

# BMJ Open Effectiveness of a complex intervention on Prioritising Multimедication in Multimorbidity (PRIMUM) in primary care: results of a pragmatic cluster randomised controlled trial

**To cite:** Muth C, Uhlmann L, Haefeli WE, *et al.* Effectiveness of a complex intervention on Prioritising Multimедication in Multimorbidity (PRIMUM) in primary care: results of a pragmatic cluster randomised controlled trial. *BMJ Open* 2018;8:e017740. doi:10.1136/bmjopen-2017-017740

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2017-017740>).

Oral presentations and posters: Muth C, Uhlmann L, Haefeli WE, Rochon J, van den Akker M, Beyer M, Perera R, Knottnerus A, Gerlach FM, Harder S (2014). PRIorisierung von MULTimedikation bei Multimorbidität (PRIMUM)\* Cluster-RCT in Hausarztpraxen zeigte keine Effekte auf die Angemessenheit der Verschreibung. 48. Kongress der DEGAM, 18; 20 September 2014, Hamburg; Abstractband V3c I 3, S. 88. Muth C, Rochon J, Namyst A, Fullerton B, Harder S, van den Akker M, Perera-Salazar R, Gerlach FM, Beyer M. Anwendung der MRC Guidance in der allgemeinmedizinischen Forschung: Ergebnisse aus der PRIMUM-Studie (PRIorisierung von MULTimedikation bei Multimorbidität). Vortrag auf 13. Jahrestagung des Deutschen Netzwerks Evidenzbasierte Medizin, Hamburg, 15; 17 March 2012, Abstractband IV/1a.

Received 13 May 2017  
Revised 30 November 2017  
Accepted 6 December 2017



For numbered affiliations see end of article.

**Correspondence to**  
Dr Christiane Muth;  
[muthc@allgemeinmedizin.uni-frankfurt.de](mailto:muthc@allgemeinmedizin.uni-frankfurt.de)

Christiane Muth,<sup>1</sup> Lorenz Uhlmann,<sup>2</sup> Walter E Haefeli,<sup>3</sup> Justine Rochon,<sup>2</sup> Marjan van den Akker,<sup>4,5</sup> Rafael Perera,<sup>6</sup> Corina Gütthlin,<sup>1</sup> Martin Beyer,<sup>1</sup> Frank Oswald,<sup>7</sup> Jose Maria Valderas,<sup>8</sup> J André Knottnerus,<sup>4</sup> Ferdinand M Gerlach,<sup>1</sup> Sebastian Harder<sup>9</sup>

## ABSTRACT

**Objectives** Investigate the effectiveness of a complex intervention aimed at improving the appropriateness of medication in older patients with multimorbidity in general practice.

**Design** Pragmatic, cluster randomised controlled trial with general practice as unit of randomisation.

**Setting** 72 general practices in Hesse, Germany.

**Participants** 505 randomly sampled, cognitively intact patients (≥60 years, ≥3 chronic conditions under pharmacological treatment, ≥5 long-term drug prescriptions with systemic effects); 465 patients and 71 practices completed the study.

**Interventions** Intervention group (IG): The healthcare assistant conducted a checklist-based interview with patients on medication-related problems and reconciled their medications. Assisted by a computerised decision support system, the general practitioner optimised medication, discussed it with patients and adjusted it accordingly. The control group (CG) continued with usual care.

**Outcome measures** The primary outcome was a modified Medication Appropriateness Index (MAI, excluding item 10 on cost-effectiveness), assessed in blinded medication reviews and calculated as the difference between baseline and after 6 months; secondary outcomes after 6 and 9 months' follow-up: quality of life, functioning, medication adherence, and so on.

**Results** At baseline, a high proportion of patients had appropriate to mildly inappropriate prescriptions (MAI 0–5 points: n=350 patients). Randomisation revealed balanced groups (IG: 36 practices/252 patients; CG: 36/253). Intervention had no significant effect on primary outcome: mean MAI sum scores decreased by 0.3 points in IG and 0.8 points in CG, resulting in a non-significant adjusted mean difference of 0.7 (95% CI –0.2 to 1.6) points in favour of CG. Secondary outcomes showed non-significant changes (quality of life slightly improved in IG but continued to decline in CG) or remained stable (functioning, medication adherence).

## Strengths and limitations of this study

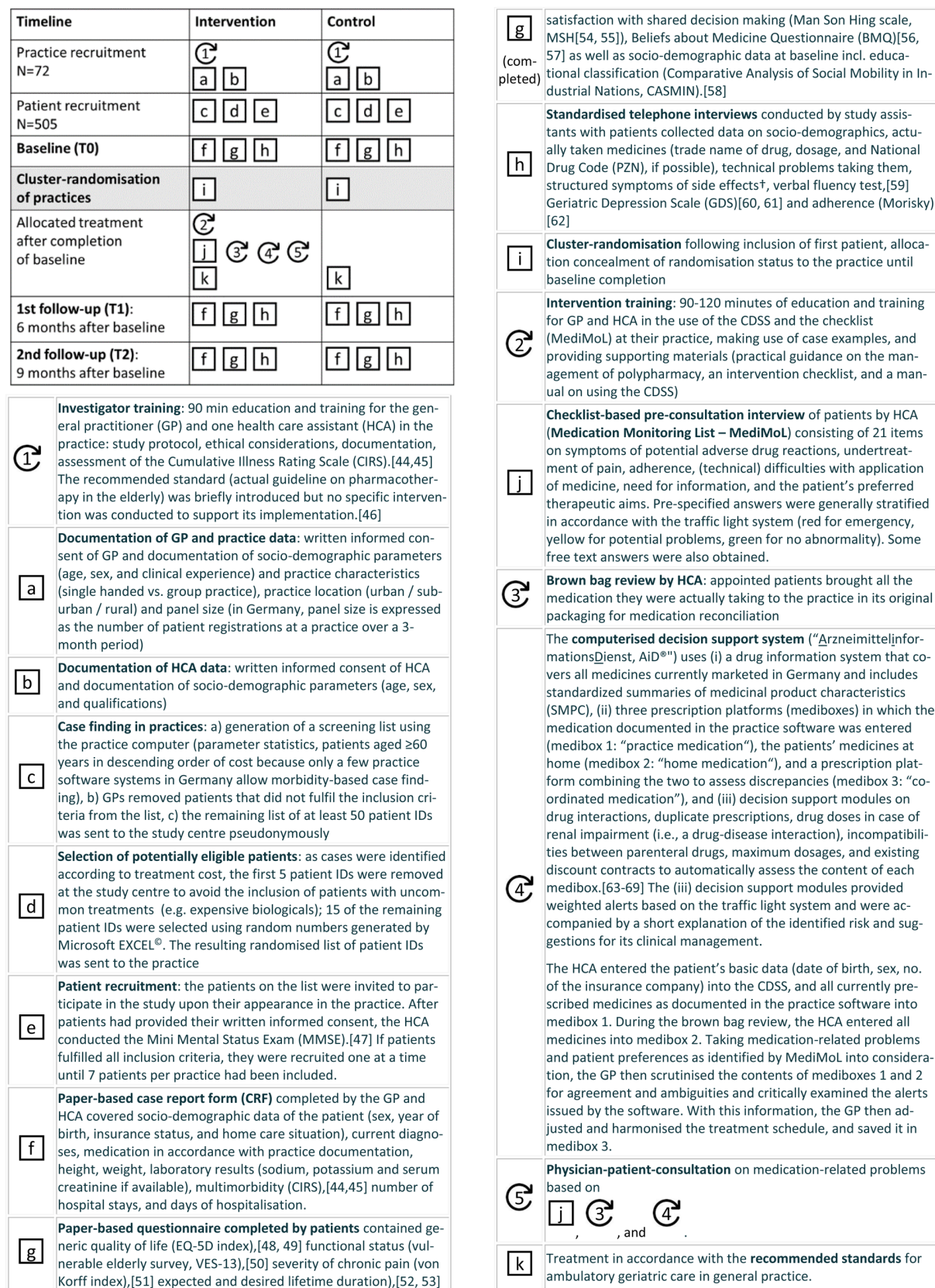
- The Prioritising Multimедication in Multimorbidity (PRIMUM) intervention was developed and piloted in accordance with the latest Medical Research Council guidance on complex interventions.
- The effectiveness of the PRIMUM intervention was evaluated in a rigorously conducted cluster randomised trial that involved random sampling of patients, disclosure of treatment allocation after baseline completion and adherence to the protocol.
- To evaluate the generic patient-centred strategy of applying PRIMUM, we used the commonly used Medication Appropriateness Index (MAI), as this implicit measure allows individualised assessments.
- We blinded both the assessment of the primary outcome MAI and the statistical analyses.
- Key limitations were that the baseline values of MAI and the secondary outcomes did not provide enough scope for improvement, and that medication underuse in polypharmacy was not sufficiently reflected in our outcome measures.

**Conclusions** The intervention had no significant effects. Many patients already received appropriate prescriptions and enjoyed good quality of life and functional status. We can therefore conclude that in our study, there was not enough scope for improvement.

