the tools and implemented the recommendations, whereas hospital-based studies very frequently involved a number of different professionals with a more distributed allocation of tasks. Primary care has first contact and continuity

functions within the health-care system and is the source of the vast majority of chronic prescriptions, thereby providing the natural environment for ensuring continuous optimisation of drug therapies. The further development of the evidence base for the relative merits of different approaches for STOPP/START implementation in primary care would inform health policy for the delivery of safe and effective care. RCTs comparing relevant outcomes for different alternatives would be particularly well suited for developing this evidence base. Of particular interest would be GP-supported enhancement of the previously discussed community pharmacy intervention.[76]

Reporting

The studies identified included little information on context. There is a need for studies of the evaluation of interventions to include greater details on context to facilitate adequate integration of the evidence. The development of consensus-based reporting standards for promoting focused and structured reporting of context for extending existing reporting guidelines, specifically to those relevant to the evaluation on interventions [mainly Consolidated Standards of Reporting Trials (CONSORT), Standards for Quality Improvement Reporting Excellence (SQUIRE) and Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)], would be an important step forward.

Chapter 8 Dissemination

General plan and dissemination to date

The purpose of realist synthesis is to improve the implementation and targeting of STOPP/START-based interventions. Accordingly, dissemination will focus on NHS care professionals and policy-makers responsible for the design and implementation of such interventions. This project will also be of interest to a broad range of stakeholders and will be disseminated through a number of different mechanisms, tailoring to the respective audiences.

We also engaged with our Patient Advisory Group who made specific recommendations. The project Patient Advisory Group had input into the dissemination plans and we explored their views on best approaches to targeting relevant audiences and the consideration of the variety of formats (see *Appendix 12, Table 23*).

Contacts networks
We will showcase our findings through existing contacts networks (including Applied Research Collaborations and Academic Health Sciences Networks) and also to the RPS through a dissemination event (see *Dissemination event*).

Dissemination through and to patients and members of the public
The *Plain English summary* has been created with the help of members of the Patient Advisory Group in the last Patient Advisory Team meeting (see *Appendix 12, Table 23*). We will offer this summary to relevant organisations (including Age UK) and networks where the summary can be fed into pamphlets, newsletters, websites and blogs.

Website
We have already set up and maintain a project webpage [URL: http://medicine.exeter.ac.uk/research/healthresearch/healthservicesandpolicy/projects/qualitysafety/stoppstart/ (accessed 4 November 2020)] on the University of Exeter Medical School website to raise awareness of the project and to publish findings following the completion of the report. This link has been shared to the project team and the scientific advisory board's affiliated universities. We plan to highlight the project findings on the websites of the institutions of the co-applicants following final submission of the report:

- Leeds (URL: www.leeds.ac.uk/)
- Bristol (URL: www.bristol.ac.uk/)
- Queen's University Belfast (URL: www.qub.ac.uk/).

Social networks
We will use social networks such as Twitter (www.twitter.com, Twitter, Inc., San Francisco, CA, USA) to showcase our findings to a wide range of audiences following completion of the report using the hashtag #STOPPSTART. In particular, we will use the accounts @josemvalderas (553 followers), @UoEAPEx (357 followers) and @uemshspr (185 followers).

Peer-reviewed academic journals
At the time of publication, we are intending to submit a manuscript for publication in an open access peer-reviewed academic journal; we have started drafting the manuscript for the journal *Age and Ageing*. Alternative journals will cover pharmacy, general practice, internal medicine and/or geriatrics.

Conference and policy presentations

Conference and policy presentations will be disseminated to audiences of pharmacists and pharmacologists, general practice and hospital-based clinicians, voluntary agencies and policy- and decision-makers. We have disseminated the results of phase 1 at the Annual Society for Academic Primary Care conference held at the University of Exeter in July 2019. An abstract has also been accepted for the International Conference for Realist Research, Evaluation and Synthesis 2021 to be held in Dublin, Ireland.

Dissemination event

A dissemination event took place on 26 February 2020 to showcase the project findings, with invitations extended to the host universities and institutions of the Scientific Expert Advisory Group, co-applicants, RPS, BGS, Collaborations for Leadership in Applied Health Research and Care, South West Academic Health Science Network and local clinicians. This event consisted of a half-day conference, with presentations from patients, an international member of the Scientific Expert Advisory Panel, a co-applicant of the project team and a member of the BGS. The project team reported the findings of the project and presented the guidance on the implementation and utilisation of STOPP/START in clinical decision-making, specifically aimed at clinicians, managers and NHS policy- and decision-makers. A workshop was conducted to explore potential uses and applications, and feedback was gathered on the project and findings throughout the day.

Chapter 9 Conclusions

A range of interventions based on the STOPP/START tools and involving a range of health professionals have been developed and tested, mainly in hospital settings but also in nursing homes and primary care. These interventions operate through three main mechanisms: (1) personalisation of the medication review, (2) systematisation of the review process and (3) implementation of evidence. The impact of the interventions suggests a decreasing effect along a cascade of effects, with strong evidence supporting the benefits of such interventions on proximal prescribing-related outcomes (e.g. appropriateness of medication and ADEs), mixed evidence for impact on falls, health service use and quality of life, and strong evidence for no benefit on mortality. Current evidence on the impact of these interventions does not enable any valid and reliable conclusions about particular configurations of the interventions producing better outcomes in given contexts or care settings. Consideration of the underlying mechanisms can contribute to designing potentially more effective configurations of STOPP/START-based interventions. In all studies that measured the cost of changes in medication, there was a cost reduction after using STOPP/START; however, few also estimated the cost of delivering STOPP/START-based interventions and so it is unclear whether or not there would have been a net saving. Potential longer-term cost impacts (e.g. adverse drug effects avoided, falls averted) or other potential health benefits (e.g. reduced pain, increased quality of life) were not evaluated, and so this further prevents reliable conclusions about cost impacts or cost-effectiveness. Although we were unable to perform a realist synthesis of economic studies from the studies reviewed, we were able to identify a range of basic factors that will influence the costs and potential cost savings of STOPP/START-based interventions.

Acknowledgements

Contributions of authors

Jaheeda Gangannagaripalli (https://orcid.org/0000-0003-1375-9406) was the lead researcher on the project and was involved in all stages of study design, the conduct of phase 1 and the drafting of the report.

Ian Porter was involved in all aspects of phase 2 and contributed to the drafting of the report.

Antoinette Davey (https://orcid.org/0000-0002-5118-6912) was involved in phase 2 and the dissemination efforts, and contributed to the drafting of the report.

Ignacio Ricci Cabello (https://orcid.org/0000-0002-4725-8274) contributed to study design, phases 1 and 2, and contributed to the drafting of the report.

Joanne Greenhalgh (https://orcid.org/0000-0003-2189-8879) contributed to study design, made very significant contributions to phase 1, contributed to phase 2 and to the drafting of the report.

Rob Anderson (https://orcid.org/0000-0002-3523-8559) contributed to study design, phases 1 and 2, led the economic analysis, contributed to the drafting of the report, and provided a critical review and final approval of the report.

Simon Briscoe (https://orcid.org/0000-0002-6982-4521) contributed to study design, designed and implemented the search strategies, contributed to the drafting of the report, and provided a critical review and final approval of the report.

Carmel Hughes (https://orcid.org/0000-0002-4656-6021) contributed to study design, phases 1 and 2, and provided a critical review and final approval of the report.

Rupert Payne (https://orcid.org/0000-0002-5842-4645) contributed to study design, phases 1 and 2, and provided a critical review and final approval of the report.

Emma Cockcroft (https://orcid.org/0000-0003-3798-9492) contributed to study design, led and managed PPI throughout the project, contributed to the drafting of the report, and provided a critical review and final approval of the report.

Jim Harris was a PPI co-applicant, contributed to study design, worked with Emma Cockcroft to run the PPI throughout the project, and provided a critical review and final approval of the report.

Charlotte Bramwell (https://orcid.org/0000-0001-8430-7867) led on the dissemination efforts, contributed to the drafting of the report, and provided a critical review and final approval of the report.

Jose M Valderas (https://orcid.org/0000-0002-9299-1555) was the chief investigator, designed and co-ordinated the delivery of the study, and drafted and finalised the application, protocol and final report.

Data-sharing statement

The authors confirm that all data underlying the findings are fully available without restriction. The authors have made the clinical and economic data sets available through the University of Exeter's Institutional Repository – Open Research Exeter (URL: https://ore.exeter.ac.uk). Access to these data are permitted by controlled requests made via the repository to the chief investigator (Professor Valderas, j.m.valderas@exeter.ac.uk). Although use will be permitted, this will be on the basis that the source of the data are acknowledged (including the funder) and reference is made to the data set name (STOPP/START) and supporting academic references.

References

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 Saf* 2008;**31**:545–56. https://doi.org/10.2165/00002018-200831060-00009

2. Kongkaew C, Noyce PR, Ashcroft DM. Hospital admissions associated with adverse drug reactions: a systematic review of prospective observational studies. *Ann Pharmacother* 2008;**42**:1017–25. https://doi.org/10.1345/aph.1L037

3. Buajordet I, Ebbesen J, Erikssen J, Brørs O, Hilberg T. Fatal adverse drug events: the paradox of drug treatment. *J Intern Med* 2001;**250**:327–41. https://doi.org/10.1046/j.1365-2796.2001.00892.x

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. *Clin Interv Aging* 2014;**9**:2079–86. https://doi.org/10.2147/CIA.S71178

5. Hajjar ER, Cafiero AC, Hanlon JT. Polypharmacy in elderly patients. *Am J Geriatr Pharmacother* 2007;**5**:345–51. https://doi.org/10.1016/j.amjopharm.2007.12.002

6. Goldberg RM, Mabee J, Chan L, Wong S. Drug–drug and drug–disease interactions in the ED: analysis of a high-risk population. *Am J Emerg Med* 1996;**14**:447–50. https://doi.org/10.1016/S0735-6757(96)90147-3

7. Hanlon JT, Pieper CF, Hajjar ER, Sloane RJ, Lindblad CI, Ruby CM, Schmader KE. Incidence and predictors of all and preventable adverse drug reactions in frail elderly persons after hospital stay. *J Gerontol A Biol Sci Med Sci* 2006;**61**:511–15. https://doi.org/10.1093/gerona/61.5.511

8. Aito S, Pieri A, D'Andrea M, Marcelli F, Cominelli E. Preventable and non-preventable risk factors for adverse drug events related to hospital admissions in the elderly: a prospective study. *Clin Drug Investig* 2002;**22**:385–92. https://doi.org/10.2165/00044011-200222060-00006

9. Chan M, Nicklason F, Vial JH. Adverse drug events as a cause of hospital admission in the elderly. *Intern Med J* 2001;**31**:199–205. https://doi.org/10.1046/j.1445-5994.2001.00044.x

10. O'Mahony D, Gallagher PF. Inappropriate prescribing in the older population: need for new criteria. *Age Ageing* 2008;**37**:138–41. https://doi.org/10.1093/ageing/afm189

11. Turner G, Clegg A, British Geriatrics Society. Best practice guidelines for the management of frailty: a British Geriatrics Society, Age UK and Royal College of General Practitioners report. *Age Ageing* 2014;**43**:744–7. https://doi.org/10.1093/ageing/afu138

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. *J Clin Pharm Ther* 2016;**41**:158–69. https://doi.org/10.1111/jcpt.12372

13. Orwig D, Rickles NM, Martin LG. Methodological issues in pharmacotherapy research in older adults. *Am J Geriatr Pharmacother* 2011;**9**:173–89. https://doi.org/10.1016/j.amjopharm.2011.04.008

14. Gallagher P, Barry P, O'Mahony D. Inappropriate prescribing in the elderly. *J Clin Pharm Ther* 2007;**32**:113–21. https://doi.org/10.1111/j.1365-2710.2007.00793.x

15. 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* 2007;**370**:173–84. https://doi.org/10.1016/S0140-6736(07)61091-5

16. 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. *Med Care* 2010;**48**:402–8. https://doi.org/10.1097/MLR.0b013e3181ca3d5d

17. van Mil JW, Westerlund LO, Hersberger KE, Schaefer MA. Drug-related problem classification systems. *Ann Pharmacother* 2004;**38**:859–67. https://doi.org/10.1345/aph.1D182

18. 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**:855–62. https://doi.org/10.1007/s11096-012-9681-8

19. 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. *Br J Clin Pharmacol* 2010;**69**:543–52. https://doi.org/10.1111/j.1365-2125.2010.03628.x

20. Lipton HL, Bero LA, Bird JA, McPhee SJ. Undermedication among geriatric outpatients results of a randomised controlled trial. *Annu Rev Gerontol Geriatr* 1992;**12**:95–108.

