

BMJ Open Development and internal validation of prognostic models to predict negative health outcomes in older patients with multimorbidity and polypharmacy in general practice

Beate S Müller ¹, Lorenz Uhlmann,² Peter Ihle,³ Christian Stock,² Fiona von Buedingen,¹ Martin Beyer,¹ Ferdinand M Gerlach,¹ Rafael Perera,⁴ Jose Maria Valderas,⁵ Paul Glasziou,⁶ Marjan van den Akker,^{1,7} Christiane Muth ¹

To cite: Müller BS, Uhlmann L, Ihle P, *et al.* Development and internal validation of prognostic models to predict negative health outcomes in older patients with multimorbidity and polypharmacy in general practice. *BMJ Open* 2020;**10**:e039747. doi:10.1136/bmjopen-2020-039747

► Prepublication history and additional material for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-039747>).

Received 24 April 2020

Revised 26 August 2020

Accepted 02 October 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Beate S Müller;
b.mueller@allgemeinmedizin.
uni-frankfurt.de

ABSTRACT

Background Polypharmacy interventions are resource-intensive and should be targeted to those at risk of negative health outcomes. Our aim was to develop and internally validate prognostic models to predict health-related quality of life (HRQoL) and the combined outcome of falls, hospitalisation, institutionalisation and nursing care needs, in older patients with multimorbidity and polypharmacy in general practices.

Methods *Design:* two independent data sets, one comprising health insurance claims data (n=592 456), the other data from the PRioritising MUltimedication in Multimorbidity (PRIMUM) cluster randomised controlled trial (n=502). *Population:* ≥60 years, ≥5 drugs, ≥3 chronic diseases, excluding dementia. *Outcomes:* combined outcome of falls, hospitalisation, institutionalisation and nursing care needs (after 6, 9 and 24 months) (claims data); and HRQoL (after 6 and 9 months) (trial data). *Predictor variables in both data sets:* age, sex, morbidity-related variables (disease count), medication-related variables (European Union-Potentially Inappropriate Medication list (EU-PIM list)) and health service utilisation. *Predictor variables exclusively in trial data:* additional socio-demographics, morbidity-related variables (Cumulative Illness Rating Scale, depression), Medication Appropriateness Index (MAI), lifestyle, functional status and HRQoL (EuroQol EQ-5D-3L). *Analysis:* mixed regression models, combined with stepwise variable selection, 10-fold cross validation and sensitivity analyses.

Results Most important predictors of EQ-5D-3L at 6 months in best model (Nagelkerke's R² 0.507) were depressive symptoms (−2.73 (95% CI: −3.56 to −1.91)), MAI (−0.39 (95% CI: −0.7 to −0.08)), baseline EQ-5D-3L (0.55 (95% CI: 0.47 to 0.64)). Models based on claims data and those predicting long-term outcomes based on both data sets produced low R² values. In claims data-based model with highest explanatory power (R²=0.16), previous falls/fall-related injuries, previous hospitalisations, age, number of involved physicians and disease count were most important predictor variables.

Conclusions Best trial data-based model predicted HRQoL after 6 months well and included parameters of

Strengths and limitations of this study

- We developed our predictive models using two completely different data sets—claims data and data primarily collected in a cluster-randomised trial.
- The claims data contained a large number of cases, enabling our models to include many possible predictors without any convergence issues.
- The trial data provided a rich set of potential predictor variables of high data quality and included data on patient-reported outcome measures, such as well-being and functional status.
- Both data sets have their own methodological limitations, such as imprecise claims data (collected for reimbursement purposes) and the trial's small sample size.
- The nature of the data meant neither data set could be used to validate a predictive model based on the other.

well-being not found in claims. Performance of claims data-based models and models predicting long-term outcomes was relatively weak. For generalisability, future studies should refit models by considering parameters representing well-being and functional status.

BACKGROUND

Currently, up to 80% of primary care consultations involve patients with multiple chronic conditions (multimorbidity).¹ A multiplicity of disorders in patients is associated with polypharmacy. Both multimorbidity and polypharmacy are recognised as a major challenge facing healthcare systems.^{2–5} Polypharmacy can increase the risk of mortality, hospitalisation^{6 7} and falls and fall-related injuries with resulting disability and loss of autonomy.^{8 9} It can also reduce cognitive and



physical function, as well as health-related quality of life (HRQoL).^{2 10}

The *number of drugs* increases the *probability* of adverse drug reactions, but the relationship is inconsistent, suggesting that the number of medications alone may not adequately indicate the quality of an individual's medication regimen.^{11 12} The kind of drugs prescribed plays an important role in the type of reaction, with certain medication classes, such as benzodiazepines, demonstrating a significant association with falls, and medications with anti-cholinergic properties being associated with impaired cognitive and physical function in elderly individuals.^{13 14} At a physician level, the cause of these negative health outcomes of polypharmacy may be inappropriate prescribing, including undertreatment.¹⁵⁻¹⁸ At a patient level, a high number of drugs and the complexity of a drug regimen is often associated with poor adherence,¹⁹ which may be exacerbated by the presence of depression and/or cognitive impairment.²⁰ Moreover, polypharmacy may also result in an accumulation of potentially inappropriate medications (PIMs).

Several complex interventions have been developed to optimise (inappropriate) polypharmacy. However, despite their evidence-based rationale, they have led to inconsistent improvements in process parameters of care and failed to impact patient-relevant outcomes.^{21 22} One possible reason for this is that the included populations are too heterogeneous in terms of their baseline risk and potentially achievable intervention effects. For example, the majority of the study population included in the PRIMUM (PRIoritising MUltimedication in Multi-morbidity) trial showed very good quality of life and functional status at baseline, even though participants had at least three chronic conditions affecting more than two organ systems, five or more chronic drug prescriptions and were 60 years of age or older. The authors therefore concluded that there was not enough room for improvement.²³ This highlights current difficulties in defining inclusion criteria in polypharmacy trials in such a way that selected populations have a considerable baseline risk and can be expected to benefit from the intervention. Moreover, as polypharmacy interventions tend to address inappropriate prescribing, healthcare coordination, and so on, they are generally complex.^{21 22} As the complex interventions are also resource-intensive, it would be preferable for a stratified approach to address patients that are at high risk of negative health outcomes and most likely to benefit from them.²⁴

The course of multimorbidity (and associated polypharmacy) has been characterised by a decline in well-being (eg, functional decline or worsening of quality of life due to inappropriate prescriptions and/or deterioration in one or more chronic diseases), interrupted by adverse events (eg, exacerbations of chronic diseases or adverse drug reactions).^{25 26} In order to identify a population at high risk, it is therefore necessary to predict a wide array of possible negative health outcomes. Several prognostic models have predicted single outcomes, mainly

mortality or unplanned hospital (re-)admission and to a lesser extent a future decline in quality of life, but no studies have investigated the risk for the above-mentioned combined endpoints, or involved polypharmacy-related predictors.²⁷

The aim of this exploratory study was to develop and internally validate prognostic models to predict the risk of adverse events or a decline in well-being in general practice patients with multimorbidity and polypharmacy, and to operationalise these negative health outcomes in terms of hospitalisation, falls, level of required nursing care, institutionalisation and HRQoL. The models were based on morbidity and medication-related variables, as well as socio-demographic characteristics and parameters of healthcare utilisation.

