BMJ Open Potentially inappropriate medications (PIMs): frequency and extent of GPrelated variation in PIMs: a registerbased cohort study

Anette Riisgaard Ribe ^(b), ¹ Line Due Christensen, ¹ Claus Høstrup Vestergaard, ¹ Anders Prior ^(b), ^{1,2} Peter Krogh Brynningsen, ³ Flemming Bro, ^{1,2} Annelli Sandbæk, ^{2,4} Peter Vedsted, ¹ Daniel R Witte, ^{2,4} Morten Fenger-Grøn¹

To cite: Ribe AR,

Christensen LD, Vestergaard CH, et al. Potentially inappropriate medications (PIMs): frequency and extent of GP-related variation in PIMs: a registerbased cohort study. *BMJ Open* 2021;**11**:e046756. doi:10.1136/ bmjopen-2020-046756

Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online. To view these files, please visit the journal online (http://dx.doi. org/10.1136/bmjopen-2020-046756).

ARR and LDC contributed equally.

Received 08 November 2020 Accepted 24 June 2021

Check for updates

© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

 ¹Research Unit for General Practice, Aarhus, Midtjylland, Denmark
²Department of Public Health, Aarhus University, Aarhus C, Denmark
³Department of Geriatrics, Aarhus University Hospital, Aarhus N, Denmark
⁴Steno Diabetes Center Aarhus, Aarhus N, Denmark

Correspondence to Anette Riisgaard Ribe; arr@ph.au.dk

ABSTRACT

Objectives Potentially inappropriate medications (PIMs) pose an increasing challenge in the ageing population. We aimed to assess the extent of PIMs and the prescriber-related variation in PIM prevalence.

Design Nationwide register-based cohort study. **Setting** General practice.

Participants The 4.2 million adults listed with general practitioner (GP) clinics in Denmark (n=1906) in 2016. Main outcome measures We estimated the patients' time with PIMs by using 29 register-operationalised STOPP criteria linking GP clinics and redeemed prescriptions. For each criterion and each GP clinic, we calculated ratios between the observed PIM time and that predicted by multivariate Poisson regressions on the patients. The observed variation was measured as the 90th/10th percentile ratios of these ratios. The extent of expectable random variation was assessed as the 90th/10th percentile ratios in randomly sampled GP populations (ie, the sampled variation). The GP-related excess variation was calculated as the ratio between the observed variation and sampled variation. The linear correlation between the observed/expected ratio for each of the criteria and the observed/expected ratio of total PIM time (for each clinic) was measured by Pearson's rho.

Results Overall, 294 542 individuals were exposed to 1 44 117 years of PIMs. The two most prevalent PIMs were long-term use (>3 months) of non-steroidal anti-inflammatory drugs (51 074 years of PIMs) or benzodiazepines (48 723 years of PIMs). These two criteria showed considerable excess variation of 2.33 and 3.05, respectively; for total PIMs, this figure was 1.65. For more than half of the criteria, we observed a positive correlation between the specific PIM and the sum of remaining PIMs. **Conclusions** This study documents considerable variations in the prescribing practice of GPs for certain PIMs. These findings highlight a need for exploring the causal explanations for such variations, which could be markers of suboptimal GP-prescribing strategies.

INTRODUCTION

Globally, the prevalence of multimorbidity, that is, the coexistence of two or more chronic conditions, is expected to rise in the

Strengths and limitations of this study

- This study is the first to describe the general practitioner (GP)-related variations in potentially inappropriate medications (PIMs), which could serve as a marker of suboptimal prescribing behaviour of the GPs.
- We developed a comprehensive algorithm for software application to measure PIMs (based on the STOPP/START criteria, version 2) using data from the nationwide Danish registers.
- We developed a novel analysis approach for measuring variations in PIMs, which accounted for differences between practice populations and for the fact that some extent of variation is expected due to randomness.
- Although each of the PIM criteria represents 'a red flag' for the prescriber, some PIMs may be the best treatment option for certain patients if administered with care, such as frail patients with high complexity of diseases for whom there may be a need for balancing conflicting demands.
- Residual confounding cannot be excluded as some data were not available in the Danish registers, for example, disease severity measures, measures of frailty, estimated life expectancy and over-thecounter medications.

future. This development is due to improved diagnostics, enhanced survival and an ageing population.^{1 2} However, the complexity of managing multiple concurrent chronic conditions challenges modern healthcare. Multimorbidity may require polypharmaceutical treatment (>5 medications),³ and this holds a risk of both undertreatment and overtreatment.⁴ Potentially inappropriate medications (PIMs)⁵ can be defined as medications in which the risks of use are likely to outweigh the benefits, or safer alternatives exist, and PIMs can thus serve as a proxy for overtreatment.^{5–7} PIMs are common in older

individuals. The prevalence ranges from 21% to 79% in different healthcare contexts, such as hospitals and general practice,^{8–10} and PIMs may have serious consequences for the individual in terms of lower quality of life, higher risk of emergency visits and higher number of hospital admissions.^{8 11} Moreover, PIMs may generate substantial financial costs for both the individual and society.⁴

In Denmark, the general practitioners (GPs) handle over 80% of all prescription medications.¹² They are central players in the management of chronic diseases and thus have a favourable position to assess the patient's medications across diseases. Yet, evidence suggests that barriers to reducing PIMs exist.¹³ Furthermore, considerable variation has been reported for other treatmentrelated activities in general practice;¹⁴¹⁵ this could indicate heterogeneous quality of the provided healthcare. Nevertheless, little is known on whether the prevalence of PIMs varies between GPs, and whether such potential variation may exceed the expected variation caused by differences in patient populations and by random variation. Such excess variation could be a marker of differences in the treatment strategies of the GPs and thus of suboptimal prescribing practices.

Using a coding algorithm to identify PIMs in the Danish nationwide registers, we aimed to explore the prevalence of PIMs and the GP-related variation in the prevalence of PIMs, while accounting for differences between practice populations and for the fact that some variation is expected due to randomness alone.

