

PHARMACOEPIDEMOLOGY

Potentially inappropriate prescribing according to STOPP and START and adverse outcomes in community-dwelling older people: a prospective cohort study

Correspondence Mr Frank Moriarty, HRB Centre for Primary Care Research, Department of General Practice, Royal College of Surgeons in Ireland, 123 St Stephens Green, Dublin 2, Ireland. Tel.: +353 1402 2720; Fax: +353 1402 2724; E-mail: frankmoriarty@rcsi.ie

Received 2 March 2016; **revised** 21 April 2016; **accepted** 28 April 2016

Frank Moriarty¹, Kathleen Bennett², Caitriona Cahir², Rose Anne Kenny³ and Tom Fahey¹

¹HRB Centre for Primary Care Research, Department of General Practice, Royal College of Surgeons in Ireland, Dublin, ²Division of Population Health Sciences, Royal College of Surgeons in Ireland, Dublin, and ³The Irish Longitudinal Study on Ageing, Trinity College Dublin, Dublin, Ireland

Keywords activities of daily living, health care utilization, potentially inappropriate prescribing, quality of life, START, STOPP

AIMS

This study aims to determine if potentially inappropriate prescribing (PIP) is associated with increased healthcare utilization, functional decline and reduced quality of life (QoL) in a community-dwelling older cohort.

METHOD

This prospective cohort study included participants aged ≥ 65 years from The Irish Longitudinal Study on Ageing (TILDA) with linked administrative pharmacy claims data who were followed up after 2 years. PIP was defined by the Screening Tool for Older Persons Prescriptions (STOPP) and Screening Tool to Alert doctors to Right Treatment (START). The association with number of emergency department (ED) visits and GP visits reported over 12 months was analyzed using multivariate negative binomial regression adjusting for confounders. Marginal structural models investigated the presence of time-dependent confounding.

RESULTS

Of participants followed up ($n = 1753$), PIP was detected in 57% by STOPP and 41.8% by START, 21.7% reported an ED visit and 96.1% visited a GP (median 4, IQR 2.5–6). Those with any STOPP criterion had higher rates of ED visits (adjusted incident rate ratio (IRR) 1.30, 95% confidence interval (CI) 1.02, 1.66) and GP visits (IRR 1.15, 95%CI 1.06, 1.24). Patients with two or more START criteria had significantly more ED visits (IRR 1.45, 95%CI 1.03, 2.04) and GP visits (IRR 1.13, 95%CI 1.01, 1.27) than people with no criteria. Adjusting for time-dependent confounding did not affect the findings.

CONCLUSIONS

Both STOPP and START were independently associated with increased healthcare utilization and START was also related to functional decline and QoL. Optimizing prescribing to reduce PIP may provide an improvement in patient outcomes.

WHAT IS ALREADY KNOWN ABOUT THIS SUBJECT

- Potentially inappropriate medicines and potential prescribing omissions are common issues in older people.
- Evidence of a link between these process measures of care and patient outcomes is important.
- Many studies to date have been hospital-based and cross-sectional and evidence of an effect on patient-centred outcomes is less clear.

WHAT THIS STUDY ADDS

- In this community-dwelling older cohort, potentially inappropriate medicines were associated with increased emergency department and GP visits.
- Patients with multiple potential prescribing omissions had higher healthcare utilization, increased chance of functional decline and reduced quality of life.
- Optimizing treatment to address potentially inappropriate prescribing may improve outcomes for older people

Introduction

Older people are particularly vulnerable to adverse effects from medicines, partly due to pharmacokinetic and pharmacodynamic changes in ageing, and also because multimorbidity and complex drug regimens involving multiple medicines (polypharmacy) are common in this age group [1–3]. This has led to concerns regarding potentially inappropriate prescribing (PIP) in older people, including both errors of commission and omission [4]. The first form of PIP refers to potentially inappropriate medicines (PIMs), the use of medicines in circumstances where the risks outweigh the benefits or where a safer or better alternative exists. The second form of PIP is potential prescribing omissions (PPOs), medications which are clinically indicated for a patient not being prescribed. A number of explicit criteria have been developed to identify PIP and two commonly used measures are the Screening Tool for Older Persons' Prescriptions (STOPP), which focuses on PIMs, and the Screening Tool to Alert doctors to Right Treatment (START) to screen for PPOs [5].

While STOPP and START can be considered process measures of medication safety [6], it is important to establish that such prescribing does have an effect on patient outcomes, such as adverse drug events (ADEs), hospitalizations or quality of life (QoL). A number of studies have assessed the association of STOPP PIMs with such outcomes. However the impact of START PPOs on patients has received little attention [7]. Much of this research has been cross-sectional with limited capacity to determine the prospective relationship of PIMs and PPOs with patient outcomes. It is difficult to establish whether an association is causal using such study designs due to potential bias and confounding. Longitudinal cohort studies provide a more robust method to assess the impact of medication exposure as they can account for confounding by time varying factors using appropriate methods and may allow for inference of causal effects [8].

