NIHR Health Services Research programme

NHS National Institute for Health Research

HSR Protocol - project ref: 10/2000/29
Version: 1
Date: 20th July 2011

Measuring prevalence, reliability and variation in high risk prescribing in general practice using multilevel modelling in a population database

Chief investigator	Professor Bruce Guthrie, Professor of Primary Care Medicine, University of Dundee
Sponsor	University of Dundee
Funder	NIHR Health Services Research Programme

Measuring prevalence, reliability and variation in high risk prescribing in general practice using multilevel modelling in a population database

1. Aims/Objectives:

The aim of this study is to define a set of prescribing safety indicators that can be operationalised in existing electronic clinical data, and to examine how high-risk prescribing varies between patients, GPs and practices in order to determine the validity of these indicators for safety improvement, clinical governance and appraisal/revalidation purposes.

Objective 1. To define and operationalise prescribing safety indicators that can be applied at individual prescriber and practice level.

Objective 2. To examine the prevalence of individual indicators and appropriate composites, associations with patient, prescriber and practice variables, and the relative importance of variation at prescriber and practice levels before and after adjustment for patient level variables.

Objective 3. To measure the reliability of individual and composite indicators at prescriber and practice level.

2. Background:

Prescribing is a high benefit, high cost and high risk activity. The sum of choices made by individual prescribers therefore significantly influences the creation of an effective, efficient and safe National Health Service (NHS). National clinical guidelines focus on effective prescribing, identifying broad groups of patients in whom particular drugs are clearly indicated, and several such recommendations are embedded in the Quality and Outcomes Framework (QOF), both explicitly (eg use of ACE inhibitors after myocardial infarction) and implicitly (blood pressure , cholesterol and glycated haemoglobin control).[1] Historically, primary care prescribing improvement activity has focused on cost, partly because existing national datasets derived from pharmacy payment data do not allow examination of the drugs prescribed to particular patients, which is usually necessary to measure quality and safety. However, although there an increasing number of systems for warning prescribers of prescribing risks (such as Drug Safety Update,[2] cascaded drug warnings, National Patient Safety Agency alerts,[3] National Prescribing Centre educational material[4]), there are currently few systematic mechanisms for measuring or improving primary care prescribing safety.

Primary care prescribing is a frequent source of harm. Adverse drug events (ADEs) account for ~6.5% of all hospital admissions,[5] and at least half of these are preventable.[6] The most frequent classes of drug implicated are anti-platelet drugs including aspirin, non-steroidal anti-inflammatory drugs (NSAIDs), warfarin, diuretics, opioids, beta-blockers and ACE inhibitors/angiotensin receptor blockers.[4 6-8] Deaths are most frequently associated with NSAIDs and anti-platelet prescribing.[5] Other less commonly prescribed drugs have also been implicated in causing preventable deaths by national safety alerts, including community-prescribed methotrexate,[3] and antipsychotics in older people with dementia.[2] Existing measures of potentially inappropriate or high-risk prescribing include relevant indicators from the Assessing Care of Vulnerable Elders (ACOVE) project,[9] the Screening Tool of Older Persons' Potentially inappropriate Prescriptions (STOPP),[10] and the Beers Criteria.[11] However, ACOVE and STOPP rely on manual record review and are therefore not easily applied on a large scale. The Beers Criteria are easily measured using routine healthcare data, but the majority of the drugs most commonly implicated in serious harm are not included, which makes them less useful in safety improvement.