METHODS

Setting and study population

The study population included all persons aged ≥ 18 years who were listed with a Danish GP clinic and who had at least five consecutive years of residence in Denmark (to allow for the collection of register-based data in the 5years preceding the inclusion). Eligible persons were followed from 1 January 2016 until death, emigration, exit from the GP listing system for other reasons, or 31 December 2016, whichever came first. We excluded GP clinics accumulating less than 500 patient years in 2016 (table 1, online supplemental material 1).

Denmark has free tax-funded healthcare, and more than 98% of the population are listed with a specific GP clinic, which serves as their primary entry point to the healthcare system and as a gatekeeper to specialised care.¹⁶ Around half of the Danish GP clinics are solo practices, around 40% of clinics consist of two–three GPs, and approximately 10% of clinics have more than three GPs.¹⁷

Data sources

The project was carried out as a population-based cohort study using the nationwide Danish registers.¹⁸ We obtained data on sex, age, emigration and death from the Danish Civil Registration System,¹⁹ data on marital status, education, income and ethnicity from Statistics Denmark

(Integrated Database for Longitudinal Labour Market Research),²⁰ data on hospital diagnoses and procedure codes from the Danish National Patient Register,²¹ data on psychiatric hospital diagnoses from the Danish Psychiatric Central Research Register,²² data on redemption dates and redeemed volumes of prescribed medications classified according to the Anatomical Therapeutic Chemical Classification codes from the Danish National Prescription Register,²³ and data on the link between the patient and the GP clinic from the Patient List Database. All data on diagnoses were obtained from public hospitals in Denmark, which must all report to the Danish National Patient Register. Private hospitals account for less than 1% of the total number of beds and do not provide acute care in Denmark.²⁴ All of these data were linked at the individual level through the unique personal identification number provided by the Danish Civil Registration System, which also keeps track of vital status and migrations for all residents in Denmark.²⁵

Identification of PIMs

PIMs can be identified through the formal STOPP criteria for potentially inappropriate drugs, drug-drug combinations and drug-disease combinations that are included in the second version of the validated STOPP/START list, which was developed by O'Mahony and colleagues.²⁶ In this study, we identified PIMs using a coding algorithm²⁷ based on these STOPP criteria,²⁶ which we adapted for a Danish register-based context through iterative consensus group discussions (see details in Supplemental methods, online supplemental material 1). This process led to the selection of 29 STOPP criteria, which could be operationalised for application in the Danish registers. Thus, 47 of the 76 STOPP criteria were excluded due to unavailable data (ie, laboratory measurements, procedures or diagnoses made in primary care, severity of diseases, dosage of medications and some specific medications), due to lack of clinical relevance (generally or in Denmark specifically) or due to overlap between criteria (in some cases, two criteria were collapsed into one after relevant modifications) (see coding details for the STOPP criteria in online supplemental material 2 and details on consensus modifications in online supplemental material 3).

After redeeming a prescription of a drug in any of the PIM categories, individuals were considered continuous 'users' of that drug for a period of time corresponding to the number of redeemed defined daily doses (DDD), that is, the assumed average maintenance dose per day for a drug used for its main indication in adults, plus a grace period of 25% to allow for some leeway when redeeming prescriptions.²⁸ This period of use was extended each time a user redeemed a new prescription. For the PIM criteria defined by two concurrent treatments, some individuals were excluded during quarantine periods in-between treatments, as we could not distinguish between overlapping treatments and treatment shifts before the first prescription expired (see details in Supplemental methods, online supplemental material 1).

Table 1 Extent of selected PIMs* and any PIM† among adults in Denmark in 2016							
STOPP criterion	Persons‡ at risk§	Risk time§, years	Persons‡ with PIMs	PIM time, years (PIM rate¶, %)			
B3: Beta blocker combined with verapamil or diltiazem	4244308	4 207 285.4	3090	1332.8 (0.0)			
B4: Beta blocker combined with bradycardia or heart block	25321	22010.0	10084	5113.8 (23.2)			
B5: Amiodarone as first-line antiarrhythmic therapy in supraventricular tachyarrhythmias	50406	45 528.7	1165	633.0 (1.4)			
B6: Loop diuretic as first-line treatment for hypertension	40440	34969.6	3800	2264.7 (6.5)			
B13: Phosphodiesterase type-5 inhibitors in severe heart failure or concurrent nitrate therapy for angina	69148	65115.3	3186	666.5 (1.0)			
C3: Aspirin, clopidogrel, dipyridamole, vitamin K antagonists, direct thrombin inhibitors or factor Xa inhibitors with concurrent significant bleeding	32829	25931.4	13136	8510.3 (32.8)			
C4: Aspirin plus clopidogrel as secondary stroke prevention, unless the patient has had a coronary stent(s) inserted in the previous 12 months, concurrent acute coronary syndrome, or a high-grade symptomatic carotid arterial stenosis	147722	138 090.5	8149	3712.8 (2.7)			
C6: Antiplatelet agents with vitamin K antagonist, direct thrombin inhibitor or factor Xa inhibitors in patients with stable coronary, cerebrovascular or peripheral arterial disease	4243867	4 195 000.3	20428	8126.5 (0.2)			
C8: Vitamin K antagonist, direct thrombin inhibitor or factor Xa inhibitors for first deep venous thrombosis without continuing provoking risk factors for >6 months	39815	36658.6	3963	2532.4 (6.9)			
C9: Vitamin K antagonist, direct thrombin inhibitor or factor Xa inhibitors for first pulmonary embolus without continuing provoking risk factors for >12 months	12811	11218.6	2703	1675.5 (14.9)			
C10: NSAID and vitamin K antagonist, direct thrombin inhibitor or factor Xa inhibitors in combination	4244304	4 206 445.8	8904	2276.7 (0.1)			
D1: Tricyclic antidepressants (TCAs) with dementia, narrow angle glaucoma, cardiac conduction abnormalities, prostatism or prior history of urinary retention	85667	73481.5	2022	804.3 (1.1)			
D3: Neuroleptics with moderate-marked anticholinergic effects with a history of prostatism, or previous urinary retention	11605	3742.0	67	9.3 (0.2)			
D5: Long-term use of benzodiazepines	4244310	4 207 443.1	81621	48722.9 (1.2)			
D6: Antipsychotics (ie, other than quetiapine or clozapine) in those with parkinsonism or Lewy body dementia	10495	9711.0	849	303.9 (3.1)			
D8: Anticholinergics in patients with delirium or dementia	43757	36604.8	8493	4459.2 (12.2)			
D9: Neuroleptic antipsychotic in patients with behavioural and psychological symptoms of dementia (BPSD)	38832	34661.5	4624	786.7 (2.3)			
D11: Acetylcholinesterase inhibitors with a known history of persistent bradycardia, heart block or recurrent unexplained syncope, or concurrent treatment with drugs that reduce heart rate	12238	6478.4	3747	1808.1 (27.9)			
D12: Phenothiazines as first-line treatment	4156076	4 105 909.6	5047	729.5 (0.0)			
G1: Theophylline as monotherapy for COPD	4 137 262	4 005 654.4	581	251.5 (0.0)			
H6: Long-term use of NSAIDs	4244310	4,207,443.1	125945	51074.4 (1.2)			
H7: COX-2 selective NSAIDs with concurrent cardiovascular disease	391248	380518.5	360	96.1 (0.0)			
I1: Antimuscarinic drugs with dementia, chronic cognitive impairment, narrow-angle glaucoma or chronic prostatism	60600	46522.8	1928	1106.4 (2.4)			
12: Selective alpha-1 selective alpha blockers in those with symptomatic orthostatic hypotension or micturition syncope	20535	6659.3	825	200.2 (3.0)			
J2: Thiazolidenediones in patients with heart failure	68932	65030.6	10	2.4 (0.0)			