21. Cherubini A, Corsonello A, Lattanzio F. Underprescription of beneficial medicines in older people: causes, consequences and prevention. *Drugs Aging* 2012;**29**:463–75. https://doi.org/10.2165/11631750-000000000-00000

22. Byrne S, O'Mahoney D, Hughes C, Parsons C, Patterson S, McCormack B, Finn F. *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*. Belfast: Centre for Ageing Research and Development in Ireland; 2011.

23. Hanlon JT, Schmader KE, Ruby CM, Weinberger M. Suboptimal prescribing in older inpatients and outpatients. *J Am Geriatr Soc* 2001;**49**:200–9. https://doi.org/10.1046/j.1532-5415.2001.49042.x

24. Tangiisuran B, Wright J, Van der Cammen T, Rajkumar C. Adverse drug reactions in elderly: challenges in identification and improving preventative strategies. *Age Ageing* 2009;**38**:358–9. https://doi.org/10.1093/ageing/afp050

25. Garfinkel D, Mangin D. Feasibility study of a systematic approach for discontinuation of multiple medications in older adults: addressing polypharmacy. *Arch Intern Med* 2010;**170**:1648–54. https://doi.org/10.1001/archinternmed.2010.355

26. Fialová D, Topinková E, Gambassi G, Finne-Soveri H, Jónsson PV, Carpenter I, *et al.* Potentially inappropriate medication use among elderly home care patients in Europe. *JAMA* 2005;**293**:1348–58. https://doi.org/10.1001/jama.293.11.1348

27. van der Hooft CS, Jong GW, Dieleman JP, Verhamme KM, van der Cammen TJ, Stricker BH, Sturkenboom MC. Inappropriate drug prescribing in older adults: the updated 2002 Beers criteria – a population-based cohort study. *Br J Clin Pharmacol* 2005;**60**:137–44. https://doi.org/10.1111/j.1365-2125.2005.02391.x

28. 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. https://doi.org/10.5414/cpp46072

29. Rankin A, Cadogan CA, Patterson SM, Kerse N, Cardwell CR, Bradley MC, *et al.* Interventions to improve the appropriate use of polypharmacy for older people. *Cochrane Database Syst Rev* 2018;**9**:CD008165. https://doi.org/10.1002/14651858.CD008165.pub4

30. Patterson SM, Cadogan CA, Kerse N, Cardwell CR, Bradley MC, Ryan C, Hughes C. Interventions to improve the appropriate use of polypharmacy for older people. *Cochrane Database Syst Rev* 2014;**10**:CD008165. https://doi.org/10.1002/14651858.CD008165.pub3

31. 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. *J Clin Pharm Ther* 2013;**38**:360–72. https://doi.org/10.1111/jcpt.12059

32. National Institute for Health and Care Excellence. *Medicines Optimisation: The Safe and Effective Use of Medicines to Enable the Best Possible Outcomes*. 2015. URL: www.nice.org.uk/guidance/ng5/resources/medicines-optimisation-the-safe-and-effective-use-of-medicines-to-enable-the-best-possible-outcomes-pdf-51041805253 (accessed 15 July 2020).

33. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q* 2004;**82**:581–629. https://doi.org/10.1111/j.0887-378X.2004.00325.x

34. 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**. https://doi.org/10.1186/1745-6215-13-95

35. Pawson R. *The Science of Evaluation: A Realist Manifesto*. London: SAGE Publications Ltd; 2014. https://doi.org/10.4135/9781473913820

36. Dalkin SM, Greenhalgh J, Jones D, Cunningham B, Lhussier M. What's in a mechanism? Development of a key concept in realist evaluation. *Implement Sci* 2015;**10**:49. https://doi.org/10.1186/s13012-015-0237-x

37. Wong G, Westhorp G, Manzano A, Greenhalgh J, Jagosh J, Greenhalgh T. RAMESES II reporting standards for realist evaluations. *BMC Med* 2016;**14**:96. https://doi.org/10.1186/s12916-016-0643-1

38. Booth A, Wright J, Briscoe S. Scoping and searching to support realist approaches. *Doing Realis Res* 2018;147–66. https://doi.org/10.4135/9781526451729.n10

39. 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. *Implement Sci* 2011;**6**:69. https://doi.org/10.1186/1748-5908-6-69

40. Manzano A. The craft of interviewing in realist evaluation. *Evaluation* 2016;**22**:342–60. https://doi.org/10.1177/1356389016638615

41. NHS England. *Toolkit for General Practice in Supporting Older People Living with Frailty*. London: NHS England; 2017.

42. National Institute for Health and Care Excellence. *Membership of Multimorbidity Guideline Committee*. 2016. URL: www.nice.org.uk/guidance/ng56/documents/committee-member-list-2 (accessed 3 November 2020).

43. Danermark B, Ekström M, Karlsson JC. Explaining society: critical realism in the social sciences. *Explain Soc Crit Realis Soc Sci* 2019;1–228. https://doi.org/10.4324/9781351017831-1

44. Mogre V, Scherpbier A, Dornan T, Stevens F, Aryee PA, Cherry MG. A realist review of educational interventions to improve the delivery of nutrition care by doctors and future doctors. *Syst Rev* 2014;**3**:148. https://doi.org/10.1186/2046-4053-3-148

45. Critical Appraisal Skills Programme (CASP). *CASP Tools*. Oxford: CASP; 2013.

46. Garg P, Eastwood J, Liaw ST. A realist synthesis of literature informing programme theories for well child care in primary health systems of developed economies. *Int J Integr Care* 2019;**19**:5. https://doi.org/10.5334/ijic.4177

47. Hardwick R, Pearson M, Byng R, Anderson R. The effectiveness and cost-effectiveness of shared care: protocol for a realist review. *Syst Rev* 2013;**2**:12. https://doi.org/10.1186/2046-4053-2-12

48. Stegemann S, Ecker F, Maio M, Kraahs P, Wohlfart R, Breitkreutz J, *et al.* Geriatric drug therapy: neglecting the inevitable majority. *Ageing Res Rev* 2010;**9**:384–98. https://doi.org/10.1016/j.arr.2010.04.005

49. De Bock L, Tommelein E, Baekelandt H, Maes W, Boussery K, Somers A. The introduction of a full medication review process in a local hospital: successes and barriers of a pilot project in the geriatric ward. *Pharmacy* 2018;**6**:E21. https://doi.org/10.3390/pharmacy6010021

50. Price M, Davies I, Rusk R, Lesperance M, Weber J. Applying STOPP guidelines in primary care through electronic medical record decision support: randomized control trial highlighting the importance of data quality. *JMIR Med Inform* 2017;**5**:e15. https://doi.org/10.2196/medinform.6226

51. Campins L, Serra-Prat M, Palomera E, Bolibar I, Martínez MÀ, Gallo P. Reduction of pharmaceutical expenditure by a drug appropriateness intervention in polymedicated elderly subjects in Catalonia (Spain). *Gac Sanit* 2019;**33**:106–11. https://doi.org/10.1016/j.gaceta.2017.09.002

52. Chandrasekhar D, Samjas M, Pattani D. Evaluation of potentially inappropriate medications among hospitalized geriatric patients in tertiary care referral hospital using STOPP/START criteria. *Clin Epidemiol Glob Heal* 2019;**7**:268–73. https://doi.org/10.1016/j.cegh.2018.10.008

53. Coronado-Vázquez V, Gómez-Salgado J, Cerezo-Espinosa de Los Monteros J, Ayuso-Murillo D, Ruiz-Frutos C. Shared decision-making in chronic patients with polypharmacy: an interventional study for assessing medication appropriateness. *J Clin Med* 2019;**8**:E904. https://doi.org/10.3390/jcm8060904

54. Cossette B, Éthier JF, Joly-Mischlich T, Bergeron J, Ricard G, Brazeau S, *et al.* Reduction in targeted potentially inappropriate medication use in elderly inpatients: a pragmatic randomized controlled trial. *Eur J Clin Pharmacol* 2017;**73**:1237–45. https://doi.org/10.1007/s00228-017-2293-4

55. Dalleur O, Boland B, Losseau C, Henrard S, Wouters D, Speybroeck N, *et al.* Reduction of potentially inappropriate medications using the STOPP criteria in frail older inpatients: a randomised controlled study. *Drugs Aging* 2014;**31**:291–8. https://doi.org/10.1007/s40266-014-0157-5

56. Delgado Silveira E, Fernandez-Villalba EM, García-Mina Freire M, Albiñana Pérez MS, Casajús Lagranja MP, Peris Martí JF. [The impact of pharmacy intervention on the treatment of elderly multi-pathological patients.] *Farm Hosp* 2015;**39**:192–202. https://doi.org/10.7399/fh.2015.39.4.8329

57. Deliens C, Deliens G, Filleul O, Pepersack T, Awada A, Piccart M, *et al.* Drugs prescribed for patients hospitalized in a geriatric oncology unit: Potentially inappropriate medications and impact of a clinical pharmacist. *J Geriatr Oncol* 2016;**7**:463–70. https://doi.org/10.1016/j.jgo.2016.05.001

58. Fernández Regueiro R, Estrada Menéndez C, Morís de la Tassa J. Impact of an intervention program to improve potentially inappropriate prescription in hospitalized elderly patients. *Rev Clin Esp* 2019;**219**:375–85. https://doi.org/10.1016/j.rce.2018.12.012

59. Fog AF, Kvalvaag G, Engedal K, Straand J. Drug-related problems and changes in drug utilization after medication reviews in nursing homes in Oslo, Norway. *Scand J Prim Health Care* 2017;**35**:329–35. https://doi.org/10.1080/02813432.2017.1397246

60. Garay-Bravo C, Peña A, Molina M, Sanfeliu J, Piles P, Blasco P, *et al.* Application of the STOPP criteria in hospitalised elderly patients to detect and optimise inappropriate psychopharmaceutical prescriptions. *Eur Geriatr Med* 2018;**9**:597–602. https://doi.org/10.1007/s41999-018-0091-x

61. García-Caballero TM, Lojo J, Menéndez C, Fernández-Álvarez R, Mateos R, Garcia-Caballero A. Polimedication: applicability of a computer tool to reduce polypharmacy in nursing homes. *Int Psychogeriatr* 2018;**30**:1001–8. https://doi.org/10.1017/S1041610217002411

62. Gaubert-Dahan ML, Sebouai A, Tourid W, Fauvelle F, Aikpa R, Bonnet-Zamponi D. The impact of medication review with version 2 STOPP (Screening Tool of Older Person's Prescriptions) and START (Screening Tool to Alert doctors to Right Treatment) criteria in a French nursing home: a 3-month follow-up study. *Ther Adv Drug Saf* 2019;**10**:2042098619855535. https://doi.org/10.1177/2042098619855535

63. Gibert P, Cabaret M, Moulis M, Bosson JL, Boivin JE, Chanoine S, *et al.* Optimizing medication use in elderly people in primary care: impact of STOPP criteria on inappropriate prescriptions. *Arch Gerontol Geriatr* 2018;**75**:16–19. https://doi.org/10.1016/j.archger.2017.10.022

64. Grion AM, Gallo U, Tinjala DD, Daragjati J, Loreggian M, Cardaci G, *et al.* A new computer-based tool to reduce potentially inappropriate prescriptions in hospitalized geriatric patients. *Drugs Aging* 2016;**33**:267–75. https://doi.org/10.1007/s40266-015-0340-3