Funded by NHS Quality Improvement Scotland and the Scottish Government Chief Scientist Office, BG and TD have recently used RAND and modified Delphi consensus methods to define a new set of evidence based indicators of high risk prescribing which can be measured using routine electronic data. In terms of prescribing safety, this includes a set of indicators of high risk prescribing focusing on drugs commonly causing serious harm.[5 6] We have applied a subset of 15 of these indicators to routine GP clinical data for 1.75 million patients. The indicators used in this study related to NSAIDs (prescription in people at high risk of GI bleeding without gastroprotection, in renal impairment, co-prescription with ACEI/ARB and diuretics), warfarin (co-prescription of NSAIDs, antiplatelets, high risk antibiotics, and oral azole antifungals), methotrexate (co-prescription of both 10mg and 2.5mg tablets), antispsychotic prescription in older people with dementia,[12] and prescription of drugs to avoid in heart failure (NSAIDs, glitazones, verapamil, tricyclics and others). 139,404 patients were defined as 'at-risk' by being included in one or more denominators, and 19,308 (13.9%) of these patients had received one or more high-risk prescriptions in the previous year (the composite outcome). Crude rates of high-risk prescribing varied significantly between practices from 3.6% of patients at risk receiving at least one high-risk prescription to 31.3% (median 12.5%, interquartile range 10.1% to 15.3%). After adjustment in a multilevel logistic regression model for patient casemix, there remained considerable variation between practices in rates of high risk prescribing. Reliability for individual indicators at practice level was variable, but the composite reasonably reliably distinguished high and low risk practices (reliability for median size practice = 0.92; 95.6% of practices had reliability greater than 0.7, and 88.2% greater than 0.8). High risk prescribing is therefore common, variable between practices, and the composite examined at practice level in this analysis is a reasonably reliable discriminator. Although not all of these prescriptions will be inappropriate, the high prevalence and high variation indicates that this prescribing is very likely to be improvable (paper currently out to peer review with BMJ).

We have subsequently used a subset of these measures to feed *practice* rates of high risk prescribing back to *general practitioners* (GPs) in a revalidation pilot, where GPs received a range of external performance data to reflect on with their appraiser. The prescribing safety data was highly valued by participants, but a common response was to question whether practice level feedback was appropriate, essentially assuming that prescribing is an individual decision, and therefore that variation between GPs was more important. Although it seems likely that individual GPs will vary in their high-risk prescribing, the sharing of patient care across primary care clinical teams makes it likely that individual prescribing will be significantly influenced by practice culture or safety climate ("the way things are done around here"), in terms of some kinds of high risk prescription becoming 'normal' or accepted practice in some settings.[13]

The literature and our previous work raises three important questions which this study will address.

1. Can high risk prescribing be feasibly measured at individual prescriber level? Operationalising indicators at prescriber level is more challenging than at practice level because patients can be prescribed the same drug by different practitioners. Many of our existing practice level indicators will therefore need modifying and re-testing for feasibility.
2. Is high risk prescribing driven more by individual prescribers (variation between prescribers within practices) or by wider practice prescribing and safety culture (variation between practices).[13] Both are likely to be important, but their relative importance is uncertain.
3. Is measurement at practice and/or prescriber level reliable enough to distinguish practices and/or prescribers? Our previous work has shown reasonable reliability at practice level

even for high-stakes evaluation (requiring reliability>0.8 and ideally >0.9), but reliability at individual prescriber level is unknown. This makes it difficult to define how suitable individual prescriber measurement is for any or all of collaborative quality improvement through clinical governance signal generation to high-stakes evaluation for revalidation.

3. Need:

As described above, prescribing is an important cause of preventable harm, and high-risk prescribing is both common and highly variable between practices (*health need*). Reducing variation and improving safety are key priorities for NHS management, and both practice and individual focused work are ways of trying to achieve these. Although the shape of revalidation remains uncertain, aligning safety improvement with it is an attractive mechanism to engage individual practitioner's attention, but it is unclear whether it is feasible to use routine data to examine high-risk prescribing at practitioner level or whether it is best considered at practice level. Prescribing is central to healthcare, and prescribing improvement will remain a high priority for the NHS. Historically, 'prescribing improvement' in the community has been focused on cost since available data could only support measurement of cost and volume of prescribing at practice level (although data was originally produced at GP level, this was always problematic because of the way that repeat prescribing systems were run). With the growth of electronic patient-level datasets, there are clear opportunities to extend surveillance and governance to quality and safety, but uncertainty about how to use this data, and whether it can be considered data for judgement or data for improvement at both practice and GP level (*expressed need* and *sustained interest and intent*).