Continued

STOPP criterion	Persons‡ at risk§	Risk time§, years	Persons‡ with PIMs	PIM time, years (PIM rate¶, %)
J4: Oestrogens with a history of breast cancer or venous thromboembolism	106975	104283.3	1727	1225.6 (1.2)
J6: Androgens in the absence of primary or secondary hypogonadism	2079319	2 060 386.4	1774	1009.0 (0.0)
K3: Vasodilator drugs with persistent postural hypotension	19534	5964.2	7504	1962.3 (32.9)
M: Concomitant use of ≥2 drugs with anticholinergic properties	4244295	4 206 255.3	15986	6529.2 (0.2)
Any PIM (unique persons and unique person time)†	4244310	4 207 443.1	294542	144 116.5 (3.4)

*PIMs were identified through our coding algorithm for selected STOPP criteria. The algorithm was modified in accordance with the STOPP/START criteria as developed by O'Mahony and colleagues²⁶ and Huibers and colleagues.²⁷

†The measure 'any PIM' included unique individuals and unique PIM time only, which implied that individuals contributed only once with PIM time, even when having two or more concurrent PIMs. Thus, this measure estimated the time spent with at least one PIM. ‡The number of persons represented unique individuals.

§In some cases, the risk population (and their risk time of relevance) included all persons in the study population (ie, B3, C10, D5, H6 and M). Although the risk population for each of these criteria included *all* individuals, the criteria in which PIMs were defined by two or more concurrent treatments (ie, B3, C10 and M) had risk populations that differed slightly from the entire study population due to exclusion of individuals during the quarantine periods in-between treatments. In other cases, the risk population included only patients with a given condition or combination of conditions.

The PIM rate was calculated as the total time spent with PIM divided by the total time spent at risk.

NSAIDs, non-steroidal anti-inflammatory drugs; PIMs, potentially inappropriate medications.

Additional covariates

In addition to the specific medications and medical or psychiatric conditions mentioned in the criteria, all persons under study were characterised in terms of age, sex, cohabitation status, education, income and ethnicity (Appendix I, online supplemental appendix I, online supplemental material 1) as well as morbidity status according to a modified version of the Danish Multimorbidity Index (DMI), which included 36 physical and mental disorders identified from a combination of diagnosis and prescription data from the above-mentioned sources (online supplemental appendix II, online supplemental material 1).²⁹

Statistical analyses

For all PIM criteria, both overall and subgroup-specific crude counts of persons with PIMs were calculated as was the potential time spent with PIMs (PIM time). Furthermore, the PIM rates in the population for each of the criteria were calculated as the PIM time divided by the total time spent at risk (risk time), corresponding to the average proportion of the risk population (specific for the criterion in question) that was exposed to PIMs at a given time during the study period. In some cases, the risk population included all persons in the study population. In other cases, it included only persons with a given condition or combination of conditions.

The analyses of the GP-related variation involved several steps. First, a series of multivariate Poisson regressions were used to estimate the expected PIM time for each of the PIM criteria for all persons according to their covariate status and time at risk (included with the regression parameter restricted to 1). The expected total PIM time for each person was calculated as the sum of expected time with each criterion (0 for criteria that were irrelevant for the given person). Subsequently, these expected PIM times were compared with the observed PIM times for each person and aggregated for each GP clinic, allowing for the calculation of clinic-specific observed/expected ratios. For each criterion, these ratios were sorted and subsequently depicted graphically (in groups of five clinics to comply with data deidentification regulations) to describe the variation between clinics, and the 90th/10th percentile ratio (equaling the ratio between the highest vs the lowest quintile medians) was calculated as a measure of the 'observed variation'. To assess the extent of variation in the GP clinic-specific observed/expected ratios that would be expectable even if no systematic/true differences between the GPs' treatment strategies exist (ie, the anticipatable random variation), we calculated a 'sampled variation' by essentially repeating all of the above steps in a second round, in which the GP clinics were assigned reference populations of matched patients (from the same risk population) sampled from the other clinics. Matching was performed for persons of the same age and sex, who had the best match of estimated propensity for that specific PIM on 1 January 2016. The propensity score included information on cohabitation status, income, education, ethnicity and the 36 comorbidities included in the modified version of the DMI. Finally, we calculated the 'excess variation' as the ratio between the observed variation and the sampled variation. Thus, the excess variation represents a factor measure of the extent to which the observed variation exceeds the level of variation that could be expected

due to chance alone; that is, the excess variation would be one if the observed variation did not exceed the expected random variation, whereas it would be equal to the observed variation in a (hypothetical) situation with no contribution from random variation. For PIM criteria with a 10th percentile equal to zero, excess variation was calculated as the ratio between the 90th percentiles alone to avoid division by zero.