65. Ilić D, Bukumirić Z, Janković S. Impact of educational intervention on prescribing inappropriate medication to elderly nursing homes residents. *Srp Arh Celok Lek* 2015;**143**:174–9. https://doi.org/10.2298/sarh1504174i

66. Kimura T, Ogura F, Yamamoto K, Uda A, Nishioka T, Kume M, *et al.* Potentially inappropriate medications in elderly Japanese patients: effects of pharmacists' assessment and intervention based on Screening Tool of Older Persons' Potentially Inappropriate Prescriptions criteria ver.2. *J Clin Pharm Ther* 2017;**42**:209–14. https://doi.org/10.1111/jcpt.12496

67. McNicholl IR, Gandhi M, Hare CB, Greene M, Pierluissi E. A pharmacist-led program to evaluate and reduce polypharmacy and potentially inappropriate prescribing in older HIV-positive patients. *Pharmacotherapy* 2017;**37**:1498–506. https://doi.org/10.1002/phar.2043

68. Mekdad SS, Alsayed AA. Quality improvement project to reduce drug-related problems (DRPs) and potentially inappropriate medications (PIMs) in geriatrics cardiac clinic in Saudi Arabia. *Can Geriatr J* 2019;**22**:49–54. https://doi.org/10.5770/cgj.22.338

69. Mulvogue K, Roberts JA, Coombes I, Cottrell N, Kanagarajah S, Smith A. The effect of pharmacists on ward rounds measured by the STOPP/START tool in a specialized geriatric unit. *J Clin Pharm Ther* 2017;**42**:178–84. https://doi.org/10.1111/jcpt.12489

70. O'Connor MN, O'Sullivan D, Gallagher PF, Eustace J, Byrne S, O'Mahony D. Prevention of hospital-acquired adverse drug reactions in older people using screening tool of older persons' prescriptions and screening tool to alert to right treatment criteria: a cluster randomized controlled trial. *J Am Geriatr Soc* 2016;**64**:1558–66. https://doi.org/10.1111/jgs.14312

71. O'Sullivan D, O'Mahony D, O'Connor MN, Gallagher P, Gallagher J, Cullinan S, *et al.* Prevention of adverse drug reactions in hospitalised older patients using a software-supported structured pharmacist intervention: a cluster randomised controlled trial. *Drugs Aging* 2016;**33**:63–73. https://doi.org/10.1007/s40266-015-0329-y

72. Rossi M, Musacchio C, Mello AM, Cericola A, Ferelli E, Arena V, *et al.* Monitoraggio dell'appropriatezza prescrittiva nell'anziano fragile ospedalizzato presso la S.C. Geriatria dell'e.o. Ospedali Galliera di Genova. *G Ital di Farm Clin* 2017;**31**:88–94.

73. Santolaya-Perrín R, Jiménez-Díaz G, Galán-Ramos N, Moreno Carvajal MT, Rodríguez-Camacho JM, Sierra-Sánchez JF, *et al.* A randomised controlled trial on the efficacy of a multidisciplinary health care team on morbidity and mortality of elderly patients attending the emergency department. Study design and preliminary results. *Farm Hosp* 2016;**40**:371–84. https://doi.org/10.7399/fh.2016.40.5.10465

74. Sennesael AL, Dalleur O, Henrard S, Artoisenet C, Schoevaerdts D, Spinewine A. Implementing a screening tool to improve prescribing in hospitalized older patients: a pilot study. *Int J Clin Pharm* 2018;**40**:15–19. https://doi.org/10.1007/s11096-017-0563-y

75. Silva C, Ramalho C, Luz I, Monteiro J, Fresco P. Drug-related problems in institutionalized, polymedicated elderly patients: opportunities for pharmacist intervention. *Int J Clin Pharm* 2015;**37**:327–34. https://doi.org/10.1007/s11096-014-0063-2

76. Twigg MJ, Wright D, Barton GR, Thornley T, Kerr C. The four or more medicines (FOMM) support service: results from an evaluation of a new community pharmacy service aimed at over-65s. *Int J Pharm Pract* 2015;**23**:407–14. https://doi.org/10.1111/ijpp.12196

77. Unutmaz GD, Soysal P, Tuven B, Isik AT. Costs of medication in older patients: before and after comprehensive geriatric assessment. *Clin Interv Aging* 2018;**13**:607–13. https://doi.org/10.2147/CIA.S159966

78. Van der Linden L, Decoutere L, Walgraeve K, Milisen K, Flamaing J, Spriet I, Tournoy J. Combined use of the rationalization of home medication by an adjusted STOPP in older patients (RASP) list and a pharmacist-led medication review in very old inpatients: impact on quality of prescribing and clinical outcome. *Drugs Aging* 2017;**34**:123–33. https://doi.org/10.1007/s40266-016-0424-8

79. Weeks WB, Mishra MK, Curto D, Petersen CL, Cano P, Hswen Y, *et al.* Comparing three methods for reducing psychotropic use in older demented Spanish care home residents. *J Am Geriatr Soc* 2019;**67**:1444–53. https://doi.org/10.1111/jgs.15855

80. Whitman A, DeGregory K, Morris A, Mohile S, Ramsdale E. Pharmacist-led medication assessment and deprescribing intervention for older adults with cancer and polypharmacy: a pilot study. *Support Care Cancer* 2018;**26**:4105–13. https://doi.org/10.1007/s00520-018-4281-3

81. Zaal RJ, Ebbers S, Borms M, Koning Bd, Mombarg E, Ooms P, *et al.* Medication review using a Systematic Tool to Reduce Inappropriate Prescribing (STRIP) in adults with an intellectual disability: a pilot study. *Res Dev Disabil* 2016;**55**:132–42. https://doi.org/10.1016/j.ridd.2016.03.014

82. Frankenthal D, Lerman Y, Kalendaryev E, Lerman Y. Intervention with the screening tool of older persons potentially inappropriate prescriptions/screening tool to alert doctors to right treatment criteria in elderly residents of a chronic geriatric facility: a randomized clinical trial. *J Am Geriatr Soc* 2014;**62**:1658–65. https://doi.org/10.1111/jgs.12993

83. 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. https://doi.org/10.1038/clpt.2011.44

84. García-Gollarte F, Baleriola-Júlvez J, Ferrero-López I, Cuenllas-Díaz Á, Cruz-Jentoft AJ. An educational intervention on drug use in nursing homes improves health outcomes resource utilization and reduces inappropriate drug prescription. *J Am Med Dir Assoc* 2014;**15**:885–91. https://doi.org/10.1016/j.jamda.2014.04.010

85. Gramage Caro T, Crespillo Romeo F, Alvarez Diaz A, Perez Menendez-Condec C, Delgado Silveira E, Munoz Garcia M, Bermejo Vicedo T. Application of electronic alerts of the STOPP criteria in an assisted electronic prescription program. *Eur J Clin Pharma* 2014;**16**:137–41.

86. Naveiro-Rilo JC, Diez-Juárez D, Flores-Zurutuza ML, Molina-Mazo R, Alberte-Pérez C. Intervención en ancianos con multimorbilidad y polimedicados: Resultados en la prescripción y en la calidad de vida. *Rev Calid Asist* 2014;**29**:256–62. https://doi.org/10.1016/j.cali.2014.06.002

87. Senin Loreto D, Franco Bosco B, Baena Mercedes DS. Improvement of the prescription in polymedicated elderly. *Eur J Clin Pharma* 2013;**15**:339–44.

88. Momblona JMS, Ferrer GM, Illamola MR. Aplicación de los nuevos criterios de prescripción inadecuada STOPP-START a pacientes geriátricos institucionalizados. *Farm Atención Primaria* 2011;**9**:2–7. https://doi.org/10.1016/S2172-3761(11)70012-7

89. Gallagher J, O'Sullivan D, McCarthy S, Gillespie P, Woods N, O'Mahony D, Byrne S. Structured pharmacist review of medication in older hospitalised patients: a cost-effectiveness analysis. *Drugs Aging* 2016;**33**:285–94. https://doi.org/10.1007/s40266-016-0348-3

90. O'Connor PJ, Crain AL, Rush WA, Sperl-Hillen JM, Gutenkauf JJ, Duncan JE. Impact of an electronic medical record on diabetes quality of care. *Ann Fam Med* 2005;**3**:300–6. https://doi.org/10.1370/afm.327

91. 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. *Int J Technol Assess Health Care* 2005;**21**:240–5. https://doi.org/10.1017/S0266462305050324

92. Gallagher J, McCarthy S, Byrne S. Economic evaluations of clinical pharmacist interventions on hospital inpatients: a systematic review of recent literature. *Int J Clin Pharm* 2014;**36**:1101–14. https://doi.org/10.1007/s11096-014-0008-9

93. Anderson R, Hardwick R. Realism and resources: Towards more explanatory economic evaluation. *Evaluation* 2016;**22**:323–41. https://doi.org/10.1177/1356389016652742

94. Drummond M, Sculpher M, Torrance G, O'Brien B, Stoddart G. *Methods for the Economic Evaluation of Health Care Programmes*. New York, NY: Oxford University Press; 2005.

Appendix 1 STOPP/START tools version 2

Screening Tool of Older Persons' Prescriptions (STOPP) version 2

The following prescriptions are potentially inappropriate to use in patients aged 65 years and older.

Section A: Indication of medication

1. Any drug prescribed without an evidence-based clinical indication.

2. Any drug prescribed beyond the recommended duration, where treatment duration is well defined.

3. Any duplicate drug class prescription e.g. two concurrent NSAIDs, SSRIs, loop diuretics, ACE inhibitors, anticoagulants (optimisation of monotherapy within a single drug class should be observed prior to considering a new agent).

Section B: Cardiovascular System

1. Digoxin for heart failure with normal systolic ventricular function (no clear evidence of benefit)

2. Verapamil or diltiazem with NYHA Class III or IV heart failure (may worsen heart failure).

3. Beta-blocker in combination with verapamil or diltiazem (risk of heart block).

4. Beta blocker with bradycardia (< 50/min), type II heart block or complete heart block (risk of complete heart block, asystole).

5. Amiodarone as first-line antiarrhythmic therapy in supraventricular tachyarrhythmias (higher risk of side-effects than beta-blockers, digoxin, verapamil or diltiazem)

6. Loop diuretic as first-line treatment for hypertension (safer, more effective alternatives available).

7. Loop diuretic for dependent ankle oedema without clinical, biochemical evidence or radiological evidence of heart failure, liver failure, nephrotic syndrome or renal failure (leg elevation and/or compression hosiery usually more appropriate).

8. Thiazide diuretic with current significant hypokalaemia (i.e. serum K+ < 3.0 mmol/l), hyponatraemia (i.e. serum Na+ < 130 mmol/l) hypercalcaemia (i.e. corrected serum calcium > 2.65 mmol/l) or with a history of gout (hypokalaemia, hyponatraemia, hypercalcaemia and gout can be precipitated by thiazide diuretic)

9. Loop diuretic for treatment of hypertension with concurrent urinary incontinence (may exacerbate incontinence).

10. Centrally-acting antihypertensives (e.g. methyldopa, clonidine, moxonidine, rilmenidine, guanfacine), unless clear intolerance of, or lack of efficacy with, other classes of antihypertensives (centrally-active antihypertensives are generally less well tolerated by older people than younger people)

11. ACE inhibitors or Angiotensin Receptor Blockers in patients with hyperkalaemia.

12. Aldosterone antagonists (e.g. spironolactone, eplerenone) with concurrent potassium-conserving drugs (e.g. ACEI's, ARB's, amiloride, triamterene) without monitoring of serum potassium (risk of dangerous hyperkalaemia i.e. > 6.0 mmol/l – serum K should be monitored regularly, i.e. at least every 6 months).

13. Phosphodiesterase type-5 inhibitors (e.g. sildenafil, tadalafil, vardenafil) in severe heart failure characterised by hypotension i.e. systolic BP < 90 mmHg, or concurrent nitrate therapy for angina (risk of cardiovascular collapse)

Section C: Antiplatelet/Anticoagulant Drugs

1. Long-term aspirin at doses greater than 160mg per day (increased risk of bleeding, no evidence for increased efficacy).