There have been previous studies (including our own) using consensus methods to define evidence based prescribing safety indicators. A smaller number of studies have examined population prevalence of selected indicators, and our own work has measured prevalence of multiple indicators and variation between practices (*builds on existing work*). However, to our knowledge, the analysis proposed here has never been done before (*capacity to generate new knowledge*), and the proposed study will examine the relative contribution to rates of high-risk prescribing of individual prescribers and groups of prescribers in practices (*organisational focus consistent with the NIHR HSR mission*). Study outputs will include new knowledge about variation in high-risk prescribing that will inform the focus of prescribing safety improvement, and a set of practice and GP-level indicators that can be operationalised in existing clinical data with evidence of what kinds of use such indicators can support (*generalisable findings and prospects for change*).

4. Methods:

a. Setting

The analysis will use data extracted from general practice clinical IT systems and held by the Primary Care Clinical Informatics Unit at University of Aberdeen. The practices providing data care for patients representative of the wider population in terms of demography and provide representative quality of care as measured by the Quality and Outcomes Framework.[14] Data to be extracted for this study will include:

- Fully anonymised unique patient, prescriber and practice identifiers.
- Prescribing data from April 2004 to March 2007 for patients defined as 'at risk' of a receiving a high-risk medication by being included in one or more denominators for selected indicators (defined by disease, prescription [eg currently prescribed warfarin], or

age).
- Available patient (eg age, sex, deprivation, relevant co-morbidity, drug allergy/intolerance), prescriber (eg GP or non-medical) and practice (eg listsize, remoteness) characteristics.

b. Design

Retrospective primary care database study using multilevel regression analysis to examine prescribing variation using routine electronic clinical data extracted from General Practice clinical IT systems.

Objective 1. To define and operationalise prescribing safety indicators that can be applied at individual prescriber and practice level.

Methods for objective 1. We will draw on the evidence based set of practice level indicators developed in a previous study, with indicator choice and modification balancing ideal measures and feasibility of implementation in routine data. This will be achieved by iteratively cycling between discussion with an external expert advisory group and operationalisation in real-life GP electronic clinical data extracted by the Primary Care Clinical Informatics Unit at University of Aberdeen. Longitudinal data from approximately 200 general practices representative of the wider population is available.

Objective 2. To examine the prevalence of individual indicators and appropriate composites, associations with patient, prescriber and practice variables, and the relative importance of variation at prescriber and practice levels before and after adjustment for patient level variables (casemix).

Methods for objective 2. Descriptive epidemiology of prevalence individual and composite indicators, and univariate analysis of how prevalence varies between patient groups, prescribers and practices. Multilevel logistic regression modelling to partition variance in outcome between patients, prescribers and practices, and examine associations of the outcome with patient, prescriber and practice characteristics.

Objective 3. To measure the reliability of individual and composite indicators at prescriber and practice level.

Methods for objective 3. Reliability will be initially estimated using the Spearman-Brown Prophecy Formula. Generalisability theory will then be applied following an analysis of variance (ANOVA) of prescribing patterns.