The linear correlation between the observed/expected ratio for each of the criteria and the observed/expected ratio of the total PIM time (for each GP clinic) was measured by Pearson's rho.³⁰ For this calculation, we used a jackknife approach, which implied subtraction of the contribution made by the given criterion from the overall score.

All analyses were performed in Stata V.16, College Station, TX.

Patient and public involvement

We analysed deidentified population-based healthcare register data. Patients were not involved in the development of the research question, the outcome measures or the study design. We plan to disseminate the results of the research to Danish GPs, the general public, and policy-makers.

RESULTS

The study population included a total of 4244310 unique individuals who were at risk of PIMs during 4207443 risk years in 2016; these individuals were distributed across 1906 clinics, that is, approximately 2200 patients per GP clinic (table 1). Here, 27% were older than 60 years of age, 51% were female and 17% had two or more comorbidities (figure 1, online supplemental material 1). However, 20 out of 29 criteria yielded risk populations with substantially higher ages and higher prevalence of multimorbidity (figure 1, online supplemental material 1).

Extent of PIMs

Overall, 294542 unique individuals were exposed to 144 117 PIM years (ie, time spent with at least one PIM), which corresponded to a PIM incidence rate of 3.4% (table 1).

The most prevalent PIM criteria were: B4 (ie, beta blocker with bradycardia or heart block) (5114 years), C3 (ie, aspirin, clopidogrel, dipyridamole, vitamin K antagonists, direct thrombin inhibitors or factor Xa inhibitors with concurrent significant bleeding risk) (8510 years), C6 (ie, antiplatelet agents with vitamin K antagonist, direct thrombin inhibitor or factor Xa inhibitors) (8127 years), D5 (ie, long-term use of benzodiazepines defined as use of more than 90 DDDs within 6 months) (48723 years), D8 (ie, anticholinergics in patients with delirium or dementia) (4459 years), H6 (ie, long-term use of non-steroidal anti-inflammatory drugs (NSAIDs) defined as use of more than 90 DDDs within 6 months) (51074 years) and M (ie, concomitant use of two or more drugs



Figure 1 Variation in the seven most prevalent PIM criteria and total PIMs in adults in 2016. Abbreviations: Observed/expected: observed/expected ratio of PIM time. Obs: observed variation in the actual GP populations (the 90th/10th percentile ratio of the observed/expected ratio of PIM time). Samp: sampled variation in the randomly sampled GP populations (the 90th/10th percentile ratio of the observed/expected ratio of PIM time). Excess: excess variation (the observed variation/sampled variation ratio). *Content for each of the criteria is detailed in Table 1. **Total PIMs represents the sum of all PIM time. Thus, persons can contribute with PIM time more than once if they have multiple PIMs at the same time. Therefore, the number of PIM years for this aggregated measure was different from the number in Table 1, which measured 'any PIM'. In all other analyses, including in this figure, 'total PIMs' was measured. GP, general practitioner; PIMs, potentially inappropriate medications.

with anticholinergic properties) (6529 years) (table 1). Among these seven criteria with the highest prevalence, the PIM rate was 23.2% for B4, 32.8% for C3 and 12.2% for D8, but 0.2% to 1.2% for the remaining four criteria. This was, however, explained by the fact that these criteria included (basically) all individuals from the study population in their risk populations (table 1).

GP-related variation in PIMs

Variation in total PIMs

Across GP clinics, a factor 2.18 variation in total PIMs was observed for the 90th/10th percentile ratio ('observed variation'), whereas a factor 1.32 variation was expected due to random variation alone ('sampled variation'), that is, if no true/systematic differences between GPs existed (figure 1 and table 2). This corresponded to a 1.65 times higher variation than expected when taking into account differences in patient characteristics and randomness alone ('excess variation') (figure 1 and table 2).

Variation in specific criteria

The observed variation was considerable for all criteria. However, only for a few cases, this variation was higher than what could be explained by differences in patient characteristics and randomness. Still, long-term use of NSAIDs, which was the most prevalent type of PIMs, showed an observed variation of 3.59 when differences

Table 2 Variation in all 29 PIM criteria and total PIMs in adults in Denmark in 2016

