

# Computerized decision support improves medication review effectiveness: an experiment evaluating the STRIP Assistant's usability

Michiel Meulendijk (corresponding author)<sup>1</sup>, Marco Spruit<sup>1,4</sup>, Clara Drenth-van Maanen<sup>2</sup>, Mattijs Numans<sup>3</sup>, Sjaak Brinkkemper<sup>1</sup>, Paul Jansen<sup>2</sup>, Wilma Knol<sup>2</sup>

E-Mail (corresponding author): [m.c.meulendijk@uu.nl](mailto:m.c.meulendijk@uu.nl), tel. (corresponding author): +31 (30) 253 5433, fax. (corresponding author): +31 (30) 253 4619

<sup>1</sup>Department of Information and Computing Sciences, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands.

<sup>2</sup>Department of Geriatric Medicine and Expertise Centre Pharmacotherapy in Old Persons, University Medical Center, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands.

<sup>3</sup>Department of Public Health and Primary Care, LUMC, Albinusdreef 2, 2333 ZA Leiden, the Netherlands.

<sup>4</sup>Spru IT B.V., Meidoornlaan 22, 1231 CB Loosdrecht, The Netherlands.

Short title: Computerized decision support improves medication review effectiveness

## ABSTRACT

**Background:** Polypharmacy poses threats to patients' health. The Systematic Tool to Reduce Inappropriate Prescribing (STRIP) is a drug optimization process in primary care for conducting medication reviews. To effectively and efficiently incorporate this method into daily practice, the STRIP Assistant has been developed, a decision support system that aims at assisting physicians with the pharmacotherapeutic analysis of patients' medical records. It generates context-specific advice based on clinical guidelines.

**Objective:** The aim of this study was to validate the STRIP Assistant's usability as a tool for physicians optimizing medical records of polypharmacy patients.

**Methods:** In an online experiment, 42 physicians were asked to optimize two comparable medical records of polypharmacy patients, one in their usual manner and one using the STRIP Assistant. Changes in effectiveness were measured by comparing respondents' optimized medicine prescriptions with medication prepared by an expert panel of two geriatrician-pharmacologists. Efficiency was operationalized by recording the time

respondents took to optimize the two cases. Satisfaction was measured with the System Usability Scale (SUS). Independent and paired t-tests were used for analysis.

**Results:** Medication optimization significantly improved with the STRIP Assistant. Appropriate decisions increased from 58% without the STRIP Assistant to 76% with it;  $p < 0.0001$ . Inappropriate decisions decreased from 42% without the STRIP Assistant to 24% with it;  $p < 0.0001$ . Participants spent significantly more time optimizing medication with the STRIP Assistant (24 minutes) than without it (13 minutes);  $p < 0.0001$ . They filled out a below-average SUS-score of 63.25.

**Conclusion:** The STRIP Assistant improves the effectiveness of medication reviews with polypharmacy patients.

## KEY POINTS

- Clinical decision support systems significantly improve the number of appropriate decisions made in medication reviews and decrease the number of inappropriate choices.
- Users spend significantly more time optimizing health records with (unfamiliar) clinical decision support systems than without any digital assistance.
- This study confirms the results of previous studies that structured methods for medication review significantly improve medication appropriateness of prescriptions.

# 1 Background

## 1.1 Polypharmacy & Inappropriate Prescribing

Polypharmacy, or the chronic use of multiple medicines, poses significant threats to patients' health. A consensual definition of polypharmacy is lacking, but it is often described as the concurrent use of five or more different chronically used drugs [1]. Polypharmacy has been associated with negative health consequences. Drugs may cause clinical interactions or adverse effects that may aggravate patients' symptoms instead of relieving them. Medicine issues including underprescribing, overtreatment, and decreased drug adherence have been associated with polypharmacy [2,3,4,5,6,7,8,9]. A 2008 study showed that in the Netherlands, 5.6% of all acute hospital admissions have medication-related causes [10]. For elderly patients, who constitute half of all chronically ill polypharmacy patients, this figure was twice as high [11].

The concurrent use of multiple medications is not entirely undesirable, as in many patient cases polypharmacy is indicated or even unavoidable. However, inappropriate prescribing of medications is prevalent amongst elderly patients [12]. An incidence-focused study found that inappropriate medication use increased elderly persons' risks of hospitalization and mortality [13]. Geriatric assessment and medication review have been shown to be effective methods in aiding prescribers with optimizing polypharmacy [14,15].