The practice level analyses we have previously conducted are methodologically relatively straightforward, since they are cross-sectional (the outcome being receipt of one or more high-risk prescriptions in the previous year), and there is a clear denominator for each practice (the number of patients registered on a particular date who are defined as being ‘at risk’, for example because they are currently being prescribed warfarin). The analysis therefore used two-level multilevel logistic regression, analysing patients clustered within practices. Analysis at GP level is more complicated since patients can be treated by any GP in virtually all practices. The data is therefore not neatly hierarchical, and the denominator at GP level is harder to define. Additionally, different types of prescription vary in how strongly they can be attributed to the GP signing the prescription. Legally, the GP signing takes full responsibility for the prescription, but currently most ‘repeat’ prescriptions are signed without significant oversight, with the signing

GP trusting the decision of whichever colleague authorised the prescription as a repeat, and assuming that the IT system will flag when the prescription is due for review (typically annually). Our approach will therefore be to focus on prescribing decisions in two contexts:

1. The decision to stop a high-risk drug at medication review, which is relatively strongly attributable to individual GPs and can be analysed using a simpler, strictly hierarchical multilevel model.
2. Acute prescribing of high-risk medication, which is relatively strongly attributable to individual GPs but where analysis is more complicated and will require fitting cross-classified multilevel models.

c. Data collection and analysis

Methods for objective 1: defining appropriate indicators to use as outcomes

Indicators examined will all relate to high-risk medications defined by our previous work. Few of these indicators measure prescribing that is *absolutely* contraindicated or a "never event" (such as surgery on the wrong leg),[15] but all are clearly identified as high-risk for particular types of patient in national guidance and/or the British National Formulary. However, such prescriptions will sometimes be justified, and they are therefore best conceived of as high-risk rather than clearly unsafe or definitely inappropriate. Examples of indicators are:

- Prescription of a non-selective beta-blocker to a patient with asthma and recent asthma treatment before beta-blocker prescription (high risk because of existing morbidity)
- Prescription of an oral non-steroidal anti-inflammatory drug to patient aged >74 years without use of gastro-protection (high risk because of age)
- Prescription of an antiplatelet drug without gastro-protection to a patient being prescribed warfarin (high risk because of other drug prescription)

As described above, we intend to examine high-risk prescribing in two contexts (stopping high-risk medications at medication review; and acute prescription of a high-risk medication). Initial analysis will focus on prescription of oral non-steroidal anti-inflammatory drugs (NSAIDs) to patients at high risk of adverse events (eg NSAID prescription to patients co-prescribed warfarin; or co-prescribed an antiplatelet drug without gastro-protection; or with previous peptic ulcer without gastro-protection; or with renal impairment; or prescribed ACE inhibitors/angiotension receptor blockers *and* diuretics; or with heart failure). Oral NSAIDs cause considerable harm in terms of both emergency hospital admission and death,[5 6] and our previous work has shown that they are relatively commonly prescribed to patients at high risk of adverse events. Most NSAID prescriptions are initiated in primary care, meaning that attribution is also relatively straightforward.

We will additionally explore the feasibility of other indicators including antipsychotic prescription in older people with dementia;[2] high-risk antibiotic (macrolides, quinolones, metronidazole), antiplatelet and oral azole antifungal prescription in patients on warfarin;[8] beta-blocker prescription in patients with active asthma; use of long-acting beta-agonists in asthma without use on inhaled corticosteroids;[16 17] prescription of glitazones and other high risk drugs in heart failure;[18] and use of mixed 10mg and 2.5mg strengths of oral methotrexate.[3] Decisions about which of these and other potential measures will be used will be made by the advisory group, based on considerations of likely harm caused and issues of attribution (for example, drugs like antipsychotics in older people with dementia[2 12] may often be initiated by specialists, making attribution of initiation to GPs and practices problematic, although this would not preclude examining stopping of drugs after medication review).

The final choice of indicators will depend on feasibility and will be decided by an expert advisory group convened for this study. For each measure, we will iteratively work with the advisory group to define an appropriate specification for each of the two contexts described above, examine the feasibility of implementing this specification in the data, and bring the findings back to the advisory group for acceptance or modification. We will additionally explore with the advisory group whether it is feasible to weight prescriptions by their degree of risk when creating composite measures, which will depend on the precision with which risk has been estimated in the literature.