METHODS

We developed and internally validated prognostic models to identify key health problems linked with multimorbidity and associated polypharmacy (decline in well-being and adverse events: [figure 1](#)). (1) Based on claims data, we predicted the combined endpoint of hospitalisation, falls/fall-related injuries, need for nursing care, deterioration in the required level of care (nursing level) or institutionalisation, after 6, 9 and 24 months. (2) We predicted HRQoL after 6 and 9 months based on data from a cluster-randomised trial.²³

Design and setting/study samples

Two data sets were used in modelling:

Claims data obtained from the *Techniker Krankenkasse* (TK) statutory health insurance company between January 2012 and December 2014. TK is the largest statutory health insurer in Germany and provided health insurance to 8.1 million persons in 2012.²⁸ In accordance with Social Code book V, all statutory health insurance companies in Germany collect basic data on socio-demographics, details of pharmacological and non-pharmacological prescriptions and information on other health services utilisation and data on morbidity.

Trial data from the cluster-randomised PRIMUM trial²³ conducted in general practices in Hesse, Germany, from August 2010 to February 2012.

Population

Claims-based models: We aimed to use the same inclusion criteria for both data sets as far as possible. We therefore included health insurance claims data of older patients (≥ 60 years) with multimorbidity (at least three documented chronic diseases, from a list of 46 diagnoses and conditions, from 01 January 2012 to 31 December 2012)²⁹ and polypharmacy (at least five documented and concurrent prescriptions from 01 July 2012 to 31 December 2012). Included patients had to have been continuously insured by TK from 01 January 2012 to 31 December 2014 (except in case of death at any time after 31 December 2012) and had to have contacted a primary care provider

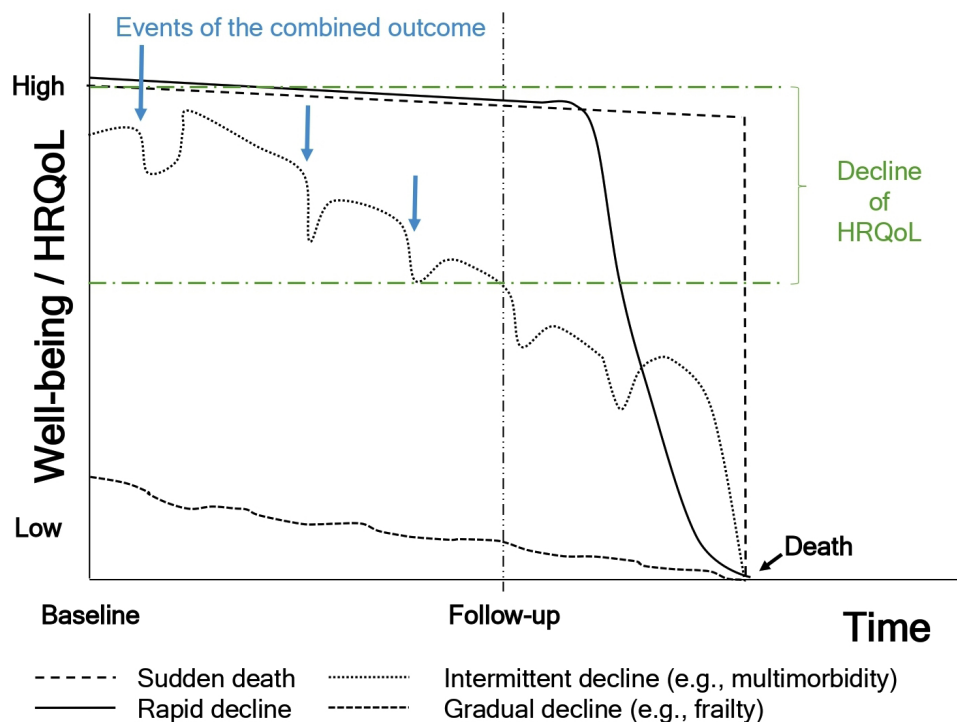


Figure 1 Predicted outcomes with regard to general trajectories of well-being and quality of life over time. HRQoL, health-related quality of life.

at least once in 2012. Patients were excluded if they were diagnosed with dementia (International Classification of Diseases, 10th Edition (ICD-10): F00-03, F05.1, G30-31, R54) or under guardianship from 01 January 2012 to 31 December 2012.

Trial data-based models: We included data from patients that participated in the cluster-randomised PRIMUM trial (n=502, intervention group: n=252, control group n=250).²³ Patients with multimorbidity and polypharmacy were included in the study if they were at least 60 years old, had at least three chronic diseases from two or more chapters of ICD-10 and at least five prescriptions. Patients were excluded if they were cognitively impaired (defined as a score lower or equal to 26 on the Mini-Mental Status Exam³⁰), had an alcohol or drug addiction or were not able to participate in telephone interviews, fill in questionnaires or express their own free will. Four out of the 502 patients (0.79%) died during the 9-month follow-up period.

Outcomes

Models based on claims data: We predicted the combined endpoint of hospitalisation, falls/fall-related injuries or institutionalisation in a long-term care facility, or if the need for nursing care was recognised, or the level of care ('Pflegestufe') had worsened at 6-month, 9-month, 24-month follow-up. We treated the parameters of health service use (hospitalisation, level of nursing care and institutionalisation) as surrogate parameters for a decline in functional status and well-being, as details of these are not included in German claims data. Outcomes were operationalised as follows:

- ▶ Hospitalisation: We included all-cause hospitalisations, as our data did not permit us to differentiate between unplanned and elective hospitalisations.
- ▶ Falls and fall-related injuries: We included all fractures and injuries coded in ICD-10 chapters 'S' and 'T'. We excluded ICD codes for severe body injuries such as S31 ('open wound of abdomen, lower back and pelvis'), which we assessed as related to severe bodily impact, rather than drug-related falls (see online supplemental additional file 1 for all excluded ICD codes). We also excluded osteoporosis-related fractures (ICD-10 M80).
- ▶ Institutionalisation was defined as the admission of a patient to a long-term care facility for at least 28 days (in Germany, this is the maximum length of time considered as 'short-term care' in such facilities).
- ▶ Level of (nursing) care ('Pflegestufe') referred to dependency on care. In the period under review, the German nursing care insurance system recognised four levels of care ('1' – lowest level to '3' – highest level, and 'H', which was mainly used for people with mental illnesses who are in need for support). The onset of care and any increase in care level were taken into consideration.