**Trial registration number** ISRCTN99526053. NCT01171339; Results.

## INTRODUCTION

The prevalence of multimorbidity, that is, the co-occurrence of multiple chronic or acute diseases and medical conditions in one person,<sup>1</sup> increases with age, and most primary care consultations currently involve



**Figure 1** PaT Plot<sup>70</sup> of the Prioritising Multimедication in Multimorbidity (PRIMUM) trial. †Structured symptoms of side effects: dizziness, dyspnoea, tachycardia/palpitations, nausea/vomiting, abdominal pain, bleeding diathesis, difficulties urinating, ankle oedema—frequency expressed as occurrence on 1 day/several days/almost every day during the past 2 weeks. CDSS, computerised decision support system.

patients with multiple conditions.<sup>2–4</sup> Multiple disorders in patients are likely to result in multiple drug prescriptions. This increases the risk of drug–drug and drug–disease interactions, inappropriate dosages or drug selection and non-adherence of patients. They may, however, also result in undertreatment.<sup>5–10</sup> Inappropriate prescriptions may result in hospitalisations, falls and related injuries, decreased quality of life, cognitive and physical dysfunction, loss of autonomy and increased mortality, particularly in the elderly.<sup>6–8 11–14</sup> Negative health outcomes caused by inappropriate polypharmacy are responsible for high outlays for hospital treatment, home care and nursing homes.<sup>15–17</sup> Much morbidity and many costs may be preventable—for instance, 20%–50% of medication-related hospitalisations on internal wards have been estimated to be avoidable.<sup>13 16 18–20</sup> Recently, Dreischulte and coresearchers observed a reduction in hospital admission rates for gastrointestinal ulcers or bleeding in their trial evaluating a complex intervention addressing nine specific high-risk prescribing patterns such as non-steroidal anti-inflammatory drugs in renal failure, or in combination with oral anticoagulants.<sup>21</sup> Further trials also evaluated interventions addressing safety indicators and achieved a reduction in high-risk prescribing through adherence to explicit criteria that are relevant to public health.<sup>22 23</sup> However, ‘the range of reported effect sizes was modest, and it is unclear whether such interventions can result in clinically significant improvements in patient outcomes.’<sup>24</sup>

Furthermore, considering there are more than 10 000 known diseases, the number of possible interactions between diseases and treatments in patients with multimorbidity is vast, and patients may not be able to cope with the treatment burden.<sup>25</sup> Generic patient-centred strategies to assess potential interactions and to prioritise and individualise management in accordance with patients’ preferences and shared treatment goals have been recommended for patients with multimorbidity and polypharmacy.<sup>26–32</sup> In these patients, evidence of interventions with proven effectiveness on clinical outcomes remains scarce. However, recent Cochrane reviews have identified strategies that appear to be beneficial in terms of reducing inappropriate prescribing.<sup>33 34</sup> Based on promising strategies to combat inappropriate polypharmacy and in accordance with guiding principles to manage patients with multimorbidity, we developed and piloted a complex intervention.<sup>35</sup> As the prevalence of multimorbidity and polypharmacy in older people is high, they made up the target population. To reduce the workload on the general practitioner (GP), the intervention also involved a healthcare assistant (HCA) from the practice.<sup>35</sup> In Germany, HCAs receive less training than nurses and are comparable to certified medical assistants in the USA. In usual care, HCAs work as receptionists, assist GPs (eg, in diagnostic procedures or wound management) and conduct, for instance, dietary counselling. On many occasions, HCAs have successfully participated in chronic care interventions where they have, for example,

surveyed patients by following protocols with fixed interview questions for conditions such as osteoarthritis, major depression and chronic heart failure, under the supervision of GPs.<sup>36–40</sup>

In accordance with the Medical Research Council (MRC) guidance on developing and evaluating complex interventions, we tested the feasibility of the complex intervention in a pilot study.<sup>35</sup> On the basis of overall feasibility findings, we improved the intervention and trial design. To compare the effectiveness of the complex Prioritising Multimедication in Multimorbidity (PRIMUM) intervention with usual care in older patients with multimorbidity and polypharmacy in general practice, we used the Medication Appropriateness Index (MAI) as primary outcome. This implicit (non-criteria-based) measure allows an individualised assessment of medication appropriateness.<sup>41–43</sup> We investigated whether the appropriateness of drug prescriptions changed after 6 months’ follow-up measured as a difference in the MAI score 6 months from baseline minus baseline (MAI T1–T0).

## METHODS

### Study design

The study was a pragmatic, cluster randomised controlled trial (RCT) with the general practice as the unit of randomisation. To further reduce contamination of the control group and unlike the pilot study, detailed information on the intervention treatment was only provided to the intervention group.<sup>35</sup> Primary and secondary outcomes were measured at patient level (figure 1<sup>44–69</sup> and online supplementary appendix 1: study protocol).

### Setting and participants

*General practices* in the German state of Hesse were eligible if they provided primary care under the German statutory health insurance system, and if at least one of the HCA staff members was able to access the internet in the practice. Practices specialising in unconventional treatments or in special indications (eg, HIV) were excluded. To recruit practices, we sent letters to about 1600 practice addresses provided by the Association of Statutory Health Insurance Physicians of Hesse—addressees were not exclusively active GPs. We checked inclusion and exclusion criteria for those who were interested by phone and agreed upon a time for investigator training (figure 1: icon ‘1’). In both groups, GPs and HCAs received a lump sum of €300 in recompense for the work involved in documenting results. In the intervention group, GPs and HCAs received an additional €150 for the extra work that the intervention entailed.

GPs who did not respond to the original letter received a reminder phone call. We phoned a random 10% sample of those who did not respond to either the letter or the reminder up to three times in order to collect data on inclusion and exclusion criteria, practice characteristics and reasons for non-participation.



**Table 1** Baseline characteristics of practices and patients

	Control group	Intervention group
<b>Practices</b>	<b>n=36</b>	<b>n=36</b>
<i>Practice characteristics</i>		
Location (number, percentage)		
City (>100 000 inhabitants)	16 (44)	6 (17)
Mid-sized town (20 000–100 000)	6 (17)	10 (28)
Small town (5000–20 000)	10 (28)	15 (41)
Rural area (<5000 inhabitants)	4 (11)	5 (14)
Single-handed practices (number, percentage)	21 (58)	20 (56)
Panel size* (number, percentage)		
Fewer than 1000	11 (31)	12 (33)
1000–1499	14 (39)	11 (31)
1500 or more	11 (31)	13 (36)
<i>General practitioners</i>		
Age (mean, SD)	50.2±7.6	51.9±7.0
Male sex (number, percentage)	21 (58)	20 (56)
Board certificate GP (number, percentage)	30 (83)	30 (83)
Years of clinical experience (mean, SD)	22.6±8.6	23.3±7.9
Years at practice site (mean, SD)	14.3±9.1	15.7±8.4
<i>Healthcare assistants</i>		
Age (mean, SD)	40.1±8.8	37.8±12.6
Female sex (number, percentage)	36 (100)	35 (97)
Fully qualified HCA (number, percentage)	25 (69)	27 (75)
Years of professional experience (mean, SD)	18.4±9.3	15.9±10.6
Years at practice site (mean, SD)	10.4±8.2	9.6±8.5
Full-time employment (number, percentage)	17 (47)	20 (56)
Cluster size (median number of patients, range)	7 (6–8)	7 (6–8)
<b>Patients</b>	<b>n=253</b>	<b>n=252</b>
<i>Sociodemographics</i>		
Age (mean, SD)	71.7±7.4	72.5±6.5
Female sex (number, percentage)	131 (52)	133 (53)
Covered by statutory health insurance (number, percentage)	243 (96)	243 (96)
Participation in a DMP (number, percentage)	139 (55)	153 (61)
Consultation with specialists in previous 6 months (number, percentage)	222 (88)	227 (90)
Living with spouse: yes (number, percentage)	166 (67)	152 (61)
Funding for themselves (number, percentage)	236 (94)	237 (94)
Home care situation rated as 'good' or 'very good' in GP assessment (number, percentage)	233 (92)	239 (95)
CASMIN educational classification (number, percentage)		
High	25 (10)	14 (6)
Middle	80 (32)	66 (27)
Low	144 (58)	169 (68)
<i>Morbidity and medication</i>		
Charlson Comorbidity Score (mean, SD)	3.2±2.4	3.0±2.0
CIRS sum score (mean, SD)	7.3±4.3	8.1±4.8

Continued

Table 1 Continued

	Control group	Intervention group
CIRS number of affected organ systems (mean, SD)	4.4±2.3	4.6±2.4
Potential ADR symptoms† (number, percentage)		
Bleeding diathesis‡	44 (17)	33 (13)
Ankle oedema	78 (31)	84 (33)
Dizziness‡	54 (21)	54 (21)
Dyspnoea‡	86 (34)	70 (28)
Difficulties urinating	51 (20)	64 (25)
Abdominal pain‡	36 (14)	37 (15)
Tachycardia or palpitation‡	36 (14)	36 (14)
Nausea or vomiting‡	16 (6)	11 (4)
Others		
BMI (mean, SD)	30.3±7.5	30.1±5.6
Geriatric Depression Scale (mean, SD)	2.4±2.3	2.3±2.2
Verbal fluency test (mean, SD)	19.1±5.6	18.6±5.8

\*The number of patient registrations in a practice over a 3-month period.

†For details see figure 1, item 'h'.

‡Symptoms appeared on at least several or almost every day.

ADR, adverse drug reaction; BMI, body mass index; CASMIN, Comparative Analysis of Social Mobility in Industrial Nations; CIRS, Cumulative Illness Rating Scale; DMP, disease management programme; GP, general practitioner; HCA, healthcare assistant.

**Patients:** A random sample of seven patients per practice were included (figure 1, patient recruitment, icons 'c' to 'e'). Patients were required to be ≥60 years old, have ≥3 chronic conditions under pharmacological treatment, ≥5 long-term prescriptions of drugs with systemic effects (the medication regimen may have included drugs with local effects but these did not fulfil the inclusion criterion), have made ≥1 practice visit during the past quarter and be able to fill in questionnaires and participate in telephone interviews. To include a greater number of patients at risk of (manageable) interactions than in the pilot study,<sup>35</sup> patients had to have diseases affecting at least two different organ systems operationalised as two different chapters of the International Classification of Diseases, 10th Revision. The chapters 'H' (diseases of the eyes and ears) and 'E00' to 'E04' (diseases of the thyroid gland without hyperthyroidism) were not counted because their potential for systemic interactions was considered to be low. We excluded patients with dementia and cognitive impairment (Mini-Mental Status Examination <26),<sup>47</sup> because we designed our intervention for cognitively intact patients and did not target caregivers. Further exclusion criteria were a life expectancy ≤12 months, alcohol and drug abuse (based on the GP's assessment) or participation in another clinical trial 30 days prior to inclusion.

### Randomisation, allocation concealment and blinding

The first patient from each practice served as the basis for randomisation (figure 1, icon 'i'). Patients registered thereafter were treated according to practice status (control or intervention), which was assigned in

an allocation ratio of 1:1 using a block randomisation of variable block length. At the study centre, an external researcher generated the allocation sequence using the random number generator of Microsoft EXCEL. Treatment allocation was disclosed to the practice after baseline completion. Owing to the nature of the intervention, it was not possible to blind GPs, HCAs, patients and the study team. Treatment allocation was blinded to the clinical pharmacologist conducting medication reviews for the primary outcome (MAI) and to the statistician.