Methods for objective 2: univariate analysis and multilevel logistic regression modelling

Stopping high risk medications at medication review.
Medication reviews will be identified by the presence of an appropriate Read Code defined by Quality and Outcomes Framework (QOF) Business Rules.[1] For repeat prescribing, patients will be defined as 'currently prescribed' a high-risk medication if there is an active repeat prescription for the drug and they have received a prescription in the last 84 days. Stopping will be defined as the repeat medication being inactivated. For acute prescribing, patients will be defined as 'currently prescribed' if they have received two or more prescription for the drug in the previous 168 days. Stopping of acute prescriptions will be defined as not receiving the drug in the subsequent 168 days.

For each individual measure, and for a composite, univariate analysis will measure the prevalence of stopping high-risk drugs at medication review, and the proportion of each GP's and each practice's patients who stop a high risk drug following medication review. Univariate associations with patient, GP and practice characteristics using appropriate statistical tests for continuous and categorical data will be examined. At patient level, characteristics to be examined for associations will include age, sex, socioeconomic status, number of drugs prescribed in the last 3 months, co-morbidity, whether the high-risk medication is an acute or repeat prescription, whether the review took place during a face to face consultation. At prescriber level, data available through the clinical IT system is limited, although we expect to be able to identify whether the prescriber is a GP or not, whether they are a locum, and how full-time they are (based on annual numbers of consultations). At practice level, available data includes listsize, remoteness/rurality, whether the practice is accredited for training or a dispensing practice, and QOF achievement (total and for the medicines management domain), and we will explore linking data used for existing prescribing cost improvement such as the percentage of prescriptions that are generic.

Multilevel modelling will be carried out in MLWin or SAS (depending on the employed researcher's existing skills – the rest of the proposal assumes MLWin). Intra-cluster correlation coefficients will be estimated in empty models to partition variance across patient, GP and practice, and to inform reliability calculations. For each individual measure and for a composite outcome, the main analysis will use a three level hierarchical logistic regression model (patients with a high-risk prescription *within* GPs carrying out most recent medication review *within* practices), with the outcome being whether or not the high-risk medication was stopped. Patient level and then practice level variables will be fitted using 2nd order penalised quasi-likelihood, with final models estimated using MCMC. If the advisory group considers that weighting by degree of risk is feasible, both unweighted and weighted models will be examined for the composite outcome. Model assumptions will be checked graphically, and using the diagnostics available in MLWin.

Acute prescription of a high-risk medication
Although an acute prescription being issued is an unambiguous event, attribution of the decision to the individual GP is more complex. Options to maximise attribution include distinguishing 'first' acute prescriptions from 'repeated' ones, and distinguishing acute prescriptions issued during a face to face encounter from those that may be telephone 'special requests' (ie varying the numerator definition). The appropriate denominator for indicators also has to be defined (in the sense of 'opportunities to prescribe' since GPs vary in how full or part time they are, and GPs who never encounter a patient at risk have no opportunity to write a high risk prescription). Denominator definition options include all encounters with patients at risk (face to face and telephone consultations and other file openings by particular GPs), encounters where any acute prescription is issued, and (for high-risk NSAID prescribing) encounters with any analgesic prescribed.

For each individual measure, and for a composite, univariate analysis will examine the acute prescription of high risk medications in the period April 2004 to March 2007. Initial analysis will focus on 'new' high risk acute prescriptions, defined as one where the same drug has not been prescribed in the previous 12 months, since these are the most clearly attributable to the individual prescriber. Subsequently, all high risk acute prescriptions will be analysed. At patient and practice level, the proportion of patients receiving one or more of each kind of high risk acute prescription in a defined period will initially be measured, and associations with patient and practice characteristics examined. We will then examine rates of high-risk prescribing at patient, GP and practice level, and explore the validity of different numerators ('new' or 'repeated' acute prescriptions or both) and denominators (rate/100 encounters, rate/100 encounters with an acute prescription, rate/100 encounters with an analgesic prescription). Associations with patient, GP and practice characteristics will be examined. The impact (if any) of using different numerators and denominators will be explored, and decisions about the most appropriate to use for different purposes made by the expert advisory group.