Models based on trial data: We predicted HRQoL 6 and 9 months after baseline. HRQoL was measured using the EQ-5D-3L index score.^{31–33} The EQ-5D-3L index score is a weighted summary score of five different dimensions of health (mobility, self-care, usual activities, pain/discomfort and anxiety/depression). Each dimension has three levels. The index score is calculated based on time

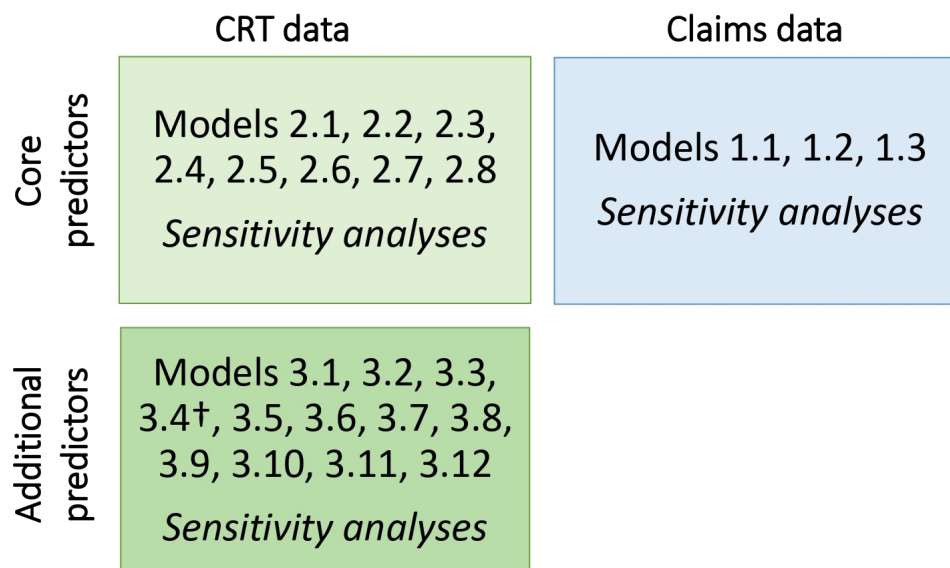


Figure 2 Models and sensitivity analyses with regard to data source and predictor set. CRT, cluster-randomised controlled trial; †Best Model.

trade-off (TTO) norm values and ranges from 0 to 1, with '0' signifying death and '1' in full health. Patients who died during follow-up were assigned the value '0'.

Potential predictors

The potential predictors that were initially used in the two modelling approaches were available in both claims and trial data (*core predictors*, see [figure 2](#)): To compare the two models, we first used these 'core predictors' (all variables were continuous variables, if not stated otherwise).

- ▶ Socio-demographics: Age (in years), sex (male/female, binary)
- ▶ Morbidity-related (excluding dementia): Number of chronic diseases (based on a modified list of 46 diagnoses and conditions),²⁹ Charlson comorbidity index,³⁴ number of specific chronic conditions according to Diederichs' list³⁵ consisting of 17 chronic diseases identified in a systematic review of existing comorbidity indices. As dementia was excluded, the final list contained 16 diagnoses. (All instruments including ICD-10 codes are provided in online supplemental additional file 2)
- ▶ Medication: Number of prescriptions (defined as Anatomical Therapeutic Chemical (ATC) agents using fifth-level coding, ATC version 2011 to 2014), excluding drugs for topical applications and drug groups that were irrelevant to our research question, for example, contrast agents (ATC V-08, three-digit level).
- ▶ Potentially inappropriate medication: We constructed two patient co-variables: (1) exposure to any PIM (yes/no) and (2) number of PIMs between 01 July 2012 and 31 December 2012 (*claims-based models*) and at baseline (*trial data-based models*). We used the following two lists to identify PIMs:

- *Modified EU-PIM list*³⁶: The list of PIMs for the elderly contains 282 chemical substances or drug classes divided into 34 therapeutic groups.
- *Modified PRISCUS list*³⁷: The German list of PIMs for the elderly includes 83 chemical substances from a total of 18 drug classes.

We excluded from the lists PIMs that referred to specific doses, treatment duration and disease severity, as valid information on these could not be obtained from the claims data. (All instruments including ATC codes are provided in online supplemental additional file 3)

- ▶ Anticholinergic drug burden: Scores were calculated based on all prescribed drugs with anticholinergic properties per patient. Despite substantial differences between existing scales, associations with adverse clinical outcomes, such as hospital admissions, fall-related hospitalisations, length of stays in hospital, and general practitioner (GP) visits, have been found for all of them.³⁸ As the evidence does not support the preferred use of any particular scale, we tested the following (all instruments including ATC codes are provided in online supplemental additional file 3):
 - *Anticholinergic Drug Scale (ADS)*³⁹: The ADS weights anticholinergic properties per drug from '0' – no anticholinergic activity, '1' – mild, '2' – moderate and '3' – strong anticholinergic activity. The overall anticholinergic burden per patient was calculated as a sum score for the entire medication regimen.
 - *Modified Anticholinergic Drug Burden Index (DBI)*¹³: The DBI comprises drugs with sedative effects (which form the sedative burden (B_S)), and drugs with anticholinergic or both sedative and anticholinergic effects (which form the anticholinergic burden (B_{AC})). As claims data do not provide

dosages, the cumulative number of sedative and anticholinergic drugs was calculated (modified DBI score).

- ▶ **Healthcare utilisation:** For each patient, we obtained information on all-cause hospitalisations (yes/no), falls and fall-related injuries (yes/no) and the number of physicians involved in ambulatory healthcare, between 01 January 2012 and 31 December 2012 for models based on *claims data*, and in the 6 months previous to baseline for models based on *trial data*.