The aim of this study was to determine the association of potentially inappropriate prescribing detected by STOPP and START with healthcare utilization, functional decline and QoL in a cohort of community-dwelling people aged ≥ 65 years in a longitudinal study.

Methods

Study design

This study included participants from The Irish Longitudinal Study on Ageing (TILDA), a nationally-representative cohort study charting the health, economic and social circumstances of community dwellers aged ≥ 50 years. Some TILDA participants also consented to use of their administrative

pharmacy claims data from the Health Service Executive Primary Care Reimbursement Service (HSE-PCRS). Participants were included in the present study if they were aged ≥ 65 years at baseline TILDA interview, had been followed up after 2 years, were eligible for the General Medical Services (GMS) scheme and provided an identifier which was successfully linked to their pharmacy claims data. The GMS scheme provides free health services and prescribed medicines to eligible persons in Ireland. However a small monthly co-payment of €0.50 per prescription item has applied since October 2010. Eligibility for the GMS scheme is based on means testing, although all people aged over 70 years were eligible until December 2008 when a higher income threshold was introduced for this age group compared with the general population. However approximately 96% of this age group were still eligible in 2012 [9]. The STROBE standardized reporting guidelines for cohort studies have been followed in the reporting of this research [10].

Data collection for TILDA is conducted in waves every 2 years, including a face-to-face interview and self-completion questionnaire. Baseline data collection was carried out between 2009 and 2011 and participants were followed up from 2012 to 2013. Medication dispensing data were extracted from the HSE-PCRS pharmacy claims database for each participant in the present study from 15 months before the date of their TILDA baseline interview up to their follow-up interview to determine PIP exposure and all data were anonymized after extraction. Ethical approval for TILDA has been granted by Trinity College Dublin Faculty of Health Sciences research ethics committee.

Outcomes

The primary outcome under investigation was healthcare utilization, including both hospital emergency department (ED) visits and general practitioner (GP) visits. Healthcare utilization was assessed during TILDA interview by asking participants in the previous 12 months how many times did they visit a hospital ED as a patient and about how often they visited their GP. Regression models were fitted for ED visits and GP visits separately using the numbers of visits reported by participants in the 12 months preceding their follow-up interview as the dependent variables.

Two secondary outcomes were also analyzed. The first was decline in physical functioning. Physical functioning is assessed during TILDA interview by asking participants if they have difficulty doing any of six named activities of daily living (ADLs) due to a health or memory problem (including dressing, eating and using the toilet). The outcome variable used was binary, classified as an increase in the number of ADLs a participant reported difficulty with between baseline

and follow-up (functional decline) or no increase (no decline). Secondly QoL was investigated, which was assessed in the TILDA self-completion questionnaire using the CASP (control, autonomy, self-realization, pleasure), a measure designed for use in middle-aged and older people. In this analysis, participants' QoL score at follow-up measured using the CASP-R12, a revised 12 item version of CASP with a possible range from 0 (worst QoL) to 36 (best) [11], was included as the continuous outcome variable.

Exposure

Exposure to PIMs measured by STOPP and PPOs measured by START was determined in this cohort of TILDA participants with linked pharmacy claims data and was reported previously [12]. Briefly, 45 of 65 (69%) STOPP criteria and 15 of 22 (68%) START criteria were applied to determine the prevalence of PIP in the 12 months preceding baseline and follow-up interviews of TILDA (applicable criteria are listed in Table S1). The number of criteria that a participant was exposed to in each time period was determined separately for both screening tools. Exposure to STOPP for each participant in the 12 months preceding outcome measurement and START exposure were included as the two main independent variables of interest.

Statistical analysis

Negative binomial regression models for the reported number of ED visits in 12 months and number of GP visits in 12 months were fitted, including two binary variables for the presence of any STOPP criteria and any START criteria (model 1). Further analysis investigated a dose–response relationship by replacing these binary variables with categorical variables for 0, 1 or ≥ 2 criteria (model 2). Results are presented as incident ratio ratios (IRR) with 95% confidence intervals (CI). This approach was then used for the secondary outcomes of functional decline using logistic regression models and QoL using linear regression models. Results of these analyses are presented as odds ratios (OR) and β regression coefficients, respectively, with 95% CI. Models were adjusted for age group, gender, level of educational attainment as a measure of socioeconomic status, living arrangements, number of repeat medicines and number of doctor-diagnosed chronic conditions (detailed variable description in Table S2). Specific covariates were also adjusted for in each model relating to the outcome of interest, for example, private health insurance status and number of ED/GP visits reported at baseline interview in the analysis of healthcare utilization. The possibility of an additional effect in individuals exposed to both PIMs and PPOs was assessed by the addition of an interaction term to each model and likelihood ratio tests were used to evaluate if this improved model fit.