Multilevel modelling of indicators selected by the advisory group will be carried out in MLWin. Choosing the 'best' model is likely to involve trade-offs between full modelling of the data structure, practicality (whether models can actually be fitted using existing software) and interpretability. We will therefore explore a range of model specifications including strictly hierarchical Poisson models (observed vs expected rates of high risk prescribing by GPs within practices) and cross-classified logistic models (where the outcome is receipt or not of a high risk prescription measured at encounter level, with encounters cross-classified between GPs and patients, all nested within practices). Model fitting and checking will be as described above.

Methods for objective 3: examining reliability at practice and prescriber level
The analysis in objective 2 will provide detailed data about how high-risk prescribing varies between patients, GPs and practices. However, multilevel modelling is very unlikely to be used to routinely measure primary care prescribing safety in the NHS, and we will therefore examine the reliability of simpler measures such as the proportion of each GP's and each practice's patients who stop high-risk drugs following medication review, and GP and practice rates of 'new' and 'repeated' high risk acute prescription. The multi-level modelling will estimate how variation in patient outcome is partitioned between GP and practice. If variation is found almost exclusively at either GP or practice level, then objective 3 will focus on analysis at that level. If there is significant variation at both GP and practice level, then we will examine both. Based on the multilevel modelling findings as to which patient characteristics are associated with high-risk prescribing, we will calculate crude and case-mix adjusted measures for GPs and/or practices.

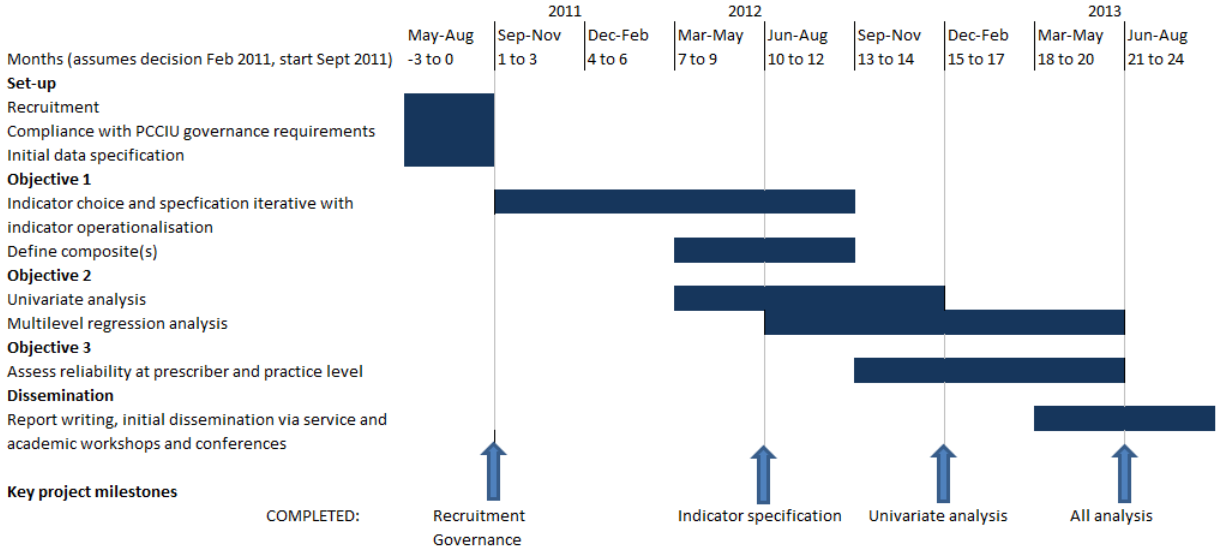
Initial analysis will examine whether the conclusions about whether GPs or practices are outliers differ significantly between simpler but more practical measures and multilevel model residuals.[19] This process will be important in assessing the validity of the simpler measures, and their suitability as 'data for judgement' or 'data for improvement'.