A multitude of initiatives has been developed to assess the appropriateness of drugs prescribed for individual patients. These approaches can be divided in implicit and explicit methods. The former implicit methods use patient-specific information, combined with medical knowledge, to determine medication appropriateness, while the latter explicit methods provide screening tools, containing lists of clinical interactions or contraindications [1]. Amongst the explicit methods are the Beers criteria and the Screening Tool to Alert to Right Treatment (START) and Screening Tool of Older People's Prescriptions (STOPP) criteria, while the implicit methods include the Medication Appropriateness Index and the pharmacotherapy review focused on drugs' use, indication, safety, and effectiveness (GIVE) [16,17,18,19]. The effectiveness of these interventions varies; generally they appear beneficial in terms of reducing inappropriate prescribing and medication-related problems, but they have not been proven to lead to clinically significant improvement [20].

In order to improve medication prescribing in primary care, several implicit and explicit methods have been combined into an all-encompassing systematic medication review approach, the Polypharmacy Optimization Method (POM). It has been shown to significantly improve GPs' prescriptions for polypharmacy patients in an experimental setting [14].

A variety of barriers is impeding the widespread adoption of structured medication reviews in daily practice. Recently Anderson et al. conducted a systematic literature review on enablers and barriers to minimizing potentially inappropriate medications by GPs. Most factors revolved around physicians, and included inertia (his or

her attitudes towards discontinuation, such as fearing negative consequences), self-efficacy (his or her knowledge and available information on the topic), and awareness (his or her having poor insight or discrepant beliefs). Barriers that were not physician-related included a lack of resources, patients resisting changes to their medication, and practical and cultural factors [21]. A separate study focusing on barriers regarding pharmacist-led medication reviews reported lack of time and lack of self-confidence as the most commonly perceived barriers [22].

### 1.1.1 STRIP

**Figure 1: The five steps of the STRIP method depicted as a yearly repeating cycle.**

Recently, the POM, GIVE, and START and STOPP criteria have been combined into the Systematic Tool to Reduce Inappropriate Prescribing (STRIP), which has consequently been included as part of a Dutch multidisciplinary guideline on polypharmacy in elderly patients [23]. The STRIP has been designed to be an all-encompassing drug optimization process in primary care, focusing not just on the pharmacotherapeutic analysis but also on patients' medication histories and preferences; Figure 1 shows the STRIP method's different steps.

The STRIP analysis is more extensive than its predecessors [14,19,17]. It combines both the implicit approaches of the POM and the GIVE, and the explicit lists of the first version of the START and STOPP criteria. The pharmacotherapeutic analysis in the STRIP includes checks on underprescribing, overtreatment, recommended dosage adjustments, drug effectiveness, potential adverse effects, dose frequency, clinical interactions, and medication adherence, including practical problems with medication use. The START and STOPP criteria are implemented in the pharmacotherapeutic analysis. This extensive medication review results in a patient-specific treatment plan in which new drugs are gradually added and superfluous ones discontinued. This approach to conducting structured medication reviews is based on consensus rather than evidence, synthesizing the results of the earlier optimization methods mentioned above. Currently, solid evidence for choosing specific strategies for the optimization of pharmacotherapy in the elderly over others is lacking [24].

Involvement of patients in the medication review is emphasized to guarantee their therapy adherence; patients' preferences are taken into account as much as possible. The pursuit of the treatment plan is monitored through regular communication between practitioner, pharmacist, and patient. The involvement of pharmacists in medication reviews, as part of multidisciplinary teams, has been shown to lead to improved pharmacotherapy for

older patients [25]. Educating patients on their medication use and treatment goals, simplifying their drugs regimens, and preventing adverse drug reactions have all been identified as factors influencing patients' adherence to their treatments [24].

## **1.2 Clinical Decision Support Systems**

In recent years, computerized physician order entry (CPOE) systems have gradually changed in terms of functionality. From systems that were traditionally organizational in nature, they have been enhanced to facilitate management of electronic medical records and clinical decision support [26]. There is consensus in literature that clinical decision support has the potential to improve GPs' and pharmacists' decision-making [27]: "Both commercially and locally developed CDSSs are effective at improving health care process measures across diverse settings". Evidence for concurrent improvement in efficiency, cost-effectiveness, or clinical effectiveness is inadequate or ambiguous. A study investigating the attitudes of Dutch GPs to the introduction of a decision support system specifically aiding them with conducting medication reviews revealed that the majority were positively inclined to using such a system [28].

### **1.2.1 STRIP Assistant**

In order to enable GPs and pharmacists to effectively and efficiently incorporate the STRIP method into their daily practice, the STRIP Assistant has been developed. The STRIP Assistant has been designed as a stand-alone web application that aims at assisting GPs and pharmacists with the pharmacotherapeutic analysis of patients' medical records. Based on patients' records and the decisions that GPs and pharmacists make during the medication review, the application generates context-specific advice. The STRIP Assistant's design decisions adhere to best practices in information science research; user interface conception and decision rule implementation have been designed to balance efficiency and information completeness, aiming to minimize previously mentioned barriers such as users' lack of confidence and lack of time [29].