2. Aspirin with a past history of peptic ulcer disease without concomitant PPI (risk of recurrent peptic ulcer).

3. Aspirin, clopidogrel, dipyridamole, vitamin K antagonists, direct thrombin inhibitors or factor Xa inhibitors with concurrent significant bleeding risk, i.e. uncontrolled severe hypertension, bleeding diathesis, recent non-trivial spontaneous bleeding) (high risk of bleeding).

4. Aspirin plus clopidogrel as secondary stroke prevention, unless the patient has a coronary stent(s) inserted in the previous 12 months or concurrent acute coronary syndrome or has a high grade symptomatic carotid arterial stenosis (no evidence of added benefit over clopidogrel monotherapy)

5. Aspirin in combination with vitamin K antagonist, direct thrombin inhibitor or factor Xa inhibitors in patients with chronic atrial fibrillation (no added benefit from aspirin).

6. Antiplatelet agents with vitamin K antagonist, direct thrombin inhibitor or factor Xa inhibitors in patients with stable coronary, cerebrovascular or peripheral arterial disease (No added benefit from dual therapy).

7. Ticlopidine in any circumstances (clopidogrel and prasugrel have similar efficacy, stronger evidence and fewer side-effects).

8. Vitamin K antagonist, direct thrombin inhibitor or factor Xa inhibitors for first deep venous thrombosis without continuing provoking risk factors (e.g. thrombophilia) for > 6 months, (no proven added benefit).

9. Vitamin K antagonist, direct thrombin inhibitor or factor Xa inhibitors for first pulmonary embolus without continuing provoking risk factors (e.g. thrombophilia) for > 12 months (no proven added benefit).

10. NSAID and vitamin K antagonist, direct thrombin inhibitor or factor Xa inhibitors in combination (risk of major gastrointestinal bleeding).

11. NSAID with concurrent antiplatelet agent(s) without PPI prophylaxis (increased risk of peptic ulcer disease).

Section D: Central Nervous System and Psychotropic Drugs

1. TriCyclic Antidepressants (TCAs) with dementia, narrow angle glaucoma, cardiac conduction abnormalities, prostatism, or prior history of urinary retention (risk of worsening these conditions).

2. Initiation of TriCyclic Antidepressants (TCAs) as first-line antidepressant treatment (higher risk of adverse drug reactions with TCAs than with SSRIs or SNRIs).

3. Neuroleptics with moderate-marked antimuscarinic/anticholinergic effects (chlorpromazine, clozapine, flupenthixol, fluphenzine, pipothiazine, promazine, zuclopenthixol) with a history of prostatism or previous urinary retention (high risk of urinary retention).

4. Selective serotonin re-uptake inhibitors (SSRI's) with current or recent significant hyponatraemia i.e. serum Na+ < 130 mmol/l (risk of exacerbating or precipitating hyponatraemia).

5. Benzodiazepines for ≥ 4 weeks (no indication for longer treatment; risk of prolonged sedation, confusion, impaired balance, falls, road traffic accidents; all benzodiazepines should be withdrawn gradually if taken for more than 4 weeks as there is a risk of causing a benzodiazepine withdrawal syndrome if stopped abruptly).

6. Antipsychotics (i.e. other than quetiapine or clozapine) in those with parkinsonism or Lewy Body Disease (risk of severe extra-pyramidal symptoms).

7. Anticholinergics/antimuscarinics to treat extra-pyramidal side-effects of neuroleptic medications (risk of anticholinergic toxicity).

8. Anticholinergics/antimuscarinics in patients with delirium or dementia (risk of exacerbation of cognitive impairment).

9. Neuroleptic antipsychotic in patients with behavioural and psychological symptoms of dementia (BPSD) unless symptoms are severe and other non-pharmacological treatments have failed (increased risk of stroke).

10. Neuroleptics as hypnotics, unless sleep disorder is due to psychosis or dementia (risk of confusion, hypotension, extra-pyramidal side effects, falls).

11. Acetylcholinesterase inhibitors with a known history of persistent bradycardia (< 60 beats/min.), heart block or recurrent unexplained syncope or concurrent treatment with drugs that reduce heart rate such as beta-blockers, digoxin, diltiazem, verapamil (risk of cardiac conduction failure, syncope and injury).

12. Phenothiazines as first-line treatment, since safer and more efficacious alternatives exist (phenothiazines are sedative, have significant anti-muscarinic toxicity in older people, with the exception of prochlorperazine for nausea/vomiting/vertigo, chlorpromazine for relief of persistent hiccoughs and levomepromazine as an anti-emetic in palliative care).

13. Levodopa or dopamine agonists for benign essential tremor (no evidence of efficacy).

14. First-generation antihistamines (safer, less toxic antihistamines now widely available).

Section E: Renal System. The following drugs are potentially inappropriate in older people with acute or chronic kidney disease with renal function below particular levels of eGFR (refer to summary of product characteristics datasheets and local formulary guidelines)

1. Digoxin at a long-term dose greater than 125µg/day if eGFR < 30 ml/min/1.73m2 (risk of digoxin toxicity if plasma levels not measured).

2. Direct thrombin inhibitors (e.g. dabigatran) if eGFR < 30 ml/min/1.73m2 (risk of bleeding).

3. Factor Xa inhibitors (e.g. rivaroxaban, apixaban) if eGFR < 15 ml/min/1.73m2 (risk of bleeding).

4. NSAIDs if eGFR < 50 ml/min/1.73m2 (risk of deterioration in renal function).

5. Colchicine if eGFR < 10 ml/min/1.73m2 (risk of colchicine toxicity)

6. Metformin if eGFR < 30 ml/min/1.73m2 (risk of lactic acidosis).

Section F: Gastrointestinal System

1. Prochlorperazine or metoclopramide with Parkinsonism (risk of exacerbating Parkinsonian symptoms).

2. PPI for uncomplicated peptic ulcer disease or erosive peptic oesophagitis at full therapeutic dosage for > 8 weeks (dose reduction or earlier discontinuation indicated).

3. Drugs likely to cause constipation (e.g. antimuscarinic/anticholinergic drugs, oral iron, opioids, verapamil, aluminium antacids) in patients with chronic constipation where non-constipating alternatives are available (risk of exacerbation of constipation).

4. Oral elemental iron doses greater than 200 mg daily (e.g. ferrous fumarate> 600 mg/day, ferrous sulphate > 600 mg/day, ferrous gluconate> 1800 mg/day; no evidence of enhanced iron absorption above these doses).

Section G: Respiratory System

1. Theophylline as monotherapy for COPD (safer, more effective alternative; risk of adverse effects due to narrow therapeutic index).

2. Systemic corticosteroids instead of inhaled corticosteroids for maintenance therapy in moderate-severe COPD (unnecessary exposure to long-term side-effects of systemic corticosteroids and effective inhaled therapies are available).

3. Anti-muscarinic bronchodilators (e.g. ipratropium, tiotropium) with a history of narrow angle glaucoma (may exacerbate glaucoma) or bladder outflow obstruction (may cause urinary retention).

4. Non-selective beta-blocker (whether oral or topical for glaucoma) with a history of asthma requiring treatment (risk of increased bronchospasm).

5. Benzodiazepines with acute or chronic respiratory failure i.e. pO2 < 8.0 kPa ± pCO2 > 6.5 kPa (risk of exacerbation of respiratory failure).

Section H: Musculoskeletal System

1. Non-steroidal anti-inflammatory drug (NSAID) other than COX-2 selective agents with history of peptic ulcer disease or gastrointestinal bleeding, unless with concurrent PPI or H2 antagonist (risk of peptic ulcer relapse).

2. NSAID with severe hypertension (risk of exacerbation of hypertension) or severe heart failure (risk of exacerbation of heart failure).

3. Long-term use of NSAID (>3 months) for symptom relief of osteoarthritis pain where paracetamol has not been tried (simple analgesics preferable and usually as effective for pain relief)

4. Long-term corticosteroids (>3 months) as monotherapy for rheumatoid arthritis (risk of systemic corticosteroid side-effects).

5. Corticosteroids (other than periodic intra-articular injections for mono-articular pain) for osteoarthritis (risk of systemic corticosteroid side-effects).

6. Long-term NSAID or colchicine (>3 months) for chronic treatment of gout where there is no contraindication to a xanthine-oxidase inhibitor (e.g. allopurinol, febuxostat) (xanthine-oxidase inhibitors are first choice prophylactic drugs in gout).

7. COX-2 selective NSAIDs with concurrent cardiovascular disease (increased risk of myocardial infarction and stroke).

8. NSAID with concurrent corticosteroids without PPI prophylaxis (increased risk of peptic ulcer disease).

9. Oral bisphosphonates in patients with a current or recent history of upper gastrointestinal disease i.e. dysphagia, oesophagitis, gastritis, duodenitis, or peptic ulcer disease, or upper gastrointestinal bleeding (risk of relapse/exacerbation of oesophagitis, oesophageal ulcer, oesophageal stricture).

Section I: Urogenital System

1. Antimuscarinic drugs with dementia, or chronic cognitive impairment (risk of increased confusion, agitation) or narrow-angle glaucoma (risk of acute exacerbation of glaucoma), or chronic prostatism (risk of urinary retention).

2. Selective alpha-1 selective alpha blockers in those with symptomatic orthostatic hypotension or micturition syncope (risk of precipitating recurrent syncope).

Section J. Endocrine System

1. Sulphonylureas with a long duration of action (e.g. glibenclamide, chlorpropamide, glimepiride) with type 2 diabetes mellitus (risk of prolonged hypoglycaemia).

2. Thiazolidenediones (e.g. rosiglitazone, pioglitazone) in patients with heart failure (risk of exacerbation of heart failure)

3. Beta-blockers in diabetes mellitus with frequent hypoglycaemic episodes (risk of suppressing hypoglycaemic symptoms).

4. Oestrogens with a history of breast cancer or venous thromboembolism (increased risk of recurrence).

5. Oral oestrogens without progestogen in patients with intact uterus (risk of endometrial cancer).

6. Androgens (male sex hormones) in the absence of primary or secondary hypogonadism (risk of androgen toxicity; no proven benefit outside of the hypogonadism indication).

Section K: Drugs that predictably increase the risk of falls in older people

1. Benzodiazepines (sedative, may cause reduced sensorium, impair balance).

2. Neuroleptic drugs (may cause gait dyspraxia, Parkinsonism).

3. Vasodilator drugs (e.g. alpha-1 receptor blockers, calcium channel blockers, long-acting nitrates, ACE inhibitors, angiotensin I receptor blockers) with persistent postural hypotension i.e. recurrent drop in systolic blood pressure $\geq$ 20mmHg (risk of syncope, falls).

4. Hypnotic Z-drugs e.g. zopiclone, zolpidem, zaleplon (may cause protracted daytime sedation, ataxia).

Section L: Analgesic Drugs

1. Use of oral or transdermal strong opioids (morphine, oxycodone, fentanyl, buprenorphine, diamorphine, methadone, tramadol, pethidine, pentazocine) as first line therapy for mild pain (WHO analgesic ladder not observed).

2. Use of regular (as distinct from PRN) opioids without concomitant laxative (risk of severe constipation).

3. Long-acting opioids without short-acting opioids for break-through pain (risk of persistence of severe pain).

Section N: Antimuscarinic/Anticholinergic Drug Burden

Concomitant use of two or more drugs with antimuscarinic/anticholinergic properties (e.g. bladder antispasmodics, intestinal antispasmodics, tricyclic antidepressants, first generation antihistamines) (risk of increased antimuscarinic/anticholinergic toxicity).

Screening Tool to Alert to Right Treatment (START) version 2

Unless an elderly patient's clinical status is end-of-life and therefore requiring a more palliative focus of pharmacotherapy, the following drug therapies should be considered where omitted for no valid clinical reason(s). It is assumed that the prescriber observes all the specific contraindications to these drug therapies prior to recommending them to older patients.

Section A: Cardiovascular System

1. Vitamin K antagonists or direct thrombin inhibitors or factor Xa inhibitors in the presence of chronic atrial fibrillation.

2. Aspirin (75 mg – 160 mg once daily) in the presence of chronic atrial fibrillation, where Vitamin K antagonists or direct thrombin inhibitors or factor Xa inhibitors are contraindicated.