Reliability for GP and/or practice individual and composite measures will be initially formally estimated using the Spearman-Brown Prophecy Formula, based on the intra-cluster correlation coefficients estimated in the appropriate multilevel regression models, and the denominator size at each level. Reliability for the median denominator size at GP and+/or practice level will be calculated, as will the proportion of practices and GPs that can be measured with reliability >0.7 (minimum acceptable) and >0.8 and >0.9 (preferable for high stakes evaluation). Generalisability theory will then be applied following an analysis of variance (ANOVA) of prescribing patterns. From this perspective, each indicator will be treated as a separate test of the GP's and/or practice's prescribing, with rates transformed into standardised z-scores, and the z-score variance used to calculate reliabilities based on each individual measure of prescribing, and on the whole set of measures. Using this data, we will estimate the number of different indicators required to reliably distinguish GPs and practices at given levels of reliability (eg 0.7 minimum, 0.9 for high stakes evaluation). An initial power calculation indicates that if actual reliability using five indicators is 0.8, then analysis requires at least 127 practices to measure this reliability with a 95% confidence interval 0.75-0.85. We expect to include ~200 practices, and correspondingly more GPs, and power is therefore unlikely to be a limitation in this analysis.

5. Contribution of existing research:

The research contribution will be to rigorously demonstrate how feasible measurement of prescribing safety at GP level is, and to significantly extend understanding of how common high risk prescribing is and how it varies between and within practices. We will disseminate the research findings via a comprehensive final report, presentations at academic conferences and articles in high-impact peer-reviewed journals. The research will also have significant impact on health service policy and management, in terms of clarifying whether primary care organisation (PCO) led prescribing safety improvement can or should focus on practices or GPs or both. Critical to this is the feasibility of reliably measuring prescribing safety at GP level. This will inform whether GP level data can support external measurement for high stakes judgement, or is best used in a more exploratory and collaborative way to help understand and respond to more reliable practice level measures. To ensure dissemination and impact to this audience, we will work through our existing policy and NHS networks (including BG's membership of the NICE QOF Indicators Advisory Committee, the NHS Quality Improvement Scotland National Clinical Data for Quality Improvement Advisory Group, and the Scottish Safety Improvement in Primary Care programme, and DM's links to Scottish and UK appraisal and revalidation networks), produce and disseminate short summaries for clinicians, managers and policymakers, and offer seminars to relevant policy stakeholders including departments of health, the General Medical Council (building on links via our colleague Dr Mairi Scott who is a GMC council member), the Royal College of General Practitioners and the National Patient Safety Agency.

6. Plan of Investigation:



7. Project Management:

BG will be responsible for day to day management and supervision of the researcher, and will meet with him/her weekly to ensure progress and resolve problems. The employed researcher will take the main responsibility for project organisation (eg liaison with PCCIU over data, organisation of meetings, with support from the Quality, Safety and Informatics core funded administrative staff) and data management and analysis. All applicants will form the core project team who will meet as a group every four weeks, with individual meetings with the researcher as appropriate (for example, meetings of the employed researcher with TD relating to indicator definition; with PD and BG for statistical and modelling advice; with DM for reliability calculation).

We will additionally convene an advisory group comprising academic experts in prescribing safety and senior managers and clinicians with an interest in using prescribing safety measures for improvement, governance and revalidation. Membership will be finalised if funding is obtained, but will draw on advisory group members for the Better Measures and TIPP revalidation projects who included both academics and individuals in senior NHS positions (including two Health Board Medical Directors, an assistant Chief Medical Officer, and a GMC council member). We have also already approached Professor Tony Avery at University of Nottingham, who has considerable experience of prescribing safety measurement and is PI of the recently completed PINCER trial examining the impact of a pharmacist led intervention on a set of safety indicators. The advisory group will have a particularly active role early in the project to help determine which measures to focus on, and later in the project to consider the implications of the findings and aid dissemination

8. Service users/public involvement:

The primary aim of service user involvement in this project is to ensure that the findings are appropriately interpreted and disseminated, particularly as regards their use in framing safety improvement and supporting appraisal and revalidation. We will therefore seek to appoint up to two service users or public representatives to the advisory group, drawing either from existing groups associated with the Scottish Patient Safety Alliance or the Scottish Safety Improvement in Primary Care Collaborative.