	Observed vari	Observed variation+		tiont	Excess variation (observed variation/sampled variation ratio)S	Correlation	
	90th/10th	Onth/10th			<u>vanation ratio/3</u>	Pearson's of	
Criterion*	percentile	Ratio	percentile	Ratio	Ratio	(p-value)	
B3	2.79/0	**	2.51/0		1.11	0.04 (0.124)	
B4	1.63/0.30	5.38	1.61/0.32	5.01	1.07	0.00 (0.982)	
B5	3.06/0		3.12/0		0.98	0.01 (0.779)	
B6	2.29/0		2.10/0		1.09	0.13 (<0.001)	
B13	2.74/0		2.62/0		1.05	0.05 (0.033)	
C3	1.49/0.50	2.99	1.46/0.53	2.74	1.09	0.01 (0.551)	
C4††	2.05/0.06	37.17	2.02/0.10	20.52	1.81	0.17 (<0.001)	
C6	1.70/0.30	5.72	1.57/0.42	3.74	1.53	0.06 (0.016)	
C8	2.17/0		2.18/0		0.99	0.02 (0.353)	
C9	2.25/0		2.22/0		1.01	0.03 (0.207)	
C10††	2.27/0.03	72.03	1.88/0.11	17.92	4.02	0.42 (<0.001)	
D1	2.89/0		2.82/0		1.02	0.02 (0.346)	
D3	0/0		0/0		•	0.04 (0.078)	
D5	1.91/0.37	5.11	1.27/0.76	1.68	3.05	0.36 (<0.001)	
D6	2.96/0		3.00/0		0.99	0.03 (0.201)	
D8	1.86/0.12	15.25	1.80/0.15	12.18	1.25	0.06 (0.012)	
D9	2.39/0		2.43/0		0.98	0.05 (0.043)	
D11	2.09/0		2.14/0		0.98	-0.01 (0.671)	
D12	2.85/0		2.69/0		1.06	0.09 (<0.001)	
G1	3.23/0		3.42/0		0.94	0.09 (<0.001)	
H6	1.64/0.46	3.59	1.21/0.79	1.54	2.33	0.45 (<0.001)	
H7	0.81/0		1.97/0		0.41	0.05 (0.024)	
11	2.84/0		2.75/0		1.03	0.07 (0.002)	
12	3.11/0		3.26/0		0.95	0.03 (0.156)	
J2	0/0		0/0			-0.00 (0.962)	
J4	2.89/0		2.75/0		1.05	0.11 (<0.001)	
J6	2.96/0		2.82/0		1.05	0.09 (<0.001)	
K3	1.69/0.28	6.04	1.61/0.15	10.49	0.58	0.03 (0.216)	
Μ	1.76/0.31	5.67	1.60/0.31	5.11	1.11	0.09 (<0.001)	
Total PIMs‡‡	1.44/0.66	2.18	1.11/0.84	1.32	1.65	-	

The bold values highlighted the observed variations and the excess variations.

*Content for each of the criteria is detailed in table 1.

†The observed variation measured the 90th/10th percentile ratio (of the observed/expected ratio of PIM time) in the studied GP populations (actual GP populations).

[‡]The sampled variation measured the 90th/10th percentile ratio (of the observed/expected ratio of PIM time) in the randomly sampled GP populations.

\$To avoid division by zero, excess variation was calculated as the ratio between the 90th percentiles alone for the PIM criteria with 10th percentile equal to zero.

¶The linear correlation between the observed/expected ratio for each of the criteria and the observed/expected ratio of the total PIM time (for each GP clinic) was measured by Pearson's rho.

**Full stop denoted missing data.

††The observed variation for these criteria should be interpreted with caution, as the 10th percentiles were close to zero.

^{‡‡}Total PIMs estimated a composite measure of all STOPP criteria (calculated by summing up observed and expected PIM time for each PIM criterion). The figure for total PIMs was 157 926 PIM years. All analyses on observed/expected PIM time were based on the total PIM measure.

GP, general practitioner; PIMs, potentially inappropriate medications.

in patient characteristics were taken into account, and the excess variation was 2.33 times higher than expected when randomness was also taken into account. Similarly, long-term use of benzodiazepines showed an observed variation of 5.11 and an excess variation of 3.05. Furthermore, the remaining five most prevalent criteria showed excess variations ranging from 1.07 to 1.53 (figure 1 and table 2).

The calculation of the 90th/10th percentile ratio (observed variation) was only feasible for three of the remaining criteria (C4, C10 and K3), as the denominator (10th percentile) was zero for the rest of the criteria. Although all three criteria showed substantial observed variation (C4: 37.17, C10: 72.03 and K3: 6.04), only C4 and C10 showed considerable excess variation when randomness was taken into account (ie, excess variation of 1.81 and 4.02, respectively) (figure 1 and table 2, online supplemental material 1).

For more than half of the criteria, we found a statistically significant positive correlation between the observed/expected ratio (of PIM time) for each of the criteria and the observed/expected ratio of total PIMs (table 2). Among the seven most prevalent criteria, such positive correlation was found for the five criteria with highest excess variation (C6: 0.06 (p-value: 0.016); D5: 0.36 (p-value:<0.001), D8: 0.06 (p-value: 0.012), H6: 0.45 (p-value:<0.001) and M: 0.09 (p-value:<0.001()) (table 2).

DISCUSSION

Principal findings

Almost 300000 individuals were exposed to around 150 000 PIM years during 2016 in Denmark. The highest prevalence of PIMs was found for long-term use of NSAIDs and benzodiazepines, which accounted for approximately 50 000 PIM years each. The observed variation was substantial for virtually all criteria, although for a minority only, this variation was substantially higher than expected due to randomness. Yet, for more than half of the criteria, we observed a statistically significant positive correlation between the specific PIM and the sum of the remaining PIMs. Across GP clinics, we found an observed variation of 2.18, which exceeded the expected random variation by a factor 1.65. Long-term use of NSAIDs and long-term use of benzodiazepines showed the most pronounced variations, and this could not be explained by differences in patient characteristics or by randomness.

Interpretation of results

The substantial prevalence of certain PIMs might be explained by the existence of barriers for deprescribing PIMs caused by patient-related factors, provider-related factors and the organisational, structural and sociocultural context.^{13 31 32} Yet, the reasons for such barriers remain multifactorial; they are highly interdependent and are characterised by considerable clinical complexity.¹³ The patients' strong belief in the necessity of their medications is often reported as a central barrier for

deprescribing.³³ Furthermore, GPs find it challenging to balance benefits and harms across treatments in patients with polypharmacy due to pharmacological complexity and lack of guidelines for multimorbidity.^{13 31} From an organisational point of view, time is a limited resource, and many GPs are short of time to prioritise comprehensive medication reviews.^{13 31} Structurally, some GPs point to the fragmented healthcare system and the unclear treatment responsibility as complicating barriers.³¹ From a sociocultural point of view, GPs report a hierarchy in which their professionalism is occasionally challenged by the specialist physicians, leading to avoidance of deprescribing medications initiated during hospital contact.³²