Additional potential predictor variables were used exclusively to re-fit models based on trial data, as they were only available in these data ('additional predictors', see figure 2; all variables were continuous variables unless stated otherwise):

1. Socio-demographics: Education (CASMIN⁴⁰) and number of persons living in the household.
2. Lifestyle: Alcohol consumption (audit-C, categorical variables on number of drinking occasions and amount of alcohol consumed),⁴¹ smoking status (smoker/non-smoker, binary) and body mass index.
3. Inappropriateness of medication: MAI 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).⁴² 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 a trained clinical pharmacologist (SH), who rated nine items for each prescription. Values ranged from '0' (appropriate) to '2' (inappropriate) whereby '1' represented a middle rating of uncertain appropriateness. The assigned values were summed to give an MAI score between 0 and 18 for each prescription and across the entire medication regimen of the patient.²³
4. Morbidity-related: Severity of multimorbidity, as measured using the CIRS (the CIRS differentiates between 14 organ systems, which are assessed on a 5-point Likert scale according to severity of impairment, with the ratings ranging from no impairment to extreme impairment),⁴³ with scores calculated as the total sum score, the number of affected organ systems and the HRQoL-CI (HRQoL-CI consists of a mental and a physical subscale, whereby the presence of certain diseases are assigned weights from '1' to '3', see online supplemental additional file 2).⁴⁴
5. Depressive symptoms, as measured using the GDS with 15 items.⁴⁵
6. HRQoL at baseline, as measured using the EQ-5D-3L index score.^{31–33}

Missing values and imputation

There were no missing values in the claims data, so no imputation was carried out in models that were based on them. In models based on trial data, imputation of missing values in predictors and outcomes was conducted

using multiple imputation via chained equations.^{46 47} We used a fully conditional specification approach by setting up an appropriate conditional density for each variable. In the imputation process, we included all variables that were used in each model. We imputed $m=50$ data sets and combined the results using 'Rubin's rules'.⁴⁶

Statistical analyses

In both models, we first investigated the core predictors that were available in both data sets, including socio-demographics, morbidity-related and medication-related variables and variables for healthcare utilisation. We then refitted the trial data-based models using the additional predictors that were exclusively available for trial data, such as variables for lifestyle and well-being (figure 2).

Models based on claims data: In order to develop a prediction model for the binary combined outcome (containing all-cause hospitalisation, falls/fall-related injuries, institutionalisation or level of (nursing) care required) at 6-month, 9-month and 24-month follow-up, we performed multiple logistic regression analyses with the occurrence of at least one of the components at 6-month, 9-month and 24-month follow-up as the dependent variable. As patients were not always assigned a single general practice,⁴⁸ we did not perform cluster analysis on the claims data.

Models based on trial data: In order to develop a prediction model for the continuous outcome HRQoL at 6-month and 9-month follow-up, we performed multiple linear regression analyses using the EQ-5D-3L index score at 6-month and 9-month follow-up as the dependent variable. The cluster structure of the data was taken into account by including a random intercept to produce a mixed regression model. We assumed a compound symmetry structure when estimating the covariance matrix.

Univariate analyses in both claims and trial data: Prior to conducting regression analyses, we performed univariate analyses to identify any associations between our potential predictors (at baseline) and the outcomes (at 6-month, 9-month and 24-month follow-up).

Regression analyses and variable selection: To find out which predictor variables influence the outcome variables, we used a stepwise variable selection procedure (combining forward and backward steps). We started with the full model and all potential predictor variables. After this, we used a selection procedure based on p values.⁴⁹ In the backward selection step, we deleted the variable with the highest p value from the model if its p value was greater than 0.157. In the forward selection step, the variable with the lowest p value was included in the model if its p value was less than 0.156. As long as each covariate had only one df, the use of these boundaries led to the same results as variable selection using the Akaike Information Criterion.⁵⁰ The resulting models are presented by providing the estimated regression coefficients (models based on trial data) or ORs (models based on claims) with 95% CIs and corresponding p values. As we expected



the large sample size of claims-based models to result in low p values, we calculated additional z values and continuous net reclassification indices to gain information on the predictive power of each variable.⁵¹ Multi-collinearity was assessed using the variance inflation factor (VIF).⁵² In the models based on trial data, we did not account for the clustering structure when we calculated the VIF.

Performance of the models

We calculated R^2 for linear models based on trial data (according to Nakagawa and Schielzeth⁵³), and Nagelkerke's R^2 for logistic models (according to Steyerberg and Nagelkerke^{54 55}) based on claims data. Furthermore, in order to assess performance more realistically and to internally validate the models, we used the AUC (area under the receiver operator curve, equivalent to the concordance index) to validate the logistic regression model based on claims data, and R^2 to validate the linear regression model based on randomised controlled trial data, in combination with 10-fold cross-validation.⁵⁶ R^2 and Nagelkerke's R^2 are measures of the overall model's ability to assess explained variance. The AUC provides a measure of the model's discriminatory ability to distinguish patients at risk from those that are not.

Sensitivity analyses

Using sensitivity analysis, we applied two further modelling approaches (at first separately and then in combination): (1) modelling without multiple imputation and (2) modelling without variable selection.

Software: We made use of different statistical packages to analyse the data in R.^{47 57–63}

We used TRIPOD reporting guidelines (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis) in the preparation of this manuscript.⁶⁴

Patient and public involvement statement

Neither patients nor the public were involved in this study.

RESULTS

Participants

Claims data

The total sample of those ≥ 60 years that were continuously insured by TK from 01 January 2012 to 31 December 2014, and had at least one primary care contact during 2012, amounted to 1 377 917 persons. Overall, 592 456 patients met the pre-specified criteria and were included in the analyses (see study flow-chart, online supplemental additional file 4).

Trial data

Of the 505 patients that participated in the PRIMUM trial, all but 3 were 60 years or older. The final analyses therefore included 502 patients.

Key characteristics of study participants are shown in table 1.

Univariate analyses

In the claims data, univariate analyses revealed significant associations between the combined outcome and the following predictors: Age, sex, disease count, Charlson Comorbidity Index, EU-PIMs, ADS, DBI, previous hospitalisations, previous falls and number of physicians involved in the patient's care at all follow-ups (after 6, 9 and 24 months) (online supplemental additional file 5). In the trial data, HRQoL was significantly correlated with the shared predictor variables disease count, number of chronic prescriptions, previous falls and sex and the additional predictors depression and HRQoL at baseline (online supplemental additional file 6).

Prognostic models

Claims data

The model predicting the combined endpoint at 6 months had the highest C-statistic (AUC with 10-fold cross validation: 0.71, see table 2), but a low explanation of variance (Nagelkerke's R^2 without cross validation: 0.16). Variables in the model with the highest predictive power were previous falls/fall-related injuries and previous hospitalisations, as well as age, number of involved physicians, and number of chronic diseases ('disease count') (table 3). The models predicting the combined outcome at 9 and 24 months had AUCs calculated with 10-fold cross validation of 0.68 (R^2 without cross validation: 0.15) and 0.69 (R^2 without cross validation: 0.13) respectively. The VIF (to assess any multi-collinearity) showed moderate values (maximum 7.5).