### Intervention and control groups

#### Intervention group

The PaT Plot<sup>70</sup> (figure 1, icons 'j' and '3' to '5') shows the four elements of the complex intervention. It consists of (1) a brown bag review and (2) a checklist-based preconsultation interview with the patient that is conducted by the HCA (online supplementary appendix 2), (3) a computerised decision support system (CDSS)-assisted medication review carried out by the GP, and (4) a GP-patient consultation to optimise and prioritise medication. GPs had the option to use the CDSS to help prepare the medication review with the patient, and during the consultation itself. Trained HCAs and GPs (figure 1, item '2') implemented the intervention on a single occasion, which took the GP and the HCA a per-patient average of 35 and 45 min, respectively.<sup>35</sup> The practice team for the intervention group received the GP guidelines for ambulatory geriatric care prepared by the Hesse Guideline Group (figure 1, item 'k'). Recommendations in the guideline focus on primary and secondary prevention (eg, physical exercise, fall assessment and prevention).<sup>46</sup>

**Table 2** Intention-to-treat analysis of primary and secondary outcomes and sensitivity analyses

	Control group		Intervention group		Adjusted difference (95% CI)	ICC/ICC <sub>adj</sub>	P
	n <sub>c</sub>	Mean (SD)	n <sub>i</sub>	Mean (SD)			
Medication Appropriateness Index (MAI)							
MAI, baseline (T0)	253	4.6 (5.8)	252	4.8 (5.4)	–	–	–
Number of prescriptions rated with MAI, baseline*	253	7.8 (2.3)	252	8.0 (2.6)			
Primary outcome							
MAI, 6 months (T1)	243	3.8 (4.3)	241	4.6 (5.5)	MD: 0.7 (–0.2 to 1.6)†	0.016/0.017	0.137
Number of prescriptions rated with MAI, 6 months*	243	7.6 (2.2)	241	8.1 (2.8)	RR: 1.0 (1.0 to 1.1)	0.067/-	0.354
Secondary outcome							
MAI, 9 months (T2)	228	3.9 (4.9)	238	4.8 (5.2)	MD 0.6 (–0.5 to 1.7)†	0.000/0.000	0.272
Number of prescriptions rated with MAI, 9 months*	228	7.7 (2.3)	238	8.1 (3.0)	RR: 1.0 (1.0 to 1.1)	0.075/-	0.497
Sensitivity analysis							
DMP non-participants							
MAI, baseline	114	4.1 (5.2)	99	3.8 (3.8)	–	–	–
MAI, 6 months	110	3.5 (4.2)	92	4.2 (4.7)	MD: 0.7 (–0.4 to 1.9)†	0.000/0.000	0.200
MAI, 9 months	103	4.5 (5.7)	91	4.5 (5.1)	MD: 0.1 (–1.5 to 1.6)†	0.000/0.000	0.939
DMP participants							
MAI, baseline	139	5.1 (6.2)	153	5.4 (6.1)	–	–	–
MAI, 6 months	133	4.0 (4.5)	149	4.8 (5.9)	MD: 0.7 (–0.6 to 1.9)†	0.006/0.010	0.295
MAI, 9 months	125	3.5 (4.0)	147	4.9 (5.3)	MD: 1.1 (0.0 to 2.2)†	0.000/0.000	0.049
Secondary outcomes on quality of life-related measures							
EQ-5D index (percentage)							
Baseline	240	74.9 (23.0)	241	73.9 (24.4)	–	–	–
6 months	225	73.2 (24.8)	229	73.9 (23.8)	MD: 1.4 (–2.5 to 5.3)	0.080/0.082	0.471
9 months	214	72.8 (25.1)	222	74.8 (23.4)	MD: 2.3 (–1.6 to 6.2)	0.049/0.048	0.247
Expected life duration (years)							
Baseline	200	11.6 (6.9)	209	10.3 (6.9)	–	–	–
6 months	200	12.0 (7.1)	202	11.0 (7.3)	MD: 0.0 (–1.1 to 1.1)	0.000/0.000	0.987
9 months	184	12.3 (7.0)	195	11.7 (7.9)	MD: 0.5 (–1.3 to 2.4)	0.185/0.192	0.588
Desired life duration (years)							
Baseline	207	16.5 (9.1)	218	15.2 (8.9)	–	–	–
6 months	196	16.6 (9.1)	200	15.2 (8.7)	MD: –0.4 (–1.6 to 0.7)†	0.000/0.000	0.423
9 months	180	16.8 (9.2)	195	16.4 (9.8)	MD: 0.5 (–0.9 to 1.8)	0.078/0.081	0.479
Secondary outcomes on functional status, pain and hospitalisation							
Functional status (VES-13)							
Baseline	228	3.0 (2.9)	223	2.6 (2.7)	–	–	–
6 months	217	3.0 (2.9)	222	2.6 (2.8)	MD: 0.1 (–0.3 to 0.5)	0.000/0.000	0.681
9 months	199	2.7 (2.8)	204	2.8 (2.8)	MD: 0.4 (0.0 to 0.8)	0.051/0.043	0.047
Pain (von Korff Index)							
Baseline	197	1.7 (1.3)	204	1.7 (1.2)	–	–	–
6 months	184	1.7 (1.4)	198	1.8 (1.2)	MD: 0.2 (–0.1 to 0.4)†	0.000/0.000	0.135
9 months	168	1.6 (1.2)	194	1.7 (1.2)	MD: 0.0 (–0.2 to 0.3)	0.004/0.006	0.782
Number of hospital stays							
Baseline	40	1.4 (0.7)	42	1.7 (1.0)		–	–

Continued

Table 2 Continued

	Control group		Intervention group		Adjusted difference (95% CI)	ICC/ICC <sub>adj</sub>	P
	n <sub>c</sub>	Mean (SD)	n <sub>i</sub>	Mean (SD)			
6 months	45	1.4 (0.7)	34	1.4 (0.5)	RR: 1.2 (0.6 to 2.3)	0.000/ -	0.646
9 months	25	1.2 (0.4)	28	1.3 (0.6)	RR: 1.0 (0.3 to 3.1)	0.000/ -	0.949
Number of days spent in hospital							
Baseline	40	14.9 (12.9)	42	19.0 (12.2)	–	–	–
6 months	45	13.1 (11.5)	34	9.8 (8.9)	RR: 1.1 (0.5 to 2.3)	0.894/ -	0.850
9 months	25	9.7 (8.2)	28	28 (11.6)	RR: 0.4 (0.1 to 2.8)	0.859/ -	0.336
<b>Secondary outcomes of adherence and related measures</b>							
Self-reported adherence							
Baseline	252	3.7 (0.8)	250	3.7 (0.6)	–	–	–
6 months	238	3.8 (0.5)	237	3.6 (0.8)	MD: –0.1 (–0.2 to 0.0)	0.005/0.002	0.044
9 months	225	3.7 (0.6)	231	3.7 (0.7)	MD: 0.0 (–0.2 to 0.1)	0.005/0.007	0.629
Observed adherence: drug score (number and percentage of deviating patients)							
Baseline	251	101 (40.2%)	250	87 (34.8%)	–	–	–
6 months	237	101 (42.6%)	237	78 (32.9%)	OR: 0.7 (0.5 to 1.0)	0.000/0.000	0.051
9 months	224	88 (39.3%)	231	85 (36.8%)	OR: 0.9 (0.6 to 1.4)	0.010/0.009	0.736
Observed adherence: dose score (number and percentage of deviating patients)							
Baseline	251	125 (49.8%)	248	134 (54%)	–	–	–
6 months	235	128 (54.5%)	236	136 (57.6%)	OR: 1.1 (0.7 to 1.6)†	0.000/0.000	0.756
9 months	222	121 (54.5%)	229	145 (63.3%)	OR: 1.4 (0.9 to 2.0)†	0.013/0.005	0.119
Observed adherence: regimen score (number and percentage of deviating patients)							
Baseline	251	124 (49.4%)	249	131 (52.6%)	–	–	–
6 months	235	117 (49.8%)	236	134 (56.8%)	OR: 1.3 (0.8 to 2.0)†	0.057/0.051	0.297
9 months	222	114 (51.4%)	229	137 (59.8%)	OR: 1.4 (0.9 to 2.1)†	0.050/0.042	0.148
Number of prescriptions							
Baseline	253	8.0 (2.4)	252	8.1 (2.8)	–	–	–
6 months	242	7.8 (2.3)	241	8.4 (3.0)	RR: 1.0 (1.0 to 1.1)†	0.097/ -	0.183
9 months	227	7.8 (2.2)	238	8.4 (3.2)	RR: 1.0 (1.0 to 1.1)†	0.100/ -	0.310
Number of single doses							
Baseline	253	9.2 (3.5)	252	9.4 (4.1)	–	–	–
6 months	242	8.9 (3.3)	241	9.4 (4.1)	RR: 1.0 (1.0 to 1.1)†	0.183/ -	0.573
9 months	227	9.0 (3.6)	238	9.4 (4.4)	RR: 1.0 (0.9 to 1.1)†	0.212/ -	0.761
MRCI							
Baseline	253	26.9 (12.3)	252	28.4 (14.3)	–	–	–
6 months	242	26.3 (12.2)	241	28.6 (14.3)	MD: 0.7 (–0.7 to 2.1)†	0.030/0.032	0.308
9 months	227	26.3 (11.9)	238	29.1 (15.6)	MD: 1.0 (–0.6 to 2.5)†	0.042/0.042	0.212
Man-Son-Hing Scale							
Baseline	241	8.4 (3.4)	246	8.6 (3.4)	–	–	–
6 months	233	8.6 (3.2)	233	8.4 (3.4)	MD: –0.1 (–0.7 to 0.5)	0.047/0.050	0.789
9 months	219	8.8 (3.5)	231	8.7 (3.7)	MD: –0.2 (–1.0 to 0.5)	0.041/0.041	0.519
BMQ, specific necessities							
Baseline	233	22.1 (3.3)	240	22.1 (3.1)	–	–	–
6 months	219	22.0 (2.9)	230	21.8 (3.5)	MD: –0.2 (–0.8 to 0.4)	0.043/0.046	0.557
9 months	207	21.6 (3.6)	226	21.9 (3.4)	MD: 0.3 (–0.4 to 1.0)	0.000/0.000	0.349
BMQ, specific concerns							

Continued

Table 2 Continued

	Control group		Intervention group		Adjusted difference (95% CI)	ICC/ICC <sub>adj</sub>	P
	n <sub>c</sub>	Mean (SD)	n <sub>i</sub>	Mean (SD)			
Baseline	229	13.4 (5.2)	238	13.4 (5.2)	–	–	–
6 months	223	13.1 (4.8)	227	12.8 (4.8)	MD: –0.2 (–1.0 to 0.7)	0.021/0.023	0.714
9 months	211	12.6 (5.0)	226	12.5 (5.1)	MD: 0.1 (–0.8 to 1.0)	0.044/0.047	0.838
BMQ, general overuse							
Baseline	237	10.5 (3.5)	241	10.5 (3.7)	–	–	–
6 months	229	10.4 (3.6)	226	10.4 (3.4)	MD: –0.2 (–0.8 to 0.5)	0.048/0.050	0.637
9 months	213	10.5 (3.6)	225	10.6 (3.6)	MD: 0.0 (–0.7 to 0.6)	0.054/0.057	0.917
BMQ, general harms							
Baseline	239	8.0 (3.0)	245	7.9 (3.0)	–	–	–
6 months	229	7.9 (2.8)	234	7.9 (3.2)	MD: 0.1 (–0.4 to 0.6)	0.000/0.002	0.631
9 months	214	8.2 (3.1)	232	8.0 (3.2)	MD: –0.2 (–0.8 to 0.4)	0.045/0.047	0.602

n<sub>c</sub>/n<sub>i</sub> (number of patients in control group/intervention group), SD, MD (mean differences), OR and RR (relative risk) are provided with 95% CIs, and adjusted for clustering effects and baseline. ICCs are provided as crude values using a mixed model without any adjustment (either group or baseline). The adjusted values use a mixed model that includes the group variable. P values are adjusted for cluster effects and baseline.