The knowledge used to generate the STRIP Assistant's advice consists of well-established guidelines on clinical interactions, double-medication, contra-indications, dosage strength and frequency, and specific implementations of version 1 of the START- and STOPP-criteria [30,31]. The rules incorporate not only patients' diseases and drugs, but also their contra-indications, complaints, and relevant physical properties (such as renal functioning, and weight). This results in items of advice that recommend users to add new drugs or to remove superfluous ones, or to change dosages of existing medicines.

In the future, the STRIP Assistant has been planned to integrate with existing CPOE systems, thereby increasing the efficiency with which the method can be performed. Additionally, use of datamining techniques on historical data should reveal patterns in users' behavior towards generated advice, which could be used to improve recommendations [32].

A video demonstrating the use of the STRIP Assistant can be viewed here:

<http://videodemo.stripa.eu/english/> [33].

### **1.3 Usability**

Usability has since long been regarded as an essential factor for the success of software applications. In the widely used definition by the International Organization for Standardization (ISO), usability is defined as ‘the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use’ [34]. In this context, effectiveness is understood as the ‘accuracy and completeness with which users achieve specified goals’, while efficiency exists of ‘resources expended in relation to the accuracy and completeness with which users achieve goals’. Satisfaction, finally, is the subjective ‘degree to which user needs are satisfied when a product or system is used in a specified context of use’.

A recent systematic literature review on clinical decision support systems showed that there is ample evidence that these systems can improve effectiveness; not enough research on efficiency and satisfaction is available to make generalizations regarding these aspects [27]. In technology adoption literature, it has been shown that systems’ perceived usefulness and ease-of-use, aspects closely related to usability, are the major determinants of people’s attitudes towards using technology [35,36,37].

Hornbaek described the current practices in evaluating usability [38]. A multitude of metrics and instruments has been used to measure the three main factors of usability **identified in the ISO-definition**. Measurements of effectiveness usually involve the degree to which a task has been successfully completed, leading to metrics such as accuracy, recall and completeness. Efficiency metrics mostly revolve around the time spent completing a task, but can also involve mental efforts. The subjective satisfaction criterion is often measured through standardized questionnaires or interface ranking.

## **2 Objectives**

The above-mentioned considerations lead us to believe clinical decision support has the potential to successfully aid GPs and pharmacists in incorporating structured medication reviews in daily practice. Therefore, we aim to validate the STRIP Assistant instrument’s usability as a tool for physicians optimizing medical records of polypharmacy patients in an experimental setting. This main research question has been divided in the following sub questions:

- Q1. Do GPs and pharmacists make significantly more appropriate decisions when optimizing the medical records of polypharmacy patients with the STRIP Assistant than without it?
- Q2. Do GPs and pharmacists make significantly fewer inappropriate decisions when optimizing the medical records of polypharmacy patients with the STRIP Assistant than without it?

Q3. Do GPs and pharmacists take significantly less time to optimize the medical records of polypharmacy patients with the STRIP Assistant than without it?

Q4. Do GPs and pharmacists perceive using the STRIP Assistant for the optimization of medical records of polypharmacy patients as satisfactory?

In this context *appropriate decisions* means decisions that correspond to those agreed upon by an expert panel.

### 3 Methods

In order to explore to what degree the STRIP Assistant is usable for aiding GPs and pharmacists with performing medication reviews, an experiment was conducted.

#### 3.1 Participants

The experiment was aimed at GPs and pharmacists. 52 respondents were selected through opportunity sampling, as the researchers lacked the resources to guarantee participants' cooperation through reimbursement. All participants were required to be either general practitioners or pharmacists in Dutch primary care, and had to fully complete both parts of the experiment to warrant inclusion. Of the 52 respondents, 9 had to be discarded because of corruptions in the data: 3 participants did not fill out the unassisted first part of the experiment, 5 did not assign drugs to diseases or did not respond to advice during the assisted part, and 1 record was a duplicate. Finally, 43 participants' results were eligible for inclusion in the data analysis.

Respondents were recruited through the researchers' personal networks (i.e. symposia, conferences, and (training) conventions). They were briefly informed about the experiment's goal and ensured that their anonymity would be guaranteed. As an incentive, respondents were offered three months of use of the software application with their own patients free of charge.

#### 3.2 Study Design

The experiment took the form of a pre-experiment with a one-group pre-test post-test design, as described by 't Hart et al. [39]. Respondents were placed in a single research group; an initial test was performed, after which a stimulus was applied and the test was repeated.

