The medical perspective in prescribing: measurements and educational strategies

Johan Lönnbro

Department of Pharmacology
Institute of Neuroscience and Physiology
Sahlgrenska Academy, University of Gothenburg

UNIVERSITY OF GOTHENBURG

Gothenburg 2021

The medical perspective in prescribing:
measurements and educational strategies

johan.lonnbro@gu.se

Cover illustration: Johan Lönnbro

ISBN 978-91-8009-368-2 (PRINT)
ISBN 978-91-8009-369-9 (PDF)

http://hdl.handle.net/2077/68319

Printed in Sweden 2021
Printed by Stema Specialtryck AB, Borås

Trycksak
3041 0234

‘It is much more important to know
what sort of patient has a disease
than what sort of a disease a patient has’
– William Osler

Abstract

Background and aims: Drug treatment is a keystone in modern healthcare and for safe prescribing of drugs, an overall medical perspective has to be applied, taking patient characteristics into account. For each medication, and all medications combined, expected benefits must be weighed against the risks of harms; adverse drug reactions, sometimes leading to hospitalization, is a matter of concern. In striving for rational use of medicines, evidence regarding measures of prescribing quality is needed, as well as knowledge regarding how to facilitate the acquisition of prescribing skills in junior physicians. The aims of this thesis were (i) to investigate the clinical relevance of a set of prescribing indicators often used in research to reflect quality of drug treatment, and (ii) to evaluate inter-rater agreement in assessments of drug treatment quality, as well as (iii) to appraise the effects of an educational intervention on pharmacotherapy towards interns, and (iv) to explore what aspects junior physicians find important when being taught the art of prescribing.

Methods and results: (i) Assessing the drug treatment of 200 hip fracture patients from a medical perspective, we show that half of the STOPP/START criteria (Screening Tool of Older Persons' Prescriptions/Screening Tool to Alert to Right Treatment) are clinically relevant when individual characteristics of the patient are considered, and that suboptimal prescribing occur more often in those with a large number of drugs in the medication list and those who have their drugs provided by the multi-dose drug dispensing system. (ii) Investigating the inter-rater agreement regarding pharmacotherapeutic assessments, based on five physicians specialized/specializing in internal medicine assessing the drug treatment of 30 randomly selected inpatients, the reliability using such assessors seems to be very good for quality-related aspects, whereas the relationship between the drug treatment and the admission appears to be more challenging to determine, especially among residents. In one randomized controlled study (iii) and one qualitative study (iv), including 57 and 34 interns, respectively, at Sahlgrenska University Hospital, we show that a short, structured and discussion-based education based on clinical cases increases the confidence in performing basic medication reviews, and that discussion-based, small-group teaching on authentic patient cases with an atmosphere allowing questions, together with assignments and feedback, are important educational aspects for the interns.

Conclusion: This thesis illustrates the value of including an overall medical perspective in measures of prescribing quality and that reliability may be an important issue to consider in pharmacotherapeutic assessments. Further, this work provides evidence on how to educate junior physicians in order to make them more confident in the art of prescribing i.e., facilitating their acquisition of skills to apply an overall medical perspective in pharmacotherapy.

Keywords: Prescribing quality, adverse drug reactions, clinical relevance, inter-rater reliability, education

Sammanfattning på svenska

Läkemedelsbehandling är en central del i modern sjukvård. För effektiv och säker användning av läkemedel krävs ett medicinskt helhetsperspektiv, med kunskap om såväl sjukdomar i sig som olika behandlingsmöjligheter, liksom kännedom om den individuella patienten. För varje läkemedel, och för kombinationer av flera, behöver nyttan med behandlingen vägas mot eventuella risker. Det är känt att läkemedel kan ge biverkningar, oftast lindriga, men ibland allvarliga, vilka till exempel kan bidra till inläggning på sjukhus. I strävan för rationell och säker läkemedelsanvändning behövs ökad evidens kring hur vi mäter läkemedelsbehandlingens kvalitet, samt eventuella biverkningar. Likaså behövs ytterligare kännedom om hur vi på bästa sätt kan förbereda yngre läkare för den komplexa uppgiften att behandla patienter med läkemedel på ett säkert sätt.

I den första delen av avhandlingen undersöktes (i) i hur stor utsträckning indikatorer för god läkemedelsbehandling är kliniskt relevanta för den enskilde patienten, samt (ii) graden av samstämmighet mellan läkare när de utifrån ett övergripande medicinskt perspektiv gör bedömningar av läkemedelsbehandlingens kvalitet på individnivå och relationen mellan läkemedel och inläggningar på sjukhus. Frågeställningarna undersöktes i en deskriptiv studie, inkluderande 200 patienter med höftfraktur, och i en reliabilitetsstudie med 30 slumpvist utvalda patienter inlagda på Sahlgrenska Universitetssjukhuset.

I avhandlingens andra del testades hypotesen (iii) att en kort, strukturerad och diskussionsbaserad utbildningsinsats utgående från patientfall, riktade mot AT-läkare, kan öka tryggheten i läkemedelsarbetet, jämfört med ingen sådan utbildning. Vidare undersöktes (iv) vilka aspekter yngre läkare tycker är viktiga när de utbildas i konsten att behandla patienter med läkemedel. Metodiken bestod av en randomiserad kontrollerad studie och en kvalitativ studie, där 57 respektive 34 AT-läkare deltog.

Resultaten i den första delen av avhandlingen visar att STOPP/START-kriterier (Screening Tool of Older Persons' Prescriptions/Screening Tool to Alert to Right Treatment) är kliniskt relevanta i ungefär hälften av fallen när man tar hänsyn till den enskilde patientens förhållanden, samt att antal ordinerade läkemedel och användning av dospåsar ökar risken för kliniskt relevanta STOPP/START-utfall. Vidare visas att samstämmigheten är mycket god mellan läkare med internmedicinsk bakgrund när det gäller bedömning av läkemedelsbehandlingens kvalitet, men att bedömningar av biverkningar är mer komplext, framför allt för ST-läkare.

Resultaten i den andra delen av avhandlingen visar att en kort, diskussionsbaserad utbildningsinsats för AT-läkare ger ökad trygghet att genomföra enkla läkemedelsgenomgångar, liksom ökad trygghet även vad gäller andra viktiga aspekter av

läkemedelsarbetet. Den kvalitativa analysen visar att diskussionsbaserad smågruppsundervisning, baserad på autentiska patientfall, med högt i tak för olika frågor och funderingar, samt med möjlighet till feedback, är viktiga aspekter för AT-läkarnas lärande.

Sammanfattningsvis belyser avhandlingen vikten av medicinska bedömningar utifrån ett helhetsperspektiv i studier av kvalitet av läkemedelsbehandling och biverkningar. Vidare bidrar avhandlingen med evidens avseende hur vi kan utbilda yngre läkare så att de blir trygga i det läkemedelsarbete som ingår i yrkesrollen, det vill säga att behandla patienter utifrån en medicinsk helhetssyn.

LIST OF PAPERS

This thesis is based on the following studies,
referred to in the text by their Roman numerals.

I.

Lönnbro, J., Wallerstedt, S.M.

Clinical relevance of the STOPP/START criteria in hip fracture patients.
European Journal of Clinical Pharmacology. 2017;73:499-505

II.

Lönnbro, J., Holmqvist, L., Persson, E., Thysell, P.,
Åberg, N.D., Wallerstedt, S.M.

Inter-rater reliability of assessments regarding the quality of drug treatment, and drug-related hospital admissions.
British Journal of Clinical Pharmacology. 2021.
PMID: 33609324

III.

Lönnbro, J., Nylén, K., Wallerstedt, S.M.

Developing professional confidence in the art of prescribing – a randomized controlled study on structured collegial discussions during internship.
European Journal of Clinical Pharmacology. 2019;75(5):687-696

IV.

Lönnbro, J., Wallerstedt, S.M.

"It's helpful to get the time and opportunity to discuss drug treatment; that's what I think is the most important thing." - A qualitative study on prescribing education in junior physicians.
European Journal of Clinical Pharmacology. 2020;76(2):249-255

Contents

Abbreviations

AC_1	Agreement coefficient (Gwet's 1st)
ADE	Adverse drug event
ADR	Adverse drug reaction
CI	Confidence interval
IQR	Interquartile range
OR	Odds ratio
PIM	Potentially inappropriate medication
PPO	Potential prescribing omission
SD	Standard deviation
START	Screening Tool to Alert to Right Treatment
STOPP	Screening Tool for Older Persons' Prescriptions
WHO	World Health Organization
WHO-UMC	World Health Organization – Uppsala Monitoring Centre

Introduction

Drug treatment is a keystone in modern healthcare, and with an increasing use of pharmaceutical drugs [1] and an ageing population with often multiple diseases, combined with an increasing number of therapeutic alternatives, use of medicines is a challenge. According to the World Health Organization (WHO), rational use of medicines means that "each patient shall receive medications appropriate to their clinical needs, in doses that meet their own individual requirements, for an adequate period of time, and at the lowest cost to them and their community". Over the last decades, numerous articles and reviews have been published on adverse drug reactions, inappropriate medications and drug-related hospital admissions [2-7], indicating that the need for safe and rational prescribing may not be met. Besides descriptive statistics on suboptimal drug treatment, apprehensions have been raised that medical students and junior physicians may be inadequately prepared in prescribing and pharmacotherapy, as an explanation to why junior doctors make more mistakes than their senior colleagues when they prescribe medications [8-12].

Prescribing of drugs is a complex task that requires several levels of knowledge and experience. Scientific development may add enhanced precision in future drug treatment [13], however, independent of scientific advances in diagnostics and treatment, there are basic aspects of drug treatment that needs to be fulfilled at all times. The most obvious part is deciding on the most suitable drug for treating a certain condition, considering the specific patient's comorbidities while weighing the benefits against the risks, taking into account the patient's wishes, i.e., the medical perspective of pharmacotherapy. Further, with the initiation of medicines comes the responsibility of managing continuous treatment, which includes monitoring effects and adverse effects, and deciding, in collaboration with the patient, on continuing, changing or discontinuing the treatment as time goes on. Individual treatment strategies may also need to be revised according to advances in drug development or due to emerging evidence, for instance reflected in updated guidelines. Last but not least, the patients' conditions evolve and change over time, which requires ongoing reconsideration of their drug treatment. Continuous patient involvement in treatment decisions, including education and assessing attitudes towards drug treatment, may not only be important to find a suitable treatment plan, but also to address non-adherence concerns – which has been reported as a non-negligible factor in drug treatment [14-17]. In summary, appropriate prescribing of medications is a demanding task for physicians, and the importance of preparing junior physicians in the best possible ways is being discussed in the scientific literature [12].

Measuring drug treatment quality

Quality

Although the WHO definition of rational use of medicines may reflect high quality prescribing, there is yet no generally accepted gold standard to determine the quality of drug treatment. There have been attempts to establish core outcome sets, including for example potentially inappropriate medications, drug-drug interactions, drug overuse/underuse, as suggested outcome measures to be used in clinical trials [18, 19]. One widely spread approach has been to target polypharmacy among older persons [20], commonly defined as the use of five or more drugs, although not all studies employ the same definition [21]. However, polypharmacy in general has been shown not to reflect suboptimal treatment in older patients [22], and the number of drugs in the medication list may rather be a surrogate measure of burden of disease [23]. Further, a recent review of the effects of pharmaceutical care interventions to reduce inappropriate polypharmacy concluded that no patient benefit could be observed with such an approach [24]. Although numerous studies have described drug treatment quality in terms of potentially inappropriate medications or drug-related problems, few studies have applied a broader overall medical perspective in the assessments, including, for instance, the expected benefit-risk balance for the specific patient that has to be considered in the prescribing process.

Tools

Over the last decades there have been several attempts to create tools to assess medication appropriateness, mainly regarding treatment of older persons [25]. The general idea has been to find a structured, easy-to-follow method to detect potentially inappropriate medications (PIMs) and potential prescribing omissions (PPOs), the latter pointing out medications that are indicated but for some reason not prescribed. Two major approaches have arisen in the development of such tools, consisting of either implicit listing tools that are based on medical judgement requiring knowledge about the individual patient, or explicit criterion-based tools to be used without intricate information about the patients [26]. In a recent review of listing tools by Pazan et al., a different categorization is suggested, dividing the tools into (1) drug-oriented listing approaches, with or without suggestions linked to indications, and (2) patient-oriented listing approaches, where knowledge about the patients is needed [25]. With this classification, the well-known medication appropriateness index, known as the MAI score [27], the updated Beers criteria [28] and the STOPP (Screening Tool of Older Persons' Prescriptions) criteria are classified as drug-oriented, whereas the START (Screening Tool to Alert to Right Treatment) criteria [29, 30] are classified as patient-oriented. Other tools, such as the Quality Indicators for Drug Use in Elderly People, by the Swedish National Board of Health and Welfare [31], also apply a patient-oriented approach, requiring knowledge

about the patients. These help systems to detect PIMs/PPOs have been created to facilitate for physicians in clinical practice, as well as in research to provide prevalence figures of inappropriate medications [32-35]. In the research setting, application of tools aiming to find *all potentially* inappropriate medications may be problematic, as the therapy may not be inappropriate from a medical perspective for the specific patient. Low application rates of pharmacist recommendations to the attending physicians after application of the STOPP/START criteria suggest that the clinical relevance of PIMs/PPOs may be an issue [36]. There are several studies providing prevalence figures of PIMs and PPOs, but less is known about the clinical relevance of these tools at the individual level.

Adverse consequences

An adverse drug reaction (ADR) is defined as "a response to a drug that is noxious and unintended and occurs at doses normally used in man for the prophylaxis, diagnosis or therapy of disease, or for modification of physiological function". Since 2012, this regulatory definition includes not only noxious and unintended effects from authorized use at normal doses, but also from medication errors and from misuse and abuse of drugs [37, 38]. Adverse drug events (ADE) are sometimes used in parallel to ADR. However, an important difference between an ADR and an ADE is that the latter includes all adverse events occurring during or after the time a patient uses a medicine, while for ADRs, causality or suspected causality connecting the event and the drug is required [39]. The prevalence and incidence of ADRs, as well as ADEs, has been explored in numerous studies during the last decades [40-42], with concurrent concern of the lack of consistency in terminology [43], possibly partly explaining the variations in prevalence regarding various measures in patient safety matters.

Another important concept in drug safety terminology is pharmacovigilance, which is defined by the WHO as the "science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any other medicine related problem" [44]. From a pharmacovigilance perspective, it is important to detect all possible symptoms and conditions that may be related to the use of a drug. Further assessments will then have to be carried out to determine whether or not the event was actually caused by the drug. There are several instruments for such assessments of causality, for example the criteria defined by the WHO-Uppsala Monitoring Centre (WHO-UMC) [45], the Naranjo algorithm [46] and the updated logistic probabilistic method [47]. The WHO causality system is a widely accepted method although dependent on assessor expertise and judgement, whereas the Naranjo algorithm, commonly used in observational studies, while showing high inter-rater reliability, has had issues with low sensitivity [48]. The updated logistic probabilistic method, showing similar reliability as the Naranjo score, is more complex to use and time consuming [49]. In summary, there is yet no gold standard for assessing causality of adverse drug reactions, and commonly used instruments all have

their pros and cons. One thing they seem to have in common though, is that they all include some degree of subjective assessments. Indeed, the relationship between a drug and an event is rarely crystal clear. As a consequence, an ADR may be considered a differential diagnosis that can be hard to distinguish from worsening and emerging morbidities and disorders. However, despite the lack of a gold standard for assessing causality, the use of an instrument could help to structure the assessments. In a recent review article where the prevalence of drug-related hospital admissions was investigated, surprisingly, only ten out of sixteen studies reported about the causality assessments performed and inter-rater agreement was reported in merely three studies [7].