\*Phytopharmaceutical, homeopathic and other complementary medicine products were excluded from rating.

†Control group tended to perform better.

BMQ, Beliefs about Medicine Questionnaire; CIRS, Cumulative Illness Rating Scale; DMP, disease management programme; EQ-5D, EuroQol five dimensions; ICC, intracluster correlation coefficient; MAI, medication appropriateness index; MRCI, Medication Regimen Complexity Index; VES-13, Vulnerable Elderly Survey-13 items.

### Control group

The control group continued to receive usual care but the practice team also received the GP guidelines for ambulatory geriatric care (figure 1, item 'k')<sup>46</sup> to harmonise usual care in both groups.

### Outcomes

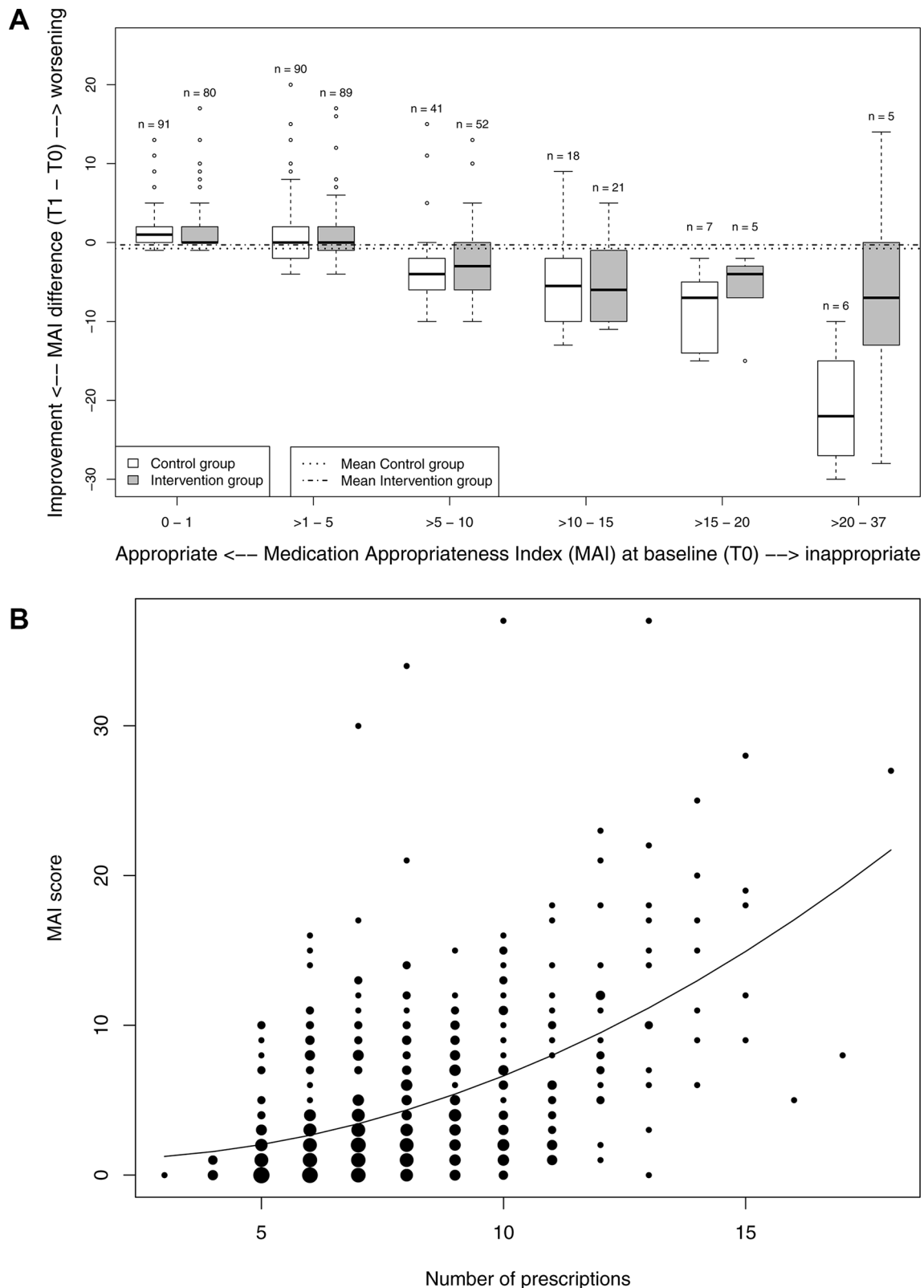
The *primary outcome* was the difference in MAI sum score<sup>41 71</sup> at 6 months minus the corresponding baseline score (MAI T1–T0). The MAI is commonly used in RCTs<sup>42 43</sup> and consists of 10 items: indication, effectiveness, correctness of dosage, correctness of direction, practicality of direction, drug–drug interactions, drug–disease interactions, unnecessary drug duplications, correctness of treatment duration and costs. The MAI item on cost was omitted because variable discount contracts between pharmaceutical companies and statutory health insurers preclude cost comparisons in Germany. The medication reviews were conducted by the same clinical pharmacologist (SH) who performed the pilot study. He rated nine items per prescription from '1' (appropriate) to '3' (inappropriate) where '2' represents a middle rating of uncertain appropriateness in a blinded chart review. In line with the piloted procedures,<sup>35</sup> he coded the MAI according to the GP's prescriptions, renal function, electrolytes, multimorbidity (diagnoses, Cumulative Illness Rating Scale)<sup>44 45</sup> (figure 1, icon 'f') and symptoms of adverse drug reactions (ADR) (figure 1, icon 'h'). Phytopharmaceutical, homeopathic and other complementary medicine products were excluded from the rating. MAI sum scores for the entire medication regimen were calculated on the basis of these ratings. Based on

the intrarater reliability of the MAI ratings in the pilot study (B statistics: the intrarater reliability for the nine MAI items ranged from 0.90 to 0.99 and was slightly better than inter-rater reliability),<sup>35</sup> we did not perform a duplicate MAI rating. MAI ratings were transformed by subtracting 1 from the original rating, resulting in values ranging from '0' (best rating) to '2' (worst rating), and summed to give an MAI score per prescription (theoretically ranging from 0 to 18) and across the entire medication regimen of the patient. Lower MAI sum scores denoted better prescribing appropriateness. A negative difference in MAI sum scores (MAI T1–T0) therefore reflected an improvement in prescribing quality.

*Secondary outcomes (6 vs 9 months):* we measured the change in the MAI score after 9 months (MAI T2–T0). On the assumption, improved medication appropriateness would result in improved health-related quality of life and functional status, we measured the differences in the EuroQol five dimensions (EQ-5D) index score,<sup>48 49</sup> changes in perceived future life expectancy (a quality of life-related concept indicating well-being and positive life evaluation measured in years of expected and desired lifetime duration),<sup>52 53</sup> functional status (differences in Vulnerable Elderly Survey-13 items),<sup>50</sup> all-cause hospitalisation and severity of chronic pain (von Korff Index)<sup>51</sup> after 6 and 9 months (T1–T0 and T2–T0).

To explain intervention effects, we also measured changes in satisfaction with shared decision-making (Man-Son-Hing Scale)<sup>54 55</sup> and medication adherence after 6 and 9 months (T1–T0 and T2–T0). We investigated (A) self-reported adherence in accordance with Morisky (low scores indicating good

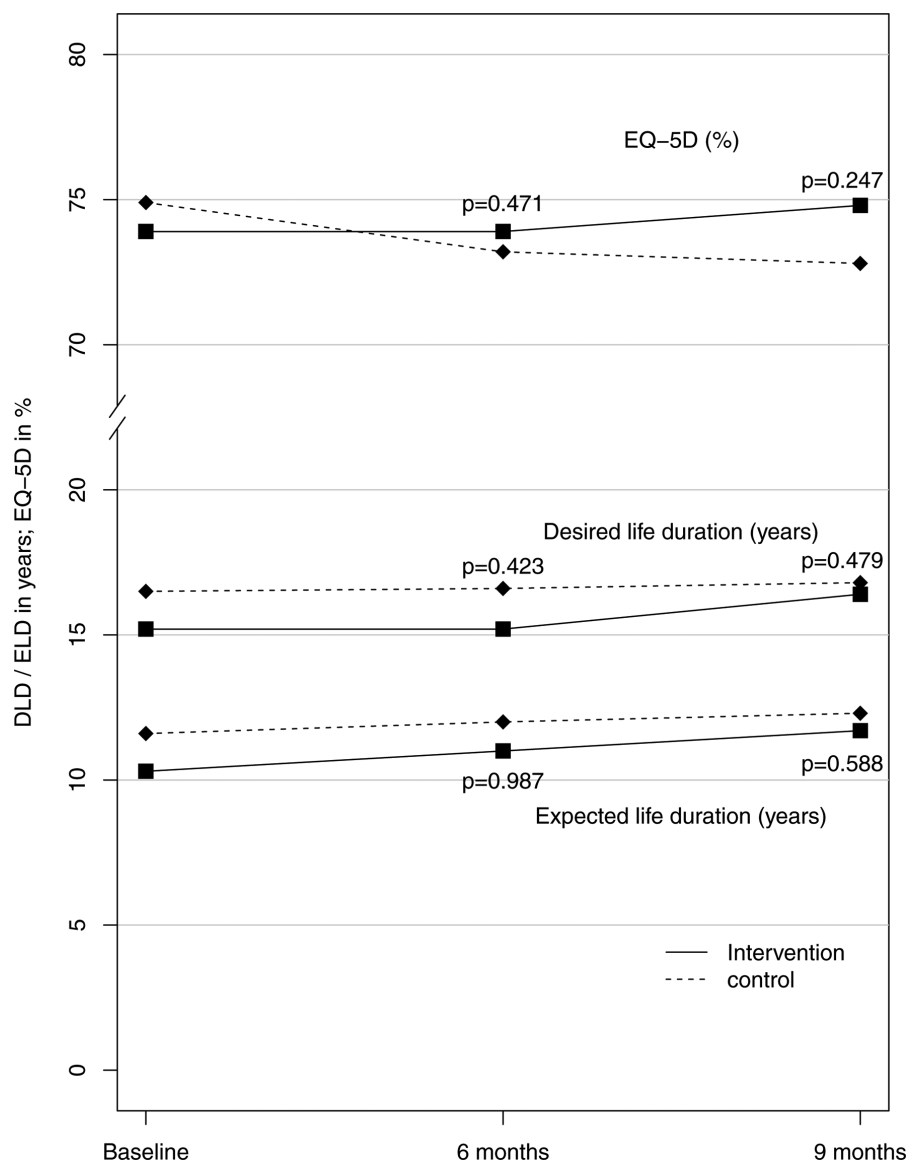




**Figure 2** Distribution and changes in the MAI using baseline values and number of prescriptions. (A) Changes in MAI scores in intervention and control groups 6 months after baseline compared with baseline values (absolute numbers of study participants and boxes and whiskers per subgroup are provided). (B) MAI scores at baseline in terms of the number of prescriptions (higher diameters of drops represent higher numbers of study participants).