In the test, two medical records of polypharmacy patients, having been earlier selected from the geriatric ward of an academic medical center for the study by Drenth-van Maanen et al., were used [14]; they were actualized (i.e. drugs that were no longer available were replaced by their contemporary counterparts) and confirmed to be of comparable difficulty by an expert panel of geriatricians specialized in clinical pharmacology (P.J. and W.K.) During the experiment, respondents were asked to optimize the first case in their usual manner and the second one using the STRIP Assistant.

The three usability aspects of effectiveness, efficiency and satisfaction were operationalized in the experiment as follows: effectiveness was measured by recording the respondents' medicine prescriptions, after their optimization. The decisions made by the respondents were then compared with the medication list that the aforementioned expert panel of two geriatrician-pharmacologists prepared. They reached consensus on the pharmacotherapeutic changes that should be made in the records optimized by the respondents, and classified the decisions as correct, neutral or potentially harmful. Efficiency was operationalized by recording the time respondents took to optimize the two cases. Finally, satisfaction was measured through a standardized questionnaire, the System Usability Scale, consisting of ten statements with which respondents had to indicate their (dis)agreement on a Likert scale.

All data was gathered between November 2013 and June 2014. During this time, no changes of any kind were made to the software.

### **3.3 Outcome Measures**

The main outcome measure was the difference in the percentage of appropriate decisions made by the participants without and with use of the STRIP Assistant. Secondary outcome measures were the difference in the number of inappropriate decisions taken by participants without and with use of the STRIP Assistant, the difference in time needed to perform the medication review without and with use of the STRIP Assistant, and the extent to which participants experienced their use of the STRIP Assistant as satisfactory.

### **3.4 Instrument**

The STRIP Assistant has been designed as a stand-alone web application that aims at assisting GPs and pharmacists with the pharmacotherapeutic analysis of patients' medical records. The user interface accommodates the six phases of the STRIP medication review (i.e. drugs-disease assignment, undertreatment, overtreatment, side effects-drugs assignment, clinical interactions, and dosage frequency). In most phases, users are shown advice on missing, superfluous, or incompatible drugs. The items of advice are patient-specific, incorporating their diseases, drugs, side effects, and users' actions up to that point. The STRIP Assistant's rule base consists of a combination of well-established clinical rule databases and specific implementations of the START and STOPP criteria.

For the experiment, the user interface was enhanced to first display one of the patient cases in a bulleted list, summing up his/her diseases, drugs, side effects, complaints, measurements, and lab values.

### **3.5 Procedure**

Respondents were asked to optimize the first case in their usual manner, specifying in an adjacent text field which drugs should be added or removed for optimal treatment. They were then shown a 1.5 minute video explaining the use of the STRIP Assistant, after which they were presented with the second patient case in the STRIP Assistant user interface. Respondents were asked to optimize this case through the STRIP process, reacting to the advice

generated by the application. Each screen contained a help button explaining what was expected of the respondents.

After optimizing the second case, respondents were presented with the SUS, consisting of ten statements with which they had to indicate their (dis)agreement on a Likert scale. Finally, respondents' demographics (age and sex) were collected, alongside their experience with medication reviews and CPOE systems. In a text field, respondents could optionally leave their comments.

### 3.6 Statistical Analysis

In all cases, an expert panel determined the correctness of the decisions made by the participants. Slight corrections to the data had to be made to account for the differences in the potential number of appropriate decisions respondents could make in each case: respectively 17 in the unassisted case and 20 in the assisted one. Similar corrections were applied to account for differences in the possible number of inappropriate decisions: respectively 30 in the unassisted case and 40 in the assisted one. Paired t-tests were used to analyze the data pertaining to appropriateness and inappropriateness of decisions, and the differences in time spent.

The results of the SUS were formatted in the manner described by Brooke [40]: the values of the odd questions were subtracted by 1 to get the corrected score; the values of the even questions were in turn subtracted from 5 to get their corrected score. The sum of all questions was multiplied by 2.5 to calculate the final score ranging from 0 to 100.

## 4 Results

### 4.1 Descriptive Statistics

**Table 1: Overview of participants' characteristics.**

		Frequency	Percentage
Sex	Male	12	27.9
	Female	27	62.8
	No data	4	9.3
Age	<= 30	5	11.6
	31-40	7	16.3
	41-50	8	18.7
	51-60	14	32.5
	=> 61	5	11.6
	No data	4	9.3
Function	GP	31	72.1

	Pharmacist	4	9.3
	Dispensing GP	2	4.7
	GP in training	2	4.7
	No data	4	9.3
Experience with medication reviews	STRIP	12	27.9
	Other medication review method	18	41.9
	None	9	20.9
	No data	4	9.3

Of the 43 respondents whose answers were valid, all but 4 filled out the questions pertaining to their personal characteristics. The majority of these were female (62.8%). Most respondents were in their fifties (32.5%) or forties (18.7%). 7 participants (16.3%) were aged between 31 and 40, and 5 (11.6%) were in their twenties. 5 (11.6%) were over 60 years of age. Most were either general practitioners (72.1%) or pharmacists (9.3%). 2 were dispensing GPs (4.7%), and 2 were GPs in training (4.7%). Most were experienced with performing medication reviews: 18 participants (41.9%) did not use STRIP for their reviews, while 12 (27.9%) did. 9 (20.9%) had no experience performing medication reviews at all.