The impact of time-dependent confounding (by number of regular medicines or chronic conditions, for example) was investigated using marginal structural models (MSMs), a two step estimation strategy which separates confounding control for covariates that vary with time from parameter estimation [8, 13]. For each participant two weights were calculated, (i) stabilized inverse probability weights, the inverse of the probability of having the PIP exposure they did conditional on past PIP exposure and covariate history

(including measurements from both baseline and follow-up) and (ii) censoring weights, the probability of remaining uncensored given past PIP exposure and covariate history. Weighted regression analyses (i.e. MSMs) were performed using the product of these weights for each outcome for STOPP PIM exposure and separately for START PPO exposure, adjusting for baseline covariates only. Statistical significance was assumed at $P < 0.05$. Analyses were performed using Stata version 13 (Stata Corporation, College Station, TX, USA).

Results

Participant inclusion in this study is shown in Figure 1 and a description of these individuals is included in Table 1. Those followed up ($n = 1753$) were mainly female (54.5%), had a mean age of 76.5 years (standard deviation (SD) 6), a median of 6 regular dispensed medicines and three reported doctor-diagnosed chronic conditions. Regarding PIP, 57% of participants had a STOPP PIM in the 12 months preceding follow-up (of these 30.1% had one and 26.9% had two or more PIMs) and the prevalence of START PPOs was 41.8% (with 29.2% having one PPO and 12.6% having multiple PPOs) [12]. The most common STOPP criteria in the cohort were proton pump inhibitors at maximal dose for >8 weeks, aspirin with no history of coronary, cerebral or peripheral arterial symptoms or occlusive arterial event and non-steroidal anti-inflammatory drugs (NSAIDs) with moderate to severe hypertension, while prevalent START omissions were calcium and vitamin D supplements in osteoporosis and anticoagulation in cases of atrial fibrillation or arrhythmia [12]. The percentage of individuals with both STOPP and START criteria was 24.8%.

In the 12 months preceding follow-up interview, 16.1% of participants reported one ED visit, 3.8% reported two visits and 1.8% reported three or more, while 96.1% of participants reported visiting a GP (median 4 visits, interquartile range (IQR) 2.5–6). Results of the healthcare utilization analysis are presented in Table 2. In the multivariate model for ED visits adjusted for covariates, presence of any STOPP PIM was significantly associated with higher rates of visits while the presence of a START PPO was not significantly associated. When number of criteria was considered, there was a statistically significant increase in the rate of ED visits for those with two or more STOPP criteria (adjusted IRR 1.42, 95% CI 1.06, 1.91) as well as for multiple START criteria (adjusted IRR 1.45, 95% CI 1.03, 2.04) relative to those with no criteria. For GP visits, having any STOPP PIM was associated with an increased rate of visits and having any PPO determined by START was not associated with a significant increase. In the model including number of criteria, the relationship of STOPP persisted regardless of number of PIMs while two or more START PPOs were also significantly associated with increased GP visits (adjusted IRR 1.13, 95% CI 1.01, 1.27).

Difficulties with ADLs were reported by 7.7% of participants at baseline and 8.3% of participants reported an increase in ADLs which caused difficulty at follow-up. In the multivariate logistic regression analysis having any START PPO was significantly associated with functional decline, with a larger effect in the dose–response model for those with multiple criteria (adjusted OR 2.06, 95% CI 1.25, 3.39).

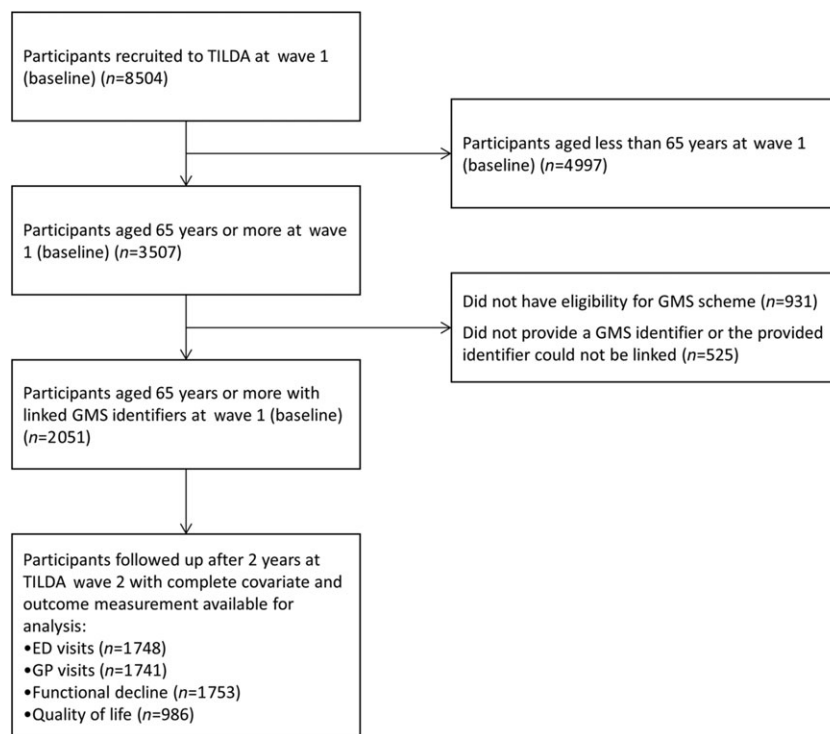


Figure 1

Flow diagram of study participants. The Irish Longitudinal Study on Ageing (TILDA)