adherence)<sup>62</sup>; (B) ‘observed adherence’ measured in terms of discrepancies between medicines actually taken (reported during patient interviews) and medicines prescribed

(reported by GP), as expressed in the three scores developed by Barat *et al.*<sup>72</sup> The scores were based on ratios calculated as follows:



**Figure 3** Secondary outcomes related to patients' self-reported quality of life measures. DLD, desired life duration; ELD, expected life duration; EQ-5D, EuroQol five dimensions.

1. The drug score representing the ratio of the number of drugs reported by the patients to the number of drugs reported by the GP.
2. The dose score ( $\text{DoS} = d_1(a_1) + d_2(a_2) + d_3(a_3) + \dots / n$ ), where  $d_i$  is the drug used by the patients (value 0 or 1),  $n$  is the number of drugs in the GP's report, and  $a_i$  is the dose-deviation ratio calculated by dividing the patient's reported daily dose by the daily dose prescribed by the GP.
3. The regimen score ( $\text{RS} = d_1(b_1) + d_2(b_2) + d_3(b_3) + \dots / n$ ), where  $b_i$  is the regimen-deviation ratio and calculated by dividing the patient's reported daily intake frequency (once daily, twice daily, and so on) by the corresponding frequency prescribed by the GP.<sup>72</sup> Scores outside an interval of 0.8–1.2 were considered to be divergent.

Further adherence-related measures assessed the complexity of the medication (total number of

prescriptions, number of single doses/day and Medication Regimen Complexity Index),<sup>73</sup> patients' beliefs and attitudes towards medication (Beliefs about Medicine Questionnaire),<sup>56 57</sup> cognitive function (verbal fluency test, VFT)<sup>59</sup> and depressive symptoms (Geriatric Depression Scale, GDS).<sup>60 61</sup> GDS and VFT will be reported elsewhere.

### Sample size

Based on the results obtained in previous studies,<sup>35 74</sup> a difference in the change values (MAI T1–T0) of at least 2 units between the treatment groups was considered clinically relevant. Based on the pilot study, an SD of 6 units was expected, resulting in a Cohen's effect size  $d$  of 0.3 and representing a small effect size.<sup>75</sup> Assuming an intraclass correlation coefficient (ICC) of 0.03 at practice level<sup>76</sup> and an average cluster size of seven patients, a total of 62 practices and 434 patients (31

practices and 217 patients per treatment arm) were required to detect such an effect with 80% power using a two-sample t-test at a two-sided significance level of  $\alpha=0.05$ . The sample size calculation was performed using NCSS Statistical Software 'PASS 2008'<sup>TM</sup> (inequality tests for two means in a cluster randomised trial). On the basis of an assumed drop-out rate of approximately 10%, the sample size was adjusted to a total of 70 practices and 490 patients (35 practices and 245 patients in each treatment group).

### Statistical analysis

We performed descriptive analyses of the primary endpoint, the secondary endpoints, and all patient and practice characteristics (separately for patients in both groups) and calculated mean and SD for continuous variables, and relative and absolute frequencies for categorical data.

In the primary analysis and using a two-sided significance level of  $\alpha=0.05$ , we tested the null hypothesis  $H_0: \mu_1=\mu_2$  (the mean difference MAI T1-T0 is the same in both groups) against the alternative hypothesis  $H_1: \mu_1\neq\mu_2$  (the mean MAI T1-T0 differs). Because of cluster randomisation, we used a multilevel regression approach with patients at level 1 and practices at level 2. The primary model included treatment group and MAI baseline as fixed factors and practice as a random factor. In a mixed model, estimates are adjusted for the correlation of observations on the same level, where a specific structure has to be chosen. We applied the compound symmetry correlation structure on the assumption that a correlation exists between patients from the same practice and that a specific numerical value can be attached to this correlation. We assumed the value was 0 for the correlation with patients from other practices. The results are presented as the adjusted mean between-group difference in MAI T1-T0 with the corresponding 95% CI. In addition, the practice-related ICC was estimated. The primary analysis was performed in accordance with the intention-to-treat (ITT) principle,<sup>77</sup> and additional sensitivity analyses were conducted on a per-protocol analysis set. In the multilevel approach, we made use of the missing at random assumption that the baseline or the treatment variable can explain missing data in the response. No additional imputation of missing data was conducted. In a sensitivity analysis, we replaced missing values for the primary endpoint using the baseline observation carried forward (BOCF) approach. The statistical analyses of the secondary endpoints used the same multilevel approach as the primary analysis. A linear, binary or Poisson mixed model was fitted in accordance with the scaling of the considered endpoint. The obtained P values in the secondary analyses are only interpreted exploratively. All evaluations were carried out using software package R (V.2.15.0 and higher),<sup>78</sup> in combination with the R packages xtable,<sup>79</sup> nlme,<sup>80</sup> lme4,<sup>81</sup> multilevel<sup>82</sup> and psychometric.<sup>83</sup>

## RESULTS

### Participant flow and non-responders

Of the 1662 practice addresses we sent letters to (1332 of them also received a phone call reminder), 1325 did not reply at all, 102 answered but were not interested in further information and 235 general practices asked for further details and were assessed for eligibility. Of those, 153 practices finally declined to participate, 3 did not meet inclusion criteria and 7 were not able to create screening lists using their practice computer. Of the 72 included practices, 3478 IDs for potentially eligible patients were provided, from which a random sample of 1346 IDs was drawn at the study centre and sent to the practices. In total, 505 patients were consecutively included from the random sample and 465 completed the study (intervention group 238/252, control group 227/253) (flow chart: online supplementary appendix 3).

Of the 1325 practices that did not reply, we called 132 randomly selected practices. Six practices did not answer the phone, 51 were willing to answer all questions and 75 provided partial information. Sixty-one interviewed practices (48%) were not eligible (7 were not active GPs, 51 had no internet access and 3 declined to say). Practice characteristics and reasons for not responding are provided in online supplementary appendix 3.

### Baseline characteristics of participants

Most practices were single handed (57%), medium sized (64%) and located in small to mid-sized towns (57%). Slightly more male GPs (57%) participated; they were either specialists in general practice (83%) or in internal medicine. On average, they were 51 years of age, had more than 23 years of clinical experience and had worked in private practice for about 15 years. With one exception, HCAs were female. They averaged about 40 years of age, had about 17 years of clinical experience and had worked in the practice at various employment levels (49% less than full time) for an average of 10 years. About three-quarters were qualified HCAs. Patients were slightly more often female (52%), had a median age of 72 years and averaged eight prescriptions in nine single doses per day. Almost all patients were covered by statutory health insurance (96%), and looked after themselves (94%). Fifty-eight per cent participated in one of the national disease management programmes (DMP). Overall, baseline characteristics were well balanced in both groups (table 1).

### Outcomes

Our study found the intervention to have no significant effect. The mean MAI sum scores had decreased minimally in both groups 6 months after baseline—by 0.3 points in the intervention group and 0.8 points in the control group—revealing a non-significant adjusted mean difference of 0.7 (95% CI -0.2 to 1.6) points in favour of the control group (ITT, per-protocol analysis and BOCF approach did not differ). To control for the effects of oversampled patients registered in a DMP, we

compared DMP participants with non-participants, which revealed no effects on MAI. Furthermore, sociodemographic factors did not have an influence (table 2).

To explore our results, we conducted additional, non-prespecified analyses. As the sample size was not sufficiently large to perform subgroup analyses, we calculated multilevel models, which revealed strong effects of the baseline values of MAI sum scores on the primary outcome MAI T1–T0 ( $P<0.001$ ) (figure 2A). The figure also shows the low proportion of patients with high inappropriateness at baseline, and the size and direction of the MAI changes in both groups after 6 months. To explain the relationship between the number of prescriptions and MAI values, we conducted exploratory regression analysis, which approximately revealed a square function (figure 2A).

Secondary outcomes showed small, non-significant changes. In the intervention group, patients' self-reported quality of life improved minimally (about 2.3% in EQ-5D, 0.5 years in both expected and desired lifetime) after 6 and 9 months, whereas it continued to decline in the control group (figure 3). Additionally, in the intervention group the mean number of hospital stays decreased and the mean number of days spent in hospital had dropped by half after 6 months, but in both groups the event rate was too small to show significant differences (ITT analyses of the primary and secondary outcomes: table 2, descriptive analysis of symptoms for potential ADRs: online supplementary appendix 4).

## DISCUSSION

### Key findings of the study

This study found the complex PRIMUM intervention to have no significant effects in older patients with multimorbidity and polypharmacy in general practice. At baseline, many patients already received appropriate prescriptions and enjoyed good quality of life and functional status. We can therefore conclude that in our study, there was not enough scope for improvement.

### Strengths and limitations of study

The systematic development and stepwise evaluation of the PRIMUM intervention in accordance with MRC guidance on complex interventions<sup>84</sup> was a strength as demonstrated by refinements in the design of the main trial, based on the results of pilot testing.<sup>35</sup> Recruitment to target, random sampling of patients, minimal attrition (we lost one cluster to follow-up because the GP moved to another town) and adherence to the protocol are additional strengths when compared with previous studies.<sup>85 86</sup> However, our study also had several limitations.