3. Antiplatelet therapy (aspirin or clopidogrel or prasugrel or ticagrelor) with a documented history of coronary, cerebral or peripheral vascular disease.

4. Antihypertensive therapy where systolic blood pressure consistently > 160 mmHg and/or diastolic blood pressure consistently >90 mmHg; if systolic blood pressure > 140 mmHg and /or diastolic blood pressure > 90 mmHg, if diabetic.

5. Statin therapy with a documented history of coronary, cerebral or peripheral vascular disease, unless the patient's status is end-of-life or age is > 85 years.

6. Angiotensin Converting Enzyme (ACE) inhibitor with systolic heart failure and/or documented coronary artery disease.

7. Beta-blocker with ischaemic heart disease.

8. Appropriate beta-blocker (bisoprolol, nebivolol, metoprolol or carvedilol) with stable systolic heart failure.

Section B: Respiratory System

1. Regular inhaled β2 agonist or antimuscarinic bronchodilator (e.g. ipratropium, tiotropium) for mild to moderate asthma or COPD.

2. Regular inhaled corticosteroid for moderate-severe asthma or COPD, where FEV1 <50% of predicted value and repeated exacerbations requiring treatment with oral corticosteroids.

3. Home continuous oxygen with documented chronic hypoxaemia (i.e. pO2 < 8.0 kPa or 60 mmHg or SaO2 < 89%)

Section C: Central Nervous System & Eyes

1. L-DOPA or a dopamine agonist in idiopathic Parkinson's disease with functional impairment and resultant disability.

2. Non-TCA antidepressant drug in the presence of persistent major depressive symptoms.

3. Acetylcholinesterase inhibitor (e.g. donepezil, rivastigmine, galantamine) for mild-moderate Alzheimer's dementia or Lewy Body dementia (rivastigmine).

4. Topical prostaglandin, prostamide or beta-blocker for primary open-angle glaucoma.

5. Selective serotonin reuptake inhibitor (or SNRI or pregabalin if SSRI contraindicated) for persistent severe anxiety that interferes with independent functioning.

6. Dopamine agonist (ropinirole or pramipexole or rotigotine) for Restless Legs Syndrome, once iron deficiency and severe renal failure have been excluded.

Section D: Gastrointestinal System

1. Proton Pump Inhibitor with severe gastro-oesophageal reflux disease or peptic stricture requiring dilatation.

2. Fibre supplements (e.g. bran, ispaghula, methylcellulose, sterculia) for diverticulosis with a history of constipation.

Section E: Musculoskeletal System

1. Disease-modifying anti-rheumatic drug (DMARD) with active, disabling rheumatoid disease.

2. Bisphosphonates and vitamin D and calcium in patients taking long-term systemic corticosteroid therapy.

3. Vitamin D and calcium supplement in patients with known osteoporosis and/or previous fragility fracture(s) and/or (Bone Mineral Density T-scores more than -2.5 in multiple sites).

4. Bone anti-resorptive or anabolic therapy (e.g. bisphosphonate, strontium ranelate, teriparatide, denosumab) in patients with documented osteoporosis, where no pharmacological or clinical status contraindication exists (Bone Mineral Density T-scores -> 2.5 in multiple sites) and/or previous history of fragility fracture(s).

5. Vitamin D supplement in older people who are housebound or experiencing falls or with osteopenia (Bone Mineral Density T-score is > -1.0 but < -2.5 in multiple sites).

6. Xanthine-oxidase inhibitors (e.g. allopurinol, febuxostat) with a history of recurrent episodes of gout.

7. Folic acid supplement in patients taking methotexate.

Section F: Endocrine System

1. ACE inhibitor or Angiotensin Receptor Blocker (if intolerant of ACE inhibitor) in diabetes with evidence of renal disease i.e. dipstick proteinuria or microalbuminuria (>30mg/24 hours) with or without serum biochemical renal impairment.

Section G: Urogenital System

1. Alpha-1 receptor blocker with symptomatic prostatism, where prostatectomy is not considered necessary.

2. 5-alpha reductase inhibitor with symptomatic prostatism, where prostatectomy is not considered necessary.

3. Topical vaginal oestrogen or vaginal oestrogen pessary for symptomatic atrophic vaginitis.

Section H: Analgesics

1. High-potency opioids in moderate-severe pain, where paracetamol, NSAIDs or low-potency opioids are not appropriate to the pain severity or have been ineffective.

2. Laxatives in patients receiving opioids regularly.

Section I: Vaccines

1. Seasonal trivalent influenza vaccine annually.

2. Pneumococcal vaccine at least once after age 65 according to national guidelines.

Section I: Vaccines

Appendix 2 Search strategies for phase 1 (identification of theories)

MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations (via Ovid)

Date searched: 7 November 2018.

Date range searched: 1946 to 6 November 2018.

Number of documents: 29.

TABLE 8 Phase 1 MEDLINE database search strategy

#	Terms
1	(STOPP adj2 (criter* or tool*)).tw.
2	(START tool* or START criter*).tw.
3	STOPP START.tw.
4	"Screening Tool of Older Persons' Potentially Inappropriate Prescriptions".tw.
5	"Screening Tool to Alert Doctors to Right Treatment".tw.
6	or/1-5
7	Comment/
8	Letter/
9	Editorial/
10	news/or newspaper article/
11	"Comment on".tw.
12	(letter* adj3 editor*).ti.
13	opinion*.tw.
14	(view or views).tw.
15	comment.cm.
16	or/7-15
17	6 and 16
18	limit 17 to yr = "2008 -Current"

EMBASE (via Ovid)

Date searched: 7 November 2018.

Date range searched: 1974 to 6 November 2018.

Number of documents: 37.

TABLE 9 Phase 1 EMBASE database search strategy

#	Terms
1	(STOPP adj2 (criter* or tool*)).tw.
2	(START tool* or START criter*).tw.
3	STOPP START.tw.
4	"Screening Tool of Older Persons' Potentially Inappropriate Prescriptions".tw.
5	"Screening Tool to Alert Doctors to Right Treatment".tw.
6	or/1-5
7	Comment/
8	Letter/
9	Editorial/
10	publications/
11	"Comment on".tw.
12	(letter* adj3 editor*).ti.
13	opinion*.tw.
14	(view or views).tw.
15	or/7-14
16	6 and 15
17	limit 16 to yr = "2008 -Current"

Health Management Information Consortium (via Ovid)

Date searched: 7 November 2018.

Date range searched: 1979 to July 2018.

Number of documents: 5.

TABLE 10 Phase 1 HMIC database search strategy

#	Terms
1	(STOPP adj2 (criter* or tool*)).tw.
2	(START tool* or START criter*).tw.
3	STOPP START.tw.
4	"Screening Tool of Older Persons' Potentially Inappropriate Prescriptions".tw.
5	"Screening Tool to Alert Doctors to Right Treatment".tw.
6	or/1-5
7	limit 6 to yr = "2008 -Current"

Cumulative Index to Nursing and Allied Health Literature (via EBSCO*host*)

Date searched: 7 November 2018.

Date range searched: 1979 to July 2018.

Number of documents: 5.

TABLE 11 Phase 1 CINAHL database search strategy

#	Terms
1	TI ((STOPP N1 (criter* or tool*))) OR AB ((STOPP N1 (criter* or tool*)))
2	TI (("START tool*" or "START criter*")) OR AB (("START tool*" or "START criter*"))
3	TI "STOPP START" OR AB "STOPP START"
4	TI "Screening Tool of Older Persons' Potentially Inappropriate Prescriptions" OR AB "Screening Tool of Older Persons' Potentially Inappropriate Prescriptions"
5	TI "Screening Tool to Alert Doctors to Right Treatment" OR AB "Screening Tool to Alert Doctors to Right Treatment"
6	S1 OR S2 OR S3 OR S4 OR S5
7	MH "News"
8	MH "Reports"
9	MH "Blogs"
10	TI "comment on" OR AB "comment on"
11	TI (letter N2 editor*) OR AB (letter N2 editor*)
12	TI opinion* OR AB opinion*
13	TI (view or views) OR AB (view or views)
14	S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13
15	S6 AND S14

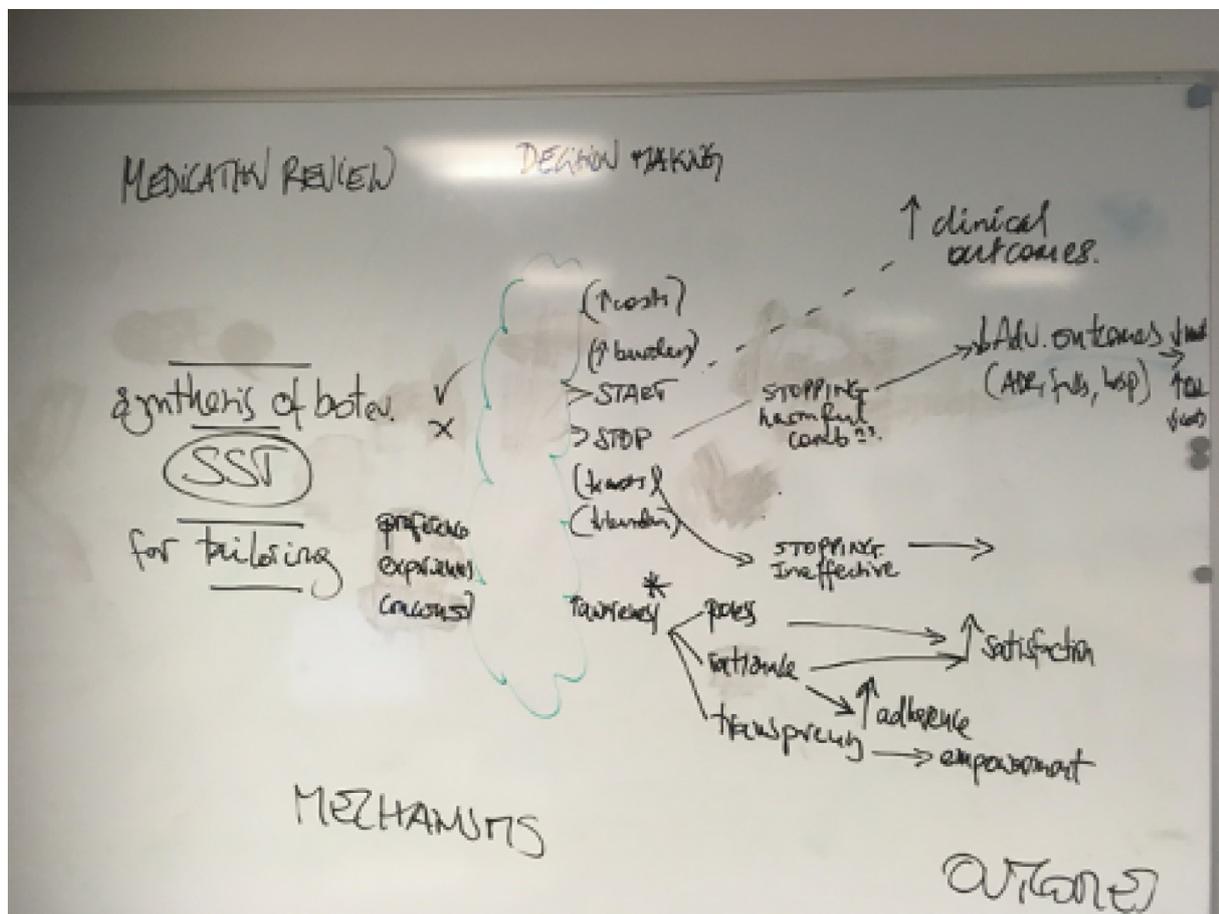# Appendix 3 Single logic model

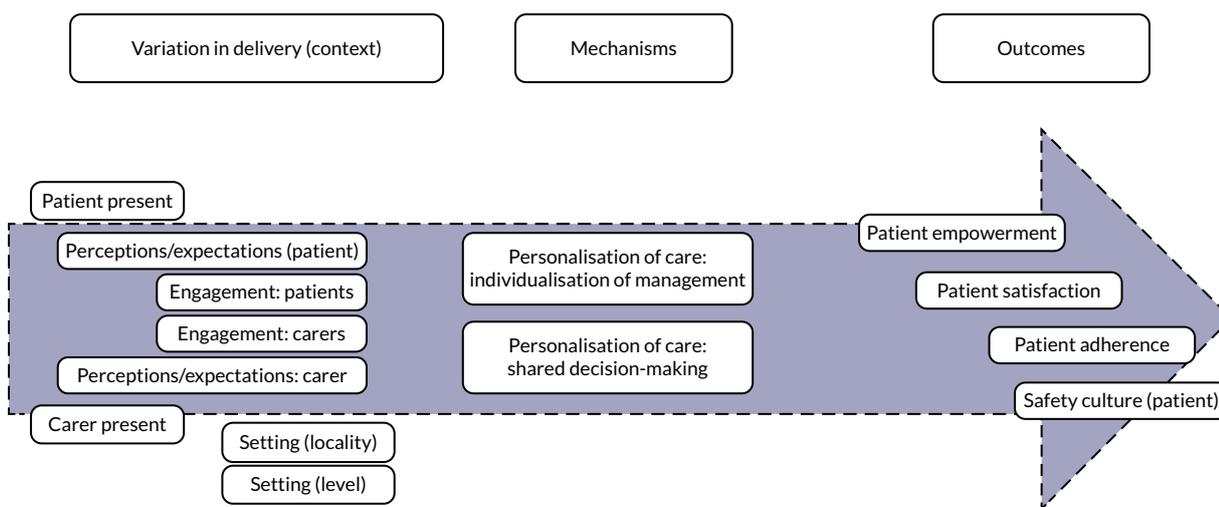FIGURE 16 Single logic model first draft.