However no evidence of an effect due to STOPP was found (Table 3). CASP-R12 scores at follow-up ranged from 5 to 36 (mean 26.2, SD 5.2). Multivariate linear regression found that neither presence of any STOPP nor any START criteria was significantly associated with CASP-R12 score (Table 4). In model 2, exposure to two or more START PPOs was associated with a small but statistically significant reduction in QoL (adjusted β coefficient -1.05 , 95% CI -1.83 , -0.26).

Variables for the interaction between STOPP and START showed no statistically significant association ($P > 0.05$) with any of the outcomes and likelihood ratio tests provided no evidence of improved model fit and, therefore, interactions terms were not included. In the MSMs analysis weighted by the inverse probability of exposure to a STOPP PIM (binary) and the probability of censoring to account for loss to follow-up at wave 2, the IRR for ED visits decreased in magnitude and became marginally non-significant (adjusted IRR 1.27, 95% CI 0.99, 1.64) and for GP visits the estimate also decreased slightly, but remained significant. For START, the adjusted odds ratio for functional decline increased slightly to 1.61 (95% CI 1.10, 2.34) in the MSM which may suggest a degree of confounding by indication by a time-dependent covariate. Results from the MSMs and standard analyses for each outcome are presented in Table S3.

Discussion

Older people in this study who were prescribed a STOPP PIM visited ED and their GP more often (for those with two or

more PIMs, 42% and 16% increases in rate, respectively). However no evidence of a relationship with functional decline and QoL was found. Participants with multiple PPOs had higher rates of healthcare utilization (45% higher rate of ED visits and 13% more GP visits) and a small reduction in QoL. Having a START PPO was also associated with higher odds of functional decline over a 2 year period. Time varying confounding did not appear to play a role in these associations.

Studies on the impact of PIP have predominantly used cross-sectional or retrospective cohort designs so this study is one of few to examine the prospective relationship between STOPP and START and patient outcomes [7]. One prospective study of older hospitalized patients found a significant association between STOPP and avoidable ADEs [14] and this is supported by other work on ADEs [15–18]. For the outcomes examined in the present study, the findings appear to be consistent with previous research, in that the weight of evidence supports an association of STOPP with hospital visits [16, 19], while fewer studies have shown an effect of STOPP on health-related QoL [16], vulnerability [19] and functional decline during hospital stay [17] and START on non-cardiovascular mortality [20].

All studies that have applied STOPP and START together in the same study have been hospital based with limited research on older populations in the primary care setting. Secondary analysis of data from a trial of a hospital pharmacist intervention found the only significant association was between number of STOPP criteria and number of medication-related hospital readmissions [21] and that both STOPP and START had poor discriminative ability to identify

Table 1

Descriptive statistics for participants at baseline (wave 1) and follow-up (wave 2)

	Baseline (Wave 1) (n = 2051)	Follow-up (Wave 2) (n = 1753)
Age (years, mean (SD))	74.8 (6.17)	76.5 (6.04)
Age group (years, n (%))		
65–74	1087 (53.0)	754 (43.0)
≥75	964 (47.0)	999 (57.0)
Gender (Female, n (%))	1107 (54.0)	953 (54.4)
Number of repeat drug classes (median (IQR))	5 (3–8)	6 (3–9)
Number of reported conditions (n (%))		
0	214 (10.4)	88 (5.0)
1	423 (20.6)	268 (15.3)
2	498 (24.3)	370 (21.1)
3 or more	916 (44.7)	1027 (58.6)
Level of education attainment (n (%))		
None/primary	1056 (51.5)	879 (50.2)
Secondary	642 (31.3)	565 (32.3)
Third/higher	351 (17.1)	308 (17.6)
Living arrangements (n (%))		
Living alone	718 (35.0)	626 (30.5)
Living with spouse	965 (47.1)	793 (38.7)
Living with others	368 (17.9)	632 (30.8)
Private health insurance (n (%))	891 (43.4)	760 (43.4)
Diagnosed mental health condition (n (%))	129 (6.3)	157 (9.0)
Any hospital admission (n (%))	354 (17.3)	366 (20.9)
Moderate activity (n (%))	799 (39.0)	751 (42.8)
Depressive symptoms (n (%))*		
None	1172 (58.2)	1248 (74.9)
Sub-clinical	613 (30.4)	277 (16.6)
Clinical	230 (11.4)	141 (8.5)
Social participation (n (%))	943 (46.0)	826 (47.1)