#### 4.2 Usability Hypotheses

In total, 86 medication reviews were performed by the participants; in half of these cases they used their usual care methods to perform the optimization, in the other half they were aided by the software application. An overview of all tested hypotheses is shown in Table 2. On average, the participants prescribed 8 drugs for the unassisted case, and 14 for the assisted one.

A paired t-test showed a statistical difference in the appropriateness of the decisions made without the STRIP Assistant (M = 11.44, SD = 2.63) and with the STRIP Assistant (mean (M) = 15.26, standard deviation (SD) = 2.05);  $t(42) = 8.80$ ,  $p < 0.0001$ . A Wilcoxon signed-rank test showed similar results;  $Z = -5.40$ ,  $p < 0.0001$ . From a total of 418 unassisted correct decisions and 656 aided ones, on a decision total of respectively 720 and 866, follows that the proportion of appropriate decisions increased from 58% without help to 76% with the STRIP Assistant.

**Table 2: Overview of the tested hypotheses and their statistical outcomes.**

	Usual care	STRIP Assistant	Statistics
The STRIP Assistant positively influences the number of <i>appropriate decisions</i> made in a	418 (58%; M=11.44;	656 (76%; M=15.26;	Paired t-test: $t(42)=8.80$ , $p<0.0001$

medication review: accepted.	SD=2.63)	SD=2.05)	
The STRIP Assistant negatively influences the number of <i>inappropriate decisions</i> made in a medication review: accepted.	302 (42%; M=9.36; SD=2.53)	210 (24%; M=4.88; SD=2.23)	Paired t-test: t(42)=8.93, p<0.0001
The STRIP Assistant negatively influences the <i>time</i> taken to perform a medication review: rejected.	13 minutes (M=0.94; SD=0.40)	24 minutes (M=1.34; SD=0.20)	Paired t-test: t(42)=7.07, p<0.0001
Users perceive using the STRIP Assistant as <i>satisfactory</i> : rejected.		63.25 (SUS-score)	Quality consensus test: 63.25 < 70

A paired t-test showed a statistical difference in the inappropriateness of the decisions made without the STRIP Assistant (M = 9.36, SD = 2.53) and with the STRIP Assistant (M = 4.88, SD = 2.23);  $t(42) = 8.93$ ,  $p < 0.0001$ . The percentage of inappropriate decisions decreased from 42% in the unassisted case to 24% in the assisted one.

On average, participants took 13 minutes to complete the unassisted part of the experiment, and 24 minutes to complete the assisted medication review. A paired t-test of the base 10 logarithm of these values showed a statistical difference in the time taken without the STRIP Assistant (M = 0.94, SD = 0.40) and with the STRIP Assistant (M = 1.34, SD = 0.20);  $t(42) = 7.07$ ,  $p < 0.0001$ . This indicates that participants spent significantly more time optimizing medication with the STRIP Assistant.

On average, the respondents filled out a SUS-score of 63.25, **out of a possible maximum of 100**. This value is lower than the quality threshold of 70 arrived at by Bangor et al. and corresponds with a marginal acceptance rate in a later paper by the same authors [41,42].

## 5 Discussion

### 5.1 Effectiveness

This study has shown that a decision support system can make general practitioners and pharmacists perform better medication reviews, **albeit in an experimental setting with preselected patient cases**. This is in line with the consensus on the effectiveness of health recommendation systems in literature [27]. More specifically, the results indicate that the choice for a recommender based on a pre-determined explicit knowledge base yields viable results in a complex domain with potentially far-reaching implications. Rather than relying solely on collaborative or content-based filtering, a knowledge base guarantees a minimal quality level when generating recommendations [43].

Even though the medication reviews performed with the STRIP Assistant are significantly better than those without, a non-negligible number of mistakes respondents made (15%) can be attributed to software suggestions.

In this experiment, each START advice was presented as an alphabetically ordered list of medicines that users had the possibility to prescribe. In practice, many users picked the first item in the list, resulting in an overabundance of suboptimal choices; when adding a vitamin D supplement, for example, many users picked alfacalcidol instead of cholecalciferol, even though the former has fewer and more specific indications. Few publications have touched upon the subject of decision support systems generating incorrect recommendations, and consequently strategies to prevent them are lacking [44,45,46]. Hybrid recommendation systems, combining an explicit knowledge base with content-based or collaborative filtering, have been shown to outperform their simpler counterparts [43]. As long as the risk associated with automatic learning systems in a precarious domain such as health care is accounted for, a hybrid approach may prove beneficial to improving recommenders' effectiveness.