Trial data

All results presented in this section are based on the modelling approach and involve multiple imputation of missing values and the variable selection procedure. Models predicting the HRQoL endpoint at 6 months that were based on core predictors available in both claims and trial data showed low predictive accuracy (R^2 with 10-fold cross validation: 0.111) (table 3, model 2.4). HRQoL at 6 months was best predicted when additional predictors that were exclusively available in the trial data were also included (R^2 with 10-fold cross validation: 0.507). The variables with the highest predictive power were depressive symptoms (GDS) and EQ-5D-3L Index Score (Baseline). MAI was also predictive (table 3, model 3.4). The VIF showed small values (maximum 2.2).

Comparison of model quality and sensitivity analyses

The shorter the time span of the prediction, the better the explained variance and hence, the performance of the model. However, model performance remained fair to poor when it only included predictor variables that were available for both claims and trial data. Sensitivity analyses confirmed these results (table 2).

Table 1 Characteristics of study participants

Characteristic	Claims data* n=592 456	CRT data* n=502
Data collection period	January 2012 to December 2014†	August 2010 to February 2012
Study design	Cohort study	Cluster-randomised controlled trial
Setting	Claims data from the TK health insurance fund. TK serves about 10 million people in Germany	72 general practices in Hesse, Germany
Inclusion criteria	≥60 years ≥3 chronic diseases ≥5 prescriptions ≥1 GP visit Continuously insured (except in case of death in follow-up period)	≥60 years ≥3 chronic diseases ≥5 prescriptions ≥1 GP visit
Exclusion criteria	Person under legal guardianship Diagnosed dementia	Person under legal guardianship Cognitive dysfunction including dementia (MMSE ≤26)
Outcomes to be predicted	Combined‡ binary outcome after 6-month, 9-month, 24 month follow-up	HRQoL (continuous outcome) after 6-month and 9-month follow-up
Potential predictors in both samples at baseline¶		
Age (years)	71.3 (7.06)	72.2 (6.86)
Male sex (n, %)	319 453 (54)	240 (48)
Morbidity		
Disease count	9.7 (3.75)	9.6 (3.25)
No. of specific chronic diseases (Diederichs)	4.3 (1.97)	4.1 (1.60)
CCI	3.0 (2.54)	2.6 (1.92)
HRQoL-CI, mental	2.8 (2.12)	2.1 (1.81)
HRQoL-CI, physical	8.0 (3.57)	7.6 (3.12)
Medication		
No. of drugs	8.6 (3.80)	8.1 (2.57)
No. of PIM (EU-PIM)	1.1 (1.15)	0.9 (0.96)
ACh burden (ADS)	1.0 (1.45)	0.8 (1.21)
Modified Drug Burden Index	0.8 (1.03)	0.5 (0.77)
No. of involved physicians	9.95 (5.26)	2.6 (1.77)
Previous hospitalisation:		
Patients that have undergone hospital treatment (n, %)	194 984 (33)	81 (16)§
No. of hospitalisations	1.67 (1.25)	1.5 (0.86)§
No. of days in hospital	14.5 (18.20)	17 (12.66)§
Patients with previous falls/fall-related injuries (n, %)	163 387 (28)	83 (17)§
Patients requiring nursing care		
Any nursing level (n, %)	28 310 (5)	–
Nursing level 1 (n, %)	19 030 (3)	–
Nursing level 2 (n, %)	7968 (1)	–
Nursing level 3 (n, %)	1273 (0.2)	–
Nursing level H (n, %)	39 (0.007)	–
Additional predictor variables in CRT data at baseline¶		
Socio-demographics		

Continued



Table 1 Continued

Characteristic	Claims data* n=592 456	CRT data* n=502
Educational level (CASMIN)	–	1.4 (0.66)
No. of persons living in household	–	1.8 (0.70)
Lifestyle		
Alcohol intake (AUDIT C)	–	1.9 (1.96) (mv: 39)
Smoker (n, %)	–	46 (10) (mv: 25)
Body mass index	–	30.1 (6.58)
Morbidity		
CIRS sum score		7.7 (4.56)
CIRS, no. of organ systems		4.5 (2.35)
Depressive Symptoms (GDS)		2.4 (2.29) (mv: 8)
Medication		
MAI	–	4.7 (5.56)
HRQoL		
EQ-5D-3L Index Score	–	74.3 (23.72) (mv: 24)

*Values are arithmetic means and SD unless otherwise indicated.

†The anamnestic period for baseline data ran from 01 January 2012 to 31 December 2012, except for medication data, for which it ran from 01 July 2012 to 31 December 2012. The follow-up period started on 01 January 2013.

‡Combined outcome included hospitalisation, fall/fall-related injuries, institutionalisation and care level.

§6 months before study entry.

¶Number of patients with missing values (mv) is zero unless indicated in square parentheses.

ACh burden, anticholinergic drug burden; ADS, Anticholinergic Drug Scale; AUDIT, Alcohol Use Disorders Identification Test (WHO); CASMIN, Comparative Analysis of Social Mobility in Industrial Nations; CCI, Charlson Comorbidity Index; CIRS, Cumulative Illness Rating Scale; CRT, cluster-randomised controlled trial; GDS, Geriatric Depression Scale; GP, general practitioner; HRQoL, health-related quality of life; HRQoL-CI, HRQoL comorbidity index; MAI, Medication Appropriateness Index; MMSE, Mini-Mental Status Exam; PIM, potentially inappropriate medication; TK, *Techniker Krankenkasse*.

DISCUSSION

Main results

Our best overall prognostic model predicted HRQoL after 6 months in older general practice patients with multimorbidity and polypharmacy. It performed well, was based on trial data and explained more than half of the variance. The most important predictors were depressive symptoms, the initial level of HRQoL and MAI—all of which were only available as ‘additional predictors’ in trial data. Prognostic models in trial data, which were exclusively developed from ‘core predictors’ (available in both data sets) performed worse, as well as claims based models and models based on both data sets that had longer forecast periods (9 months or more). In both trial data-based and claims-based models, outcome components at baseline had a relatively high impact (ie, HRQoL at baseline in the trial data-based model and previous hospitalisation and previous falls/fall-related injuries in claims-based models). Although this is unsurprising and is often the case in prognostic models,⁶⁵ it nonetheless seems reasonable to retain the variables in the model. Furthermore, we identified further predictors, such as depressive symptoms and medication appropriateness, which had a relatively high predictive power.