*Depressive symptoms measured by Centre for Epidemiological Studies Depression scale missing for 36 participants at baseline and 87 participants at follow-up.

older patients at risk of unplanned rehospitalization or death [22]. A study of patients following hip fracture showed higher all cause mortality among patients with a greater combined number of STOPP and START criteria [23]. A case-control study of medication-related hospital admissions found an association with STOPP criteria and a composite of STOPP and START [24]. Studies that have used different measures of inappropriate prescribing have also found an association with adverse outcomes in community-dwelling older people, such as the Medication Appropriateness Index and high risk prescribing classified using the Drug Burden Index [25, 26]. A recent trial in general practice targeting high risk use of NSAIDs and antiplatelet drugs significantly reduced not only the targeted prescribing but also the rate of hospitalizations for related adverse events [27].

Patients with either STOPP PIMs or START PPOs appear to have poorer outcomes, so incorporating review of these criteria into the care of older people and acting to rectify situations defined as inappropriate may benefit patients. When screening tools such as STOPP and START were developed, criteria were included if deemed by expert consensus to be

potentially inappropriate with a marginally unfavourable risk : benefit ratio. This study provides evidence to support that this is the case and that there is an association between such prescribing and harm for patients. This is independent of the effect of number of medications, lending credence to the view that polypharmacy itself is not necessarily detrimental, but can be if it includes inappropriate prescribing [28]. However given the limited time available to healthcare professionals to review and optimize treatment, the modest size of the effect of PIP should be considered when prioritizing issues to spend time on with patients. If reviewing PIP can be incorporated easily into routine clinical practice, for example through clinical decision support systems or by streamlining explicit measures to focus on fewer high risk criteria, using these screening tools may be an efficient way to avoid extra healthcare utilization, functional decline or reduced QoL. Further research should consider the cost-effectiveness of such approaches and large scale prospective cohort studies or economic modelling would provide evidence to identify the most clinically significant prescribing issues to focus on in practice.

Table 2

Number (percentage) with an emergency department (ED) visit and mean (SD) GP visits the 12 months preceding follow-up by subgroup and adjusted incident rate ratios (95% CI) for ED visits ($n = 1748$) and GP visits ($n = 1741$)

	<i>n</i> (%)	Emergency department visits Adjusted IRR (95% CI)*	
		Model 1†	Model 2‡
Any STOPP PIM (vs. none)	246 (25.6)	1.30 (1.02, 1.66)§	–
Number of STOPP PIMs			
0 (reference)	134 (17.0)		
1	121 (22.4)	–	1.23 (0.94, 1.62)
≥2	125 (29.6)	–	1.42 (1.06, 1.91)§
Any START PPO (vs. none)	174 (26.1)	1.23 (0.98, 1.53)	–
Number of START PPOs			
0 (reference)	206 (19.0)		
1	118 (24.3)	–	1.15 (0.90, 1.46)
≥2	56 (30.9)	–	1.45 (1.03, 2.04)§
	Mean (SD)	GP visits Adjusted IRR (95% CI)*	
		Model 1†	Model 2‡
Any STOPP PIM (vs. none)	6.3 (6.4)	1.15 (1.06, 1.24)§	–
Number of STOPP PIMs			
0 (reference)	4.5 (4.9)		
1	5.9 (6.3)	–	1.14 (1.05, 1.25)§
≥2	6.8 (6.6)	–	1.16 (1.06, 1.28)§
Any START PPO (vs. none)	5.9 (6.1)	1.04 (0.97, 1.12)	–
Number of START PPOs			
0 (reference)	5.2 (5.7)		
1	5.6 (5.8)	–	1.01 (0.93, 1.09)
≥2	6.9 (6.6)	–	1.13 (1.01, 1.27)§

*Adjusted for age group, gender, number of repeat drug classes, number of reported conditions, level of educational attainment, living arrangements, private health insurance status and number of ED/GP visits reported at baseline; †PIP exposure assessed using binary variables for presence or absence of STOPP and START; ‡PIP exposure assessed using categorical variables for presence of 0, 1 and ≥2 STOPP and START criteria; § $P < 0.05$.

Table 3

Number (percentage) with an increase in ADL difficulties (functional decline) between baseline and follow-up by subgroup and adjusted odds ratios (95% CI) for functional decline compared with no functional decline ($n = 1753$)

	<i>n</i> (%)	Functional decline Adjusted OR (95% CI)*	
		Model 1†	Model 2‡
Any STOPP PIM (vs. none)	110 (11.0)	1.23 (0.78, 1.92)	–
Number of STOPP PIMs			
0 (reference)	35 (4.6)		
1	45 (8.5)	–	1.23 (0.75, 2.02)
≥2	65 (13.8)	–	1.25 (0.75, 2.06)
Any START PPO (vs. none)	84 (11.5)	1.55 (1.07, 2.25)§	–
Number of START PPOs			
0 (reference)	61 (6.0)		
1	50 (9.8)	–	1.35 (0.89, 2.04)
≥2	34 (15.4)	–	2.06 (1.25, 3.39)§

*Adjusted for age group, gender, number of repeat drug classes, number of reported conditions, level of educational attainment, living arrangements, reporting diagnosis of a mental health conditions, reporting a hospital admission in the 12 months preceding follow-up and reporting moderate activity at baseline; †PIP exposure assessed using binary variables for presence or absence of STOPP and START; ‡PIP exposure assessed using categorical variables for presence of 0, 1 and ≥2 STOPP and START criteria; § $P < 0.05$.