## **5.2 Efficiency**

Contrary to our assumption, performing medication reviews with the STRIP Assistant is less efficient (i.e. takes more time) than optimizing drugs manually. Traditionally, the three aspects of usability are assumed to be positively correlated [47]. However, a different perspective viewing effectiveness and efficiency as conflicting requirements in a project has been proposed by Nilsson & Følstad [48]. In an experiment such as the one in this study, where respondents either use their habitual approach, or have to learn a new structured method, a drop in efficiency can be reasonably attributed to effectiveness and efficiency conflicting. Due to the experimental setting, unfamiliarity with the method and the user interface is likely to play a role as well.

Conducting experiments in which more gradual changes to the method are applied may result in improvements in both effectiveness and efficiency; in a study related to this one, a paper version of the earlier polypharmacy optimization method was tested in an experiment [14]. It, too, proved to be less efficient than performing a medication review manually. However, the software-aided reviews performed in this study took less time than the paper-based ones in the previous study. This lends credibility to the assumption that gradual changes may improve all aspects of usability simultaneously.

## **5.3 Satisfaction**

Respondents perceived using the STRIP Assistant as only marginally acceptable. The SUS-average of 63.25 was lower than the commonly accepted quality indicator of 70 [41,42]. This aspect, too, can be understood by viewing the usability aspects as conflicting requirements [48]. The suboptimal prototypical design of the software's user interface, and the unfamiliarity of the respondents with the application may explain this inconsistency with consensus in literature.

## **5.4 Clinical Relevance**

Methods for medication review have been proven to be valuable in improving prescriptions of polypharmacy patients. The Polypharmacy Optimization Method, which served as a foundation for the STRIP method, has led to improvements in appropriate decisions in medication reviews [14]. The START and STOPP criteria, which constitute

a major part of the STRIP Assistant's knowledge base, have been shown to be associated with improvements in medication appropriateness, reductions in adverse drug reactions, and decreases in drug use and costs [17,49].

The two patient cases used in this experiment were comparable in complexity and number of medicines, but there could, for reasons of validity, not be a complete overlap of diseases and drugs. This makes determining the clinical relevance of the intervention difficult. Nevertheless, the most noticeable improvements in the adequate prescribing of drugs were the treatment of osteoporosis with bisphosphonates, calcium and vitamin D, and the treatment of systolic heart failure with ACE inhibitors. The most important improvement relating to stopping medicine use is the discontinuation of digoxin when atrial fibrillation is adequately treated with beta blockers. These interventions correspond with guidelines of the START and STOPP criteria [17].

Thus, the results in this study confirm the results of previous studies that structured methods for medication review significantly improve medication appropriateness of prescriptions.

### **5.5 Limitations**

When interpreting these results, the experimental nature of the method should be taken into account. The STRIP Assistant's usability has been tested and validated with real patient cases in a controlled environment, but has not been validated in practice with users reviewing their own patients. While the results lend credibility to the STRIP method being useful in practice, this study does not prove its clinical relevance.

When generalizing the results of this study, the limited number of participants should be considered, as well as the sampling method. 42 GPs and pharmacists participated voluntarily, enabling the possibility that they were positively biased towards use of a clinical decision support system aiding them with medication reviews.

### **5.6 Further Research**

A randomized controlled trial incorporating a large representative sample should be conducted to conclude the STRIP Assistant's effectiveness, efficiency, and satisfaction. Further research should focus on its usability through evaluation in a real life setting over a longer period of time, exploring to what extent experience influences users' effectiveness and efficiency in working with the software. Furthermore, longitudinal research could show if the STRIP Assistant is clinically relevant in practice and evaluate its impact on adverse effects and medicine costs.

## **6 Conclusion**

In this study, a clinical decision support system (the STRIP Assistant) designed to aid GPs and pharmacists with conducting medication reviews was validated in an experimental setting. The results show that using the STRIP Assistant positively influences the number of appropriate decisions made in a medication review of elderly polypharmacy patients and decreases the number of inappropriate choices. Contrary to our assumptions, users spent more time optimizing health records with the STRIP Assistant than without it. Users perceive using the

software only as marginally acceptable. Further research is needed to determine whether optimization of polypharmacy with the help of the STRIP Assistant is clinically beneficial.

## ACKNOWLEDGEMENTS

We would like to thank the participating GPs and pharmacists for their contribution to the experiment.