Comparison with the literature

The presented results are consistent with results from other studies. The AUC values in our claims-based models (AUC 0.68 to 0.71) are comparable to those of 23 prognostic models for Case Finding conducted in elderly patients in primary care. These models predicted (re) hospitalisation, functional impairment, institutionalisation and death.⁶⁵ The quality of models with a low risk of bias was AUC 0.60 to 0.78, but no explanation of variance was provided. The best model for predicting death within 4 years (AUC: 0.82) included 12 predictors comprising age, sex, body mass index, chronic diseases, smoking status and functional parameters.⁶⁵ Models that included additional trial data (eg, clinical data) predicted endpoints better than models based only on claims data.^{65–67} In many models described in other studies, healthcare utilisation parameters, and especially previous hospitalisations, were predictive of (re)hospitalisations, emergency admissions and functional impairment.^{66 68 69} The predictive power of sex is inconsistent: in 18/27 risk models, sex was included in the final model;⁶⁶ in 7/23 risk models, male sex was predictive,⁶⁵ while a further 25 studies found sex to have no influence.^{68 69} Model quality also improved in studies that included multimorbidity and polypharmacy parameters.^{66 68 70} However, the parameters and instruments used

Table 2 Comparison of models

Models based on claims data: core predictors		AUC*	R ²	
1.1	Combined outcome after 6 months	0.71 (0.70)	0.16 (0.16)	
1.2	Combined outcome after 9 months	0.69 (0.69)	0.15 (0.14)	
1.3	Combined outcome after 24 months	0.68 (0.68)	0.13 (0.12)	
Models based on CRT data: core predictors*		AIC	R ²	R ² (10x)
EQ-5D-3L after 6 months				
2.1	No imputation, no variable selection	4138.86 (4069.41)	0.155 (0.159)	0.112 (0.103)
2.2	No imputation, with variable selection	4138.81 (4068.69)	0.150 (0.155)	0.129 (0.122)
2.3	With imputation, no variable selection	4582.30 (4507.71)	0.159 (0.163)	0.094 (0.108)
2.4	With imputation, with variable selection	4583.15 (4507.47)	0.919 (0.925)	0.111 (0.128)
EQ-5D-3L after 9 months				
2.5	No imputation, no variable selection	3917.75 (3917.75)	0.150 (0.150)	0.030 (0.030)
2.6	No imputation, with variable selection	3921.95 (3921.95)	0.146 (0.146)	0.053 (0.053)
2.7	With imputation, no variable selection	4540.58 (4505.52)	0.156 (0.152)	0.090 (0.093)
2.8	With imputation, with variable selection	4546.42 (4511.10)	0.221 (0.218)	0.107 (0.106)
Models based on CRT data: core predictors and additional predictors*				
EQ-5D, after 6 months				
3.1	No imputation, no variable selection	3205.13 (3205.13)	0.034 (0.034)	0.442 (0.442)
3.2	With imputation, no variable selection	4308.94 (4308.94)	0.538 (0.538)	0.481 (0.481)
3.3	No imputation, with variable selection	3197.37 (3197.37)	0.526 (0.526)	0.483 (0.483)
3.4	With imputation, with variable selection†	4307.47 (4307.47)	0.677 (0.677)	0.507 (0.507)
Models with 'fixed variables'				
3.5	No imputation, no variable selection	3208.58 (3208.58)	0.514 (0.514)	0.468 (0.468)
3.6	With imputation, with variable selection	4308.90 (4308.90)	0.665 (0.665)	0.499 (0.499)
EQ-5D, after 9 months				
3.7	No imputation, no variable selection	3061.06 (3113.53)	0.042 (0.028)	0.411 (0.409)
3.8	With imputation, no variable selection	4307.28 (4361.36)	0.498 (0.477)	0.433 (0.404)
3.9	No imputation, with variable selection	3062.03 (3108.61)	0.490 (0.485)	0.448 (0.443)
3.10	With imputation, with variable selection	4309.88 (4360.32)	0.453 (0.346)	0.455 (0.431)
Models with fixed variables				
3.11	No imputation, no variable selection	3064.76 (3113.08)	0.490 (0.485)	0.439 (0.434)
3.12	With imputation, with variable selection	4310.92 (4363.62)	0.113 (0.071)	0.447 (0.423)

(sensitivity analyses)

*Models based on randomised controlled trial data: fixed effects.

†Best overall model.

AIC, Akaike Information Criterion; AUC, area under the curve after 10-fold cross validation; R², Nagelkerke's R²; R² (10x), Nagelkerke's R² with 10-fold cross validation.

in modelling (eg, CIRS, Charlson Comorbidity Index and disease count, as reported here) varied considerably among studies. They were neither consistently predictive, nor were certain parameters or instruments better than others.^{66 69 70}

Most published models were developed to predict the risk of hospitalisation.^{66 68–74} Other models predicted functional outcomes,⁷⁰ while four models predicted adverse drug reactions.⁷⁴ So far, little is known about the predictive power of polypharmacy parameters and the

appropriateness of prescriptions, especially the MAI has never been used in prognostic models. Furthermore, no models have yet been developed to predict HRQoL in patients with multimorbidity and polypharmacy in general practice.^{27 70}

Strengths and limitations

One strength of our study is that we could use two data sources with differing advantages in our exploratory analysis: claims data contained a large number of cases, and

**Table 3** Best performing models per data set and set of predictors

Best model based on claims data: core predictors (model 1.1)	OR	(95% CI)	P value	z-value	NRI
Age (years)	1.02	(1.02 to 1.02)	<0.001	45.4	0.11
Sex (female)	0.99	(0.97 to 1.00)	0.025	-2.2	-0.03
Disease count	1.02	(1.02 to 1.03)	<0.001	19.2	0.05
CCI	1.03	(1.03 to 1.04)	<0.001	22.7	0.02
No. of specific chronic diseases (Diederichs)	1.01	(1.00 to 1.01)	<0.001	3.9	0.01
No. of PIM (EU-PIM)	1.03	(1.02 to 1.03)	<0.001	8.9	0.02
ACh burden (ADS)	1.04	(1.03 to 1.05)	<0.001	14.9	0.04
Modified Drug Burden Index	1.08	(1.07 to 1.08)	<0.001	20.1	0.08
Previous hospitalisations	1.67	(1.65 to 1.70)	<0.001	82.3	0.34
Previous falls/fall-related injuries	3.29	(3.25 to 3.34)	<0.001	188.6	0.55
No. of involved physicians	1.02	(1.02 to 1.02)	<0.001	29.0	0.08
Best model based on CRT data: core predictors (model 2.4)	Coefficient	(95% CI)	P value		
Intercept	101.18	(93.11 to 109.25)	<0.001		
Sex (female)	-11.26	(-15.59 to -6.94)	<0.001		
No. of specific chronic diseases (Diederichs)	-2.18	(-3.64 to -0.73)	0.004		
No. of drugs	-1.28	(-2.23 to -0.32)	0.010		
Modified Drug Burden Index	-5.19	(-8.26 to -2.12)	0.001		
Previous falls	-6.11	(-12.07 to -0.15)	0.045		
Best model based on CRT data: core predictors and additional predictors (model 3.4, best overall model)	Coefficient	(95% CI)	P value		
Intercept	51.74	(38.91 to 64.57)	<0.001		
Sex (female)	-3.61	(-6.96 to -0.27)	0.036		
No. of specific chronic diseases (Diederichs)	-1.03	(-2.08 to 0.01)	0.055		
No. of involved physicians	0.80	(-0.13 to 1.74)	0.093		
Body mass index	-0.28	(-0.53 to -0.03)	0.031		
Medication Appropriateness Index	-0.39	(-0.70 to -0.08)	0.015		
Depressive symptoms (GDS)	-2.73	(-3.56 to -1.91)	<0.001		
EQ-5D Index Score (baseline)	0.55	(0.47 to 0.64)	<0.001		