Substantial PIM prevalence and GP-related variation in long-term use of benzodiazepines might be caused by complex interaction of drug, patient, physician and organisational barriers.³⁴ Although recent systematic guidelines have concluded that deprescribing appears to be relatively safe and feasible; in the case of benzodiazepines, it is particularly challenging.³⁵ First, psychological and physical dependence may compound the problem of inadequate patient adherence.³⁵ Second, lack of equally effective alternatives to treat the patients' insomnia or anxiety may play a role. Although the licensing of selective serotonin reuptake inhibitors for anxiety disorders and melatonin agonists for insomnia has allowed for therapeutic choices of prescribers,³⁶ there is no convincing evidence to recommend substitution.³⁵ Third, limited evidence exists on how to best conduct the process of deprescribing, although patient involvement and shared decision-making have been suggested to be important for facilitating deprescribing.³⁵ Specifically, the optimal tapering schedule remains unknown.^{35 36}

Similarly, the substantial prevalence of and variation seen for long-term use of NSAIDs might be explained by the fact that it is complicated to deprescribe NSAIDs as no safer and equally effective alternative analgesic exists.³⁷

Apart from long-term use of benzodiazepines and NSAIDs, five other PIM criteria were common (B4, C3, C6, D8 and M). In some cases, this could partly be explained by large risk populations (C6 and M). In other cases, this could be explained by high rates (B4, C3 and D8), which are likely to be caused by significantly older populations with more multimorbidity; this entails higher degree of complexity in the clinical care and higher risk of overseeing needs for reducing medications.

Although the observed variation was substantial for virtually all PIM criteria, this variation did not substantially exceed what could be expected due to random variation for the majority of the studied criteria. Hence, these criteria cannot be used as independent performance measures of the GP clinics. Nevertheless, more than half of the studied criteria had a statistically significant positive correlation to the sum of remaining PIMs. This suggests that the majority of the criteria constitute markers of an underlying GP-related variation in prescribing strategy and may thus contribute with relevant information to the aggregate measure of total PIM time.

Comparison with other studies

Over the past decades, it has been increasingly recognised that the variations in the use of healthcare services, for example, between geographical areas, periods or organisational units like GP clinics, are often higher than expected based on the health status of the population.^{14 15 38-40} Few studies have explored the GP-related variation in the prescribing of PIMs,⁴¹ high-risk prescribing,^{42–44} polypharmacy^{3 45} and overall prescribing,⁴⁶ but they only take into account differences in the characteristics of the GP-clinics or of the patient population when concluding that the observed variation is a marker of inefficient or inappropriate prescribing. This approach is likely to overestimate the variation and thus to overestimate the extent of suboptimal GP prescribing. To the best of our knowledge, our study is the first to describe GP-related variations in PIMs, while accounting for differences between GP populations and for the fact that some degree of variation is to be expected due to randomness.

Strengths and weaknesses of the study

This study has several strengths, including the large data set, complete follow-up and the population-based setting, in which more than 4 million adults were followed essentially without loss to follow-up. This practically eliminated the risk of selection bias and allowed for the identification of even rare PIM outcomes. Furthermore, a couple of aspects of the analysis approach deserve mentioning, as these facilitated the estimation of the GP-related variation in PIMs, which could be interpreted as a marker of suboptimal treatment performance. First, we had comprehensive high-quality data at the patient level. This allowed for thorough matching and adjustments, which took into account in-between GP differences in the patient characteristics. This enabled us to characterise the GP-clinics by qualified observed/expected ratios rather than by crude frequency measures. Second, we applied a matching/ sampling procedure to quantify the extent of variation, which could be expected to be due to randomness alone, and used this as a reference for the observed variation. Noteworthy, without these approaches, the GP-related variation would have been grossly overestimated, as we demonstrated that although the observed variation was substantial, for the majority of criteria, this variation did not exceed that expected due to randomness.

However, some important limitations pertain to this study. First, the perception of PIMs as markers of suboptimal treatment quality does not necessarily hold true in all cases. Although each of the PIM criteria certainly represents 'a red flag' for the prescriber, some PIMs may be the best treatment option for certain patients if administered with care, such as those who do not fit the guide-lines (eg, frail patients with high complexity of diseases for whom there may be a need for balancing conflicting demands).⁴⁷ Yet, exact identification of these patient would require more detailed clinical information than available from the registers. However, as we were able to characterise the study participants comprehensively,

allowing for an extensive patient profiling, we find it reasonable to assume that additional characteristics, such as frailty measures, are randomly distributed across the GP clinics. This implies that the considerable excess variation, as calculated in this paper, is likely, at least to some degree, to be a marker of uneven and thus suboptimal GP performance. Second, although the excess variation is a valuable marker of the existence of 'true variation' in GP performance, it may be a highly conservative measure. The sampled variation quantifies the inflation of the observed variation that should be expected due to randomness under the null hypothesis (ie, if no true differences between GP clinics exist). However, if the true between-GP variation is non-null, the inflation due to random noise could be lower or even negative. Third, we studied GP clinics as the unit of variation, which may entail that we have underestimated the variation in-between GPs, although it is likely that GPs in the same clinic have reasonably homogeneous prescribing preferences. Fourth, some data were not available, including data on laboratory measurements, clinical observations (eg, blood pressure profiles), untreated diagnoses exclusively managed in primary care (eg, alcohol abuse), disease severity measures, measures of frailty, estimated life expectancy, medication adherence, over-the-counter medications, medications administered during in-patient hospitalisations and some of the medication data. This had several implications. Primarily, we cannot exclude residual confounding, which could have impacted our findings in either direction. Furthermore, leaving noncodeable elements out of the algorithm led to the simplification of certain criteria. Finally, the lack of information on over-the-counter NSAIDs might imply that the use of these drugs could be underestimated. Fifth, we used the original Irish STOPP/START criteria²⁶ and modified these for a Danish register-based setting. This could have reduced the generalisability, as some countries may rely on other local versions of these criteria and adhere to slightly different clinical guidelines. Yet, by providing our coding algorithms, this algorithm can easily be adapted for application in other settings.