## COMPLIANCE WITH ETHICAL STANDARDS

The manuscript does not contain clinical studies or patient data. No sources of funding were used to assist in the conduct or preparation of this study. The authors have no conflicts of interest that are directly relevant to the content of this study. All procedures were performed in accordance with the ethical standards of the 1964 Helsinki declaration and its later amendments. Since the Medical Research Involving Human Subjects Act (WMO) did not apply to this study, approval of the medical ethical committee of University Medical Center Utrecht was not required.

## 7 Bibliography

1. Jansen PAF, Brouwers JRBJ. Clinical pharmacology in old persons. *Scientifica (Cairo)*. 2012.
2. Björkman I, Fastbom J, Schmidt I, et al. Drug-drug interactions in the elderly. *Ann Pharmacother*. 2002; 36(11): p. 1675-81.
3. Claxton A, Cramer J, Pierce C. A systematic review of the association between dose regimens and medication adherence. *Clin Ther*. 2001; 23(8): p. 1296-310.
4. Frazier SC. Health outcomes and polypharmacy in elderly individuals; an integrated literature review. *J Gerontol Nurs*. 2005; 31(9): p. 4-11.
5. Kuyuma M, Endo H, Umegaki H. Factors influencing nonadherence with medication regimens in the elderly. *Nihon Ronen Igakkai Zasshi*. 2000; 37: p. 363-70.
6. Shi S, Mörike K, Klotz U. The clinical implications of ageing for rational drug therapy. *Eur J Clin Pharmacol*. 2008; 64(2): p. 183-99.
7. Sloane P, Gruber-Baldini A, Zimmerman S, et al. Medication undertreatment in assisted living settings. *Arch Intern Med*. 2004; 164(18): p. 2031-2037.

8. Steinman M, Landefeld C, Rosenthal G, et al. Polypharmacy and prescribing quality in older people. *J Am Geriatr Soc.* 2006; 54(10): p. 1516-23.
9. Wright R, Sloane R, Pieper C, et al. Underuse of indicated medications among physically frail older US veterans at the time of hospital discharge: results of a cross-sectional analysis of data from the geriatric evaluation and management drug study. *Am J Geriatr Pharmacother.* 2009; 7(5): p. 271-280.
10. Leendertse A, Egberts A, van den Bemt P, et al. Frequency of and risk factors for preventable medication-related hospital admissions in the Netherlands. *Arch Intern Med.* 2008; 168(17): p. 1890-1896.
11. Stichting Farmaceutische Kengetallen. Polyfarmacie. *Pharmaceutisch Weekblad.* 2005 augustus 12; 140(32).
12. Gallagher PF, O'connor MN, O'Mahony D. Prevention of potentially inappropriate prescribing for elderly patients: a randomized controlled trial using STOPP/START criteria. *Clinical Pharmacology & Therapeutics.* 2011; 89(6): p. 845-854.
13. Dedhiya SD, Hancock E, Craig BA, et al. Incident use and outcomes associated with potentially inappropriate medication use in older adults. *Am J Geriatr Pharmacother.* 2010; 8(6): p. 562-570.
14. Drenth-van Maanen AC, Van Marum RJ, Knol W, et al. Prescribing Optimization Method for improving prescribing in elderly patients receiving polypharmacy. *Drugs Aging.* 2009; 26(8): p. 1-15.
15. Sergi G, De Rui M, Sarti S, et al. Polypharmacy in the elderly: can comprehensive geriatric assessment reduce inappropriate medication use? *Drugs Aging.* 2011; 28(7): p. 509–518.
16. Campanelli CM. American Geriatrics Society updated Beers criteria for potentially inappropriate medication use in older adults: the American Geriatrics Society 2012 Beers criteria update expert panel. *J Am Geriatr Soc.* 2012; 60(4): p. 616-631.
17. Gallagher P, Ryan C, Byrne S, et al. 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): p. 72-83.
18. Hanlon JT, Schmader KE, Samsa GP, et al. A method for assessing drug therapy appropriateness. *J Clin Epidemiol.* 1992; 45(10): p. 1045-1051.
19. Leendertse AJ, De Koning FH, Goudswaard AN, et al. Preventing hospital admissions by reviewing medication (PHARM) in primary care: design of the cluster randomised, controlled, multi-centre PHARM-study. *BMC*

Health Serv Res. 2011; 11(1).