ACh burden, anticholinergic drug burden; ADS, Anticholinergic Drug Scale; CCI, Charlson Comorbidity Index; GDS, Geriatric Depression Scale; NRI, continuous net reclassification index; PIM, potentially inappropriate medication.

trial data provided additional high-quality patient data including functional status and HRQoL. Both data sets also have their limitations, since claims are documented for billing purposes and are therefore imprecise, whereas our trial data set consisted of only a limited number of observations. Thus, each data set allows its own endpoints to be modelled. Risk modelling is especially complex in multimorbid patients with polypharmacy, as predictor variables in this patient collective are often associated with one another (eg, diagnoses and prescriptions). In addition, comparable risk situations can lead to different endpoints, as risk often depends on context. For example, a drug-induced fall may have no health-related consequences or may lead to impairment and institutionalisation.

Further to these key limitations, our results need careful interpretation for several reasons: First, the combined endpoint in the claims-based models yielded a high event rate, which may have resulted in overoptimistic results in our logistic regression. However, other approaches would not have resolved this problem to suit our purposes either. Additionally, we still have enough cases in both categories of the dependent variable to conduct a valid model estimation. Nonetheless, the low performance of the claims model may have been because predictors acted in different ways on the different elements of the combined outcome, thus resulting in greater heterogeneity.⁷⁵ Second, the small sample size of the trial population may have led to some overfitting of the model. At the same time, the VIF (to assess any

multi-collinearity) showed only up to moderate values. The application of shrinkage methods would have been a possible alternative to address this limitation.⁷⁶ However, there is an ongoing debate whether it solves such problems, and a recent study has suggested that although shrinkage can result in improved calibration, it may not be superior in terms of reducing overfitting.⁷⁷ Furthermore, shrinkage models lead to biased estimates of the regression coefficients, thus making results more difficult to interpret. Third, in our modelling approach we tested disease-based indicators such as the Charlson Comorbidity Index and CIRS that were developed and validated for other purposes. However, we chose indicators that showed a strong association with negative health outcomes.³⁵

Relevance for primary care and research implications

As the models derived in our study have not been externally validated and our methods have some limitations, we do not claim to have developed comprehensive prognostic models to identify older general practice patients with multimorbidity and polypharmacy at risk of negative health outcomes. For this reason, we plan to conduct an individual patient data-based meta-analysis to further develop and externally validate the models presented here (PROSPERO ID: CRD42018088129).

It is, however, very likely that baseline components of our predicted endpoints are important predictors, especially considering these results are unsurprising and entirely plausible. A decline in HRQoL, a previous hospitalisation and a previous falls/fall-related injury can therefore be seen as a warning parameter ('red flag') that may help general practitioners in recognising older patients with multimorbidity and polypharmacy at high risk of adverse health outcomes. These patients are therefore more likely to benefit from an intervention than others with low or no risk.²⁴ Hence, researchers evaluating polypharmacy interventions, such as medication reviews, may like to consider our models when deciding on selection and inclusion criteria for a study population.

CONCLUSIONS

This study provides prognostic models to identify older general practice patients with multimorbidity and polypharmacy at high risk of deterioration in HRQoL, hospitalisation, falls/fall-related injuries, institutionalisation and a need of nursing care. Outcome components, such as previous falls, hospital stays, reduced HRQoL and depression, were important predictors of these negative health outcomes in our models. They can be seen as warning signs of future worsening and an indication that these patients are likely to benefit from interventions to optimise their medication. Future studies should externally validate the models and evaluate the effectiveness of polypharmacy interventions in high-risk patients.

Ethics approval and consent to participate

Claims may be analysed by statutory health insurance companies in accordance with § 284 of Social Code Book V. For the research questions of this project, claims data were analysed by Cologne University, Goethe-University Frankfurt and Heidelberg University collaboratively with TK. When claims are anonymously analysed in accordance with Good Practice in Claims Data Analysis,⁷⁸ no further ethics vote is required. Regarding our trial data, the ethics commission of the medical faculty of the Johann Wolfgang Goethe University, Frankfurt/Main approved the PRIMUM trial (resolution number E 46/10, file number 123/10, date: 20 May 2010) and all of the participants gave their written informed consent before taking part.

Author affiliations

- ¹Institute of General Practice, Goethe University Frankfurt, Frankfurt am Main, Hessen, Germany
²Institute of Medical Biometry and Informatics, University of Heidelberg, Heidelberg, Baden-Württemberg, Germany
³PMV Research Group, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Nordrhein-Westfalen, Germany
⁴Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK
⁵APEx Collaboration for Academic Primary Care, University of Exeter Medical School, Exeter, UK
⁶Centre for Research in Evidence-Based Practice, Faculty of Health Sciences and Medicine, Bond University, Gold Coast, Queensland, Australia
⁷Department of Family Medicine, School CAPHRI, Maastricht University, Maastricht, Limburg, The Netherlands

Twitter Rafael Perera @rafaoxford

Acknowledgements The authors would like to thank Phillip Elliott for the language review of the paper.

Contributors MB, FMG and CM designed the study. PI, LU and CS analysed the data. BSM, LU, PI, CS, FvB, MB, FMG, RP, JMV, PG, MvdA and CM contributed to the interpretation of the data. CM and BM drafted the manuscript and all authors revised it and subsequent versions of the manuscript critically for important intellectual content. All authors approved the version to be submitted for publication. LU, CS and CM had full access to all data and are responsible for the integrity and the accuracy of the data analysis.

Funding This study was supported by the German Statutory Healthcare Insurance Company *Techniker Krankenkasse*.