First, there is no agreed definition of polypharmacy and patient inclusion at the numerical threshold of  $\geq 5$  prescriptions was somewhat arbitrary,<sup>87 88</sup> but using a higher threshold would have meant losing patients whose medication was highly inappropriate (figure 2B). Moreover, the association between the number of prescriptions

and health outcomes is not linear: Payne and coauthors found only the most extreme levels of polypharmacy to be associated with increased admission rates in patients with multimorbidity,<sup>89</sup> while Gnjidic and her coresearchers identified the best discriminating threshold to be between 4.5 and 6.5 medicines for associations with frailty, disability, mortality and falls.<sup>90</sup>

Second, our study population may limit the generalisability of the results. Our study was population based and involved no preselection, and the response rate of practices was low. We cannot rule out that relatively ambitious GPs volunteered more frequently. As far as the choice of patients is concerned, we took a random sample within each practice and our selection criteria aimed at including a broad range of diseases involving as many organ systems as possible. We applied the cognition test during recruitment and after consent. As our ultimate aim was to promote regular practice consultations, we excluded patients with dementia. The study required that patients who were unable to fill in questionnaires or to answer telephone calls should not attend (eg, some nursing home residents and migrants). These groups may therefore have been under-represented. To enable random sampling, we applied a systematic case finding using prescription costs as a proxy but oversampled DMP participants. However, German DMPs do not address multimorbidity or polypharmacy and we did not find any DMP impact on outcomes in our study.

Third, our outcome measures were slightly insensitive. In the intervention group, the increase in the average number of prescriptions indicates that GPs had more often begun to prescribe patients a new medicine. If undertreatment had been a key problem in our study, having the MAI as the main outcome variable would have led us to underestimate its impact, because it does not reliably detect underuse.<sup>42</sup> It is noteworthy that the number of medicines used in intervention and control groups had diverged after 6 and 9 months, with the adjusted mean number of drugs being 1.0 higher in the intervention group (table 2). Figure 2B shows that the more drugs a physician prescribes, the greater the chance that the MAI score will increase. The intervention may have induced increased prescribing of medicines (eg, in case of otherwise undetected underuse), which may explain the trend towards smaller reductions of the MAI scores in the intervention group.

Fourth, our efforts to reduce contamination of controls by using a cluster randomised design and withholding intervention details may have been substantially offset by a potentially important Hawthorne effect, as has been noted in other studies.<sup>85 91</sup> GPs and HCAs collected extensive data on medication, diseases and laboratory parameters (see icon 'f' in figure 1) at each study visit. It can be assumed that data collection will have had the same effect as the structured medication reviews: we also observed improvements in MAI mean values in the control group at the first follow-up (figure 2A), and a slight decrease in the average numbers of prescriptions. The net effect



was that the decrease in MAI scores in the control group was slightly larger than in the intervention group where it had been partly offset by an increase in the number of prescriptions (and higher MAI scores) resulting from identified underuse. However, the differences were very small.

### Comparison with other studies

Most primary care studies have investigated pharmacist-led interventions, and have shown inconclusive results in various outcomes.<sup>33 92–96</sup> However, pharmacist-led interventions may be difficult to implement in health-care contexts in which pharmacists have no access to clinical information (eg, patients' diagnoses, laboratory tests), patients often visit many different pharmacies and interprofessional relationships between GPs and pharmacists are not well established, as in Germany.<sup>85 86</sup> In this context in particular, information technology systems have been identified by European GPs as supporting safer prescribing.<sup>97–99</sup> Further factors that have been addressed include support from other healthcare professionals such as nurses, systematic medication reviews and greater involvement of the patient.<sup>97–99</sup> However, the efficacy of these measures is inconclusive: Olsson and coinvestigators found that a physician-led medication review had no effect on indicators of high-risk prescribing in older patients with polypharmacy.<sup>100</sup> In contrast, a large-scale cluster RCT achieved reductions in unintentional drug duplications, drug–drug interactions and new prescriptions of potentially inappropriate medications, but failed to show an impact on the discontinuation of inappropriate medicines.<sup>101</sup>

No evidence yet exists that polypharmacy interventions lead to decrease in mortality and hospitalisations,<sup>94</sup> functional decline and falls<sup>102 103</sup> and health-related quality of life.<sup>85 86 100 104–107</sup> A recent meta-analysis revealed a modest reduction in the number of drugs (on average –0.2 in the intervention group vs +0.2 in controls) but the results of the included studies differed widely<sup>94</sup> and, considering the frequency and potential impact of medication underuse,<sup>6–8</sup> a reduction in net prescription numbers is an ambiguous study endpoint.

### Possible explanations and implications of the study

Our study showed the intervention to have no significant effect. We cannot rule out that there was not enough scope for improvement in our study (figure 2A: the MAI of the patients included in the left two box plots in both groups could not improve). Additionally, there was a relevant Hawthorne effect (figure 2A: the patients included in the four box plots of the control group on the right hand side also improved). The patients depicted in the four box plots of the intervention group on the right hand side (figure 2A) improved less than corresponding patients in the control group, which probably reflects the small numbers of patients and the lack of an intervention effect. In addition, given the MAI's inability to detect changes in inappropriate underuse, it may have not been

sensitive enough for the purpose of our study. As any newly prescribed drug worsens the MAI score, unless it is completely appropriate, this may at least partially explain the difference. Ongoing process evaluation concerning medication changes may provide further explanations of the outcomes and information on the implications of the study.

Further research is needed to identify patients who stand to benefit significantly from an intervention that aims to support the care of complex patients with multimorbidity and high treatment burden.<sup>108 109</sup> Future studies may also benefit from considering a refined choice of outcome measures that adequately takes underuse into account.

### CONCLUSION

We did not find the intervention to have significant effects. Many patients already received appropriate prescriptions and enjoyed good quality of life and functional status. We can therefore conclude that in our study, there was not enough scope for improvement. Further research should seek to identify groups of patients who are most likely to benefit from such resource-intensive interventions. Outcome measures should be patient relevant and detect changes in underuse.

### Author affiliations

<sup>1</sup>Institute of General Practice, Johann Wolfgang Goethe University, Frankfurt, Germany

<sup>2</sup>Institute of Medical Biometry and Informatics, University of Heidelberg, Heidelberg, Germany

<sup>3</sup>Department of Clinical Pharmacology and Pharmacoepidemiology, University of Heidelberg, Heidelberg, Germany

<sup>4</sup>Department of Family Medicine, School CAPHRI, Maastricht University, Maastricht, The Netherlands

<sup>5</sup>Department of Public Health and Primary Care, Academic Center for General Practice, KU Leuven, Leuven, Belgium

<sup>6</sup>Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK

<sup>7</sup>Interdisciplinary Ageing Research (IAW), Faculty of Educational Sciences, Johann Wolfgang Goethe University, Frankfurt, Germany

<sup>8</sup>APEX Collaboration for Academic Primary Care, University of Exeter Medical School, Exeter, UK

<sup>9</sup>Institute for Clinical Pharmacology, Johann Wolfgang Goethe University Hospital, Frankfurt / Main, Germany

**Acknowledgements** The authors thank all participating patients and all general practice investigators and their teams. The authors are grateful to our dedicated research team Petra Lödige, Zeycan Albay, Mareike Leifermann, Anja Paesel and Natalie Aksamit, as well as to the team who was involved in developing and programming the study version of the CDSS, namely Kristina Zint, Diana Witticke, Jens Kaltschmidt and Michael Metzner. The authors also thank our practice advisory board Joachim Fessler, Joachim Seffrin, Karola Mergenthal and Vera Müller. This study was only possible thanks to their extraordinary engagement. The authors thank Cornelia Mahler for the provision of the German version of the BMQ and Roman Kaspar for his advice on quality of life measures in gerontology. The authors also thank Professor Paul Glasziou from CREBP, Bond University, QLD, Australia, for his comments on a former version of the manuscript.

**Contributors** CM drafted the manuscript, coordinated the study and contributed to the conception, design, data collection and data analyses. SH contributed to the conception and design, and conducted the MAI ratings. FMG was the guarantor of the study. JR and LU contributed to the conception and design, and conducted the data analyses. WEH contributed to the conception and design, and provided the study version of CDSS. CG, MB, FO, RP, MvdA, JAK, JMV and FMG provided specific

advice on the conception, methods and coordination of the study. All authors critically revised and agreed on the final version of the manuscript.

**Funding** Funding has been provided by the German Federal Ministry of Education and Research, BMBF (grant number 01GK0702).

**Competing interests** CM, FMG, CG, MB, SH, WEH, JR and LU report receiving grants from the German Federal Ministry of Education and Research, BMBF, grant number 01GK0702, during the course of the study. WEH reports other grants from Dosing, Heidelberg, during the course of the study; he received personal fees, non-financial support and other from Aqua Institute Göttingen; personal fees, non-financial support and other from Aspen Europe; personal fees and other from Diaplan; grants, personal fees, non-financial support and other from Actelion; personal fees and other from GSK GER/UK/Slovakia/France; personal fees from Thieme Verlag; personal fees and other from Daiichi Sankyo; personal fees and other from Bristol-Myers Squibb; personal fees and other from MSD Sharp & Dohme; personal fees and other from AstraZenica; personal fees and other from Boehringer; personal fees and other from Grünenthal; personal fees and other from KWHC; personal fees and other from Novartis; personal fees and other from Berlin-Chemie; grants, personal fees and other from Landesapothekerkammer BW; grants and other from BMBF (DZIF, ESTHER); grants, personal fees, non-financial support and other from BayerPharma; grants from CHIESI; personal fees and other from Doctrina Med; personal fees and other from GSK France, UK, Germany, Slovakia; personal fees, non-financial support and other from Pfizer; grants from Smooth ClinicalTrials; grants from Sumaya Biotech; grants from Klaus Tschira Stiftung; other from University Frankfurt; grants from Vaximm, outside the submitted work. FO, MvdA, JMV, RP and JAK have nothing to disclose.