20. Patterson SM, Hughes C, Kerse N, et al. Interventions to improve the appropriate use of polypharmacy for older people. Systematic Literature Review. ; 2012.
21. Anderson K, Stowasser D, Freeman C, et al. Prescriber barriers and enablers to minimising potentially inappropriate medications in adults: a systematic review and thematic synthesis. BMJ Open. 2014; 4(12).
22. Niquille A, Lattmann C, Bugnon O. Medication reviews led by community pharmacists in Switzerland: a qualitative survey to evaluate barriers and facilitators. Pharm Pract (Granada). 2010; 8(1): p. 35–42.
23. Dutch College of General Practitioners. Multidisciplinary guideline polypharmacy in elderly ('Multidisciplinaire richtlijn polyfarmacie bij ouderen' in Dutch). Guideline. Utrecht;; 2012.
24. Topinková E, Baeyens JP, Michel JP, et al. Evidence-based strategies for the optimization of pharmacotherapy in older people. Drugs Aging. 2012; 29(6): p. 477-494.
25. Spinewine A, Fialová D, Byrne S. The role of the pharmacist in optimizing pharmacotherapy in older people. Drugs Aging. 2012; 29(6): p. 495-510.
26. Van der Lugt J, Klapwijk S. Overzicht in beeld ('Overview in focus' in Dutch). SynthesHIS. 2008 December.
27. Bright TJ, Wong A, Dhurjati R, et al. Effect of clinical decision-support systems: a systematic review. Ann Intern Med. 2012; 157(1): p. 29-43.
28. Meulendijk MC, Spruit MR, Drenth-van Maanen AC, et al. General practitioners' attitudes towards decision-supported prescribing: an analysis of the Dutch primary care sector. Health Informatics J. 2013; 19(4): p. 247-263.
29. Meulendijk MC, Spruit MR, Jansen PAF, et al. STRIPA: a rule-based decision support system for medication reviews in primary care (accepted). In Proceedings of the European Conference on Information Systems (ECIS) 2015; 2015; Münster, Germany.
30. Z-Index. G-Standaard - Z-Index. [Online].; 2014 [cited 2014 December 17. Available from: <https://www.z-index.nl/g-standaard>.
31. De Groot DA, De Vries M, Joling KJ, et al. Specifying ICD9, ICPC and ATC codes for the STOPP/START criteria: a multidisciplinary consensus panel. Age Ageing. 2014.

32. Pachidi S, Spruit M, Weerd Ivd. Understanding users' behavior with software operation data mining. *Comput Human Behav.* 2014; 30: p. 583-594.
33. Meulendijk MC. STRIP Assistant. [Online].; 2014. Available from: <http://videodemo.stripa.eu/english/>.
34. International Organization for Standardization (ISO). ISO/IEC 25010:2011: Systems and software engineering -- Systems and software quality requirements and evaluation (SQuaRE) -- System and software quality models. Standard. ; 2011.
35. Davis F. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly.* 1989; 13(3): p. 319-340.
36. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. *Manage Sci.* 2000; 46(2): p. 186-204.
37. Venkatesh V, Morris MG, Davis GB, et al. User acceptance of information technology: toward a unified view. *MIS Quarterly.* 2003; 27(3): p. 425-478.
38. Hornbaek K. Current practice in measuring usability: challenges to usability studies and research. *Int J Hum Comput Stud.* 2006; 64: p. 79-102.
39. 't Hart H, Van Dijk J, De Goede M, et al. *Onderzoeksmethoden (Research methods, in Dutch)*. 6th ed. Amsterdam: Boom; 2003.
40. Brooke J. SUS: a quick and dirty usability scale. In Jordan PW, Thomas B, McClelland IL, et al. *Usability Evaluation In Industry.*: CRC Press; 1996. p. 189-194.
41. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. *Int J Hum Comput Interact.* 2008; 24(6): p. 574-594.
42. Bangor A, Kortum PT, Miller JT. Determining what individual SUS scores mean: adding an adjective rating scale. *Journal of Usability Studies.* 2009; 4(3): p. 114-123.
43. Adomavicius G, Tuzhilin A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Trans Knowl Data Eng.* 2005; 17(6): p. 734-749.
44. Eslami S, Abu-Hanna A, De Keizer NF, et al. Errors associated with applying decision support by suggesting default doses for aminoglycosides. *Drug Saf.* 2006; 29(9): p. 803-809.

45. Koppel R, Metlay JP, Cohen A, et al. Role of computerized physician order entry systems in facilitating medication errors. *JAMA*. 2005; 293(10): p. 1197-1203.
46. Ash JS, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. *J Am Med Inform Assoc*. 2004; 11(2): p. 104-116.
47. Hornbæk K, Law ELC. Meta-analysis of correlations among usability measures. In Proceedings of the SIGCHI conference on Human factors in computing systems; 2007. p. 617-626.
48. Nilsson EG, Følstad A. Effectiveness and efficiency as conflicting requirements in designing emergency mission reporting. *I-UxSED*. 2012;; p. 20-25.
49. O'Mahony D, O'Sullivan D, Byrne S, et al. STOPP/START criteria for potentially inappropriate prescribing in older people: version 2. *Age Ageing*. 2014.