Table 4Mean (SD) of CASP-R12 quality of life score at follow-up by subgroup and adjusted β coefficient (95% CI) for CASP-R12 score ($n = 986$)

	Mean (SD)	CASP-R12 score Adjusted β coefficient (95% CI)*	
		Model 1†	Model 2‡
Any STOPP PIM (vs. none)	25.5 (5.4)	-0.26 (-0.81, 0.29)	-
Number of STOPP PIMs			
0 (reference)	27.0 (5.0)		
1	26.2 (5.1)	-	-0.21 (-0.81, 0.39)
≥2	24.7 (5.5)	-	-0.45 (-1.16, 0.27)
Any START PPO (vs. none)	25.5 (5.4)	-0.24 (-0.75, 0.26)	-
Number of START PPOs			
0 (reference)	26.7 (5.1)		
1	26.0 (5.2)	-	0.08 (-0.48, 0.64)
≥2	24.2 (5.7)	-	-1.06 (-1.84, -0.27)§

*Adjusted for age group, gender, number of repeat drug classes, number of reported conditions, level of educational attainment, living arrangements, level of depressive symptoms, reporting social participation and CASP-R12 score at baseline; †PIP exposure assessed using binary variables for presence or absence of STOPP and START; ‡PIP exposure assessed using categorical variables for presence of 0, 1 and ≥ 2 STOPP and START criteria; § $P < 0.05$.

For patients who are identified as having PIP, discussing advantages and disadvantages of any medication change with the patient themselves is important. A recent trial in general practice to reduce PIP found that more changes were made when patients were present for medication review [29]. Any discussion should put particular emphasis on the patient's own priorities as they may place different weights on various benefits and risks. This is especially important when considering starting a new medication to address a PPO, as it may be preferable to both prescriber and patient to not start a preventive treatment in advancing older age despite it being indicated [30]. Rigidly applying treatment guidelines can be ineffective as they do not take account of either comorbidities or patients' preferences and evidence often comes from trials which did not include older patients [31, 32]. This is in contrast to the process of addressing PIMs which may require consideration of stopping a medicine [33]. Both types of PIP present distinct challenges and different approaches may be needed to address potential errors of commission and omission [34].

Adjustment for time-dependent confounding using MSMs did not alter the results here, possibly because factors such as number of medicines were relatively time stable over the study period. This may relate to therapeutic inertia, failure to start new drugs [35] and conversely due to prescriber and patient reluctance to deprescribe treatments for fear of negative consequences [36]. Although MSMs may provide better evidence for causal relationships than conventional regression analyses, associations from longitudinal studies should also be interpreted in the context of other criteria for causation such as the Bradford Hill criteria [37].

This is one of the first longitudinal studies in the community setting to determine the prospective relationship between PIP and adverse outcomes. This robust design allowed for baseline differences to be accounted for and also addressed a number of criteria for inference of causality in epidemiological studies, including temporality and biological

gradient [37]. Pharmacy claims data as used here may provide more reliable determination of medication exposure compared with self-reported drug use [38]. However as dispensing data were used, it had to be assumed that dispensed medications were actually consumed and information on non-prescription medicines use was not available.

Limitations to the study were that a proportion of criteria from STOPP and START could not be applied to the data available in this study (see Table S1 for applicable criteria) and for those that were applied, this study did not have sufficient power to determine if any individual criteria were more strongly associated with adverse patient outcomes. The outcomes of healthcare utilization and functional decline were patient reported rather than objective measures which may have affected accuracy. However the range of covariates adjusted for should have addressed any systematic reporting bias amongst participant subgroups [39, 40]. The CASP-R12 is a measure of QoL rather than health-related QoL specifically and so may not have been sensitive to changes in the health status of participants. While this cohort was well characterized, there is still potential for the presence of unmeasured or unknown confounders. Although reverse causality could explain the relationships of START with adverse outcomes, i.e. preventive treatments being omitted in frailer patients with limited life expectancy who then experience functional decline and reduced QoL, controlling for baseline outcome differences and time varying covariates should reduce this possibility [8].

PIP determined by STOPP and START was associated with adverse outcomes in a prospective older community-dwelling cohort. If application of these criteria can be integrated into routine medication review, they may help to support prescribers in optimizing treatment and improve patient outcomes. Although such prescribing is only potentially inappropriate, the independent effects identified add weight to the suggestion that PIP is a marker of healthcare quality and patient safety and should be minimized if possible.