Research during the last decades have reported widely diverging prevalence figures of hospital admissions related to drugs ranging from 0.16 to 41% [3, 7, 40, 50]. There may be several potential reasons for this considerable variation. For example, studies are conducted in different countries and include diverse patient populations. Different terminology is used to describe similar outcomes, and definitions are sometimes vague, exemplified with the term "drug-related problems". While several studies use "adverse drug reactions" as a primary outcome, others use "drug-related problems", and, in the end, the results from such studies are sometimes merged into the same category [7]. Further, different causality instruments are used and the definition of a "drug-related admission", as well as level of causality between a drug and a reaction required for an admission to be labelled "drug-related", differ among studies. For instance, while some studies require a causality level of "probable", defined by the WHO as a reaction "unlikely to be attributed to disease or other drugs", other studies allow the lower level, "possible", defined as a reaction that "could also be explained by disease or other drugs", to count as a drug-related admission [45]. As most ADRs are categorized as "possible" [47, 51], using different requirements in definitions will generate diverging results. Finally, there are also differences in how studies define the level of contribution to hospital admission. Some differentiate between "main reason" and "contributory reason" [52], while others use terms like "associated with admission" or "related to admission" [53, 54]. Altogether, this methodological variation regarding definitions may create grounds for diverging results.

Reliability

In order to provide valid results, outcomes need to be measured in a reliable way. When using assessors as instruments, the terms intra-rater and inter-rater agreement illustrate this methodological concern. While intra-rater reliability describes the agreement between two assessments on two different occasions by the same assessor, inter-rater reliability relates to agreement between multiple assessors rating the same phenomenon. The statistical methods of measuring reliability have evolved through the years, starting with simply calculating the percent agreement between the assessment sets, followed by Cohen's kappa in the sixties, which also take into consideration the agreement by chance [55]. This commonly used method was followed by variants of kappa, as for instance

Fleiss kappa and Light's kappa for multiple raters, the former used for raters randomly sampled for each subject, and the latter when all raters assess all subjects, calculating the average of Cohen's kappa between all possible pairs of raters [56]. However, in studies with high prevalence of ratings within certain categories, kappa estimates are typically disproportionally low, explained as one of the known paradoxes of kappa [57]. One way to overcome this was described by Byrt and Bishop, although their method does not take into account the other part of the paradox, concerning bias of the raters [56]. In such situations, when the marginal distribution of ratings in specific cases is substantially different between assessors, kappa estimates tend to get disproportionally high. To overcome this phenomenon, Siegel and Castellan's kappa can be used, although this method has the same weakness concerning skewed prevalence as Cohen's kappa [56]. In order to deal with both parts of the paradox, Gwet proposed a new way to calculate inter-rater agreement that is less sensitive to skewed ratings as well as bias of the raters [58, 59].

Several studies have been performed on drug-related problems and the relationship between drugs and hospital admissions [3, 4, 7], and inter-rater reliability has been addressed, although the number of assessors has been few, often restricted to two [60-62]. In studies with multiple raters, the assessed patients have commonly been pre-selected as having drug-related problems [63-66], or they have been elective [67] or fictitious [68]. Further, in a review of studies concerning preventability of ADEs, Hakkarainen et al. concluded that reliability between assessors, mostly two, varied markedly, and that the instruments of determining preventability were imprecise [60]. Studies providing agreement figures regarding causality assessments of drug-related admissions show kappa values ranging from 0.16 to 0.83, with the majority of assessors being pharmacists [7, 66, 69].

Strategies to improve prescribing

Prescribing competence

Clinical pharmacology and pharmacotherapy have been elements of education in medical school in Europe for many years, though, during the last decade concerns of the extent of the education have been brought to attention. Studies have shown that junior doctors make more prescribing errors, which perhaps reflects that managing drug treatment is a demanding task that requires adequate medical skills and knowledge [11, 12, 70]. In a recent study of 895 final-year medical students in Europe, the potential for improvements in medical school was illustrated [8]. This is in line with a study directed towards teachers responsible for clinical pharmacology and therapeutics in European medical schools, that showed that a majority of the teachers felt that they did not prepare the students sufficiently for the coming task of prescribing [71]. In a review of studies concerning prescribing competence among medical students, there was a lack of consensus among

teachers in clinical pharmacology on what skills medical students should have after graduating, and the students seem to experience a lack of self-confidence in prescribing-related matters [9]. As junior physicians are responsible for writing a substantial number of prescriptions, they may be an optimal target for educational interventions [10].

What has been done?

To target the aim of sufficient prescribing competence at graduation, measures have been taken to improve education in pharmacotherapy. One example is the Prescribing Safety Assessment (PSA) in the United Kingdom, designed to be a comprehensive test of competence in prescribing and supervising the use of medicines, directed towards final-year medical students [72]. Further, a core curriculum for education in clinical pharmacology and therapeutics has been outlined after involving teachers from 27 European countries, suggesting an entire 252 outcomes to be included in undergraduate education [73]. Similar to the efforts in the United Kingdom, medical schools in the Netherlands and Belgium have developed a pharmacotherapy test, focusing on safe prescribing, that is mandatory before graduation [74]. A recent study in Sweden showed that medical students increased their confidence in prescribing after introduction of a modified course in clinical pharmacology that included, for example, clarified learning outcomes and a list of common drugs [75].

As the art of prescribing takes time to master, actions for improved education in medical school need to be followed by continuous education during professional practice, perhaps in particular for junior physicians. Indeed, basic knowledge in prescribing, as well as indicators of prescribing quality, recommendations and guidelines, are helpful tools in decision-making [31, 76, 77], but in order to make reasonable decisions in a complex clinical context, basic knowledge and general guidelines may not be sufficient [78, 79]. Junior physicians may need a structured and supervised, continuous education in drug treatment management in order to quickly adapt the concepts of safe and rational prescribing. Active learning, as for example problem-solving and workshops, in contrast to traditional lecturing, has been shown to improve students' scores [80], and could perhaps be combined with creative and interactive discussions on clinical cases [81], in order to provide an environment where complex matters of prescribing could be elaborated upon. This might give the inexperienced physicians the opportunity to incorporate various important aspects of drug treatment. Cochrane reviews also provide insights regarding the small but possibly important positive effect of educational outreach visits regarding prescribing, as well as continuing educational meetings, including audit and feedback for improved professional practice [82, 83]. The latter, however, seem less likely to be effective on changing complex behaviors [83].

There are indications that few junior physicians find current education on prescribing sufficient, and they are asking for problem-based learning in the form of seminars [84].

Even though all attempts to improve the education in clinical pharmacology and therapeutics are for the better, they are likely to be designed by the education providers themselves. Planning educational efforts in consultation with the targeted students may be a favorable approach.

Aims

The overall aim of this thesis was divided in two. In the first part, attention was directed towards how we measure medication appropriateness, drug treatment quality and drug-related events, with the basic assumption that an overall medical perspective may add insights of importance. In the second part, we shifted focus to educational aspects of pharmacotherapy – how do we prepare junior physicians for the challenging task of prescribing from an overall medical perspective, often to older people with multiple diseases?

The specific aim for each study was as follows:

Study I
To investigate the clinical relevance of PIMs and PPOs identified by the STOPP/START criteria, and to identify predictors for clinically relevant PIMs/PPOs. We hypothesized that a significant part of detected PIMs/PPOs would not be clinically relevant at the individual level.

Study II
To investigate the inter-rater agreement regarding assessments of drug treatment quality and the relation between drug treatment and unplanned hospital admissions, including preventability. Our hypothesis was that reliability could be an important methodological issue in pharmacotherapeutic assessments.

Study III
To investigate the effects of a short and structured educational intervention, based on collegial discussions on patient cases, regarding junior doctors' confidence in prescribing. The hypothesis for this study was that an intervention could facilitate the process of acquiring improved professionalism regarding pharmacotherapy.

Study IV
To explore what aspects junior physicians find important when being taught the art of prescribing.

An overview of the four studies included in this thesis is presented in Table 1

Table 1 Overview of the four studies included in the thesis.

	Focus	Design	Patients/ Participants/Data	Analytical methods	Outcome measures
I	Clinical relevance of the STOPP/START criteria	Descriptive	200 hip fracture patients, consecutively recruited in 2009. Assessments performed in 2012.	Mann-Whitney U and Chi-square tests. Multiple logistic regression. Cohen's kappa.	Clinically relevant PIMs/PPOs at the individual level. Two specialist physicians applied STOPP/START criteria on all patients and then assessed the clinical relevance of each PIM/PPO. All assessments were performed independently, and disagreements were resolved in consensus.
II	Inter-rater reliability of assessments regarding aspects of drug treatment quality and adverse events	Reliability study	Five physicians (three specialists and two residents) assessed 30 patients, randomly selected from 864 unplanned hospital admissions in 2018. Assessments performed in 2020.	Gwet's AC_1 agreement coefficient, observed agreement, Cohen's kappa, Light's kappa.	IRR regarding assessments of the quality of drug treatment at admission and discharge, and the relationship with admission to the hospital, including preventability. Assessments were made from an overall medical perspective, independently and in consensus.
III	Intervention for improved prescribing confidence	Randomized controlled study	57 intern physicians. Intervention group (n=26) receiving structured education in pharmacotherapy in 2014-2016.	Mann-Whitney U, Chi-square, Wilcoxon's signed rank and McNemar's tests. Multiple logistic regression.	Primary outcome: Confidence in basic prescribing practices, including performing *basic medication reviews*, measured with questionnaires at baseline and after 6 months (self-estimation on a 5-point Likert scale).
IV	Key aspects to educate junior physicians	Qualitative	34 intern physicians (2014-2016). Written feedback from study III formed the data.	Manifest content analysis.	Research question: "Educating junior physicians in the art of prescribing: what aspects do they find important?".

IRR inter-rater reliability, *PIMs* Potentially Inappropriate Medications, *PPOs* Potential Prescribing Omissions, *START* Screening Tool to Alert to Right Treatment, *STOPP* Screening Tool of Older Persons' Prescriptions

Methods and Results

Measuring drug treatment quality

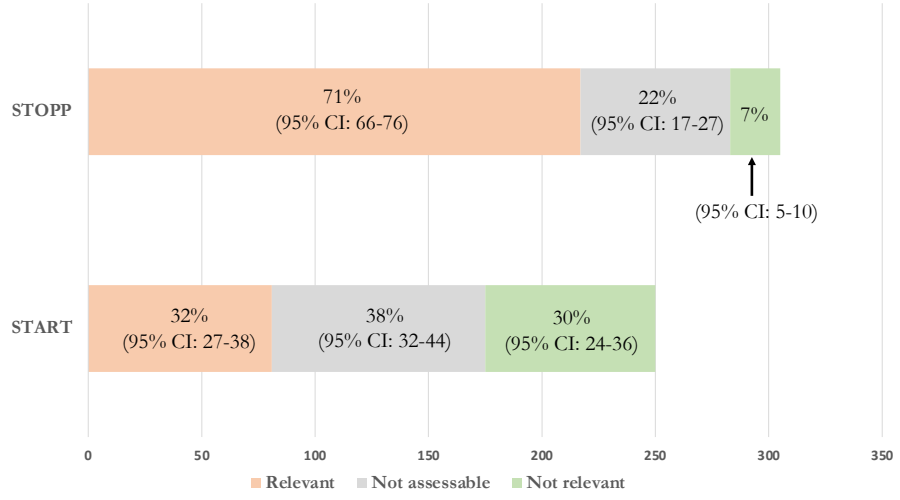
The first study included 200 hip fracture patients, originally participating in a randomized controlled study in 2009 [85]. The patients were ≥65 year of age, had undergone hip fracture surgery at Sahlgrenska University Hospital and provided informed consent. The mean age was 84.5 years, 67% were women and the mean number of drugs in the medication list was 7.2.

Methods: The STOPP/START criteria [86] were applied on the medication lists obtained in the original study in 2009, to identify PIMs and PPOs. The clinical relevance of each PIM/PPO was then assessed at the individual level by one general practitioner and one geriatrician in 2012-2013. If there was a clear benefit of treatment for a specific patient, or, conversely, a clinical reason not to treat a patient with a specific drug, the PIM or PPO was assessed as not clinically relevant. When there was not sufficient information to determine clinical relevance, the PIM/PPO was categorized as not assessable. After the initial independent assessments, the assessors reached agreement regarding identified PIMs/PPOs, including their clinical relevance, in a consensus discussion.

Results: A total of 555 PIMs/PPOs were found in 170 patients and 298 (54%) of these were assessed as clinically relevant. A greater proportion of PIMs than PPOs was clinically relevant (217 out of 305 (71%) vs. 81 out of 250 (32%); $P<0.0001$), while a greater proportion of PPOs compared to PIMs was not assessable regarding relevance (94 out of 250 (38%) vs. 66 out of 305 (22%); $P<0.0001$). In all, 22 PIMs and 75 PPOs were not clinically relevant (Figure 1).

The most frequently occurring PIMs were benzodiazepines in those prone to falls and aspirin without cardiovascular disease. For the former of the two, and for loop diuretics to manage ankle edema without clinical signs of heart failure, no PIM was assessed as not clinically relevant, though in some cases clinical relevance could not be assessed. For the STOPP criteria "long-term long-acting benzodiazepines", all PIMs were assessed as clinically relevant (Table 2).

Figure 1 Number and proportion (95% confidence interval (CI)) of *clinically relevant, not clinically relevant,* and *not assessable* PIMs (according to STOPP) and PPOs (according to START).



PIMs Potentially Inappropriate Medications, *PPOs* Potential Prescribing Omissions.
START Screening Tool to Alert to Right Treatment; *STOPP* Screening Tool of Older Persons' Prescriptions.

The most frequently occurring PPOs were osteoporosis without calcium and vitamin D supplementation, and chronic atrial fibrillation without warfarin. For the former, 77% were not assessable, while for the latter, regarding anticoagulants, almost half of the PPOs were assessed as not clinically relevant (Table 2).

The odds for ≥ 1 clinically relevant PIMs/PPOs were greater for patients with multi-dose drug dispensing (odds ratio (OR) 7.95 (95% confidence interval (CI) 2.35 to 26.9)) and increased by the number of drugs in the medication list (OR 1.28 (95% CI 1.14 to 1.43)). There was no significant association between age, sex, cognition or nursing home residence and the presence of ≥ 1 clinically relevant PIMs/PPOs. The tolerance level was >0.6 for all variables included in the model, except for multi-dose drug dispensing (0.48), indicating that multicollinearity was not a problem.

Table 2 A selection of PIMs and PPOs identified by the STOPP and START criteria.

Description		Clinically relevant	Not clinically relevant	Not assessable
STOPP	**PIM***	**PIM**	**PIM**	**PIM**
Benzodiazepines in those prone to falls	76 (38)	47	0	29
Neuroleptic drug in those prone to falls	17 (8.5)	11	0	6
Aspirin with no history of coronary, cerebral or peripheral arterial symptoms, or occlusive arterial event	27 (14)	13	12	2
Loop diuretics without clinical signs of heart failure	24 (12)	20	0	4
Long-term long-acting benzodiazepines	13 (6.5)	13	0	0
Vasodilator drugs known to cause hypotension in those with persistent postural hypotension	11 (5.5)	9	1	1
START	**PPO***	**PPO**	**PPO**	**PPO**
Osteoporosis and no calcium and vitamin D supplementation	61 (31)	10	4	47
Chronic atrial fibrillation without warfarin	29 (15)	7	14	8
Atherosclerotic disease with sinus rhythm without aspirin or clopidogrel	21 (11)	9	8	4
Chronic heart failure without ACE inhibitor	20 (10)	5	10	5
Chronic stable angina without and no beta-blocker	17 (8.5)	10	3	4

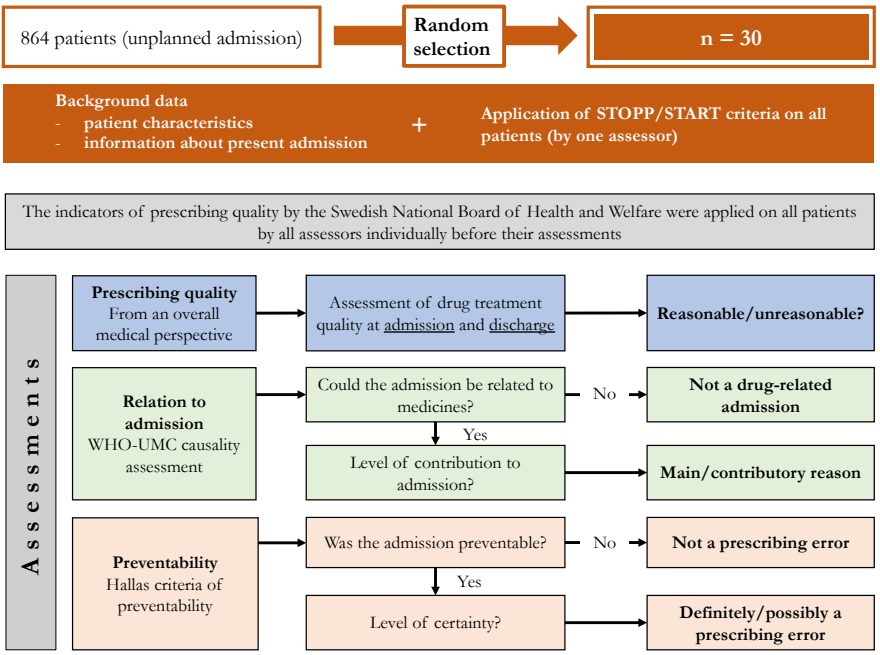
Values are presented as numbers of patients (*percentage of all patients).