CONCLUSION

Clinical implications

We identified long-term use of NSAIDs and benzodiazepines as the most prevalent PIMs with the most pronounced GP-related variations. Furthermore, our study suggests that prescribing certain types of PIMs is a marker of prescribing other PIMs. Our findings allow for the interpretation that some PIMs, such as longterm use of NSAIDs and benzodiazepines, seem to mark suboptimal GP-prescribing and targeting the use of these drugs may be a reasonable place to start in the conduct of future intervention studies. These PIMs both had a high prevalence, making such intervention clinically relevant and a substantial excess variation, suggesting room for improvements.

Future research

This study highlights the need for exploring the causal explanations for the substantial prevalence and the related variations in the treatment strategies of GPs for certain PIMs, such as long-term use of NSAID and benzodiazepines, which could represent obvious targets for future interventions aimed at reducing PIMs. Moreover, our register-based algorithm for measuring PIMs can be applied in future interventions aimed at optimising GP prescribing behaviour, including selecting patients with the greatest need for special efforts and assessing the effectiveness of the intervention.

Acknowledgements The authors would like to thank Lone Niedziella, Research Unit for General Practice in Aarhus, for language editing and Dr Simon Winther, Department of Cardiology, Regional Hospital West Jutland, Herning, Denmark, for reviewing key criteria in the field of cardiology.

Contributors ARR, MF-G, LDC, AP and FB conceived the study. ARR, MF-G, LDC, AP and CHV planned the study and developed the register-based algorithm for identifying PIMs. CHV and MF-G carried out the analyses. PKB reviewed the algorithm for consistency with existing clinical guidelines. AS and PV obtained the data. FB, ARR, MF-G and AP achieved the funding. ARR and LDC wrote the first draft of the manuscript. ARR, LDC, MF-G, CHV, AP, FB, PKB, AS, PV and DRW interpreted the results, revised the manuscript and approved the final version. ARR is the guarantor.

Funding This study was supported by grants from the Novo Nordisk Foundation donated to ARR and AP (NNF160C0019592 and NNF180C0031194) and from the General Practice Research Foundation of the Central Denmark Region to MFG. All authors are independent from the funders, and the funders had no role in the study design; in the collection, analysis and interpretation of data; in the writing of the report or in the decision to submit the article for publication.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This project was approved by the Danish Data Protection Agency, the Danish Health Data Authority and Statistics Denmark. According to Danish legislation, ethical approval and informed consent were not needed as the study was entirely register-based.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available. Data not available due to legal restrictions.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs

Anette Riisgaard Ribe http://orcid.org/0000-0003-0011-2208 Anders Prior http://orcid.org/0000-0003-4053-3701

REFERENCES

Barnett K, Mercer SW, Norbury M, *et al.* Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *Lancet* 2012;380:37–43.

- 2 Boyd CM, Fortin M. Future of multimorbidity research: how should understanding of multimorbidity inform health system design? *Public Health Rev* 2010;32:451–74.
- 3 Sinnige J, Braspenning JC, Schellevis FG, et al. Inter-practice variation in polypharmacy prevalence amongst older patients in primary care. *Pharmacoepidemiol Drug Saf* 2016;25:1033–41.
- 4 Cahir C, Fahey T, Teeling M, *et al*. Potentially inappropriate prescribing and cost outcomes for older people: a national population study. *Br J Clin Pharmacol* 2010;69:543–52.
- 5 Beers MH, Ouslander JG, Rollingher I, et al. Explicit criteria for determining inappropriate medication use in nursing home residents. UCLA division of geriatric medicine. Arch Intern Med 1991;151:1825–32.
- 6 Rankin A, Cadogan CA, Patterson SM, et al. Interventions to improve the appropriate use of polypharmacy for older people. Cochrane Database Syst Rev 2018;9:Cd008165.
- 7 Corsonello A, Onder G, Abbatecola AM, *et al.* Explicit criteria for potentially inappropriate medications to reduce the risk of adverse drug reactions in elderly people: from beers to STOPP/START criteria. *Drug Saf* 2012;35 Suppl 1:21–8.
- 8 Moriarty F, Bennett K, Cahir C, et al. Potentially inappropriate prescribing according to STOPP and start and adverse outcomes in community-dwelling older people: a prospective cohort study. Br J Clin Pharmacol 2016;82:849–57.
- 9 Hedna K, Hakkarainen KM, Gyllensten H, et al. Potentially inappropriate prescribing and adverse drug reactions in the elderly: a population-based study. Eur J Clin Pharmacol 2015;71:1525–33.
- 10 Hill-Taylor B, Sketris I, Hayden J, et al. 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.
- 11 Cahir C, Bennett K, Teljeur C, et al. Potentially inappropriate prescribing and adverse health outcomes in community dwelling older patients. Br J Clin Pharmacol 2014;77:201–10.
- 12 Vass M, Hendriksen C. Polypharmacy and older people--the GP perspective. Z Gerontol Geriatr 2005;38 Suppl 1:I14–17.
- 13 Anderson K, Stowasser D, Freeman C, et al. Prescriber barriers and enablers to minimising potentially inappropriate medications in adults: a systematic review and thematic synthesis. BMJ Open 2014;4:e006544.
- 14 Hjertholm P, Fenger-Grøn M, Vestergaard M, et al. Variation in general practice prostate-specific antigen testing and prostate cancer outcomes: an ecological study. Int J Cancer 2015;136:435–42.
- 15 Fenger-Grøn M, Kjaersgaard MIS, Parner ET, et al. Early treatment with talk therapy or antidepressants in severely bereaved people and risk of suicidal behavior and psychiatric illness: an instrumental variable analysis. *Clin Epidemiol* 2018;10:1013–26.
- 16 Pedersen KM, Andersen JS, Søndergaard J. General practice and primary health care in Denmark. *J Am Board Fam Med* 2012;25 Suppl 1:S34–8.
- 17 Nørøxe KB, Pedersen AF, Bro F, et al. Mental well-being and job satisfaction among general practitioners: a nationwide crosssectional survey in Denmark. *BMC Fam Pract* 2018;19:130.
- 18 Erlangsen A, Fedyszyn I. Danish nationwide registers for public health and health-related research. *Scand J Public Health* 2015;43:333–9.
- 19 Schmidt M, Pedersen L, Sørensen HT. The Danish civil registration system as a tool in epidemiology. *Eur J Epidemiol* 2014;29:541–9.
- 20 Denmark S. The integrated database for longitudinal labour market research (Integreret database for arbejdsmarkedsforskning (IDA)), 1991.
- 21 Schmidt M, Schmidt SAJ, Sandegaard JL, et al. The Danish national patient registry: a review of content, data quality, and research potential. *Clin Epidemiol* 2015;7:449–90.
- 22 Mors O, Perto GP, Mortensen PB. The Danish psychiatric central research register. *Scand J Public Health* 2011;39:54–7.
- 23 Kildemoes HW, Sørensen HT, Hallas J. The Danish national prescription registry. Scand J Public Health 2011;39:38–41.
- 24 Pedersen KM, Christiansen T, Bech M. The Danish health care system: evolution - not revolution - in a decentralized system. *Health Econ* 2005;14:S41–57.
- 25 Pedersen CB. The Danish civil registration system. *Scand J Public Health* 2011;39:22–5.
- 26 O'Mahony D, O'Sullivan D, Byrne S, et al. STOPP/START criteria for potentially inappropriate prescribing in older people: version 2. Age Ageing 2015;44:213–8.
- 27 Huibers CJA, Sallevelt BTGM, de Groot DA, et al. Conversion of STOPP/START version 2 into coded algorithms for software implementation: a multidisciplinary consensus procedure. Int J Med Inform 2019;125:110–7.