Competing Interests

All authors have completed the Unified Competing Interest form at http://www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare FM had support from the Health Research Board in Ireland (HRB) through the HRB PhD Scholars Programme in Health Services Research (grant no. PHD/2007/16) and TF had support from the HRB through the HRB Centre for Primary Care Research (grant no. HRC/2007/1) for the submitted work. TILDA was supported by Department of Health and Children, The Atlantic Philanthropies and Irish Life. The authors also declare no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years and no other relationships or activities that could appear to have influenced the submitted work.

The authors wish to acknowledge the Health Service Executive Primary Care Reimbursement Service (HSE-PCRS) for providing access to the administrative pharmacy claims data used in this study.

Contributors

All authors conceived and designed this study. Data were acquired by KB (HSE-PCRS) and RAK (TILDA). FM carried out the statistical analysis and all authors interpreted the data. The manuscript was drafted by FM and all authors were involved in the critical revision of this and approval of the final manuscript.

References

- Mallet L, Spinewine A, Huang A. The challenge of managing drug interactions in elderly people. *Lancet* 2007; 370: 185–91.
- Marengoni A, Angleman S, Melis R, Mangialasche F, Karp A, Garmen A, *et al.* Aging with multimorbidity: a systematic review of the literature. *Ageing Res Rev* 2011; 10: 430–9.
- Fulton MM, Riley Allen E. Polypharmacy in the elderly: a literature review. *J Am Acad Nurse Pract* 2005; 17: 123–32.
- Levy HB, Marcus E-L, Christen C. Beyond the Beers criteria: A comparative overview of explicit criteria. *Ann Pharmacother* 2010; 44: 1968–75.
- 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.
- Donabedian A. Evaluating the quality of medical care. 1966. *Milbank Q* 2005; 83: 691–729.
- 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.
- Robins JM, Hernán MÁ, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000; 11: 550–60.
- Health Service Executive. Primary Care Reimbursement Service Annual Report. Dublin, 2012.
- von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med* 2007; 4: e296.
- Sexton E, King-Kallimanis BL, Conroy RM, Hickey A. Psychometric evaluation of the CASP-19 quality of life scale in an older Irish cohort. *Qual Life Res* 2013; 22: 2549–59.
- Moriarty F, Bennett K, Fahey T, Kenny RA, Cahir C. Longitudinal prevalence of potentially inappropriate medicines and potential prescribing omissions in a cohort of community-dwelling older people. *Eur J Clin Pharmacol* 2015; 71: 473–82.
- Fewell Z, Hernán M, Wolfe F, Tilling K, Choi H, Sterne J. Controlling for time-dependent confounding using marginal structural models. *Stata J* 2004; 4: 402–20.
- Hamilton H, Gallagher P, Ryan C, Byrne S, O'Mahony D. Potentially inappropriate medications defined by STOPP criteria and the risk of adverse drug events in older hospitalized patients. *Arch Intern Med* 2011; 171: 1013–9.
- Gallagher P, O'Mahony D. STOPP (Screening Tool of Older Persons' potentially inappropriate Prescriptions): application to acutely ill elderly patients and comparison with Beers' criteria. *Age Ageing* 2008; 37: 673–9.
- Cahir C, Bennett K, Teljeur C, Fahey T. Potentially inappropriate prescribing and adverse health outcomes in community dwelling older patients. *Br J Clin Pharmacol* 2014; 77: 201–10.
- Tosato M, Landi F, Martone AM, Cherubini A, Corsonello A, Volpato S, *et al.* Potentially inappropriate drug use among hospitalised older adults: results from the CRIME study. *Age Ageing* 2014; 0: 1–7.
- Hedna K, Hakkarainen KM, Gyllensten H, Jönsson AK, Petzold M, Hägg S. Potentially inappropriate prescribing and adverse drug reactions in the elderly: a population-based study. *Eur J Clin Pharmacol* 2015; 71: 1525–33.
- Cahir C, Moriarty F, Teljeur C. Potentially inappropriate prescribing and vulnerability and hospitalization in older community-dwelling patients. *Ann Pharmacother* 2014; 48: 1546–54.
- Meid AD, Quinzler R, Freigofas J, Saum K-U, Schöttker B, Holleczer B, *et al.* Medication Underuse in Aging Outpatients with Cardiovascular Disease: Prevalence, Determinants, and Outcomes in a Prospective Cohort Study. *PLoS One* 2015; 10: e0136339.
- Gillespie U, Alsaad A, Hammarlund-Udenaes M, Mörlin C, Henrohn D, Bertilsson M, *et al.* Effects of pharmacists' interventions on appropriateness of prescribing and evaluation of the instruments' (MAI, STOPP and STARTs') ability to predict hospitalization—analyses from a randomized controlled trial. *PLoS One* 2013; 8: e62401.
- Alsaad A, Melhus H, Hammarlund-Udenaes M, Bertilsson M, Gillespie U, Sundstrom J. A tool for prediction of risk of rehospitalisation and mortality in the hospitalised elderly: secondary analysis of clinical trial data. *BMJ Open* 2015; 5: e007259–9.
- Gosch M, Wortz M, Nicholas JA, Doshi HK, Kammerlander C, Lechleitner M, *et al.* Inappropriate Prescribing as a Predictor for Long-Term Mortality after Hip Fracture. *Gerontology* 2013; 60: 114–22.