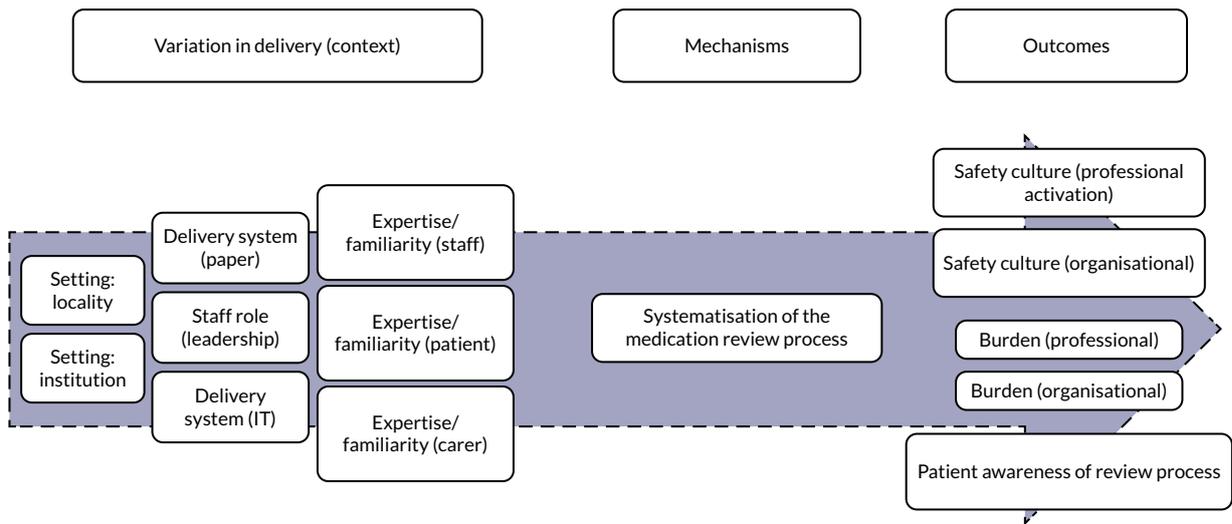FIGURE 17 Logic model for mechanism 1: personalisation.

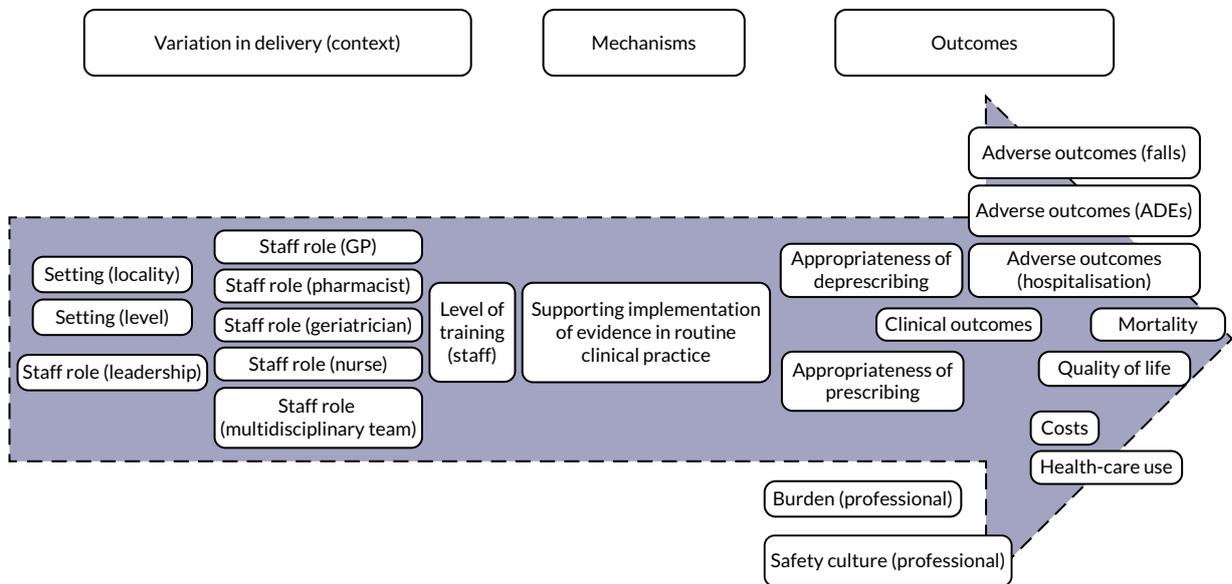FIGURE 18 Logic model for mechanism 2: systematisation. IT, information technology.

FIGURE 19 Logic model for mechanism 3: implementation of evidence.

Appendix 4 Topic guide for expert interviews for phase 1

Introduction

One of the specific objectives of this research is to identify the ideas and assumptions (programme theories) underlying how interventions based on STOPP/START tools are intended to work, for whom, in what circumstances and why. STOPP/START tools are screening tools for supporting medicines optimisation of older adults. A greater understanding of how interventions based on the use of these tools work, for whom they work, in what contexts and why is currently lacking. The purpose of this interview is to help us in understanding and identifying how the STOPP/START tools was intended to work.

Explain what will happen in the interview and afterwards.

> *This is your opportunity to have your say about your views about the STOPP/START-based interventions. The interview should take up to 30 minutes.*

> *I will be recording the interview on a digital recorder to ensure that we have an accurate and detailed record of your views. You are free to stop the discussion and/or the recording at any point. If there are questions that you would prefer not to answer, please let me know and we can move on to the next question.*

> *Are you happy to carry on now and for me to record the discussion?*

Turn on the digital recorder.

Questions

1. When did you become familiar with the STOPP/START tool?
2. What do you know about the STOPP/START tools and drug groups/conditions the tools are applied to?
3. What are your views on the use of STOPP/START tools interventions?

   - Probe: what interventions are available/you are aware of?
   - Probe: what do you think about the impact of the interventions? (For example, prescribing appropriateness?)
   - Probe: how can we overcome the blockages/barriers?

4. What are the (anticipated) practicalities of using the STOPP/START tools in clinical practice?

   - Probe: How does using it affect the way you practise?

5. Do you think the STOPP/START tools will enable you to involve patients in decision-making about their health care?

   - Probe: What do you see as key elements for the use of STOPP/START tools?

6. What are the main problems with using STOPP/START? Do you have any suggestions how we address them?
7. What are the necessary conditions that STOPP/START-based interventions result in improvement of care?
8. How do you think STOPP/START-based interventions affect the way the multidisciplinary team work?
9. Are you aware of any other similar tools? In your view how appropriate is STOPP/START compared with other tools? (For example, Beers criteria?)
10. In your experience, how do STOPP/START tools work in different settings, for different populations and when used by different professionals?
11. Do you have any suggestions on how to improve the use and implementation of STOPP/START tools?
12. Do you have any final comments on the use of STOPP/START tools?

Thank you for taking part in this interview. Your comments have been most helpful.

Appendix 5 Search strategies for phase 2 (testing of programme theories)

MEDLINE and MEDLINE In-Process & Other Non-Indexed Citations (via Ovid)

Date searched: 12 June 2019.

Date range searched: 1946 to 11 June 2019.

Number of documents: 338.

TABLE 12 Phase 2 MEDLINE database search strategy

#	Terms
1	(STOPP adj2 (criter* or tool*)).tw.
2	(START tool* or START criter*).tw.
3	STOPP START.tw.
4	"Screening Tool of Older Persons' Potentially Inappropriate Prescriptions".tw.
5	"Screening Tool to Alert Doctors to Right Treatment".tw.
6	or/1-5
7	limit 6 to yr = "2008 –Current"

EMBASE (via Ovid)

Date searched: 12 June 2019.

Date range searched: 1974 to 11 June 2019.

Number of documents: 749.

TABLE 13 Phase 2 EMBASE database search strategy

#	Terms
1	(STOPP adj2 (criter* or tool*)).tw.
2	(START tool* or START criter*).tw.
3	STOPP START.tw.
4	"Screening Tool of Older Persons' Potentially Inappropriate Prescriptions".tw.
5	"Screening Tool to Alert Doctors to Right Treatment".tw.
6	or/1-5
7	limit 6 to yr = "2008 –Current"

Health Management Information Consortium (via Ovid)

Date searched: 11 June 2019.

Date range searched: 1979 to April 2019.

Number of documents: 5.

TABLE 14 Phase 2 HMIC database search strategy

#	Terms
1	(STOPP adj2 (criter* or tool*)).tw.
2	(START tool* or START criter*).tw.
3	STOPP START.tw.
4	"Screening Tool of Older Persons' Potentially Inappropriate Prescriptions".tw.
5	"Screening Tool to Alert Doctors to Right Treatment".tw.
6	or/1-5
7	limit 6 to yr = "2008 -Current"

Social Policy & Practice database (via Ovid)

Date searched: 11 June 2019.

Data parameters: 201904.

Number of documents: 3.

TABLE 15 Phase 2 Social Policy & Practice database search strategy

#	Terms
1	(STOPP adj2 (criter* or tool*)).tw.
2	(START tool* or START criter*).tw.
3	STOPP START.tw.
4	"Screening Tool of Older Persons' Potentially Inappropriate Prescriptions".tw.
5	"Screening Tool to Alert Doctors to Right Treatment".tw.
6	or/1-5
7	limit 6 to yr = "2008 -Current"

Cumulative Index to Nursing and Allied Health Literature (via EBSCO*host*)

Date searched: 11 June 2019.

Date range searched: not applicable.

Number of documents: 179.

TABLE 16 Phase 2 CINAHL database search strategy

#	Terms
1	TI ((STOPP N1 (criter* or tool*))) OR AB ((STOPP N1 (criter* or tool*)))
2	TI (("START tool*" or "START criter*")) OR AB (("START tool*" or "START criter*"))
3	TI "STOPP START" OR AB "STOPP START"
4	TI "Screening Tool of Older Persons' Potentially Inappropriate Prescriptions" OR AB "Screening Tool of Older Persons' Potentially Inappropriate Prescriptions"
5	TI "Screening Tool to Alert Doctors to Right Treatment" OR AB "Screening Tool to Alert Doctors to Right Treatment"
6	S1 OR S2 OR S3 OR S4 OR S5
Notes: Search results limited 2008 to date.	

Cochrane Central Register of Controlled Trials (via Cochrane Library)

Date searched: 11 June 2019.

Date range searched: not applicable.