ACE Angiotensin converting enzyme, *PIMs* Potentially Inappropriate Medications, *PPOs* Potential Prescribing Omissions, *START* Screening Tool to Alert to Right Treatment, *STOPP* Screening Tool of Older Persons' Prescriptions. Figure adapted and modified from Table 2, Paper I.

In the second study, three specialists and two residents in internal medicine, with experience of patient work ranging from 2.5 to 19 years, assessed aspects of drug treatment quality and adverse events regarding 30 randomly selected patients with an unplanned admission to Sahlgrenska University Hospital during a 2-week period in 2018.

Methods: The structure of the assessments is displayed in Figure 2. In order to identify PIMs/PPOs generally considered suboptimal, one assessor applied the STOPP/START criteria on all medication lists [30] and shared with the other assessors before the individual assessments. Each of the five assessors went through a broad set of indicators of prescribing quality provided by the Swedish National Board of Health and Welfare on each case before the assessment [31]. Individual independent assessments were followed by a consensus discussion on all cases.

Figure 2 Flowchart of the assessments performed in Study II.



Assessments of (i) prescribing quality, (ii) the relationship with admission and (iii) preventability was performed individually by all five assessors, from an overall medical perspective, followed by a consensus discussion following the same assessment structure described above.
START Screening Tool to Alert to Right Treatment, *STOPP* Screening Tool of Older Persons' Prescriptions, *WHO-UMC* World Health Organization - Uppsala Monitoring Centre. Adapted and modified from Figure 1, Paper II.

The assessors used the WHO criteria for causality assessment of suspected ADRs [45], and the Hallas criteria of preventability for the assessments of potential prescribing errors [87]. All suspected ADRs were assessed regarding relation to hospital admission, and in cases where a relation was found, dichotomized into main or contributory reason. Potential omissions of indicated drugs were assessed similarly, i.e., if there was a medical condition that could have been treated, whether it contributed to or caused admission, and if a prescribing error was present.

Results: The full characteristics of the 30 patients can be viewed in the original article. In short, 50% were women and the median age was 72 years (range 25 – 93). The patients had a median of 7 drugs in their medication list at admission (range 0 to 16).

Agreement was almost perfect between the assessors regarding whether each drug individually was reasonable or not at admission and discharge, and nearly all drugs were in consensus considered reasonable (Table 3).

Table 3 Assessments of (i) the overall drug treatment quality, and (ii) the relationship between the drug treatment and hospital admission, as well as the inter-rater agreement (AC_1).

		Assessments (n)		Agreement		
		Consensus	Discordant assessments (n)	All assessors		
Quality assessment of drug treatment, individual drug level				Obs. agr.*	Kappa**	AC_1***
Admission	Reasonable	207	6	0.98	0.41	0.98
	Unreasonable	3	3			
Discharge	Reasonable	242	4	0.99	0.68	0.99
	Unreasonable	3	1			
Quality assessment of drug treatment, patient level						
Admission	Reasonable	28	5	0.90	0.42	0.88
	Unreasonable	2	2			
Discharge	Reasonable	28	3	0.95	0.63	0.94
	Unreasonable	2	1			
Hospitalization						
Not related to drug treatment		21	7	0.75	0.43	0.67
Drug treatment main/contributory reason		1/8	0/7			
	Not a prescribing error	27	10	0.77	0.17	0.73
	Definitely/possibly a prescribing error	0/3	0/3			

*Observed agreement (= percent agreement).

**Cohen's kappa was used for calculating agreement between two raters, and Light's kappa between multiple raters.

***AC_1 = Gwet's agreement coefficient, was used to calculate agreement between pairs as well as multiple raters. See Methods.

Adapted from Table II, Paper II.

In 23 cases out of 28, where the drug treatment at admission was determined in consensus to be reasonable, the individual ratings were in total concordance. The corresponding numbers regarding discharge were 25 cases out of 28. This resulted in an agreement coefficient (AC_1) of 0.88 and 0.94 for admission and discharge, respectively (Table 3). Low variation was seen in the pairwise AC_1 within the specialist group, ranging between 0.83-0.88 for assessments regarding admission and 0.92-0.96 regarding discharge. The AC_1 for residents was 0.92 (95% CI 0.82-1) for assessments of reasonableness at admission as well as at discharge.

Regarding the relationship between drug treatment and hospital admission, discordant individual assessments were seen in seven out of the nine (78%) cases where the physicians in consensus assessed that the admission could be related to the drug treatment, compared with in seven out of the 21 (33%) cases where there was no relation to the drug treatment. The corresponding inter-rater agreement was 0.74 for the specialists (95% CI 0.56-0.92) and 0.54 (0.29-0.80) for the residents. As for the quality assessments above, the pairwise variation between specialists was low, ranging between 0.69-0.78.

In all three cases where the hospital admission in consensus was considered to be *possibly* attributed to a prescribing error, the physicians' assessments were discordant. No case of hospital admission was in consensus considered to be *definitely* caused by a prescribing error. The AC_1 for the specialists was 0.76 and for the residents 0.82.

The agreement between specialist and resident assessments, respectively, and consensus is presented in Table 4.

Table 4 Agreement versus consensus for assessments of drug treatment quality, relation to hospital admission and preventability.

	Agreement vs. consensus*							
	Quality of drug treatment				Relation to hospitalization			
	At admission		At discharge		Admission caused by drug(s)?		Prescribing error?	
	Kappa**	AC_1***	Kappa	AC_1	Kappa	AC_1	Kappa	AC_1
Specialists	0.71	0.92	0.86	0.97	0.69	0.78	0.37	0.84
Residents	0.73	0.96	0.72	0.96	0.54	0.79	0.15	0.82

*Agreement vs. consensus = the average of all individually calculated agreement coefficients between assessors and consensus within the groups.

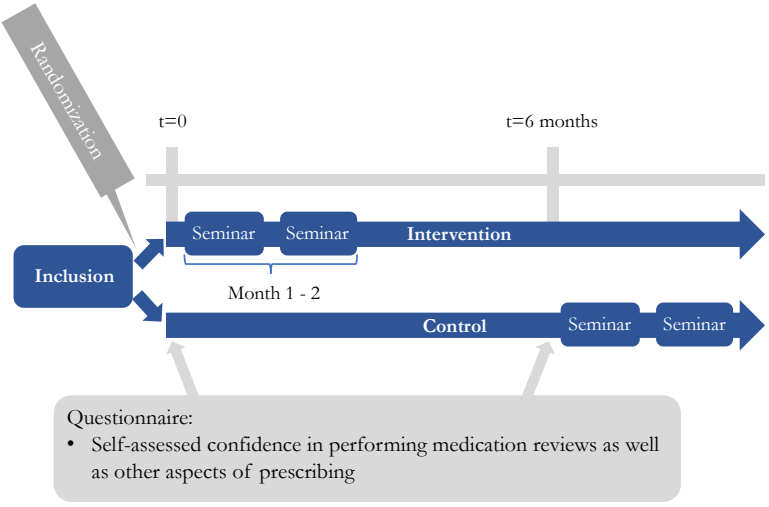
**Cohen's kappa

***AC_1 = Gwet's agreement coefficient.

Strategies to improve prescribing

In the third study, interns at Sahlgrenska University Hospital were asked to participate in a randomized controlled study between 2014 and 2016, receiving either a short, seminar-based education in prescribing and pharmacotherapy (intervention), or the standard internship without any specified educational effort on the subject in question (control) (Figure 3).

Figure 3 Outline of Study III.



Intervention: Interns in the intervention group participated in two 3-hour seminars, led by one internist and one clinical pharmacologist. The first seminar was structured to cover the different parts of a medication review, based on fictive or authentic cases. Participants were also given practical tips on how to use different decision support systems and how to use a publicly available database alerting for drug-drug interactions. For a summary of the cases used as a base for discussion, see Appendix A. As homework before the second seminar, scheduled 2-3 weeks after the first seminar, the interns were to practice performing medication reviews. The content of the second seminar was largely unspecified beforehand and focused on collegial discussions about the participants' cases and experiences from their clinical work.

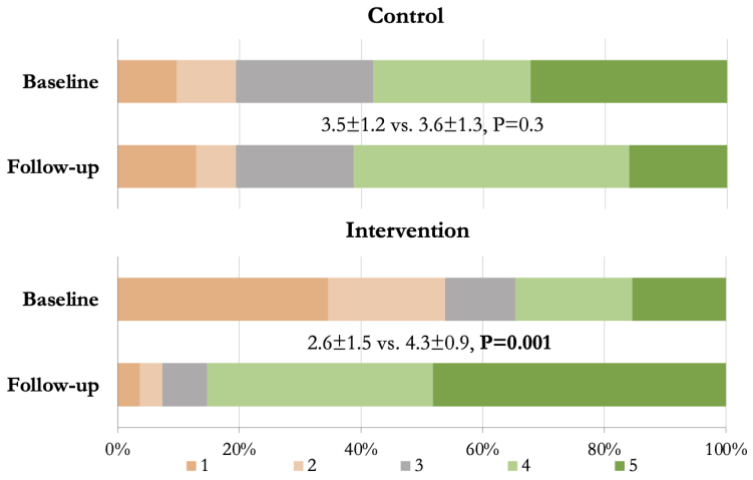
All interns in the control group were offered to participate in the seminars after completion of the follow-up questionnaire.

Outcome measures: The effects of the intervention were evaluated with a questionnaire, distributed at baseline as well as at 6-month follow-up (Appendix B). Self-assessed confidence in performing basic medication reviews was the primary outcome of the study, but we also assessed the interns' perceived knowledge of the components of basic and expanded medication reviews, as well as of other important aspects in pharmacotherapy, such as attention to kidney function, drug-drug interactions and ADRs. The participants responded to statements on a Likert scale, ranging from 1 = completely disagree, to 5 = completely agree.

Results: In total, 69 interns were included, 57 (83%) of whom completed the study (median age: 29 (interquartile range (IQR) 27-34) years, 54% female, 28% with any type of active research included in the internship). Eight participants (intervention group) did not attend the seminars and four (control group) did not complete the questionnaires.

A graphical view of the distribution of ratings in the two groups is presented in Figure 4. Although less confident at performing a basic medication review at baseline, intervention participants showed more confidence than the control group at follow-up. There were only small and insignificant changes in the control group.

Figure 4 Respondent agreement, at baseline and follow-up, with the statement "*I feel confident in performing a basic medication review*" from 1 (completely disagree) to 5 (completely agree). Significant P values (<0.05) are bolded.



In line with these results, the change in ratings between baseline and follow-up was larger for intervention interns (Δ+1.7±1.9 vs. +0.2±0.9, P=0.0002). This was also true regarding confidence in performing an expanded medication review and taking responsibility for the medication list at discharge (Δ+1.3±1.2 vs. +0.5±1.0, P=0.008 and +0.6±0.9 vs. -0.2±1.0, P=0.002, respectively). Further, regarding perceived knowledge of the components of basic and expanded medication reviews, intervention interns increased their ratings to a larger extent compared with the control group (Δ+1.6±1.7 vs. +0.4±0.8, P<0.0001 and +1.5±1.3 vs. +0.5±1.0, P<0.002, respectively).

From baseline to follow-up, intervention interns also increased their self-perceived performance in several other important aspects related to safe and rational use of medicines, for example, confidence in withdrawing medicines and reflecting on the overall reasonableness of the drug treatment (Figure 5). Control participants, somewhat unexpectedly, showed increased ratings regarding, for example, reflecting on drug-drug interactions (Δ+0.4±1.1, P=0.04) and kidney function (Δ+0.4±0.8, P=0.02) (Figure 6), which could not be seen for intervention participants. Further, intervention interns increased the number of sources used for medication reconciliation (Δ+0.6±0.9, P=0.005), while the control group showed no such increase.

After adjusting for age, sex, research experience, work experience before completing the follow-up questionnaire, and baseline confidence, the odds for an intern to feel confident in performing a basic medication review at follow-up was 8.38 (1.37 to 39.7) times higher if the intern belonged to the intervention group. The tolerance level was >0.7 for all variables included in the model, indicating that multicollinearity was not a problem.

Figure 5 *Intervention intern* agreement, at baseline and follow-up, with statements concerning important pharmacotherapeutic aspects, from 1 (completely disagree) to 5 (completely agree). Significant P values (<0.05) are bolded.

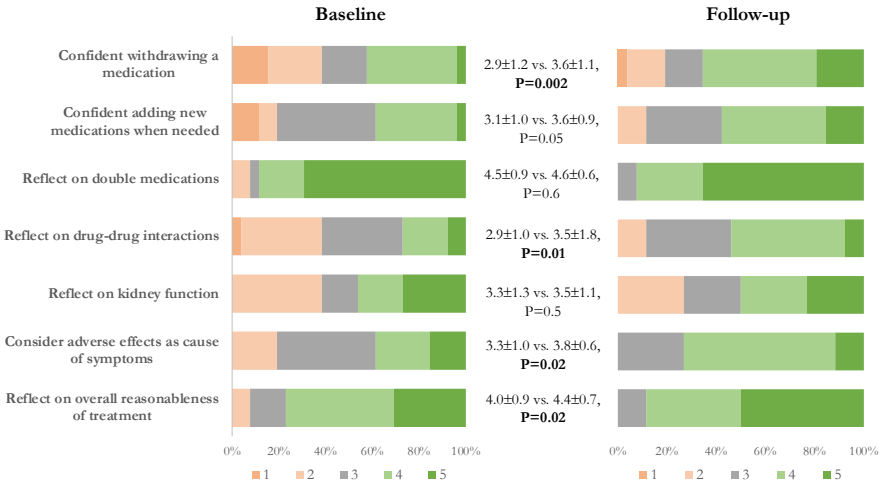
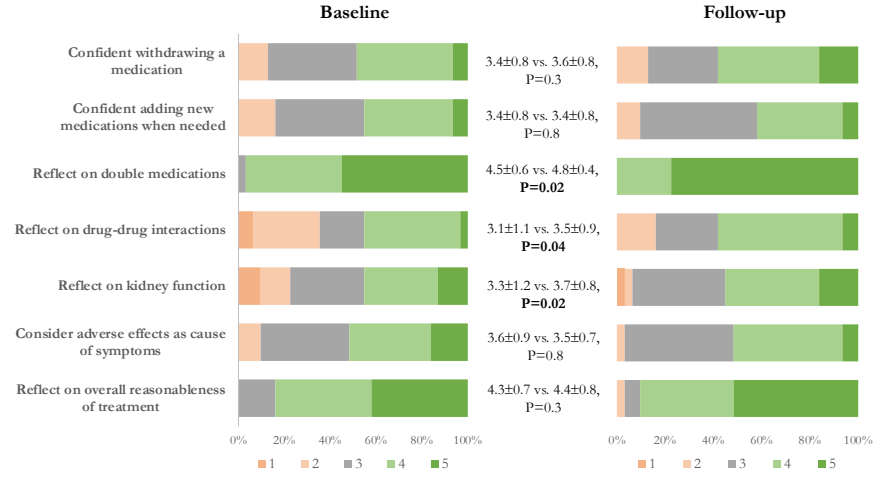


Figure 6 *Control intern* agreement, at baseline and follow-up, with statements concerning important pharmacotherapeutic aspects, from 1 (completely disagree) to 5 (completely agree). Significant P values (<0.05) are bolded.