- 24** van der Stelt CAK, Vermeulen Windsant-van den Tweel AMA, Egberts ACG, van den Bemt PMLA, Leentertse AJ, Hermens WAJJ, *et al.* The association between potentially inappropriate prescribing and medication-related hospital admissions in older patients: a nested case control study. *Drug Saf* 2016; 39: 79–87.
- 25** Lund B, Carnahan R, Egge J, Chrischilles E, Kaboli P. Inappropriate prescribing predicts adverse drug events in older adults. *Ann Pharmacother* 2010; 44: 957–63.
- 26** Gnjjidic D, Hilmer SN, Hartikainen S, Tolppanen A-M, Taipale H, Koponen M, *et al.* Impact of high risk drug use on hospitalization and mortality in older people with and without Alzheimer's disease: a national population cohort study. *PLoS One* 2014; 9: e83224.
- 27** Dreischulte T, Donnan P, Grant A, Hapca A, McCowan C, Guthrie B. Safer prescribing – a trial of education, informatics, and financial incentives. *N Engl J Med* 2016; 374: 1053–64.
- 28** Cadogan CA, Ryan C, Hughes CM. Appropriate polypharmacy and medicine safety: when many is not too many. *Drug Saf* 2016; 39: 109–16.
- 29** Clyne B, Smith SM, Hughes CM, Boland F, Bradley MC, Cooper JA, *et al.* Effectiveness of a multifaceted intervention on potentially inappropriate prescribing in older patients in primary care: a cluster randomised controlled trial (the OPTI-SCRIPT study). *Ann Fam Med* 2015; 13: 545–53.
- 30** Mangin D, Sweeney K, Heath I. Preventive health care in elderly people needs rethinking. *BMJ* 2007; 335: 285–7.
- 31** Hughes LD, McMurdo MET, Guthrie B. Guidelines for people not for diseases: the challenges of applying UK clinical guidelines to people with multimorbidity. *Age Ageing* 2013; 42: 62–9.
- 32** Scott IA, Guyatt GH. Cautionary tales in the interpretation of clinical studies involving older persons. *Arch Intern Med* 2010; 170: 587–95.
- 33** Scott IA, Hilmer SN, Reeve E, Potter K, Le Couteur D, Rigby D, *et al.* Reducing Inappropriate Polypharmacy. *JAMA Intern Med* 2015; 175: 827–34.
- 34** Patterson S, Cadogan C, Kerse N, Cardwell C, Bradley M, Ryan C, *et al.* Interventions to improve the appropriate use of polypharmacy for older people. *Cochrane Database Syst Rev* 2014; 10: CD008165.
- 35** Guthrie B, Inkster M, Fahey T. Tackling therapeutic inertia: role of treatment data in quality indicators. *BMJ* 2007; 335: 542–4.
- 36** Anderson K, Stowasser D, Freeman C, Scott I. Prescriber barriers and enablers to minimising potentially inappropriate medications in adults: a systematic review and thematic synthesis. *BMJ Open* 2014; 4: e006544–4.
- 37** Bradford Hill A. The environment and disease: association or causation? *Proc R Soc Med* 1965; 58: 295–300.
- 38** Richardson K, Kenny RA, Pecklar J, Bennett K. Agreement between patient interview data on prescription medication use and pharmacy records in those aged older than 50 years varied by therapeutic group and reporting of indicated health conditions. *J Clin Epidemiol* 2013; 66: 1308–16.
- 39** Reijneveld S, Stronks K. The validity of self-reported use of health care across socioeconomic strata: a comparison of survey and registration data. *Int J Epidemiol* 2001; 30: 1407–14.
- 40** Short M, Goetzel R, Pei X, Tabrizi M, Ozminkowski R, Gibson T, *et al.* How accurate are self-reports? An analysis of self-reported healthcare utilization and absence when compared to administrative data. *J Occup Environ Med* 2009; 51: 786–96.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

<http://onlinelibrary.wiley.com/doi/10.1111/bcp.12995/supinfo>.

Table S1 List of applicable criteria from STOPP and START used to determine exposure to PIMs and PPOs

Table S2 Description of covariates adjusted for in multivariate regression models

Table S3 Comparison of adjusted parameter estimates (95% CI) for presence of any STOPP PIM and any START PPO from unweighted multivariate regression models and marginal structural (weighted) models