Number of documents: 85.

TABLE 17 Phase 2 CENTRAL database search strategy

#	Terms
1	(STOPP NEAR/2 (criter* or tool*)):ti OR (STOPP NEAR/2 (criter* or tool*)):ab
2	("START tool*" or "START criter*"):ti OR ("START tool*" or "START criter*"):ab
3	("STOPP START"):ti OR ("STOPP START"):ab
4	("Screening Tool of Older Persons' Potentially Inappropriate Prescriptions"):ti OR ("Screening Tool of Older Persons' Potentially Inappropriate Prescriptions"):ab
5	("Screening Tool to Alert Doctors to Right Treatment"):ti OR ("Screening Tool to Alert Doctors to Right Treatment"):ab
6	#1 OR #2 OR #3 OR #4 OR #5 with Publication Year from 2008 to 2019

Cochrane Database of Systematic Reviews (via Cochrane Library)

Date searched: 11 June 2019.

Date range searched: not applicable.

Number of documents: 0.

TABLE 18 Phase 2 CDSR database search strategy

#	Terms
1	(STOPP NEAR/2 (criter* or tool*)):ti OR (STOPP NEAR/2 (criter* or tool*)):ab
2	("START tool*" or "START criter*"):ti OR ("START tool*" or "START criter*"):ab
3	("STOPP START"):ti OR ("STOPP START"):ab
4	("Screening Tool of Older Persons' Potentially Inappropriate Prescriptions"):ti OR ("Screening Tool of Older Persons' Potentially Inappropriate Prescriptions"):ab
5	("Screening Tool to Alert Doctors to Right Treatment"):ti OR ("Screening Tool to Alert Doctors to Right Treatment"):ab
6	#1 OR #2 OR #3 OR #4 OR #5 with Publication Year from 2008 to 2019

NHS Economic Evaluation Database (via Centre for Reviews and Dissemination)

Date searched: 11 June 2019.

Date range searched: not applicable.

Number of documents: 0.

Record date: 2008–19.

Publication year: 2008–19.

TABLE 19 Phase 2 NHS EED database search strategy

#	Terms
1	(Any field) "Screening Tool of Older Persons' Potentially Inappropriate Prescriptions" OR(Any field) "Screening Tool to Alert Doctors to Right Treatment" OR (Any field) "STOPP START"

Appendix 6 Web searches

Age UK

URL: www.ageuk.org.uk/search/.

Date searched: 11 October 2019.

Number of documents: 3.

Search strategy
STOPP.

OpenGrey

URL: www.opengrey.eu.

Date searched: 11 October 2019.

Number of documents: 7.

TABLE 20 OpenGrey website search strategy

#	Terms
1	STOPP/START ($n = 4$)
2	Screening Tool of Older Persons Prescriptions ($n = 1$)
3	Screening Tool to Alert to Right Treatment ($n = 2$)

British Geriatrics Society

URL: www.bgs.org.uk/search/node.

Date searched: 11 October 2019.

Number of documents: 6.

Search strategy
STOPP.

Appendix 7 Guide for patient and public involvement feedback on initial theories

Why these tools?

(Ranked from important to not important.)

- STOPP/START tools promote a person-centred approach to shared decision-making.
- STOPP/START tools give personalised recommendations to patients.
- STOPP/START tools are easy to use.
- The application of STOPP/START tools have a direct effect on the health of individuals and appropriate use of limited resources.
- STOPP/START tools are widely used/recognised.
- STOPP/START tools are comprehensive.
- STOPP/START tools facilitate audit and outcomes monitoring.
- STOPP/START tools reduce health-care costs.

To what aim should they be used?

- To be of use clinically, STOPP/START tools need to demonstrate a positive effect on adverse outcomes (i.e. reduction in ADEs, reduction in hospitalisation, reduction in health-care use, reduction in mortality).
- STOPP/START tools help clinicians avoid unnecessary prescribing cascades by encouraging them to consider medications as possible causes of non-specific symptoms, such as confusion, falls or constipation.
- STOPP is a tool that can be used as a screening/prompt/checklist/decision support tool for:

  - identifying medications with potential risks
  - screening for potentially inappropriate medications
  - rapid identification of potentially inappropriate prescribing
  - identifying adverse drug effects in older people.

- STOPP/START tools promote standardisation of clinical judgement.

How to use them

- STOPP/START tools do not take into account patient preferences or comorbidities.
- STOPP/START are tools that can facilitate making decisions about medications without specialist knowledge/with little clinical judgement.
- STOPP/START are tools aimed at supporting clinicians in making decisions about medications, but do not replace clinical judgement.
- STOPP/START tools leave room for interpretation.
- STOPP/START are tools that can facilitate making decisions about medications, without specialist knowledge/with little clinical judgement.
- Non-medical prescribers are more receptive to the use of STOPP/START tools.
- STOPP/START are tools that can improve the training of medical students in the management of polypharmacy.
- Additional training should be optional when using STOPP/START tools.

- Staff training is important to ensure consistency; the receptiveness of prescribers, patients and staff in various settings will have an impact on the uptake and effectiveness of STOPP/START tools in older people.

Implementation

- Patients may not understand the notion of deprescribing and they may think that clinicians are trying to stop them from having medications they need.
- Paper versions of the STOPP/START tools are time-consuming to use and require users to be very familiar with the content of the tool to use it, which is rarely the case.
- Computerised use would make the use of STOPP/START tools more efficient.
- Process redesign is needed to fit STOPP/START tools into the clinical workflow.
- Computerised use of STOPP/START tools should be as sensitive and specific as possible, as there will not usually be subsequent evaluation by a clinician.
- STOPP/START tools are difficult to use in everyday clinical practice.
- Is it sufficient to use the STOPP/START tools during a single episode of care, or should patients be exposed to the tools on a daily/weekly or monthly basis?
- STOPP/START tools need to be adapted to reflect local policy.

Appendix 8 Patient and public involvement

Patient and public involvement in theory development

In the first meeting, the wording of two of the programme theories and the concept of personalisation were discussed.

Changes to the wording of programme theories

The advisory group made suggestions to change the wording of the programme theories on work adherence and empowerment.

Changing the word work 'adherence' to 'taking medication as prescribed' (i.e. from 'STOPP/START tools lead to greater adherence to medication when patients are present at their medication review' to 'STOPP/START tools lead to patients taking medication as prescribed when patients are present at their medication review').

There was also discussion on the use of the word 'empowerment' (i.e. from 'STOPP/START tools increase patient empowerment when their opinions and expectations have been taken into account' to 'when patients' opinions and expectations are taken into account STOPP/START tools can increase patients' confidence and control of medications'). The advisory team suggested changing the word 'empowerment' to 'confidence' as, from the patients' perspective, different people can be empowered in different ways and it is more about confidence and control of medication.

Individualisation or personalisation

As these terms were used interchangeably in the literature, there was a consensus among the Patient Advisory Group to use the term personalisation in preference to the term individualisation, which was considered to be more clinical and less 'personal'.

What constitutes personalisation?

As the studies in the review often did not explicitly mention personalisation, the project team and the advisory members interpreted the concept of personalisation. Three studies (*Box 3*) were discussed with the advisory team, specifically on the subject of whether or not they included elements of personalisation. Following discussion, the team agreed on a typology of personalisation that may happen as a continuum rather than as a simple personalised or not personalised approach. Some examples of the personalised/not personalised approach discussed in the meeting are included below.

Personalised

- Opportunity for the patient to give input on the medications they take.
- Discussion about the possible implications of the changes to medicines and what to look out for.
- Having a follow-up after changes are made.
- Face-to-face discussion where the patient is involved in discussion about what changes to make.
- The changes made are agreed with the patient.
- The aim of making changes should be about patient outcomes rather than financial outcomes.

Not personalised

- Discussion happens without the patient, and then they are asked to agree (or not) with changes.
- When the patient agrees to proposed changes without being part of discussion about them.

A subsequent meeting focused on the results, *Plain English summary* and dissemination of the results (see *Chapter 8* and *Appendix 12* for details on input on dissemination).

BOX 3 Studies discussed with the Patient Advisory Group

Study 1: Fog *et al.*[59]

Not personalised.

The individuals have not been consulted. Although it is personalised from the professionals' perspective (i.e. they discuss the specifics of that particular patient), the patient has not been included in the discussion and are just 'accepted by the patient'. The changes are imposed on the patient rather than decisions being made with them.

Study 2: Campins *et al.*[51]

It was thought that this study has elements that showed personalisation and elements that did not.

Personalised: it included face-to-face discussions with patients, changes were agreed with patients and there was a follow-up to see how the patient was getting on with changes. The adaptions to medications were recommendations rather than decisions.

Not personalised: the outcomes were about expenditure rather than health. It was felt that the decision was made before speaking to the patients and the doctor did not agree this with the patient. It was mentioned that patients should be involved at the beginning of the discussion and not afterwards.

Study 3: De Bock[49]

Opinions differed for this study. Again, it seemed that there were elements of personalisation and elements that were not personalised. A positive highlighted in this study was the communication to the patients' general doctor about changes. This is something that has been brought up before (i.e. that it's pointless someone changing something without communication with other people involved in the care of the patient).

A factor in this study that was personalised was the opportunity for the patient to tell the doctor about the medication they were taking. There was also effort made to help patients understand what changes had been made and the potential implications of these changes. It also seems like the doctors checked that the changes made were appropriate.

It was highlighted that patients in this study attended hospital as an 'unplanned' admission. This makes personalisation hard, as carers may not be present to help, patients might not know all the medications they are on and it is potentially too much information to take in. Just knowing what medication someone is taking is not personalisation. It seems as if patients were just 'told' about changes that were made.

Patient and public involvement feedback on the *Plain English summary*

- Providing information about what STOPP/START is a dictionary explanation (e.g. stopping harmful/expensive and starting more effective medications).
- Change use of the word 'tool' to approach.
- In the section about phase 1, tidy up layout of programme theories so that the key bolded words are at the start of sentence.
- Phase 2: first paragraph needs simplification, especially the last sentence.
- Final paragraph is negative and needs rephrasing – the point made in the last sentence is important but reword so it is more about the focus of future work and not the current lack of patient focus.

Synthesis results

Results from the synthesis were shared and discussed with the Patient Advisory Group and constructive feedback was provided on implementation of the results and future research, in terms of setting, outcomes, communication between patients/providers, inherent barriers between providers and personalisation.

- Rather than just knowing how many studies in each country and how many in which setting, it would be interesting to see how these relate.
- Outcomes to consider from the patients' perspective:

  - falls
  - reduced health-care use
  - quality of life.

- Significance of communication between health-care professionals and patients (between doctor and patient, pharmacist and GP, etc.) to ensure that changes during the review are actually implemented.
- One of the recommendations is to conduct more research in the UK and in primary care, as the majority of the research is in hospital settings.
- The inherent barriers between pharmacists and GPs.
- Focus should be on studies where patients' opinions are taken into consideration because of the lack of studies in the literature.
- Highlight that current evidence suggests that these reviews are not being personalised.