**Patient consent** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** No additional data available.

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

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

## REFERENCES

- van den Akker M, Buntinx F, Knottnerus J. Comorbidity or multimorbidity: what's in a name. A review of literature. *Eur J Gen Pract* 1996;2:65–70.
- Fortin M, Stewart M, Poitras ME, et al. A systematic review of prevalence studies on multimorbidity: toward a more uniform methodology. *Ann Fam Med* 2012;10:142–51.
- Salisbury C, Johnson L, Purdy S, et al. Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study. *Br J Gen Pract* 2011;61:12–21.
- Violan C, Foguet-Boreu Q, Flores-Mateo G, et al. Prevalence, determinants and patterns of multimorbidity in primary care: a systematic review of observational studies. *PLoS One* 2014;9:e102149.
- Kuijpers MA, van Marum RJ, Egberts AC, et al. Relationship between polypharmacy and underprescribing. *Br J Clin Pharmacol* 2008;65:130–3.
- Meid AD, Quinzler R, Freigofas J, et al. Medication underuse in aging outpatients with cardiovascular disease: prevalence, determinants, and outcomes in a prospective cohort study. *PLoS One* 2015;10:e0136339.
- Meid AD, Quinzler R, Groll A, et al. Longitudinal evaluation of medication underuse in older outpatients and its association with quality of life. *Eur J Clin Pharmacol* 2016;72:877–85.
- Meid AD, Haefeli WE. Age-dependent impact of medication underuse and strategies for improvement. *Gerontology* 2016;62:491–9.
- Nobili A, Marengoni A, Tettamanti M, et al. Association between clusters of diseases and polypharmacy in hospitalized elderly patients: results from the REPOSI study. *Eur J Intern Med* 2011;22:597–602.
- Steinman MA, Landefeld CS, Rosenthal GE, et al. Polypharmacy and prescribing quality in older people. *J Am Geriatr Soc* 2006;54:1516–23.
- Campbell SE, Seymour DG, Primrose WR, et al. A multi-centre European study of factors affecting the discharge destination of older people admitted to hospital: analysis of in-hospital data from the ACMEplus project. *Age Ageing* 2005;34:467–75.
- Deandrea S, Lucenteforte E, Bravi F, et al. Risk factors for falls in community-dwelling older people: a systematic review and meta-analysis. *Epidemiology* 2010;21:658–68.
- Pirmohamed M, James S, Meakin S, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. *BMJ* 2004;329:15–19.
- Schneeweiss S, Rassen JA, Glynn RJ, et al. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology* 2009;20:512–22.
- Heinrich S, Rapp K, Rissmann U, et al. Cost of falls in old age: a systematic review. *Osteoporos Int* 2010;21:891–902.
- Rottenkolber D, Schmiedl S, Rottenkolber M, et al. Adverse drug reactions in Germany: direct costs of internal medicine hospitalizations. *Pharmacoepidemiol Drug Saf* 2011;20:626–34.
- Schneeweiss S, Hasford J, Göttinger M, et al. Admissions caused by adverse drug events to internal medicine and emergency departments in hospitals: a longitudinal population-based study. *Eur J Clin Pharmacol* 2002;58:285–91.
- Davies EC, Green CF, Taylor S, et al. Adverse drug reactions in hospital in-patients: a prospective analysis of 3695 patient-episodes. *PLoS One* 2009;4:e4439.
- Hakkaraian KM, Gyllenstein H, Jönsson AK, et al. Prevalence, nature and potential preventability of adverse drug events – a population-based medical record study of 4970 adults. *Br J Clin Pharmacol* 2014;78:170–83.
- Thomsen LA, Winterstein AG, Søndergaard B, et al. Systematic review of the incidence and characteristics of preventable adverse drug events in ambulatory care. *Ann Pharmacother* 2007;41:1411–26.
- Dreischulte T, Donnan P, Grant A, et al. Safer prescribing—a trial of education, informatics, and financial incentives. *N Engl J Med* 2016;374:1053–64.
- Avery AJ, Rodgers S, Cantrill JA, et al. A pharmacist-led information technology intervention for medication errors (PINCER): a multicentre, cluster randomised, controlled trial and cost-effectiveness analysis. *Lancet* 2012;379:1310–9.
- Guthrie B, Kavanagh K, Robertson C, et al. Data feedback and behavioural change intervention to improve primary care prescribing safety (EFIPPS): multicentre, three arm, cluster randomised controlled trial. *BMJ* 2016;354:i4079.
- Clyne B, Fitzgerald C, Quinlan A, et al. Interventions to address potentially inappropriate prescribing in community-dwelling older adults: a systematic review of randomized controlled trials. *J Am Geriatr Soc* 2016;64:1210–22.
- May C, Montori VM, Mair FS. We need minimally disruptive medicine. *BMJ* 2009;339:b2803.
- Anonymus. Guiding principles for the care of older adults with multimorbidity: an approach for clinicians: American geriatrics society expert panel on the care of older adults with multimorbidity. *J Am Geriatr Soc* 2012;60:E1–E25.
- Scottish Government Model of Care Polypharmacy Working Group. Polypharmacy guidance 2015. <http://www.sehd.scot.nhs.uk/publications/DC20150415polypharmacy.pdf> (accessed 14 Nov 2017).
- Bergert FW, Conrad D, Guidelines Group Hesse Pharmacotherapy Guidelines by Family Doctors for Family Doctors. Pharmacotherapy guidelines for the aged by family doctors for the use of family doctors – Part D basic conditions supporting drug treatment. Part E guidelines group, disclaimer, internet addresses. *Int J Clin Pharmacol Ther* 2009;47:366–73.
- Muth C, Beyer M, Fortin M, et al. Multimorbidity's research challenges and priorities from a clinical perspective: the case of 'Mr Curran'. *Eur J Gen Pract* 2014;20:139–47.
- Muth C, van den Akker M, Blom JW, et al. The Ariadne principles: how to handle multimorbidity in primary care consultations. *BMC Med* 2014;12:223.
- Valderas JM. Getting impatient about person-centred health care. *Eur J Gen Pract* 2013;19:141–2.
- Wallace E, Salisbury C, Guthrie B, et al. Managing patients with multimorbidity in primary care. *BMJ* 2015;350:h176.
- Patterson SM, Cadogan CA, Kerse N, et al. Interventions to improve the appropriate use of polypharmacy for older people. *Cochrane Database Syst Rev* 2014;10:CD008165.
- Smith SM, Wallace E, O'Dowd T, et al. Interventions for improving outcomes in patients with multimorbidity in primary

- care and community settings. *Cochrane Database Syst Rev* 2016;3:CD006560.
35. Muth C, Harder S, Uhlmann L, et al. Pilot study to test the feasibility of a trial design and complex intervention on PRIoritising MULtimedication in Multimorbidity in general practices (PRIMUMpilot). *BMJ Open* 2016;6:e011613.
  36. Bosley S, Dale J. Healthcare assistants in general practice: practical and conceptual issues of skill-mix change. *Br J Gen Pract* 2008;58:118–24.
  37. Gensichen J, von Korff M, Peitz M, et al. Case management for depression by health care assistants in small primary care practices: a cluster randomized trial. *Ann Intern Med* 2009;151:369–78.
  38. Gensichen J, Jaeger C, Peitz M, et al. Health care assistants in primary care depression management: role perception, burdening factors, and disease conception. *Ann Fam Med* 2009;7:513–9.
  39. Peters-Klimm F, Müller-Tasch T, Schellberg D, et al. Rationale, design and conduct of a randomised controlled trial evaluating a primary care-based complex intervention to improve the quality of life of heart failure patients: HICMan (Heidelberg Integrated Case Management). *BMC Cardiovasc Disord* 2007;7:25.
  40. Rosemann T, Körner T, Wensing M, et al. Rationale, design and conduct of a comprehensive evaluation of a primary care based intervention to improve the quality of life of osteoarthritis patients. The PraxArt-project: a cluster randomized controlled trial [ISRCTN87252339]. *BMC Public Health* 2005;5:77.
  41. Hanlon JT, Schmader KE, Samsa GP, et al. A method for assessing drug therapy appropriateness. *J Clin Epidemiol* 1992;45:1045–51.
  42. Hanlon JT, Schmader KE. The medication appropriateness index at 20: where it started, where it has been, and where it may be going. *Drugs Aging* 2013;30:893–900.
  43. Spinewine A, Schmader KE, Barber N, et al. Appropriate prescribing in elderly people: how well can it be measured and optimised? *Lancet* 2007;370:173–84.
  44. Hudon C, Fortin M, Soubhi H. Abbreviated guidelines for scoring the Cumulative Illness Rating Scale (CIRS) in family practice. *J Clin Epidemiol* 2007;60:212.e1–212.e4.
  45. Linn BS, Linn MW, Gurel L. Cumulative illness rating scale. *J Am Geriatr Soc* 1968;16:622–6.
  46. Bergert FW, Braun M, Clarius H, et al. Hausärztliche Leitlinie Geriatrie - Teil 1: Allgemeine Geriatrie. Teil 2: Spezielle Geriatrie. Leitliniengruppe Hessen 2008;PMV Leitliniengruppe Hessen 2017. [http://www.pmvforschungsguppe.de/pdf/03\\_publicationen/geriatrie1\\_ll.pdf](http://www.pmvforschungsguppe.de/pdf/03_publicationen/geriatrie1_ll.pdf); [http://www.pmvforschungsguppe.de/pdf/03\\_publicationen/geriatrie2\\_ll.pdf](http://www.pmvforschungsguppe.de/pdf/03_publicationen/geriatrie2_ll.pdf) (accessed 14 Nov 2017).
  47. Folstein MF, Folstein SE, McHugh PR, et al. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–98.
  48. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199–208.
  49. Graf von der Schulenburg JM, Claes C, Greiner W, et al. The German version of the EuroQol questionnaire. *Zeitschrift für Gesundheitswissenschaften (J Public Heal)* 1998;6:3–20.
  50. Saliba D, Elliott M, Rubenstein LZ, et al. The Vulnerable Elders Survey: a tool for identifying vulnerable older people in the community. *J Am Geriatr Soc* 2001;49:1691–9.
  51. Von Korff M, Ormel J, Keefe FJ, et al. Grading the severity of chronic pain. *Pain* 1992;50:133–49.
  52. Lang FR, Baltes PB, Wagner GG. Desired lifetime and end-of-life desires across adulthood from 20 to 90: a dual-source information model. *J Gerontol B Psychol Sci Soc Sci* 2007;62:P268–P276.
  53. Lawton MP, Moss M, Hoffman C, et al. Health, valuation of life, and the wish to live. *Gerontologist* 1999;39:406–16.
  54. Krones T, Keller H, Becker A, et al. The theory of planned behaviour in a randomized trial of a decision aid on cardiovascular risk prevention. *Patient Educ Couns* 2010;78:169–76.
  55. Man-Son-Hing M, Laupacis A, O'Connor AM, et al. A patient decision aid regarding antithrombotic therapy for stroke prevention in atrial fibrillation: a randomized controlled trial. *JAMA* 1999;282:737–43.
  56. Horne R, Weinman J, Hankins M. The beliefs about medicines questionnaire: the development and evaluation of a new method for assessing the cognitive representation of medication. *Psychol Health* 1999;14:1–24.
  57. Mahler C, Hermann K, Horne R, et al. Patients' beliefs about medicines in a primary care setting in Germany. *J Eval Clin Pract* 2012;18:409–13.
  58. Brauns H, Steinmann S. Educational reform in France, West-Germany, the United Kingdom and Hungary. Updating the CASMIN educational classification. *ZUMA-Nachrichten* 1999;44:7–44.
  59. Luck T, Riedel-Heller SG, Wiese B, et al. [CERAD-NP battery: Age-, gender- and education-specific reference values for selected subtests. Results of the German Study on Ageing, Cognition and Dementia in Primary Care Patients (AgeCoDe)]. *Z Gerontol Geriatr* 2009;42:372–84.
  60. Sheikh JI, Yesavage JA, Brooks JO, et al. Proposed factor structure of the geriatric depression scale. *Int Psychogeriatr* 1999;3:23–8.
  61. Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res* 1982;17:37–49.
  62. Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. *Med Care* 1986;24:67–74.
  63. Bergk V, Haefeli WE, Gasse C, et al. Information deficits in the summary of product characteristics preclude an optimal management of drug interactions: a comparison with evidence from the literature. *Eur J Clin Pharmacol* 2005;61:327–35.
  64. Bertsche T, Fleischer M, Pfaff J, et al. Pro-active provision of drug information as a technique to address overdosing in intensive-care patients with renal insufficiency. *Eur J Clin Pharmacol* 2009;65:823–9.
  65. Bertsche T, Pfaff J, Schiller P, et al. Prevention of adverse drug reactions in intensive care patients by personal intervention based on an electronic clinical decision support system. *Intensive Care Med* 2010;36:665–72.
  66. Falconnier AD, Haefeli WE, Schoenenberger RA, et al. Drug dosage in patients with renal failure optimized by immediate concurrent feedback. *J Gen Intern Med* 2001;16:369–75.
  67. Martin-Facklam M, Rengelshausen J, Tayrouz Y, et al. Dose individualisation in patients with renal insufficiency: does drug labelling support optimal management? *Eur J Clin Pharmacol* 2005;60:807–11.
  68. Martin P, Haefeli WE, Martin-Facklam M. A drug database model as a central element for computer-supported dose adjustment within a CPOE system. *J Am Med Inform Assoc* 2004;11:427–32.
  69. Quinzler R, Schmitt SPW, Szecsenyi J, et al. Optimizing information on drug exposure by collection of package code information in questionnaire surveys. *Pharmacoepidemiol Drug Saf* 2007;16:1024–30.
  70. Perera R, Heneghan C, Yudkin P. Graphical method for depicting randomised trials of complex interventions. *BMJ* 2007;334:127–9.
  71. Spinewine A, Dumont C, Mallet L, et al. Medication appropriateness index: reliability and recommendations for future use. *J Am Geriatr Soc* 2006;54:720–2.
  72. Barat I, Andreasen F, Damsgaard EMS. Drug therapy in the elderly: what doctors believe and patients actually do. *Br J Clin Pharmacol* 2001;51:615–22.
  73. George J, Phun YT, Bailey MJ, et al. Development and validation of the medication regimen complexity index. *Ann Pharmacother* 2004;38:1369–76.
  74. Hanlon JT, Weinberger M, Samsa GP, et al. A randomized, controlled trial of a clinical pharmacist intervention to improve inappropriate prescribing in elderly outpatients with polypharmacy. *Am J Med* 1996;100:428–37.
  75. Cohen J. *Statistical power analysis for the behavioral sciences*. Hillsdale: Lawrence Erlbaum Associates, 1988.
  76. Adams G, Gulliford MC, Ukoumunne OC, et al. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *J Clin Epidemiol* 2004;57:785–94.
  77. Campbell MK, Piaggio G, Elbourne DR, et al. Consort 2010 statement: extension to cluster randomised trials. *BMJ* 2012;345:e5661.
  78. R Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2012. <https://www.r-project.org/> (accessed 14 Nov 2017).
  79. Dahl DB. xtable: Export tables to LaTeX or HTML; 2012. R package version 1.7-0. 2012 <https://CRAN.R-project.org/package=xtable> (accessed 14 Nov 2017).
  80. Pinheiro J, Bates D, DebRoy S, et al. Linear and nonlinear mixed effects models; 2012. R package version 3.1-103. 2012 <https://cran.r-project.org/web/packages/nlme/index.html> (accessed 14 Nov 2017).
  81. Bates D, Maechler M, Bolker B. lme4: Linear mixed-effects models using Eigen and Eigen; 2012. R package version 0.999999-0. 2017 <https://CRAN.R-project.org/package=lme4> (accessed 14 Nov 2017).
  82. Bliese P. multilevel: Multilevel Functions; 2012. R package version 2.4. 2017 <https://CRAN.R-project.org/package=multilevel> (accessed 14 Nov 2017).
  83. Fletcher TD. Applied psychometric theory; 2010. R package version 2.2. 2017 <http://CRAN.R-project.org/package=psychometric> (accessed 14 Nov 2017).