.

In the **fourth and final study**, we wanted to explore what aspects junior physicians find important when they are educated in pharmacotherapy. To do so, we used written feedback provided by the interns participating in **study III** (n=34) and analyzed their free-text replies to the questions: "What was helpful about the seminars?" and "What could be improved?".

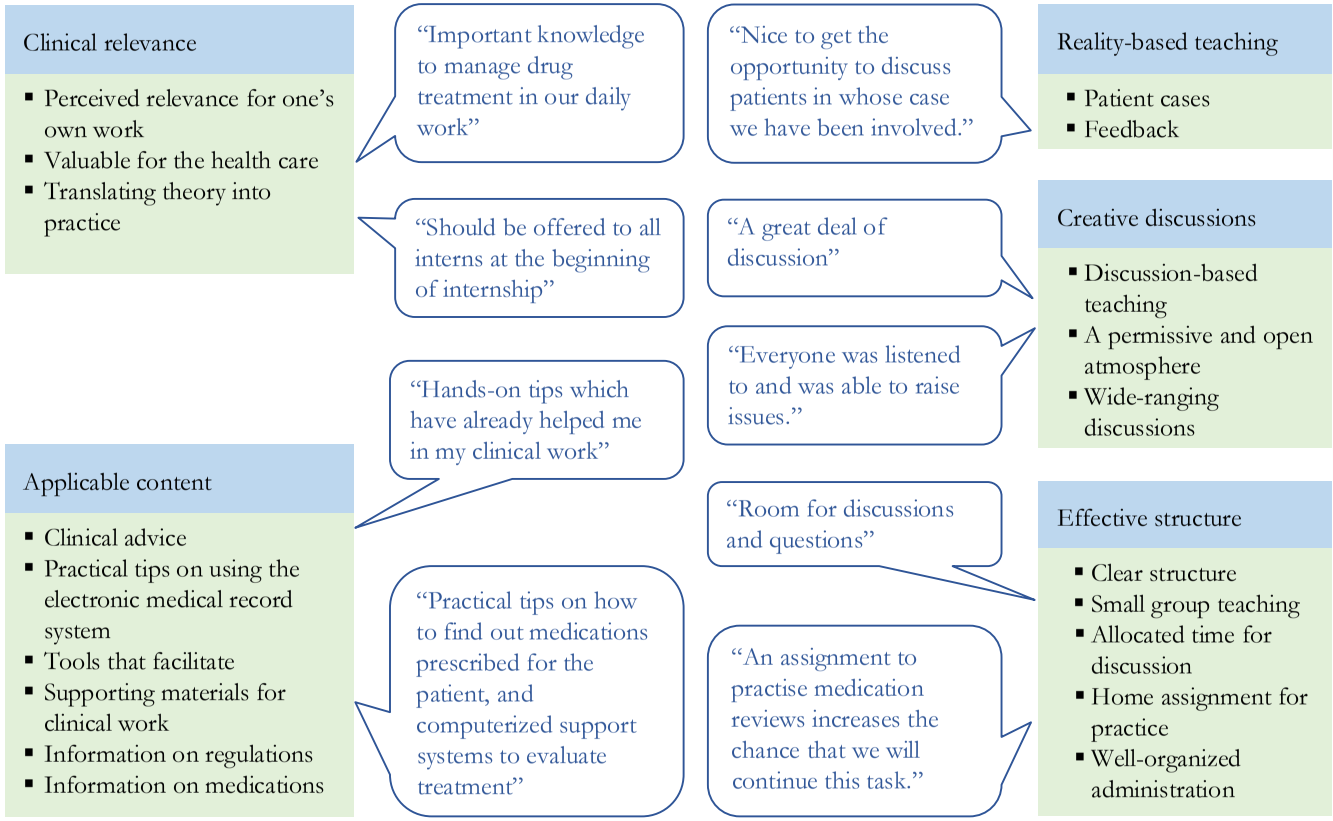
Methods: We performed a manifest content analysis guided by the research question "Educating junior physicians in the art of prescribing: what aspects do they find important?" [88].

In the first step, the author of this thesis together with a consultant in clinical pharmacology/professor in pharmacotherapy independently identified meaning units concerning the overall research question, classifying them as "appreciation statements" or "improvement statements". In the second step, we discussed all extracted units and reached consensus. At the end of this process, indications that the gathered data had reached saturation became evident, as similar codes reappeared, while there was a lack of new codes. The next step included a data-driven inductive thematic analysis, were we sorted the units into various, previously undetermined, categories. Finally, we identified emerging themes, each one of which was a thread of an underlying meaning at an interpretative level [88]. Any disagreements between the analysts during the process were resolved through consensus discussions.

Results: Five themes emerged during the content analysis: *clinical relevance; applicable content; reality-based teaching; creative discussions; and effective structure.* Themes and categories, as well as example quotes, are presented in Figure 7.

To begin with, interns seem to feel a need for education in pharmacotherapy and prescribing to manage their clinical work. It also became evident that clinical advice and practical tips, including a hands-on approach to digital support systems, are appreciated pieces of education. The interns also seem to find it helpful to get a review of common medications, including drug-drug interactions and side effects, as well as to discuss medications that are in general considered inappropriate for older persons.

Figure 7 Themes and categories that emerged from the meaning units, with example quotes.



Another aspect of importance, when educating junior physicians in the art of prescribing, appears to be to base the sessions on authentic patient cases, and to combine an interactive teaching approach with time for reflection and feedback. In addition, interns appreciate a clear structure where the core messages of the education are brought forward. Also, small group teaching seems to be important, in combination with allocated time for discussions, in a permissive and open atmosphere. The interns appreciate to practice new skills and get feedback during a second seminar.

Finally, it appeared that interns find well-organized administration important, including the possibility to have several options of when to participate in the seminars. Interns also seem to be positive about having a home assignment; it encourages them to practice.

Statistics

Statistical analyses were performed with SPSS (IBM SPSS Statistics, v 17-27) (**study I - III**) and with R (R Core Team, 2020) (**study II**). In **study IV**, NVivo 12 Pro (QSR International, Melbourne, Australia) was used for data management. P values less than 0.05 were considered significant.

An overview of the statistical methods used in the studies is presented in Table 1. Descriptive statistics were used to present the characteristics of the study populations in all papers. Discrete data are presented with means and standard deviations (SDs) or medians with range or interquartile range (IQR), and categorical data are displayed as frequencies, n (%). See Discussion for methodological considerations regarding the use of mean as a descriptive measure of Likert ratings.

In **study I** and **III**, we used the Mann-Whitney U and Chi-square tests for comparisons of discrete/ordinal and categorical data between groups. For within-group analysis of the randomization groups in **study III,** we used Wilcoxon's signed-rank and McNemar's tests for ordinal and categorical data, respectively. Logistic regression was performed in **study I** to obtain odds ratios (OR), with 95% confidence interval (CI), for ≥1 PIMs/PPOs, as well as ≥1 clinically relevant PIMs/PPOs, and in **study III** to obtain OR for the intervention to predict perceived confidence in performing basic medication reviews at follow-up. Covariates used for adjustments in logistic regression are reported with the results. To detect potential multicollinearity issues between independent variables, we used tolerance statistics.

Cohen's kappa was used in **study I** to assess the agreement between two raters.

In **study II**, we used Gwet's AC_1 coefficient to calculate inter-rater agreement for categorical, unweighted data [58]. As a reference, and to visualize the paradox of kappa (see Introduction and Discussion) [89], we also calculated Cohen's kappa between pairs and Light's kappa between multiple raters.

In **study I and II**, inter-rater agreement was interpreted as none (<0.20), minimal (0.21-0.39), weak (0.40-0.59), moderate (0.60-0.79), strong (0.80-0.89) and almost perfect (>0.90) [55].

In **study II**, we assumed a prevalence distribution of 75%, 20% and 5% for assessments within the categories "no", "contributory" and "main" regarding drug related admission, and similar within the categories "no", "possibly" and "definitely" regarding prescribing error. In order to have at least 80% power to detect a difference between kappa 0.40 and

0.70, with a margin of error of 5%, using 5 raters and up to 3 categories, we needed to include 25 patients. Sample size calculations were performed with the R package "Power3Cats" [90]. The algorithm is based on kappa, though, such methods of sample size calculations have previously been used also for AC_1 [91].

In **study III**, we needed to include 30 interns in each group (for a total of 60 participants) in order to have at least 80% power to detect a difference of 0.8 on a Likert scale from 1 to 5, with a standard deviation of 1.0. To compensate for drop - outs, another 10% were included.

Ethical considerations

All studies were approved by the Regional Ethical Review Board in Gothenburg: **Study I** (ID: 095-09 with amendment), **Study II** (ID: 463-18), **Study III** (ID: 344-14), **Study IV** (ID: 344-14 with amendment).

Patients in **study I** provided written informed consent before they were included in the original study in 2009. For the study included in this thesis, an amendment was submitted and approved, concerning ethical considerations regarding further analysis of previously collected data without renewed informed consent from the patients. Analyzing health-related data is always associated with a risk of violation of personal integrity. In this particular case, however, 20% of the patients included in the original study were deceased after 12 months, and 45% were diagnosed with dementia or impaired cognition at the initial inclusion [85]. If informed consent would have been required, many patients would probably have had to be excluded, thereby severely aggravating the external validity. Further, as the patients included in the original study received information that they were not to be contacted by the researchers again, we considered such contact to be of greater concern for personal integrity than it would have been to analyze the data already collected without further consent. The ethical review board approved the amendment without requirement of additional informed consent, thereby confirming that the benefits of these additional analyses exceeded the integrity risks.

In **study II**, which was a retrospective, descriptive study, we did not collect informed content from the patients. As in **study I**, the risks for the patients concerned the possible violation of their personal integrity associated with, in this case, accessing the health record and collection of personal data. Since it was a study about drug treatment quality and drug-related admissions, it was important to include patients with many drugs, i.e., commonly patients with multiple deceases and of old age. If informed consent would have been required, there would have been a considerable risk that patients had to be excluded from the study, thus reducing the external validity of the results. Since a deeper understanding about assessments of drug treatment quality and the relationship with adverse events is needed, the ethical review board determined that the benefits of waiving informed consent outweighed the risks. Apart from the patients, the performance of the physicians in charge of the patients during the time of hospital care could also be considered to be evaluated as part of the study, risking violation of their professional integrity as their work was scrutinized by a group of scientists and colleagues. However, as no information about the physicians was recorded, this risk was considered minor. Also, it was almost exclusively the work

of the physicians in charge *prior* to the visit to the emergency department that was in focus, and the identities of these physicians were not known.

For **study III**, written informed consent was obtained from all participants before inclusion. Eligible interns were contacted in three different ways. At first, some of them were informed about the study during an educational session at Sahlgrenska University Hospital. Second, we contacted each intern by telephone and informed them about the study, using a predetermined template that included, for instance, information that participation was fully optional and could be discontinued at any time. Finally, interns that could not be reached by telephone received the same information by post. Although no health-related personal data was collected about the participants, integrity issues are present in all studies on human subjects. In this case we gathered data by asking the participants to fill in questionnaires. The participants could choose to skip certain questions in case he or she perceived it as sensitive in one way or another. To further protect the integrity of the participants, results were presented at the group level only. The benefits of new knowledge on educational aspects of pharmacotherapy, together with the benefits for each participant to receive an educational effort, were considered to exceed the risks for the interns included in the study.

For **study IV**, we analyzed written feedback on the seminars included in **study III**, provided by the participants directly after the second seminar. The participants were not informed about this qualitative analysis when they entered the randomized controlled study. For that reason, an amendment was submitted and approved by the ethical review board before the analysis took place. To start with, feedback was given anonymously by the interns, to make sure they could express their opinions freely. As the texts were written by hand, there was a slight risk of recognition by characteristics of handwriting. However, we considered that risk to be insignificant. Importantly, there was no sensitive information about the participants in the texts. Further, if any of texts were to be identified, it would not have been of any disadvantage for the participants, since the researchers were not involved in their work in any substantial way. It could possibly cause inconvenience in case their criticism of the seminars was to have been identified by the researchers, though we believed that to be of minor concern. In the end, the benefit of learning more about what junior physicians find important when being taught pharmacotherapy was considered to exceed the risks mentioned above.

Discussion

Methodological considerations

Study design

This thesis is built upon one descriptive (**study I**) and one reliability study (**study II**), as well as one randomized, controlled study (**study III**) and one qualitative study (**study IV**). The descriptive design is commonly used for measuring prevalence figures and to generate hypotheses for further research. In study I, we also calculated odds ratios related to the presence of at least one clinically relevant PIM or PPO. This could, to some extent, be considered a cross-sectional design comparing those with and without such treatment. With the lack of random allocation to the groups, such an approach has an inherent risk of bias. However, it should be stressed that the main purpose of the study was to describe the clinical relevance of PIMs/PPOs, and associations were only explored.

Despite the fact that the study population in **Study I** had a high participation rate, 200 out of 229 consecutive eligible patients being included in the original randomized controlled study, there are still concerns of external validity. One such concern is to what extent hip fracture patients represent older persons in general. Hip fracture patients, as a group, are known to be at risk of suboptimal drug treatment [92]. However, hip fractures are also common among older persons [93], and therefore the group may represent a relevant cohort of older patients. Nevertheless, the prevalence of drugs that are related to falls can be expected to be higher in our study population than in the general population. Other studies have shown prevalence figures of benzodiazepines in those prone to falls of 5-9% [94-96], compared to 38% in our study, supporting this assumption. Our results should therefore be interpreted with this in mind and may primarily be applicable to older hip fracture patients.

In **Study II,** the patient sample was randomly selected from all patients of at least 18 years of age with an unplanned admission to a tertiary hospital during a 2-week period. The random selection of patients decreased the risk of selection bias, although, the number of selected patients was rather small (n=30), which is discussed further in the text below (lack of precision). Another possible factor of concern is the time period from when the patient sample was collected. Collecting patients from a short period of time could increase the risk of bias due to seasonal variation, or the risk of influence of individual physicians on the care of patients admitted to the hospital. We therefore chose a study period to avoid national holiday periods and the period between June and August due to summer vacation. As the patients were admitted to several different wards at the hospital, the influence by individual physicians on drug treatment quality at discharge could be expected to be small. Further, as the study outcome was inter-rater agreement between

physicians, the selection of assessing physicians may also be a matter of concern. Ideally, to be able to draw distinct conclusions on agreement, the assessors, as well as the patients, should be randomly selected from a relevant target population. For this particular study, the best solution to reduce this risk of observer bias would have been to randomly select the assessors from the department of internal medicine, or ideally from other specialties as well. However, such a random selection of assessors would also imply feasibility challenges. Assessing the drug treatment from an overall medical perspective in a standardized way in a transparent research process is time-consuming, potentially unveiling, and requires willingness as well as opportunity among the assessors.