Appendix 9 Quality appraisal of studies for phase 2 (non-randomised controlled trials)

TABLE 21 Quality appraisal of studies for phase 2 (non-RCTs)

Study	Focused question	Recruitment	Exposure measurement	Blinding	Confounding identification	Confounding methods	Follow-up completeness	Follow-up length	Believe results	Applicability local population	Other available evidence	Implications	Score
Naveiro-Rilo et al.[86]	2	0	2	2	1	1	2	0	2	1	2	0	15/M
Chandrasekhar et al.[52]	2	2	1	2	1	1	2	0	2	2	2	0	17/M
De Bock et al.[49]	2	2	1	2	1	1	2	2	2	2	2	0	19/H
García-Caballero et al.[61]	2	2	1	2	1	1	2	2	2	2	2	1	20/H
Gaubert-Dahan et al.[62]	2	2	1	2	1	1	2	1	2	2	2	0	18/H
Ilić et al.[65]	2	2	1	2	1	1	2	2	2	2	2	1	20/H
Kimura et al.[66]	2	0	2	2	1	1	2	2	2	2	0	1	17/M
McNicholl et al.[67]	2	2	1	0	1	1	2	2	2	2	2	0	17/M
Momblona et al.[88]	2	2	1	0	1	1	2	2	0	2	2	0	15/M
Grion et al.[64]	2	2	2	2	2	2	2	2	2	2	2	2	24/H
Mekdad and Alsayed[68]	2	0	2	2	2	2	2	2	2	0	2	0	18/H
Rossi et al.[72]	2	2	0	0	0	0	2	2	2	1	2	2	15/M
Senin Loreto et al.[87]	2	2	2	2	0	0	2	2	2	1	2	0	17/M
Sennesael et al.[74]	2	1	2	2	2	2	2	2	2	1	2	2	22/H
Gramage Caro et al.[85]	2	1	2	2	2	1	2	2	1	1	1	2	19/H
Delgado Silveira et al.[56]	2	2	2	2	2	2	2	1	2	1	2	1	21/H
Deliens et al.[57]	2	2	2	2	2	2	2	1	1	1	1	1	19/H
Fernández Regueiro et al.[58]	2	2	2	2	2	2	2	2	2	1	1	1	21/H
Fog et al.[59]	2	2	2	2	2	2	2	1	2	2	2	1	22/H
Garay-Bravo et al.[60]	2	2	2	2	1	2	2	2	2	1	2	1	21/H
Gibert et al.[63]	2	2	2	2	1	0	2	1	1	2	2	1	18/H
Mulvogue et al.[69]	2	2	2	2	2	0	2	2	1	1	2	1	19/H
Silva et al.[75]	2	2	2	2	1	1	2	1	1	1	2	2	19/H
Twigg et al.[76]	2	2	2	2	2	1	0	0	1	1	1	1	15/M
Unutmaz et al.[77]	2	2	1	2	2	2	2	2	2	1	2	1	21/H
Weeks et al.[79]	2	1	0	2	2	2	2	0	2	1	2	2	18/H
Whitman et al.[80]	2	1	2	1	0	0	2	2	1	1	1	1	14/M
Zaal et al.[81]	2	0	2	2	2	0	2	2	2	1	1	1	17/M

H, high; M, medium.
The score was low if half or less than half of the criteria were met, medium if more than half but less than three-quarters of the criteria were met, and high if most of the criteria were met.

Appendix 10 Quality appraisal of studies for phase 2 (randomised controlled trials)

TABLE 22 Quality appraisal of studies for phase 2 (RCTs)

Study	Focused issue	Randomisation	Attrition	Outcome measurement	Group comparability	Comparability other	Local population applicability	Clinically important outcomes	Benefits worth harms and costs	Score
Campins et al.[51]	2	2	2	2	2	2	2	2	2	18/H
Dalleur et al.[55]	2	2	2	2	2	2	1	1	2	16/H
Frankenthal et al.[82]	2	2	2	2	2	2	0	2	2	16/H
O'Connor et al.[70]	2	2	2	2	2	2	2	2	2	18/H
O'Sullivan et al.[71]	2	2	2	0	2	2	0	2	2	14/H
Price et al.[50]	2	2	2	1	1	2	2	1	2	15/H
Santolaya-Perrín et al.[73]	2	2	2	0	2	1	2	2	1	14/H
Van der Linden et al.[78]	2	0	2	2	2	1	1	0	2	12/M
Gallagher et al.[83]	2	2	2	0	2	2	2	2	2	16/H
Coronado-Vázquez et al.[53]	2	2	2	0	2	2	0	2	2	14/H
Cossette et al.[54]	2	2	2	0	2	2	2	2	2	16/H
García-Gollarte et al.[84]	2	2	2	2	2	2	2	2	2	18/H

H, high; M, medium.
The score was low if half or less than half of the criteria were met, medium if more than half but less than three-quarters of the criteria were met, and high if most of the criteria were met.

Appendix 11 Consensus Health Economic Criteria appraisal of economic evaluation

TABLE 23 Consensus Health Economic Criteria appraisal of economic evaluation

Criterion	Item	Gallagher et al.[89]	Twigg et al.[76]
Is the study population clearly described?	1	Yes	Yes
Are competing alternatives clearly described?	2	Yes: usual care described in accompanying trial effectiveness paper	Partly: intervention and context clearly described (note, no control/alternative)
Is a well-defined research question posed in answerable form?	3	Yes	Yes
Is the economic study design appropriate to the stated objective?	4	Yes: although ADRs are a partial outcome	Yes
Is the chosen time horizon appropriate to include relevant costs and consequences?	5	No	Yes
Is the actual perspective chosen appropriate?	6	Yes	Yes
Are all important and relevant costs for each alternative identified?	7	Partly: medication change costs not included	Partly: health services attendances only (no medication change costs)
Are all resources measured appropriately in physical units?	8	Yes	Yes
Are resources valued appropriately?	9	Partly: software and training costs look very crudely estimated	Yes
Are all important and relevant outcomes for each alternative identified?	10	Partly: ADRs only (trial's primary outcome)	Yes
Are all outcomes measured appropriately in physical units?	11	Yes	Yes: EQ-5D-5L QALYs measured
Are outcomes valued appropriately?	12	No	Yes: mapped from EQ-5D-3L
Is an incremental analysis of costs and outcomes performed?	13	Yes	Yes
Are all future costs and outcomes discounted appropriately?	14	NA	NA
Are all important variables, whose values are uncertain, appropriately subjected to sensitivity analysis?	15	NR: CEAC produced, but variance of input variables NR	NR: CEAC produced, but variance of input variables NR
Do the conclusions follow from the data reported?	16	Yes: with caveat about uncertainty	Yes: potential bias because of non-randomised comparator acknowledged
Does the study discuss the generalisability of the results to other settings and patient/client groups?	17	Yes	Yes

TABLE 23 Consensus Health Economic Criteria appraisal of economic evaluation (*continued*)

Criterion	Item	Gallagher *et al.*[89]	Twigg *et al.*[76]
Does the article indicate that there is not potential conflict of interest of study researcher(s) and funder(s)?	18	Yes: ICMJE declaration	Yes: journal COI declaration
Are ethics and distributional issues discussed appropriately?	19	No	No

CEAC, cost-effectiveness acceptability curve; COI, conflict of interest; EQ-5D-3L, EuroQol-5 Dimensions, three-level version; EQ-5D-5L, EuroQol-5 Dimensions, five-level version; ICMJE, International Committee of Medical Journal Editors; NA, not applicable; NR, not reported.

Note
The Consensus Health Economic Criteria list for assessing quality of economic evaluations (Evers *et al.*[91]) incorporates all but one of the widely used critical appraisal questions recommended by Drummond *et al.*[94]

Appendix 12 Patient and public recommendations for dissemination

TABLE 24 Patient and public recommendations for dissemination

Who	What	Where/how

• Patients
• Carers
• Clinicians: especially pharmacists (also consider the wider health-care community, such as dentists and physiotherapists)
	• Patient leaflets
	• Information on online pharmacy websites
	• Twitter
	• Poster/infographic of the key information: easier to digest and more appealing than a word summary
	• 'Who is in charge of your medication' as a title to draw people in to read more
		• Plan part of the dissemination when people are already together (e.g. target flu clinics in winter); this is when medication reviews might also be triggered
		• A number of organisations were suggested as described:
		  ○ CCGs – there seems to be a drive to reduce medication costs at the moment so this could fit well
		  ○ Patient participation groups
		  ○ Living Options Devon (Exeter, UK)
		  ○ Devon Council (similar in other councils) has a health and well-being board that is led by the local authority, includes groups like Healthwatch
		• Organisational newsletters (e.g. PenARC, medical school)

CCG, Clinical Commissioning Group; PenARC, National Institute for Health Research Applied Research Collaboration South West Peninsula.

Appendix 13 RAMESES II reporting standards for realist evaluations

TABLE 25 RAMESES II reporting standards for realist evaluations

Item			Reported in document: yes/no/unclear	Page(s) in document
Title				
1		In the title, identify the document as a realist evaluation	Yes	i
Summary of abstract				
2		Journal articles will usually require an abstract, while reports and other forms of publication will usually benefit from a short summary. The abstract or summary should include brief details on the policy, programme or initiative under evaluation; programme setting; purpose of the evaluation; evaluation question(s) and/or objective(s); evaluation strategy; data collection, documentation and analysis methods; key findings and conclusions Where journals require it and the nature of the study is appropriate, brief details of respondents to the evaluation and recruitment and sampling processes may also be included. Sufficient detail should be provided to identify that a realist approach was used and that realist programme theory was developed and/or refined	Yes	viii
Introduction				
3	Rationale for evaluation	Explain the purpose of the evaluation and the implications for its focus and design	Yes	3
4	Programme theory	Describe the initial programme theory (or theories) that underpin the programme, policy or initiative	Yes	9
5	Evaluation questions, objectives and focus	State the evaluation question(s) and specify the objectives for the evaluation. Describe if and how the programme theory was used to define the scope and focus of the evaluation	Yes	5
6	Ethics approval	State whether or not the realist evaluation required and has gained ethics approval from the relevant authorities, providing details as appropriate. If ethics approval was deemed unnecessary, explain why	Yes	10
Methods				
7	Rationale for using realist evaluation	Explain why a realist evaluation approach was chosen and (if relevant) adapted	Yes	3
8	Environment surrounding the evaluation	Describe the environment in which the evaluation took place	Yes	1–3
9	Describe the programme policy, initiative or product evaluated	Provide relevant details on the programme, policy or initiative evaluated	Yes	2

TABLE 25 RAMESES II reporting standards for realist evaluations (*continued*)

Item			Reported in document: yes/no/unclear	Page(s) in document
10	Describe and justify the evaluation design	A description and justification of the evaluation design (i.e. the account of what was planned, carried out and why) should be included, at least in summary form or as an appendix, in the document that presents the main findings. If this is not carried out, the omission should be justified and a reference or link to the evaluation design given. It may also be useful to publish or make freely available (e.g. online on a website) any original evaluation design document or protocol, where they exist	Yes	46–47, 54
11	Data collection methods	Describe and justify the data collection methods, which ones were used, why and how they fed into developing, supporting, refuting or refining programme theory. Provide details of the steps taken to enhance the trustworthiness of data collection and documentation	Yes	7–13
12	Recruitment process and sampling strategy	Describe how respondents to the evaluation were recruited or engaged and how the sample contributed to the development, support, refutation or refinement of programme theory	Yes	9
13	Data analysis	Describe in detail how data were analysed. This section should include information on the constructs that were identified, the process of analysis, how the programme theory was further developed, supported, refuted and refined, and (where relevant) how analysis changed as the evaluation unfolded	Yes	9, 12
Results				
14	Details of participants	Report (if applicable) who took part in the evaluation, the details of the data they provided and how the data were used to develop, support, refute or refine programme theory	Yes	19–25
15	Main findings	Present the key findings, linking them to contexts, mechanisms and outcome configurations. Show how they were used to further develop, test or refine the programme theory	Yes	56
Discussion				
16	Summary of findings	Summarise the main findings with attention to the evaluation questions, purpose of the evaluation, programme theory and intended audience	Yes	55
17	Strengths, limitations and future directions	Discuss both the strengths of the evaluation and its limitations. These should include (but need not be limited to) (1) consideration of all the steps in the evaluation processes; and (2) comment on the adequacy, trustworthiness and value of the explanatory insights which emerged. In many evaluations, there will be an expectation to provide guidance on future directions for the programme, policy or initiative, its implementation and/or design. The particular implications arising from the realist nature of the findings should be reflected in these discussions	Yes	57–59
18	Comparison with existing literature	Where appropriate, compare and contrast the evaluation's findings with the existing literature on similar programmes, policies or initiatives	No	–

TABLE 25 RAMESES II reporting standards for realist evaluations (*continued*)

Item			Reported in document: yes/no/unclear	Page(s) in document
19	Conclusion and recommendations	List the main conclusions that are justified by the analyses of the data. If appropriate, offer recommendations consistent with a realist approach	Yes	63
20	Funding and conflict of interest	State the funding source (if any) for the evaluation, the role played by the funder (if any) and any conflicts of interests of the evaluators	Yes	viii