Open access

- 28 Lund JL, Richardson DB, Stürmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. Curr Epidemiol Rep 2015;2:221-8.
- 29 Prior A, Fenger-Grøn M, Larsen KK, et al. The association between perceived stress and mortality among people with multimorbidity: a prospective population-based cohort study. Am J Epidemiol 2016;184:199-210.
- 30 Rodgers JL, Nicewander WA. Thirteen ways to look at the correlation coefficient. Am Stat 1988;42:59-66.
- 31 Sinnott C, Mc Hugh S, Browne J, et al. Gps' perspectives on the management of patients with multimorbidity: systematic review and synthesis of qualitative research. BMJ Open 2013;3:e003610.
- Laursen J, Kornholt J, Betzer C, et al. General practitioners' 32 barriers toward medication reviews in Polymedicated Multimorbid patients: how can a focus on the pharmacotherapy in an outpatient clinic support GPs? Health Serv Res Manag Epidemiol 2018;5:2333392818792169.
- 33 Clyne B, Cooper JA, Boland F, et al. Beliefs about prescribed medication among older patients with polypharmacy: a mixed methods study in primary care. Br J Gen Pract 2017;67:e507-18.
- 34 Ng BJ, Le Couteur DG, Hilmer SN. Deprescribing benzodiazepines in older patients: impact of interventions targeting physicians, pharmacists, and patients. Drugs Aging 2018;35:493-521.
- Lee JY, Farrell B, Holbrook AM. Deprescribing benzodiazepine 35 receptor agonists taken for insomnia: a review and key messages from practice guidelines. Pol Arch Intern Med 2019;129:43-9.
- 36 Lader M, Tylee A, Donoghue J. Withdrawing benzodiazepines in primary care. *CNS Drugs* 2009;23:19–34. Ong CKS, Lirk P, Tan CH, *et al.* An evidence-based update on
- 37 nonsteroidal anti-inflammatory drugs. Clin Med Res 2007;5:19-34.

- Wennberg JE. Unwarranted variations in healthcare delivery: 38 implications for academic medical centres. BMJ 2002;325:961-4.
- 39 Brownlee S, Chalkidou K, Doust J, et al. Evidence for overuse of medical services around the world. Lancet 2017;390:156-68.
- 40 Fisher ES, Wennberg DE, Stukel TA, et al. The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care. Ann Intern Med 2003;138:273-87.
- Cahir C, Fahey T, Teljeur C, et al. Prescriber variation in potentially 41 inappropriate prescribing in older populations in Ireland. BMC Fam Pract 2014:15:59.
- Byrne CJ, Cahir C, Curran C, et al. High-Risk prescribing in an Irish 42 primary care population: trends and variation. Br J Clin Pharmacol 2017;83:2821–30.
- 43 Guthrie B, McCowan C, Davey P, et al. High risk prescribing in primary care patients particularly vulnerable to adverse drug events: cross sectional population database analysis in Scottish general practice. BMJ 2011;342:d3514.
- Stocks SJ, Kontopantelis E, Akbarov A, et al. Examining variations in 44 prescribing safety in UK general practice: cross sectional study using the clinical practice research Datalink. BMJ 2015;351:h5501.
- 45 Ong SM, Lim YMF, Sivasampu S, et al. Variation of polypharmacy in older primary care attenders occurs at prescriber level. BMC Geriatr 2018:18:59.
- 46 et alDuerden M, Millson D, Avery A. The quality of GP prescribing: the kings fund, 2011. Available: https://www.kingsfund.org.uk/sites/ default/files/field/field_document/quality-gp-prescribing-gp-inquiryresearch-paper-mar11.pdf [Accessed 27 Oct 2020].
- 47 Christensen LD, Petersen J, Andersen O, et al. Physicians' Non-Uniform Approach to Prescribing Drugs to Older Patients - A Qualitative Study. Basic Clin Pharmacol Toxicol 2017;121:505-11.