Lack of blinding

In **Study III**, we performed a randomized controlled study to test the hypothesis that a short education with collegial discussions on pharmacotherapy could increase interns self-assessed confidence in prescribing. One limitation was that neither participants nor researchers (tutors) were blinded. Even though the participants did not receive explicit information about which group they belonged to, it was possible for them to figure it out. There are some methodological risks with this design, in our study mainly related to performance bias. Knowledge about being in the intervention group could have affected the participants in such a way that they *expected* to feel more confidence during follow-up. Conversely, control participants could possibly feel disappointed not to receive the intervention until after follow-up. This sort of bias is known to increase the risk of over-estimating the effect of the intervention and could be of particular concern in studies with self-reported estimations of competence or confidence. However, with an intervention like this, it was not possible to blind the participants, and with a 6-month period between baseline and follow-up questionnaires, the risk of being affected by one's previous answers was considered to be of minor concern. Secondly, there was a risk that the tutors (researchers) transferred their own expectations of increased confidence onto the intervention participants. Using blinded tutors would have decreased this risk. Though, enthusiastic tutors may also increase the participants engagement in the discussions and interest in the subject, resulting in actual increased knowledge and understanding, and in the end, greater confidence. The results of **study III** should be interpreted with these different aspects of potential bias in mind.

Lack of precision

One important limitation in **study II** was that the number of assessed patients was relatively small. As stated in the Methods and Results section, we had 80% power to detect a difference between kappa 0.40 and 0.70 with a margin of error of 5% concerning all five physicians, though we did not have enough power to detect differences between physicians with different experience. Furthermore, the fact that the group of residents included only two assessors may imply limitations regarding the external validity of our

results; more residents would have made the results less sensitive for individual differences in experience among the assessors.

Surrogate outcome measures

In **study III**, we examined self-perceived confidence regarding different aspects of prescribing to determine whether an educational intervention could increase confidence. Previous research shows a weak correlation between self-reported confidence and prescribing competence in medical students [97]. However, medical students and interns differ in several aspects, including experience in prescribing. For someone who has not encountered the responsibility of prescribing medications in daily work, it may be difficult to estimate confidence. Therefore, the correlation between confidence and competence could be particularly problematic in medical students. To learn more about this, we performed a post-hoc analysis in which we compared self-reported confidence in prescribing between students in the fourth year of medical school at Gothenburg University in 2016-2017 (n = 386, response rate 82%) and the interns in our study. It turned out that among the students, 7% were not at all confident in performing medication reviews, compared to 23% among the interns. This finding may illustrate that a certain level of experience and knowledge may be needed, under one's own responsibility, to be able to detect a lack of skills.

Instead of measuring professional confidence, there are other, perhaps more relevant, ways to measure potential effects of an intervention such as ours. The most relevant outcome would be to measure differences in outcomes for the patients treated by the two groups of physicians. However, since interns work under supervision by their senior colleagues, any differences arising from the intervention would have been difficult to distinguish. Another path could have been to measure other quality aspects of their work. As a matter of fact, we initially intended to compare the number and quality of medication reviews performed by the two groups of interns. Since all interns did not have the same clinical locations during the time between measuring points, there would have been other factors influencing the results, thus, yet again diluting any potential effects from the intervention. Finally, we could have provided the interns with a test to find out whether or not they had increased their theoretical knowledge on important aspects of prescribing. It should be noted though, that tests often tend to highlight general knowledge and guidelines are not always applicable on individual patients, especially when it comes to older persons with multiple diseases. Indeed, the median age of all 864 patients with an unplanned admission to Sahlgrenska University Hospital during the 2-week period at issue in **study II** was 70 years. A more elaborate test that measures complex skills of medical reasoning may be difficult to accomplish, although not impossible. Future research could therefore include both perspectives to fully cover the different aspects of prescribing skills.

Methodological aspects of statistics

Two studies in this thesis use statistics to measure inter-rater agreement between two or more assessors. Measures of agreement may at first glance seem straightforward, but are commonly accompanied with various methodological errors, such as choosing the wrong method for the study design, not reporting what method was used or not estimating statistical power when needed [56]. The statistical methods of inter-rater reliability have evolved during the years, from measuring percent agreement to adjusting for chance agreement in the commonly used Cohen's kappa and Fleiss kappa [55, 56]. In recent years, critique has arisen regarding the commonly used kappa statistics, as they may result in inappropriately low kappa values despite high percent agreement. The problem appears when the distribution between the outcome categories is skewed, for instance, as in **study II**, where few patients had unreasonable drug treatment. This phenomenon, referred to as the "kappa paradox", has been described in detail elsewhere [89] and is also discussed shortly in the Introduction. Gwet et al. proposed a different method of measuring inter-rater reliability that is resistant to this paradox and shows more consistent results independent of trait prevalence [58, 59]. In **study II** we therefore chose to use Gwet's AC_1 to calculate agreement between assessors, as the ratings were greatly skewed. In order to visualize the paradox, percent agreement and Cohen's kappa are presented side by side in this thesis, together with Gwet's AC_1 (Table 4).

Although the subjectivity of assessment-based outcomes needs to be quantified, illustrated by the varying inter-rater agreement reported in original studies contributing data to a recent meta-analysis on drug-related admissions [7]: 0.16 [65] to 0.90 [53], the varying results of different methods to quantify reliability imply that comparisons between studies will be difficult. Further, there is still no consensus in how to interpret calculated values of kappa. Cohen initially suggested kappa of 0.41 – 0.60 to be "moderate" [55], which was also adapted by Landis and Koch [98], whereas McHugh proposed that the same interval should be considered "weak", while kappa of 0.61 – 0.80 was suggested to represent "moderate" agreement [55]. Although all research articles are supposed to clearly present what categorization has been used, commonly, in abstracts as well as in conclusions, the interpretations are standing alone, without the corresponding values. This may have the effect that two completely different values of kappa, for example 0.41 and 0.80, could be misinterpreted as similar agreement by the reader. Therefore, taken together, as there are several different methods of how to measure agreement, with an ongoing debate on which method to prefer and how to interpret the results, we have all reasons to be careful when we use, interpret and communicate these methods. As for **study II**, we interpreted values according to the more restrictive categorization by McHugh, which could be considered proper for studies in healthcare.

In **study III**, we present the results from the 5-point Likert assessments with a combination of means, medians and bivariate descriptive statistics, the latter formed by

dichotomizing ratings 4 or 5 as "confident" and 1-3 as "not confident". There has been some controversy concerning how to analyze Likert data, reaching back over several decades [99]. Likert scales are ordinal by nature, and categories within the scale are not necessarily at equal distances from each other. When equal distances can be assumed with sufficient approximation, means and medians could be used to visualize the responses. However, whether to use means or medians in such a situation must be carefully considered from case to case. For instance, using means when Likert data is based on words that express the opinions of the respondents may be less appropriate; intuitively, opinions are difficult to convert into numbers and means. However, as our Likert items represent cognitive aspects, as for instance "self-confidence", rather than opinions, and the scale consist of five points with "completely disagree" and "completely agree" marking the extremes, we considered it acceptable to use means and standard deviations for descriptive statistics. The ratings were not normally distributed, nor were they strongly skewed. Importantly, when visualizing the ratings using histograms, we could not see any sign of "U-shapes", i.e., responses clustered at the high and low extremes. If that would have been the case, calculating the mean would have been clearly misleading. However, using the median in such a case should have been equally wrong. The fact that means and medians did not differ much in **study III** also make the presentation of means less problematic. Finally, as a method of comparing changes between baseline and follow-up, and between the two groups, mean is easier to visualize. Nevertheless, when analyzing ordinal data with statistics usually used for interval scales, caution in interpretation is required, especially when it comes to the size of mean differences.

In **study III**, we aimed at including a total of 60 participants to ensure enough power to detect differences between the groups (see Statistics in the Methods and Results section). In retrospect, the drop-out rate used for the power calculation appeared somewhat optimistic, but this was compensated by a larger than expected difference between the comparison groups.

Measuring drug treatment quality

In **study I**, we show that one in two PIMs/PPOs identified with the STOPP/START criteria were clinically relevant from an overall medical perspective. This is in line with previous research that has shown implementation rates of 56% and 39% of PIMs and PPOs, respectively [100], and recommendations followed in 87% and 34% of PIMs and PPOs, respectively [101]. In a recent study, 40% of STOPP/START (version 2) recommendations were of at least "possibly important relevance" when assessed by a pharmacist and a physician, although the inter-rater reliability was minimal, with a kappa value of 0.24 [102]. Surprisingly, this value was reported in the study as "fair", according to Landis and Koch [102]. Another study reported acceptance rates of 91% and 97% for recommendations provided by STOPP and START, respectively [94]. However, as the

recommendations were provided to the medical team by a research physician, not relevant PIMs/PPOs would probably not be forwarded to the attending physician, and may therefore not be captured in the denominator, possibly explaining the high acceptance rates.

The extent of clinical relevance reflected by our results, together with rates of implementation and recommendations followed described above, underlines the importance of caution when interpreting prevalence figures based on general criteria like STOPP/START. Our results suggest that such instruments alone should not be used to reflect prescribing quality. When used as a measure of quality, it should be kept in mind that these criteria were created as a support tool for physicians in the daily clinical work, and not as a quality measure at the individual level [29]. As stated by the inventors of the STOPP/START criteria, a close interaction between the attending physician and the trained applicator of the criteria in the clinical setting is crucial [103]. While screening tools can be helpful not to overlook prescribing that is generally considered inappropriate, they may need to be accompanied by a physician assessment to reflect quality of prescribing from a medical perspective, i.e., to ascertain clinical relevance. Indeed, also other sets of indicators of prescribing quality have been shown to have limitations when it comes to concurrent validity, i.e., their ability to reflect quality of drug treatment [104, 105]. Similarly, other approaches that aim to find potentially inappropriate prescriptions, such as pharmaceutical medication reviews, may have the same weaknesses, and without a medical assessment by a physician, such measurements of medication appropriateness or quality should also be interpreted with caution [104]. Nevertheless, PIMs, PPOs and alerted drug-drug interactions, in indicator sets like STOPP/START, have been suggested in core outcome sets to evaluate interventions for improved prescribing [18, 106].

Potential prescribing omissions were significantly more often not assessable regarding clinical relevance. This finding could reflect that whether to start a new medication or not, despite a thorough assessment of the patient, may put a higher demand on the level of information, compared to withdrawing a potentially inappropriate drug. This may, as suggested in a previous study, be partly explained by time constraints [107], or, which is a more likely explanation in our study, lack of necessary information due to the retrospective study design. Reviewing medical records in retrospect implies that the information available is restricted to the information recorded. In real-life patient work, the patients can also be asked directly about aspects relevant for decision-making regarding drug treatment initiation. In **study III**, we found that interns participating in an education in pharmacotherapy significantly increased their confidence in withdrawing a medication when needed. Confidence in taking responsibility for adding a new medication, however, had not increased at the time of follow-up. This is in line with the above reasoning about PPOs, suggesting a difference in complexity between STOPP and START criteria. Indeed, the STOPP criteria are considered drug-oriented, as opposed to the START criteria, which have a patient-oriented approach [25], which requires more of

the assessor regarding medical competence and more in-depth information about the patient, thus possibly making prescribing decisions harder.

The results of this study suggest that directing attention towards, in particular, older persons with multi-dose drug dispensing, could increase the chances of finding clinically relevant inappropriate treatment. The same may also apply to older persons with many drugs in the medication list. In previous studies, multi-dose drug dispensing has been associated with poorer quality and fewer changes in drug treatment [108-110]. As for the number of drugs in the medication list, this may be a surrogate measure reflecting burden of disease [23]. As drug treatment considerations may be complex in multimorbid patients, it could be speculated that medical attention to the pharmacotherapy could be particularly valuable in these cases.

In line with previous and recent research, as well as with clinical experience, when attempting to improve the quality of drug treatment for older persons, a first approach could be to consider stopping treatment with long-acting benzodiazepines, and benzodiazepines and neuroleptic drugs in those prone to falls [101, 111, 112]. In the first case, all of these PIMs were assessed as clinically relevant, and in the latter two cases no PIM was assessed as *not* clinically relevant, although roughly one third of the cases were not assessable. It must be acknowledged, however, that withdrawal of such drugs is not an easy task, as it may require both motivated patients and physician time for follow-up. Similarly, one could also reconsider the need of loop diuretics for isolated ankle edema, i.e., without clinical signs of heart failure, and vasodilator drugs in patients with known postural hypotension, as they were also assessed as clinically relevant to a high degree [112]. When identifying undertreatment and starting new medications, one might consider starting off to assess whether a patient with cardiovascular disease would benefit from treatment with a platelet inhibitor, a beta-blocker, or a statin. As for renin-angiotensin-aldosterone system blockade, it was two times more likely that a PPO was *not* clinically relevant than clinically relevant, perhaps reflecting the intolerance of such treatment for older persons with regard to kidney function and risk of falls due to low blood pressure.

In study II, we found that the inter-rater reliability regarding assessments of *drug treatment quality* and *adverse events* ranged from weak to almost perfect. This wide range serves as a reminder of the built-in subjectivity of assessments in prescribing-related matters.

Agreement was strong to almost perfect for assessments of *drug treatment quality* between all assessors, as well as within the groups of specialists and residents. Together with the almost perfect agreement between assessors and consensus, which was seen for both residents and specialists, it appears that quality-related aspects can be reliably assessed by two physicians with at least 2 years of clinical experience.

A similar overall strong agreement was not seen for assessments regarding the relationship between *drug treatment and hospital admissions*. While the specialists showed moderate agreement, the residents were less concordant, at a weak level, although the confidence intervals of the two groups overlapped. This points towards a higher level of complexity involved in ADR assessments. This interpretation is also illustrated by the eight cases of admission in our study in which the drug treatment in consensus was considered a contributing factor for the admission. In these cases, up to four out of five assessors made a divergent individual assessment. Decisions regarding the causal relationship between a drug and a reaction, as well as between the reaction and admission, is a complex task that requires considerable medical experience. Indeed, most cases of suspected adverse drug reactions are categorized as "possible", according to the WHO definition of causality as well as the Naranjo algorithm, which implies that the reaction could just as well be caused by a worsening disease or another drug [47, 49, 51]. As residents, although well acquainted with medical treatment and patient care in their daily work, are still at the beginning of their professional career, it may not be surprising that they are less consistent in their assessment of these complex issues, in comparison to quality-related assessments, which may be more straightforward.

Systematic reviews illustrate that there are numerous studies addressing adverse drug reactions or adverse drug events [5] as well as hospital admissions related to medicines [3, 7, 113]. For the former, measurements of inter-rater reliability have been provided. Almost all the referred original studies, however, included only two assessors [5]. A few studies have included multiple assessors, with kappa values ranging from 0.14 to 0.88, however, commonly, the patients have been pre-selected as drug-related before assessment [63, 64, 66, 114], or they have been elective surgery patients [67]. Studies on drug-related hospital admissions reporting agreement figures between at least two assessors are uncommon, showing inter-rater reliability figures of 0.01 to 0.90 on causality [53, 61, 65, 68, 115, 116]. However, the cases were either pre-selected as drug-related [65], or they were an extremely selective cohort of patients [53], fictive [68], or solely assessed by assessors without a medical profession [53, 61, 115, 116]. There is a clear gap of knowledge when it comes to studies reporting inter-rater reliability between multiple assessors with a medical profession assessing drug treatment quality from an overall medical perspective, as well as the relationship between an adverse drug reaction and hospital admission on unselected patients. One important strength of **study II** is that it brings knowledge that helps filling that gap.

Another strength of the study is that the individual assessments were preceded by the application of screening instruments of medication appropriateness. Although constructed with the intention to detect suboptimal drug treatment in older people, we considered them appropriate to ensure that general aspects of importance in prescribing were not overlooked. Nevertheless, this method of systematization may have affected the assessments of drug treatment quality, contributing to the high level of agreement.

However, as the individual assessments were less concordant in the cases in which drug treatment in consensus was considered to contribute to admission, the effects of the instruments on agreement may be less of a concern, although not negligible.