9. References:

1. NHS England: Primary Care Commissioning. QOF Implementation: Business Rules, 2009.
2. Medicines and Healthcare products Regulatory Agency. Antipsychotics: use in elderly people with dementia. *Drug Safety Update Vol. 2 Issue 8*, 2009.
3. National Patient Safety Agency. Improving compliance with oral methotrexate guidelines, 2006.
4. NHS National Prescribing Centre. Update on the prescribing of NSAIDs. *MeReC Monthly No. 2* 2008.
5. Pirmohamed M, James S, Meakin S et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. *BMJ* 2004;329(7456):15-19.
6. Howard R, Avery A, Slavenburg S et al. Which drugs cause preventable admissions to hospital? A systematic review. *British Journal of Clinical Pharmacology* 2007;63(2):136-47.
7. Medicines and Healthcare products Regulatory Agency. Non-steroidal anti-inflammatory drugs: cardiovascular risk. *Drug Safety Update Vol. 2 Issue 7*, 2009.
8. Holbrook AM, Pereira JA, Labiris R et al. Systematic Overview of Warfarin and Its Drug and Food Interactions. *Archives of Internal Medicine* 2005;165(10):1095-106.
9. Higashi T, Shekelle P, Solomon DH et al. The quality of pharmacological care for vulnerable older patients. *Archives of Internal Medicine* 2004;140:714-20.
10. Gallagher P, O'Mahony D. STOPP (Screening Tool of Older Persons' potentially inappropriate Prescriptions): application to acutely ill elderly patients and comparison with Beers' criteria. *Age and Ageing* 2008;37(6):673-79.
11. Fick DM, Cooper JW, Wade WE et al. Updating the Beers Criteria for potentially inappropriate medication use in older adults. *Archives of Internal Medicine* 2003;163 2716-24.
12. Guthrie B, Clark SA, McCowan C. The burden of psychotropic drug prescribing in people with dementia: a population database study. *Age Ageing* 2010;39(5):637-42.
13. Vincent C. *Patient Safety*. 2nd edition ed. London: Wiley-Blackwell, 2010.
14. Elder R, Kirkpatrick M, Ramsay W et al. Measuring quality in primary medical services using data from SPICE. Edinburgh: Information and Statistics Division, NHS National Services Scotland, 2007.
15. NPSA National Reporting and Learning Service. Never Events: Framework Update for 2010/11. London: National Patient Safety Agency, 2010.
16. Cates C, Lasserson T, Jaeschke R. Regular treatment with formoterol and inhaled steroids for chronic asthma: serious adverse events. *Cochrane Database of Systematic Reviews* 2009;Issue 2. Art. No.: CD006924:DOI: 10.1002/14651858.CD006924.pub2.
17. Medicines and Healthcare products Regulatory Agency. Long-acting ß2 agonists for asthma: review. *Drug Safety Update Vol. 1 Issue 6*, 2008.
18. Scottish Intercollegiate Guidelines Network. SIGN 95: Management of chronic heart failure. Edinburgh: SIGN, 2007.
19. Spiegelhalter DJ. Handling over-dispersion of performance indicators. *Quality and Safety in Health Care* 2005;14(5):347-51.

This protocol refers to independent research commissioned by the National Institute for Health Research (NIHR). Any views and opinions expressed therein are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the HSR programme or the Department of Health.