84. Craig P, Dieppe P, Macintyre S, *et al.* Developing and evaluating complex interventions: the new medical research council guidance. *BMJ* 2008;337:a1655.
85. Bryant LJM, Coster G, Gamble GD, *et al.* The General Practitioner-Pharmacist Collaboration (GPPC) study: a randomised controlled trial of clinical medication reviews in community pharmacy. *Int J Pharm Pract* 2011;19:94–105.
86. Leendertse AJ, de Koning GHP, Goudswaard AN, *et al.* Preventing hospital admissions by reviewing medication (PHARM) in primary care: an open controlled study in an elderly population. *J Clin Pharm Ther* 2013;38:379–87.
87. Duerden M, Avery T, Payne R. *Polypharmacy and medicines optimisation. Making it safe and sound*: The King's Fund, 2013. [http://www.kingsfund.org.uk/sites/files/kf/field/field\\_publication\\_file/polypharmacy-and-medicines-optimisation-kingsfund-nov13.pdf](http://www.kingsfund.org.uk/sites/files/kf/field/field_publication_file/polypharmacy-and-medicines-optimisation-kingsfund-nov13.pdf) (accessed 14 Nov 2017).
88. Sönnichsen A, Trampisch US, Rieckert A, *et al.* Polypharmacy in chronic diseases-Reduction of Inappropriate Medication and Adverse drug events in older populations by electronic Decision Support (PRIMA-eDS): study protocol for a randomized controlled trial. *Trials* 2016;17:57.
89. Payne RA, Abel GA, Avery AJ, *et al.* Is polypharmacy always hazardous? A retrospective cohort analysis using linked electronic health records from primary and secondary care. *Br J Clin Pharmacol* 2014;77:1073–82.
90. Gnjidic D, Hilmer SN, Blyth FM, *et al.* Polypharmacy cutoff and outcomes: five or more medicines were used to identify community-dwelling older men at risk of different adverse outcomes. *J Clin Epidemiol* 2012;65:989–95.
91. Allard J, Hébert R, Rioux M, *et al.* Efficacy of a clinical medication review on the number of potentially inappropriate prescriptions prescribed for community-dwelling elderly people. *CMAJ* 2001;164:1291–6.
92. Geurts MME, Talsma J, Brouwers JRB, *et al.* Medication review and reconciliation with cooperation between pharmacist and general practitioner and the benefit for the patient: a systematic review. *Br J Clin Pharmacol* 2012;74:16–33.
93. Holland R, Desborough J, Goodyer L, *et al.* Does pharmacist-led medication review help to reduce hospital admissions and deaths in older people? A systematic review and meta-analysis. *Br J Clin Pharmacol* 2008;65:303–16.
94. Johansson T, Abuzahra ME, Keller S, *et al.* Impact of strategies to reduce polypharmacy on clinically relevant endpoints: a systematic review and meta-analysis. *Br J Clin Pharmacol* 2016;82:532–48.
95. Nkansah N, Mostovetsky O, Yu C, *et al.* Effect of outpatient pharmacists' non-dispensing roles on patient outcomes and prescribing patterns. *Cochrane Database Syst Rev* 2010;7:CD000336.
96. Thomas R, Huntley AL, Mann M, *et al.* Pharmacist-led interventions to reduce unplanned admissions for older people: a systematic review and meta-analysis of randomised controlled trials. *Age Ageing* 2014;43:174–87.
97. Anthierens S, Tansens A, Petrovic M, *et al.* Qualitative insights into general practitioners' views on polypharmacy. *BMC Fam Pract* 2010;11.
98. Riordan DO, Byrne S, Fleming A, *et al.* GPs' perspectives on prescribing for older people in primary care: a qualitative study. *Br J Clin Pharmacol* 2017;83:1521–31.
99. Schuling J, Gebben H, Veehof LJG, *et al.* Deprescribing medication in very elderly patients with multimorbidity: the view of Dutch GPs. A qualitative study. *BMC Fam Pract* 2012;13:56.
100. Olsson IN, Runnamo R, Engfeldt P. Drug treatment in the elderly: an intervention in primary care to enhance prescription quality and quality of life. *Scand J Prim Health Care* 2012;30:3–9.
101. Tamblyn R, Huang A, Perreault R, *et al.* The medical office of the 21st century (MOXXI): effectiveness of computerized decision-making support in reducing inappropriate prescribing in primary care. *CMAJ* 2003;169:549–56.
102. Weber V, White A, McIlvried R. An electronic medical record (EMR)-based intervention to reduce polypharmacy and falls in an ambulatory rural elderly population. *J Gen Intern Med* 2008;23:399–404.
103. Williams ME, Pulliam CC, Hunter R, *et al.* The short-term effect of interdisciplinary medication review on function and cost in ambulatory elderly people. *J Am Geriatr Soc* 2004;52:93–8.
104. RESPECT trial team. Effectiveness of shared pharmaceutical care for older patients: RESPECT trial findings. *Br J Gen Pract* 2010;60:e10–e19.
105. Krska J, Cromarty JA, Arris F, *et al.* Pharmacist-led medication review in patients over 65: a randomized, controlled trial in primary care. *Age Ageing* 2001;30:205–11.
106. Sellors J, Kaczorowski J, Sellors C, *et al.* A randomized controlled trial of a pharmacist consultation program for family physicians and their elderly patients. *CMAJ* 2003;169:17–22.
107. Taylor CT, Byrd DC, Krueger K. Improving primary care in rural Alabama with a pharmacy initiative. *Am J Health Syst Pharm* 2003;60:1123–9.
108. Huntley AL, Johnson R, Purdy S, *et al.* Measures of multimorbidity and morbidity burden for use in primary care and community settings: a systematic review and guide. *Ann Fam Med* 2012;10:134–41.
109. Valderas JM, Starfield B, Sibbald B, *et al.* Defining comorbidity: implications for understanding health and health services. *Ann Fam Med* 2009;7:357–63.