In conclusion, to reliably assess quality-related aspects of prescribing, it seems like two physicians with at least two years of patient work would be sufficient. However, residents' assessments of drug-related admissions compared to their quality assessments were significantly less consistent; determining whether or not drugs cause or contribute to admission to hospital appears to require extensive medical competence. As the variation in agreement was low between all pairs of assessors in the specialist group, the results indicate that for assessments regarding drug-related admissions, at least two specialists in internal medicine should be needed to obtain reliable results.

Strategies to improve prescribing

In **study III** we hypothesized that a short educational intervention in a small-group seminar format, with collegial discussions on pharmacotherapy, could increase interns' professional confidence in prescribing. Despite starting at a significantly lower level of confidence in performing basic medication reviews at baseline, intervention interns were more confident than the control group at follow-up. The odds of being confident, as defined as a rating of at least 4 on a 5-point Likert scale, was substantially higher if an intern belonged to the intervention groups, which also withstood relevant adjustments. Intervention interns also increased their confidence regarding other important aspects of prescribing, such as considering adverse effects as a possible cause of symptoms and feeling confident withdrawing medications. Also, presumably due to the thorough introduction to various electronic help systems, they increased their use of multiple sources for gathering information about which medications patients were using, while the control group showed no increase at all. Control participants, on the other hand, showed a significant increase regarding their reflection on potential interactions, kidney function and avoiding double medications. For the intervention group, surprisingly, no significant changes could be seen regarding reflection on interactions or kidney function, despite focus on these matters during the seminars.

Interestingly, for control participants, confidence in performing basic medication reviews, as well as other prescribing-related matters, like for example assessing whether the drug treatment at the overall level is reasonable or feeling confident taking responsibility for the medication list at discharge, did not increase at all during the time between baseline and follow-up, despite daily clinical work. This finding is worth attention, since junior doctors have been shown to write a large amount of hospital prescriptions and have higher error rates [11]. Assumingly, as they continue their clinical training, eventually they will gain considerable knowledge and skills in prescribing. Though, from a quality and patient

safety perspective, this may not be sufficient. Our study shows the benefits of a short and structured education for junior colleagues, including collegial discussions on common as well as complex pharmacotherapeutic aspects, to provide the necessary tools for proper prescribing. This will soon become even more important in Sweden as internship is to be abolished and the license to prescribe obtained directly after medical school, in concordance with many other European countries. As previously stated, to develop advanced skills, including the ability to use prior knowledge in new situations, basic knowledge needs to be elaborated on, analyzed and understood through application [117]. Further, for efficient education, prescribing may need to be lifted as an important matter, and tailored supervision with continuous review by senior colleagues encouraged [118]. In recent years, evidence has shed light on the need to enhance the competence in prescribing among final-year medical students in Europe [8, 9]. The need for a common ground of necessary prescribing competencies has led to learning outcomes that could be included in undergraduate curricula of clinical pharmacology [73]. National examinations for final-year medical students have been developed in the Netherlands and in Belgium [74], and in the United Kingdom, the Prescribing Safety Assessment (PSA) has been running successfully for several years [72]. In contrast, in Sweden, there is currently no national final-year pharmacotherapy exam in medical school. However, recent research suggests that efforts for enhanced prescribing skills in Swedish medical students could be valuable [75, 119].

The results from **Study IV** brings further understanding as to how junior physicians could be educated. Through our qualitative analysis of interns' free-text replies it appears that creative discussions, well-structured and filled with clinically relevant, reality-based content, are key components when educating junior physicians in the art of prescribing. This evidence, in which experiences and requests of the receivers are explored, could constitute valuable information for further development of pharmacotherapeutic education for medical students and physicians early in the professional career. These findings are supported by previous studies showing positive effects of interactive learning based on clinical cases [81], active learning [80], and well-structured peer-to-peer teaching with interactive learning techniques and distribution of manuals to interns [120]. Further, small-group education has been shown beneficial in a recent study on evidence-based knowledge among healthcare workers [121], and the positive comments on getting practical tips including advice on the electronic health record and other medical help systems are in line with previous findings [122, 123]. Highlighting such knowledge to relatively young colleagues could easily be forgotten nowadays, as skills in information technology are generally high. However, as electronic medical record systems have been shown to add to physician's frustration in the daily work [124], such efforts may be valuable.

Strengths of this qualitative approach was that we could analyze the free-text replies provided by a relatively large number of interns and that we reached saturation, suggesting

that the topic was explored to a sufficient extent. However, since the participants were interns at one tertiary hospital, there may have been aspects that were not captured. Furthermore, those participating in research studies may diverge from those who do not [125], and therefore it cannot be excluded that some aspects were not fully explored. Nevertheless, as qualitative research on this topic is scarce, our results provide valuable insights.

Conclusion

In this thesis, methodological insights are provided regarding (i) the clinical relevance of instruments commonly used to describe prescribing quality and (ii) the reliability of drug treatment assessments. In short, half of the suggestions in the commonly applied STOPP/START instrument were clinically relevant for a specific patient when assessed from an overall medical perspective. Caution is therefore urged when interpreting research results based on these tools, which may have limited applicability at the individual level and reflect quality of drug treatment to a limited extent. Further, the inter-rater agreement between physicians with at least 2-year experience of patient work was fully acceptable regarding assessments of drug treatment quality, but the assessments regarding the relationship between drug treatment and hospitalization appeared more divergent; the inter-rater agreement for residents was lower for this causality issue than for quality issues. The results suggest that specialist assessments are required for reliable results regarding drug-related admissions.

The thesis also provides practical insights regarding how junior physicians can be supported to acquire basic prescribing skills; prior research shows that efforts are needed in this regard. Two 3-hour seminars for interns, led by an internist and a clinical pharmacologist, were shown to substantially increase the recipients' confidence in prescribing at 6-months follow-up, whereas little happened in the control group, merely working as junior physicians. Exploring what junior physicians find important when educated within the art of prescribing, key aspects were found to be clinical relevance, applicable content, reality-based teaching, creative discussions, and effective structure. Considering these pieces of evidence in future educational efforts may help junior physicians to be confident about their professional responsibility to ascertain that the drug treatment of every patient is appropriate from an overall medical perspective, thereby contributing to rational use of medicines.

Future perspective

Previous research on medication appropriateness and drug-related problems is extensive. However, as illustrated in this thesis, there are methodological concerns that may have implications for the validity of the results. For instance, although there are studies providing data on reliability between assessors, these results are seldom highlighted or discussed as a methodological issue, not even when the inter-rater agreement is conspicuously low. Subjectivity appears to be a non-negligible factor and may illustrate that assessment of drug treatment is complex, reflecting that patients are individuals and need to be treated accordingly. Further, prescribing of medicines for a specific patient is a difficult task, beyond what may be captured by general quality indicators or criteria-based tools. Therefore, applying an overall medical perspective in assessments of drug treatment quality and adverse drug events may be a preferable approach for future studies to better reflect the complex reality and to provide patient-relevant results in interventional studies. With such an approach, the weighing of benefits against risks, which is essential and always present in every-day prescribing of medicines, would be highlighted. Individualized decisions based on solid pharmacotherapeutic knowledge and with in-depth experience from years of clinical work, form the basis of rational use of medicines. Combined with the ongoing academic efforts for a standardized core curriculum in clinical pharmacology and therapeutics in medical schools in Europe, including structured and extensive tests before graduation, future research could target continuing medical education, directed, for instance, towards junior physicians and being provided by clinicians "on the floor", perhaps with support from clinical pharmacologists. Making the right diagnosis and deciding on the best treatment, followed by monitoring of positive effects as well as adverse events, and taking the overall responsibility for the patient, are all core aspects of modern medicine. They are also tightly connected to the attending physician, with new decisions needed to be taken on an irregular basis. Therefore, an interesting approach of future studies could be to evaluate structured, long-term bedside education.

Acknowledgements

Jag skulle vilja rikta ett stort tack till ett flertal personer som på ett eller annat sätt bidragit till att denna avhandling har kunnat genomföras:

Min huvudhandledare, Susanna Wallerstedt, för att du har introducerat mig till och guidat mig genom forskningens spännande värld. Med en tillsynes outtröttlig energi har du lotsat mig längs forskningsfältets snåriga stig, där fallgroparna och möjligheterna har väntat bakom varje krök. Du har inspirerat mig med ett osannolikt flöde av idéer, och dina alltid lika snabba återkopplingar på mina utkast har varit värda sin vikt i guld (och jag har aldrig riktigt förstått hur du hinner med så mycket, på så kort tid). Tack för alla dina ansträngningar, ditt tålamod och för all intressant forskning som jag har fått bli en del av!

Min bihandledare, David Åberg, som har visat mig sitt bästa stöd inte bara i forskningen, utan även som handledare under ST-åren, som min mentor i schemaläggningen och som uppskattad medmänniska och kollega. Tack för ditt oavbrutna stöd sedan min första dag på medicinkliniken – det har varit värt väldigt mycket.

Karin Nylén, för ditt engagemang i projektet med läkemedelsgenomgångar, vilket i förlängningen ledde till arbetet med den här avhandlingen. Tack för hjälpen med seminarierna för AT-läkarna och för dina värdefulla synpunkter på såväl manuskript som poster till delarbete III.

Carina Tukukino och Naldy Parody Lopez i forskargruppen, för ert stöd och för era värdefulla kommentarer på ramberättelsen.

Mina medförfattare på delarbete II, likaledes kollegor i kliniken, Elisabeth Persson och Per Thysell, för er hjälp med bedömningar i studien och inmatning i databasen.

Mina kollegor på medicinkliniken, och framför allt mina vänner: Lina Holmqvist, Mazdak Tavoly och Andreas Hein – tack för att ni alltid finns där. Ett speciellt tack även till Bengt Persson, min kliniska förebild framför alla; jag frågar mig själv ofta: ”Vad hade Bengt gjort?”

Min japanska familj på andra sidan jordklotet - 遠く離れていても, 身近に感じる.

Mina föräldrar, för att ni alltid stöttar och uppmuntrar mig, i vått och torrt.

Slutligen, tack Shiho, Maika och Haku – ert stöd och er kärlek betyder allt för mig.

This thesis was supported by grants from The Swedish Research Council (521-2013-2639), the Healthcare Board, Region Västra Götaland (VGFOUREG-392071 and VGFOUREG-476621), and the Swedish state under the agreement between the Swedish government and the county councils – the ALF agreement (ALFGBG-428711 and ALFGBG-716941).

References

1. World Health Organization. The World Medicines Situation Report 2011 [Internet]. Geneva: World Health Organization; 2011. [cited 2021 April 30]. Available from: http://www.who.int/medicines/areas/policy/world_medicines_situation/en/

2. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA. 1998;279(15):1200-5.

3. Kongkaew C, Noyce PR, Ashcroft DM. Hospital admissions associated with adverse drug reactions: a systematic review of prospective observational studies. Ann Pharmacother. 2008;42(7):1017-25.

4. Nivya K, Sri Sai Kiran V, Ragoo N, Jayaprakash B, Sonal Sekhar M. Systemic review on drug related hospital admissions - A pubmed based search. Saudi Pharm J. 2015;23(1):1-8.

5. Hakkarainen KM, Hedna K, Petzold M, Hägg S. Percentage of patients with preventable adverse drug reactions and preventability of adverse drug reactions - a meta-analysis. PLoS One. 2012;7(3).

6. Patel NS, Patel TK, Patel PB, Naik VN, Tripathi C. Hospitalizations due to preventable adverse reactions—a systematic review. Eur J Clin Pharmacol. 2017;73(4):385-98.

7. Ayalew MB, Tegegn HG, Abdela OA. Drug Related Hospital Admissions; A Systematic Review of the Recent Literatures. Bull Emerg Trauma. 2019;7(4):339-46.

8. Brinkman DJ, Tichelaar J, Schutte T, Benemei S, Böttiger Y, Chamontin B, et al. Essential competencies in prescribing: A first European cross-sectional study among 895 final-year medical students. Clin Pharmacol Ther. 2017;101(2):281-9.

9. Brinkman DJ, Tichelaar J, Graaf S, Otten RHJ, Richir MC, van Agtmael MA. Do final-year medical students have sufficient prescribing competencies? A systematic literature review. Br J Clin Pharmacol. 2018;84(4):615-35.

10. Ryan C, Ross S, Davey P, Duncan EM, Francis JJ, Fielding S, et al. Prevalence and causes of prescribing errors: the PRescribing Outcomes for Trainee Doctors Engaged in Clinical Training (PROTECT) study. PLoS One. 2014;9(1):e79802.

11. Dornan T, Ashcroft D, Heathfield H, Lewis P, Miles J, Taylor D, et al. An in depth investigation into causes of prescribing errors by foundation trainees in relation to their medical education: EQUIP study. Final report or the GMC, December 2009. [cited 2021 April 30]. Available from: http://www.gmc-uk.org/FINAL_Report_prevalence_and_causes_of_prescribing_errors.pdf_28935150.pdf

12. Ashcroft DM, Lewis PJ, Tully MP, Farragher TM, Taylor D, Wass V, et al. Prevalence, Nature, Severity and Risk Factors for Prescribing Errors in Hospital Inpatients: Prospective Study in 20 UK Hospitals. Drug Saf. 2015;38(9):833-43.

13. Pacanowski M, Liu Q. Precision Medicine 2030. Clin Pharmacol Ther. 2020;107(1):62-4.

14. Cornelissen D, de Kunder S, Si L, Reginster JY, Evers S, Boonen A, et al. Interventions to improve adherence to anti-osteoporosis medications: an updated systematic review. Osteoporos Int. 2020;31(9):1645-69.

15. Sjölander M, Eriksson M, Glader EL. The association between patients' beliefs about medicines and adherence to drug treatment after stroke: a cross-sectional questionnaire survey. BMJ Open. 2013;3(9):e003551.

16. Holmqvist L, Boström KB, Kahan T, Schiöler L, Qvarnström M, Wettermark B, et al. Drug adherence in treatment resistant and in controlled hypertension-Results from the Swedish Primary Care Cardiovascular Database (SPCCD). Pharmacoepidemiol Drug Saf. 2018;27(3):315-21.

17. Brown MT, Bussell J, Dutta S, Davis K, Strong S, Mathew S. Medication Adherence: Truth and Consequences. Am J Med Sci. 2016;351(4):387-99.

18. Beuscart JB, Knol W, Cullinan S, Schneider C, Dalleur O, Boland B, et al. International core outcome set for clinical trials of medication review in multi-morbid older patients with polypharmacy. BMC Med. 2018;16(1):21.

19. Rankin A, Cadogan CA, In Ryan C, Clyne B, Smith SM, Hughes CM. Core outcome set for trials aimed at improving the appropriateness of polypharmacy in older people in primary care. J Am Geriatr Soc. 2018;66(6):1206-12.

20. Gutiérrez-Valencia M, Izquierdo M, Cesari M, Casas-Herrero Á, Inzitari M, Martínez-Velilla N. The relationship between frailty and polypharmacy in older people: A systematic review. Br J Clin Pharmacol. 2018;84(7):1432-44.

21. Masnoon N, Shakib S, Kalisch-Ellett L, Caughey GE. What is polypharmacy? A systematic review of definitions. BMC Geriatr. 2017;17(1):230.

22. Belfrage B, Koldestam A, Sjöberg C, Wallerstedt SM. Number of drugs in the medication list as an indicator of prescribing quality: a validation study of polypharmacy indicators in older hip fracture patients. Eur J Clin Pharmacol. 2015;71(3):363-8.

23. Brilleman SL, Salisbury C. Comparing measures of multimorbidity to predict outcomes in primary care: a cross sectional study. Fam Pract. 2013;30(2):172-8.

24. Rankin A, Cadogan CA, Patterson SM, Kerse N, Cardwell CR, Bradley MC, et al. Interventions to improve the appropriate use of polypharmacy for older people. Cochrane Database Syst Rev. 2018;9(9):Cd008165.

25. Pazan F, Kather J, Wehling M. A systematic review and novel classification of listing tools to improve medication in older people. Eur J Clin Pharmacol. 2019;75(5):619-25.

26. O'Connor MN, Gallagher P, O'Mahony D. Inappropriate prescribing: criteria, detection and prevention. Drugs & aging. 2012;29(6):437-52.

27. Hanlon JT, Schmader KE, Samsa GP, Weinberger M, Uttech KM, Lewis IK. A method for assessing drug therapy appropriateness. J Clin Epidemiol. 1992;45.

28. By the American Geriatrics Society 2015 Updated Beers Criteria for Potentially Inappropriate Medication Use in Older Adults. J Am Geriatr Soc. 2015;63(11):2227-46.

29. 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(2):72-83.

30. O'Mahony D, O'Sullivan D, Byrne S, O'Connor MN, Ryan C, Gallagher P. STOPP/START criteria for potentially inappropriate prescribing in older people: version 2. Age Ageing. 2015;44(2):213-8.

31. The Swedish National Board of Health and Welfare (Socialstyrelsen). Indikatorer för god läkemedelsbehandling hos äldre [Internet]. Stockholm: Socialstyrelsen; 2017. [cited 2021 April 30]. Available from: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/ovrigt/2017-6-7.pdf

32. Kara Ö, Arık G, Kızılarslanoglu MC, Kılıc MK, Varan HD, Sümer F, et al. Potentially inappropriate prescribing according to the STOPP/START criteria for older adults. Aging Clin Exp Res. 2016;28(4):761-8.

33. Ryan C, O'Mahony D, Kennedy J, Weedle P, Byrne S. Potentially inappropriate prescribing in an Irish elderly population in primary care. Br J Clin Pharmacol. 2009;68(6):936-47.

34. Onatade R, Auyeung V, Scutt G, Fernando J. Potentially inappropriate prescribing in patients on admission and discharge from an older peoples' unit of an acute UK hospital. Drugs & aging. 2013;30(9):729-37.

35. Delgado Silveira E, Alvarez Díaz A, Pérez Menéndez-Conde C, Muñoz García M, Cruz-Jentoft A, Bermejo Vicedo T. [Results of integrating pharmaceutical care in an Acute Geriatric Unit]. Rev Esp Geriatr Gerontol. 2012;47(2):49-54.

36. O'Sullivan D, O'Mahony D, O'Connor MN, Gallagher P, Cullinan S, O'Sullivan R, et al. The impact of a structured pharmacist intervention on the appropriateness of prescribing in older hospitalized patients. Drugs & aging. 2014;31(6):471-81.

37. Directive 2010/84/EU of the European parliament and of the council of 15 December 2010 amending, as regards pharmacovigilance, Directive 2001/83/EC

on the Community code relating to medicinal products for human use. (cited 2021 April 30). Available from: http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:348:0074:0099:EN:PDF

38. World Health Organization. International drug monitoring: the role of national centres. Technical Report Series No 498. Geneva: WHO, 1972. (cited 2021 April 30). Available from: https://www.who-umc.org/media/2680/who-technical-report-498.pdf

39. Aronson JK, Ferner RE. Clarification of terminology in drug safety. Drug Saf. 2005;28(10):851-70.

40. Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. BMJ. 2004;329(7456):15-9.

41. Formica D, Sultana J, Cutroneo PM, Lucchesi S, Angelica R, Crisafulli S, et al. The economic burden of preventable adverse drug reactions: a systematic review of observational studies. Expert Opin Drug Saf. 2018;17(7):681-95.

42. Bouvy JC, De Bruin ML, Koopmanschap MA. Epidemiology of adverse drug reactions in Europe: a review of recent observational studies. Drug Saf. 2015;38(5):437-53.

43. Falconer N, Barras M, Martin J, Cottrell N. Defining and classifying terminology for medication harm: a call for consensus. Eur J Clin Pharmacol. 2019;75(2):137-45.

44. WHO. Pharmacovigilance: Ensuring the Safe Use of Medicines. Geneva, WHO, October 2004. [cited 2021 April 30]. Available from: https://www.who.int/teams/regulation-prequalification/pharmacovigilance

45. The WHO-UMC system for standardized case causality assessment. World Health Organization (WHO) — Uppsala Monitoring Centre. [cited 2021 April 30]. Available from: https://www.who-umc.org/media/164200/who-umc-causality-assessment_new-logo.pdf

46. Naranjo CA, Busto U, Sellers EM, Sandor P, Ruiz I, Roberts EA, et al. A method for estimating the probability of adverse drug reactions. Clin Pharmacol Ther. 1981;30(2):239-45.

47. Théophile H, André M, Miremont-Salamé G, Arimone Y, Bégaud B. Comparison of three methods (an updated logistic probabilistic method, the Naranjo and Liverpool algorithms) for the evaluation of routine pharmacovigilance case reports using consensual expert judgement as reference. Drug Saf. 2013;36(10):1033-44.

48. Thaker SJ, Sinha RS, Gogtay NJ, Thatte UM. Evaluation of inter-rater agreement between three causality assessment methods used in pharmacovigilance. J Pharmacol Pharmacother. 2016;7(1):31-3.

49. Behera SK, Das S, Xavier AS, Velupula S, Sandhiya S. Comparison of different methods for causality assessment of adverse drug reactions. Int J Clin Pharm. 2018;40(4):903-10.

50. Beijer HJ, de Blaey CJ. Hospitalisations caused by adverse drug reactions (ADR): a meta-analysis of observational studies. Pharm World Sci. 2002;24(2):46-54.

51. Schurig AM, Böhme M, Just KS, Scholl C, Dormann H, Plank-Kiegele B, et al. Adverse Drug Reactions (ADR) and Emergencies. Dtsch Arztebl Int. 2018;115(15):251-8.

52. Pedros C, Formiga F, Corbella X, Arnau JM. Adverse drug reactions leading to urgent hospital admission in an elderly population: prevalence and main features. Eur J Clin Pharmacol. 2016;72(2):219-26.

53. Repp KL, Hayes C, 3rd, Woods TM, Allen KB, Kennedy K, Borkon MA. Drug-related problems and hospital admissions in cardiac transplant recipients. Ann Pharmacother. 2012;46(10):1299-307.

54. Chan A, Soh D, Ko Y, Huang YC, Chiang J. Characteristics of unplanned hospital admissions due to drug-related problems in cancer patients. Support Care Cancer. 2014;22(7):1875-81.

55. McHugh ML. Interrater reliability: the kappa statistic. Biochemia medica. 2012;22(3):276-82.

56. Hallgren KA. Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. Tutor Quant Methods Psychol. 2012;8(1):23-34.

57. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. J Clin Epidemiol. 1990;43(6):543-9.

58. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. Br J Math Stat Psychol. 2008;61(Pt 1):29-48.

59. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. BMC Med Res Methodol. 2013;13:61.

60. Hakkarainen KM, Andersson Sundell K, Petzold M, Hägg S. Methods for assessing the preventability of adverse drug events: a systematic review. Drug Saf. 2012;35(2):105-26.

61. Marcum ZA, Amuan ME, Hanlon JT, Aspinall SL, Handler SM, Ruby CM, et al. Prevalence of unplanned hospitalizations caused by adverse drug reactions in older veterans. J Am Geriatr Soc. 2012;60(1):34-41.

62. Woo SA, Cragg A, Wickham ME, Peddie D, Balka E, Scheuermeyer F, et al. Methods for evaluating adverse drug event preventability in emergency department patients. BMC Med Res Methodol. 2018;18(1):160.

63. Howard RL, Avery AJ, Howard PD, Partridge M. Investigation into the reasons for preventable drug related admissions to a medical admissions unit: observational study. Qual Saf Health Care. 2003;12(4):280-5.

64. Olivier P, Boulbés O, Tubery M, Lauque D, Montastruc JL, Lapeyre-Mestre M. Assessing the feasibility of using an adverse drug reaction preventability scale in clinical practice: a study in a French emergency department. Drug Saf. 2002;25(14):1035-44.

65. Ahern F, Sahm LJ, Lynch D, McCarthy S. Determining the frequency and preventability of adverse drug reaction-related admissions to an Irish University Hospital: a cross-sectional study. Emerg Med J. 2014;31(1):24-9.

66. Tangiisuran B, Auyeung V, Cheek L, Rajkumar C, Davies G. Inter-rater reliability of the assessment of adverse drug reactions in the hospitalised elderly. J Nutr Health Aging. 2013;17(8):700-5.

67. de Boer M, Kiewiet JJS, Boeker EB, Ramrattan MA, Dijkgraaf MGW, Lie-A-Huen L, et al. A targeted method for standardized assessment of adverse drug events in surgical patients. J Eval Clin Pract. 2013;19(6):1073-82.

68. Thevelin S, Spinewine A, Beuscart JB, Boland B, Marien S, Vaillant F, et al. Development of a standardized chart review method to identify drug-related hospital admissions in older people. Br J Clin Pharmacol. 2018;84(11):2600-14.

69. Baker GR, Norton PG, Flintoft V, Blais R, Brown A, Cox J, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. CMAJ. 2004;170(11):1678-86.

70. Harding S, Britten N, Bristow D. The performance of junior doctors in applying clinical pharmacology knowledge and prescribing skills to standardized clinical cases. Br J Clin Pharmacol. 2010;69(6):598-606.

71. Brinkman DJ, Tichelaar J, Okorie M, Bissell L, Christiaens T, Likic R, et al. Pharmacology and therapeutics education in the European Union needs harmonization and modernization: a cross-sectional survey among 185 medical schools in 27 countries. Clin Pharmacol Ther. 2017;102(5):815-22.

72. Maxwell SRJ, Coleman JJ, Bollington L, Taylor C, Webb DJ. Prescribing Safety Assessment 2016: Delivery of a national prescribing assessment to 7343 UK final-year medical students. Br J Clin Pharmacol. 2017;83(10):2249-58.

73. Brinkman DJ, Tichelaar J, Mokkink LB, Christiaens T, Likic R, Maciulaitis R, et al. Key Learning Outcomes for Clinical Pharmacology and Therapeutics Education in Europe: A Modified Delphi Study. Clin Pharmacol Ther. 2018;104(2):317-25.

74. Kramers C, Janssen BJ, Knol W, Hessel MHM, Mulder WM, Dumont G, et al. A Licence to Prescribe. Br J Clin Pharmacol. 2017;83(8):1860-1.

75. Eriksson AL, Wallerstedt SM. Developing confidence in basic prescribing skills during medical school: a longitudinal questionnaire study investigating the effects of a

modified clinical pharmacology course. Eur J Clin Pharmacol. 2018;74(10):1343-9.

76. Axelsson MA, Spetz M, Mellen A, Wallerstedt SM. Use of and attitudes towards the prescribing guidelines booklet in primary health care doctors. BMC Clin Pharmacol. 2008;8:8.

77. O'Connor MN, O'Sullivan D, Gallagher PF, Eustace J, Byrne S, O'Mahony D. Prevention of Hospital-Acquired Adverse Drug Reactions in Older People Using Screening Tool of Older Persons' Prescriptions and Screening Tool to Alert to Right Treatment Criteria: A Cluster Randomized Controlled Trial. J Am Geriatr Soc. 2016;64(8):1558-66.

78. Brinkman DJ, van Rossem AP, Tichelaar J, Richir MC, van Agtmael MA. Does Medical Students Knowing More About Drugs Lead to Better Treatment Choices? J Clin Pharmacol. 2017;57(8):1071-2.

79. Aronsson P, Booth S, Hagg S, Kjellgren K, Zetterqvist A, Tobin G, et al. The understanding of core pharmacological concepts among health care students in their final semester. BMC Med Educ. 2015;15:235.

80. Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, et al. Active learning increases student performance in science, engineering, and mathematics. Proc Natl Acad Sci USA. 2014;111(23):8410-5.

81. Gillen JR, Farineau DW, Lau CL, Williams MD, Marzani-Nissen GR. Observations: utilizing an interactive case-based format for intern orientation. J Grad Med Educ. 2015;7(2):301-2.

82. Forsetlund L, Bjorndal A, Rashidian A, Jamtvedt G, O'Brien MA, Wolf F, et al. Continuing education meetings and workshops: effects on professional practice and health care outcomes. Cochrane Database Syst Rev. 2009(2):Cd003030.

83. O'Brien MA, Rogers S, Jamtvedt G, Oxman AD, Odgaard-Jensen J, Kristoffersen DT, et al. Educational outreach visits: effects on professional practice and health care outcomes. Cochrane Database Syst Rev. 2007(4):Cd000409.

84. Gharbi M, Moore LS, Castro-Sanchez E, Spanoudaki E, Grady C, Holmes AH, et al. A needs assessment study for optimising prescribing practice in secondary care junior doctors: the Antibiotic Prescribing Education among Doctors (APED). BMC Infect Dis. 2016;16(1):456.

85. Sjoberg C, Wallerstedt SM. Effects of medication reviews performed by a physician on treatment with fracture-preventing and fall-risk-increasing drugs in older adults with hip fracture-a randomized controlled study. J Am Geriatr Soc. 2013;61(9):1464-72.

86. O'Mahony D, Gallagher P, Ryan C, Byrne S, Hamilton H, Barry P, et al. STOPP & START criteria: A new approach to detecting potentially inappropriate prescribing in old age. Eur Geriatr Med. 2010;1(1):45-51.

87. Hallas J, Harvald B, Gram LF, Grodum E, Brosen K, Haghfelt T, et al. Drug related hospital admissions: the role of definitions and intensity of data collection, and the possibility of prevention. J Intern Med. 1990;228(2):83-90.

88. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today. 2004;24(2):105-12.

89. Zec S, Soriani N, Comoretto R, Baldi I. High Agreement and High Prevalence: The Paradox of Cohen's Kappa. Open Nurs J. 2017;11:211-8.

90. Rotondi MA, Donner A. A confidence interval approach to sample size estimation for interobserver agreement studies with multiple raters and outcomes. J Clin Epidemiol. 2012;65(7):778-84.

91. Varmdal T, Ellekjær H, Fjærtoft H, Indredavik B, Lydersen S, Bonaa KH. Inter-rater reliability of a national acute stroke register. BMC Res Notes. 2015;8:584.

92. Sjoberg C, Bladh L, Klintberg L, Mellstrom D, Ohlsson C, Wallerstedt SM. Treatment with fall-risk-increasing and fracture-preventing drugs before and after a hip fracture: an observational study. Drugs & aging. 2010;27(8):653-61.

93. Kanis JA, Johnell O, De Laet C, Jonsson B, Oden A, Ogelsby AK. International variations in hip fracture probabilities: implications for risk assessment. J Bone Miner Res. 2002;17(7):1237-44.

94. Gallagher PF, O'Connor MN, O'Mahony D. Prevention of potentially inappropriate prescribing for elderly patients: a randomized controlled trial using STOPP/START criteria. Clin Pharmacol Ther. 2011;89(6):845-54.

95. Gallagher P, O'Mahony D. STOPP (Screening Tool of Older Persons' potentially inappropriate Prescriptions): application to acutely ill elderly patients and comparison with Beers' criteria. Age and Ageing. 2008;37(6):673-9.

96. 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(11):1013-9.

97. Brinkman DJ, Tichelaar J, van Agtmael MA, de Vries TP, Richir MC. Self-reported confidence in prescribing skills correlates poorly with assessed competence in fourth-year medical students. J Clin Pharmacol. 2015;55(7):825-30.

98. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.

99. Sullivan GM, Artino AR, Jr. Analyzing and interpreting data from likert-type scales. J Grad Med Educ. 2013;5(4):541-2.

100. Verdoorn S, Kwint HF, Faber A, Gussekloo J, Bouvy ML. Majority of drug-related problems identified during medication review are not associated with STOPP/START criteria. Eur J Clin Pharmacol. 2015;71(10):1255-62.

101. Lozano-Montoya I, Vélez-Diaz-Pallarés M, Delgado-Silveira E, Montero-Errasquin B, Cruz Jentoft AJ. Potentially inappropriate prescribing detected by STOPP–START criteria: are they really inappropriate? Age and Ageing. 2015;44(5):861-6.

102. Dalton K, Curtin D, O'Mahony D, Byrne S. Computer-generated STOPP/START recommendations for hospitalised older adults: evaluation of the relationship between clinical relevance and rate of implementation in the SENATOR trial. Age and Ageing. 2020;49(4):615-21.

103. O'Mahony D. STOPP/START criteria for potentially inappropriate medications/potential prescribing omissions in older people: origin and progress. Expert Rev Clin Phar. 2020;13(1):15-22.

104. Wallerstedt SM, Belfrage B, Fastbom J. Association between drug-specific indicators of prescribing quality and quality of drug treatment: a validation study. Pharmacoepidemiol Drug Saf. 2015;24(9):906-14.

105. Wallerstedt S, Belfrage B, Koldestam A, Sjöberg C. Many patients with suboptimal drug treatment are missed. Sensitivity, specificity and prediction value of drug-specific quality indicators for the elderly. Läkartidningen. 2016;113.

106. Millar AN, Daffu-O'Reilly A, Hughes CM, Alldred DP, Barton G, Bond CM, et al. Development of a core outcome set for effectiveness trials aimed at optimising prescribing in older adults in care homes. Trials. 2017;18(1):175.

107. O'Connor J, Adabavazeh B, Choi H, Khan A, Shah S, Shah S. Use of the STOPP and START criteria to address polypharmacy for elderly patients in University Hospital Lewisham Clinical Decisions Unit. Hong Kong J Emerg Me.0(0):1024907919849358.

108. Sjoberg C, Edward C, Fastbom J, Johnell K, Landahl S, Narbro K, et al. Association between multi-dose drug dispensing and quality of drug treatment--a register-based study. PLoS One. 2011;6(10):e26574.

109. Wallerstedt SM, Fastbom J, Johnell K, Sjöberg C, Landahl S, Sundström A. Drug treatment in older people before and after the transition to a multi-dose drug dispensing system--a longitudinal analysis. PLoS One. 2013;8(6):e67088.

110. Belfrage B, Koldestam A, Sjöberg C, Wallerstedt SM. Prevalence of suboptimal drug treatment in patients with and without multidose drug dispensing--a cross-sectional study. Eur J Clin Pharmacol. 2014;70(7):867-72.

111. Díaz-Gutiérrez MJ, Martínez-Cengotitabengoa M, Sáez de Adana E, Cano AI, Martínez-Cengotitabengoa MT, Besga A, et al. Relationship between the use of benzodiazepines and falls in older adults: A systematic review. Maturitas. 2017;101:17-22.

112. de Ruiter SC, Biesheuvel SS, van Haelst IMM, van Marum RJ, Jansen R. To STOPP or to START? Potentially inappropriate prescribing in older patients with falls and syncope. Maturitas. 2020;131:65-71.

113. Oscanoa TJ, Lizaraso F, Carvajal A. Hospital admissions due to adverse drug reactions in the elderly. A meta-analysis. Eur J Clin Pharmacol. 2017;73(6):759-70.

114. van Doormaal JE, Mol PG, van den Bemt PM, Zaal RJ, Egberts AC, Kosterink JG, et al. Reliability of the assessment of preventable adverse drug events in daily clinical practice. Pharmacoepidemiol Drug Saf. 2008;17(7):645-54.

115. Karuppannan M, Nee TK, Ali SM, Thong WK, Boardman H. The prevalence of adverse drug event-related admissions at a local hospital in Malaysia. Arch Pharma Pract. 2013;4:160-7.

116. Kempen TGH, Hedström M, Olsson H, Johansson A, Ottosson S, Al-Sammak Y, et al. Assessment tool for hospital admissions related to medications: development and validation in older patients. Int J Clin Pharm. 2019;41(1):198-206.

117. Bloom BS, Englehart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. The Taxonomy of educational objectives, handbook I: The Cognitive domain. New York: David McKay Co., Inc; 1956.

118. Coombes ID, Stowasser DA, Coombes JA, Mitchell C. Why do interns make prescribing errors? A qualitative study. Med J Aust. 2008;188(2):89-94.

119. Eriksson AL, Wallerstedt SM. [Läkarstudenter anser sig behöva utbildning i läkemedelsbehandling]. Article in Swedish. Läkartidningen. 2020;117.

120. Otokiti A, Sideeg A, Ward P, Dongol M, Osman M, Rahaman O, et al. A quality improvement intervention to enhance performance and perceived confidence of new internal medicine residents. J Community Hosp Intern Med Perspect. 2018;8(4):182-6.

121. Mickan S, Hilder J, Wenke R, Thomas R. The impact of a small-group educational intervention for allied health professionals to enhance evidence-based practice: mixed methods evaluation. BMC Med Educ. 2019;19(1):131.

122. Antonoff MB, Swanson JA, Acton RD, Chipman JG, Maddaus MA, Schmitz CC, et al. Improving surgery intern confidence through the implementation of expanded orientation sessions. Surgery. 2010;148(2):181-6.

123. Wright M, Mankey CG, Miller BW. Improving upon the 'July effect': a collaborative, interdisciplinary orientation for internal medicine interns. Med Educ Online. 2013;18:23249.

124. Gardner RL, Cooper E, Haskell J, Harris DA, Poplau S, Kroth PJ, et al. Physician stress and burnout: the impact of health information technology. J Am Med Inform Assoc. 2019;26(2):106-14.

125. Carlsson T, Winder M, Eriksson AL, Wallerstedt SM. Student Characteristics Associated with Passing the Exam in Undergraduate Pharmacology Courses—a Cross-sectional Study in Six University Degree Programs. Medical Science Educator. 2020;30(3):1137-44.

Appendices

Appendix A: Description of patient cases from seminar 1 (Study III)

The overall structure of the seminar follows a *basic* medication review as defined by the Swedish National Board of Health and Welfare (Socialstyrelsen):

- **Reconciliation** (what drugs are the patient taking and for which diseases)
    - Different sources of information: patient/relatives, medical record, multi-dose drug dispensing system, Register of dispensed drugs, pharmacy register of current prescriptions available for dispensing, patient's GP
- **Plausibility assessment**
    - Is the medical treatment reasonable given the patient's medical history?
    - Is it reasonable given the current medical situation?
    - Take into account possible adverse effects, drug-drug interactions, double medications, kidney function *et cetera*, and suggest adjustments (initiate/withdraw medications, adjust doses)
- **Documentation and information**
    - Document which changes have been made and why.
    - Write medication discharge summaries and hand over to each patient at discharge, together with an updated medication list. Send to the next healthcare provider to facilitate transitions in care.

Case 1. A 50-year-old man with hypertension, type II diabetes mellitus, and chronic back pain has been admitted to an emergency ward because of high fever, chills, and urinary tract symptoms starting two days earlier. He has been given ciprofloxacin for suspected pyelonephritis. Apart from the antibiotics, he has also been prescribed the following medications which he has reported at admission that he usually takes:

Felodipine 5 mg a day
Furosemide 40 mg a day
Diclofenac 50 mg twice a day

Teaching points:

- Recommended treatment for pyelonephritis, type II diabetes mellitus, hypertension
- Using different sources of medication reconciliation, for example the Register of dispensed drugs, to find medications the patient has not reported (in this case metformin and losartan)
- How to do with certain medications in case of acute illness (antihypertensive drugs, RAAS-blockade, COX inhibitors, and metformin) and why (hypotension, renal failure, lactate acidosis)
- Risks with prolonged use of COX inhibitors for vulnerable patients (peptic ulcer, cardiovascular risk, renal failure, negative effects of antihypertensive treatment)
- Important drug-drug interactions; how to find out and how to handle (losartan/diclofenac, furosemide/diclofenac)

- How to calculate eGFR and general discussion on dose adjustment with decreasing kidney function
- General tips and tricks on how to use the medical record system efficiently (prescribing, decision support regarding interactions and calculation of eGFR)

Case 2. A 78-year-old woman with hypertension, atrial fibrillation, and osteoporosis presents at the emergency department with severe and progressive fatigue, dizziness, and several falls over the last weeks. She hands over her "medication list", which is a copy of the pharmacy register of current prescriptions available for dispensing (which is given to the patient at the pharmacy when new medications are dispensed). She reports taking the medications as described in this list:

Warfarin 2.5 mg according to regular follow-up at the anticoagulant clinic

Calcichew-D3 twice a day	Aspirin 160 mg a day
Seloken ZOC 50 mg a day	Metoprolol 200 mg a day
Amlodipine 5 mg a day	Alendronate 70 mg a week
Zolpidem 10 mg in the evening	Oxazepam 5 mg when needed

Teaching points:

- Recommended treatment for non-valvular atrial fibrillation (rhythm control and stroke prevention including new oral anticoagulants), osteoporosis, sleeping disorder (in particular for older patients)
- Pitfalls with using the pharmacy register of drugs available for dispensing (they can include duplicates, in particular when treatment is changed, and does not include medications that cannot be dispensed another time without renewal from the physician)
- Importance of written information of drug treatment changes to patients

Case 3. A 95-year-old woman, living alone at home, with hypertension, glaucoma, long term leg pain, dizziness, and a tendency of hyponatremia, is admitted because of worsened dizziness and fatigue. She has normal vital signs, including a blood pressure of 160/80 and normal CBC, electrolytes, creatinine, and thyroid hormones. She did not bring a medication list and cannot recall all her medications, but says she takes "a lot of pills". Using the Register of dispensed drugs reveals that she uses the following medications:

Omeprazole 20 mg a day	Low dose aspirin once a day
Oral cyanocobalamin 1 mg a day	Folic acid 5 mg twice a day
Furosemide 25 mg twice a day	Felodipine 5 mg a day
Acetaminophen (modified release) 665 mg 2 pills three times a day	Diazepam 10 mg four times a day
Zolpidem 10 mg 2 in the evening	Citalopram 10 mg a day
Mirtazapine 45 mg a day	Alprazolam 0.5 mg when needed
Brinzolamide/timolol 10mg/ml+5mg/ml twice a day in right eye	

Teaching points:

- This authentic case illustrates the complex reality where several physicians, in various healthcare centres without a unified drug order system, prescribe drugs not knowing about one another's prescriptions.
- The importance of reconsidering the indication of medications (in this case aspirin, vitamin B12, folic acid, loop diuretics, acetaminophen, SSRI/NaSSA, alprazolam, zolpidem, high dose diazepam)
- Attention to drugs with increased risk for the elderly, for example falls, cognitive deterioration (benzodiazepines, sleeping pills, opiates, glucocorticoids, anticholinergic drugs)
- Using correct doses of drugs (folic acid, acetaminophen, diazepam, zolpidem, antidepressants)
- Strategies for withdrawing drugs gradually (benzodiazepines)
- Choice of drugs for anxiety and sleeping disorder in the elderly
- How to check and manage interactions (in this case ASA/SSRI, omeprazole/SSRI)
- Arrhythmias with hypokalemia (loop-diuretics, drugs causing QT prolongation)

Case 4. An 84-year-old woman with atrial fibrillation, hypertension, and hypothyroidism is brought to the emergency department by her family because of nausea, fatigue, bad appetite, and sporadic vomiting over the last 2 months, starting after an episode with fever and vomiting. Her medications are as follows:

Warfarin 2.5 mg according to regular follow-up at the anticoagulant clinic
Furosemide 40 mg a day (not used for 4 weeks)
Metoprolol 100 mg twice a day
Felodipine 5 mg a day (not used for 4 weeks)
Valsartan-hydrochlorotiazide 160/12.5 mg a day
Levothyroxine 100 µg a day
Digoxin 130 microgram a day

Teaching points:

- The participants are to decide what information and examinations are important to correctly diagnose and treat the patient (hypotension, bradycardia, EGG with signs of digoxin intoxication, clinical signs of dehydration, renal failure, electrolyte balance, digoxin concentration, thyroid hormones *et cetera*)
- Pharmacodynamics and pharmacokinetics of common drugs (warfarin, loop and thiazide diuretics, ARBs/ACE-inhibitors, COX inhibitors, digoxin) and relation to kidney function
- Practicing plausibility assessment, documentation of medication changes and written information to patient

Appendix B: Questionnaire used in Study III

Appendix B Questionnaire

Approach to drug treatment management **Page 1: Start here**

Please mark to what extent you agree with the following statements:
(on a scale from "Strongly disagree" to "Strongly agree")

	Strongly disagree				**Strongly agree**
1. During hospitalisation, I usually reconcile the medications ordered for the patient	☐	☐	☐	☐	☐
2. During hospitalisation, I usually reconcile the medications actually used by the patient	☐	☐	☐	☐	☐
3. When I handle a patient, I usually assess whether the drug treatment, at the overall level, is reasonable taking the current situation into account	☐	☐	☐	☐	☐
4. When I handle a patient, I always consider adverse effects as a possible cause of symptoms	☐	☐	☐	☐	☐
5. I usually assess solely the medications related to the medical situation causing hospitalisation	☐	☐	☐	☐	☐
6. When I extend the order of a prescribed medication, I usually reflect on preventing double medications (e.g. two different acetaminophen drugs)	☐	☐	☐	☐	☐
7. When I extend the order of a prescribed medication, I usually reflect on whether the dose is plausible	☐	☐	☐	☐	☐
8.When I extend the order of a prescribed medication, I usually reflect on the importance of kidney function for the specific drug	☐	☐	☐	☐	☐
9. When I extend the order of a prescribed medication, I usually reflect on potential interactions	☐	☐	☐	☐	☐
10. When I add a new medication, I usually reflect on the aspects in questions 6 to 9	☐	☐	☐	☐	☐

11. Do you know what a *basic* medication review is?

☐ Yes, absolutely ☐ Yes, approximately ☐ No (go to question 13)

12. Over the last two weeks of clinical practice, I performed…… *basic* medication reviews
(Enter number)

13. Do you know what an *expanded* medication review is?

☐ Yes, absolutely ☐ Yes, approximately ☐ No (go to question 15)

14. Over the last 6 months of clinical practice, I referred patients to the primary health care or other physician practice for an expanded medication review …… no. of times.
(Estimated number)

15. I estimate that I use the computerized decision support system embedded in the medical records regarding interactions

☐ every day ☐ every week ☐ every month ☐ more seldom ☐ never

16. I use the following sources to find information about a patient's prescribed medications (Mark all boxes that apply. All potential sources are not presented)

☐ Register of dispensed drugs ☐ Multi-dose drug dispensing system ☐ Medical records ☐ Pharmacy register of current prescriptions available for dispensing

Perceived skills and confidence performing medication reviews

Page 2

Please mark to what extent you agree with the following statements: *(on a scale from "Strongly disagree" to "Strongly agree")*	**Strongly disagree**				**Strongly agree**
17. I know the components of a *basic* medication review	☐	☐	☐	☐	☐
18. I feel confident performing a *basic* medication review	☐	☐	☐	☐	☐
19. I know the components of an *expanded* medication review	☐	☐	☐	☐	☐
20. I feel confident performing an *expanded* medication review	☐	☐	☐	☐	☐
21. I feel confident about adding a new medication when this is required by the health condition of the patients	☐	☐	☐	☐	☐
22. I feel confident about withdrawing a medication	☐	☐	☐	☐	☐
23 I feel confident taking responsibility for the medication list at discharge	☐	☐	☐	☐	☐

Respondent characteristics

24. Before starting my internship, I worked as a physician assistant for ……months
(Enter number of months)

25. I have now completed ……..months of my internship (do not count research or parental leave) (Enter number of months)

26. I am:
- ☐ not involved in research
- ☐ not registered as Ph.D. student but involved in research
- ☐ a Ph.D. student
- ☐ a Ph.D.
- ☐ an associate professor

27. After medical school, I have completed the following courses related to drug treatment management:
- ☐ The Internet based education from the Swedish National Board of Health and Welfare "Drug treatment of the elderly"
- ☐ Other: ……..…. (enter name / content of the course)

28. Name: …….

Comments……..…

……

……

Thank you!