Competing interests FMG, BSM, MB and CM received grants from the German Statutory Healthcare Insurance Company *Techniker Krankenkasse* during the course of the study. CS has been employed by Boehringer Ingelheim GmbH & Co KG since October 2019.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement The data sets generated and analysed in the current study are not publicly available, as further analyses are ongoing.

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

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is

properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Beate S Müller <http://orcid.org/0000-0002-6745-1047>

Christiane Muth <http://orcid.org/0000-0001-8987-182X>

REFERENCES

- 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:e12–21.
- Moßhammer D, Haumann H, Mörike K, *et al*. Polypharmacy—an upward trend with unpredictable effects. *Dtsch Arztebl Int* 2016;113:627–33.
- Muth C, Blom JW, Smith SM, *et al*. Evidence supporting the best clinical management of patients with multimorbidity and polypharmacy: a systematic guideline review and expert consensus. *J Intern Med* 2019;285:272–88.
- Palmer K, Marengoni A, Forjaz MJ, *et al*. Multimorbidity care model: recommendations from the consensus meeting of the joint action on chronic diseases and promoting healthy ageing across the life cycle (JA-CHRODIS). *Health Policy* 2018;122:4–11.
- 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.
- Oscanoa TJ, Lizaraso F, Carvajal A. Hospital admissions due to adverse drug reactions in the elderly. A meta-analysis. *Eur J Clin Pharmacol* 2017;73:759–70.
- Angamo MT, Chalmers L, Curtain CM, *et al*. Adverse-Drug-Reaction-Related hospitalisations in developed and developing countries: a review of prevalence and contributing factors. *Drug Saf* 2016;39:847–57.
- 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.
- Heinrich S, Rapp K, Rissmann U, *et al*. Cost of falls in old age: a systematic review. *Osteoporos Int* 2010;21:891–902.
- Fried TR, O’Leary J, Towle V, *et al*. Health outcomes associated with polypharmacy in community-dwelling older adults: a systematic review. *J Am Geriatr Soc* 2014;62:2261–72.
- 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.
- 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.
- Hilmer SN, Mager DE, Simonsick EM, *et al*. A drug burden index to define the functional burden of medications in older people. *Arch Intern Med* 2007;167:781–7.
- Salahudeen MS, Duffull SB, Nishtala PS. Anticholinergic burden quantified by anticholinergic risk scales and adverse outcomes in older people: a systematic review. *BMC Geriatr* 2015;15:31.
- Gurwitz JH, Field TS, Harrold LR, *et al*. Incidence and preventability of adverse drug events among older persons in the ambulatory setting. *JAMA* 2003;289:1107–16.
- Kuijpers MAJ, van Marum RJ, Egberts ACG, *et al*. Relationship between polypharmacy and underprescribing. *Br J Clin Pharmacol* 2008;65:130–3.
- Steinman MA, Landefeld CS, Rosenthal GE, *et al*. Polypharmacy and prescribing quality in older people. *J Am Geriatr Soc* 2006;54:1516–23.
- 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.
- Müller BS, Uhl MC, Sophia NT, *et al*. Patienten MIT Multimedikation: Ambulante Herausforderungen und Lösungswege: Eine qualitative studie. *Zeitschrift für Allgemeinmedizin* 2018;94:396–400.
- Horne R, Weinman J, Barber N, *et al*. Concordance, adherence and compliance in medicine taking: Report for the National Co-ordinating Centre for NHS Service Delivery and Organisation R & D (NCCSDO), 2005.
- Rankin A, Cadogan CA, Patterson SM, *et al*. Interventions to improve the appropriate use of polypharmacy for older people. *Cochrane Database Syst Rev* 2018;9:CD008165.
- 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.
- 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.
- Glasziou PP, Irwig LM. An evidence based approach to individualising treatment. *BMJ* 1995;311:1356–9.
- Murray SA, Kendall M, Mitchell G, *et al*. Palliative care from diagnosis to death. *BMJ* 2017;356:j878.
- Lynn J. Living long in fragile health: the new demographics shape end of life care. *Hastings Cent Rep* 2005;Spec No:s14–18.
- National Institute for Health and Care Excellence. *Multimorbidity: clinical assessment and management*. 56. NICE guideline NG, 2016.
- Bundesministerium für Gesundheit. Km 6-Statistik, 2013. Available: <https://www.bundesgesundheitsministerium.de/themen/krankenversicherung/zahlen-und-fakten-zur-krankenversicherung/mitglieder-und-versicherte.html> [Accessed 23 Apr 2020].
- Schäfer I, von Leitner E-C, Schön G, *et al*. Multimorbidity patterns in the elderly: a new approach of disease clustering identifies complex interrelations between chronic conditions. *PLoS One* 2010;5:e15941.
- Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–98.
- Graf J-M, Claes C, Greiner W, *et al*. Die Deutsche version des EuroQol-Fragebogens. *J Public Health* 1998;6:3–20.
- Agborsangaya CB, Lahtinen M, Cooke T, *et al*. Comparing the EQ-5D 3L and 5L: measurement properties and association with chronic conditions and multimorbidity in the general population. *Health Qual Life Outcomes* 2014;12:74.
- EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199–208.
- Charlson ME, Pompei P, Ales KL, *et al*. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;40:373–83.
- Diederichs C, Berger K, Bartels DB. The measurement of multiple chronic diseases—a systematic review on existing multimorbidity indices. *J Gerontol A Biol Sci Med Sci* 2011;66:301–11.
- Renom-Guiteras A, Meyer G, Thürmann PA. The EU(7)-PIM list: a list of potentially inappropriate medications for older people consented by experts from seven European countries. *Eur J Clin Pharmacol* 2015;71:861–75.
- Holt S, Schmiedl S, Thürmann PA. Potentially inappropriate medications in the elderly: the PRISCUS list. *Dtsch Arztebl* 2010;107:31–2.
- Salahudeen MS, Hilmer SN, Nishtala PS. Comparison of anticholinergic risk scales and associations with adverse health outcomes in older people. *J Am Geriatr Soc* 2015;63:85–90.
- Carnahan RM, Lund BC, Perry PJ, *et al*. The anticholinergic drug scale as a measure of drug-related anticholinergic burden: associations with serum anticholinergic activity. *J Clin Pharmacol* 2006;46:1481–6.
- Brauns H, Steinmann S. *Educational reform in France, West-Germany, the United Kingdom and Hungary: updating the CASMIN educational classification*. ZUMA-Nachrichten, 1999: 7–44. www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_44.pdf
- Saunders JB, Aasland OG, Babor TF, *et al*. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption—II. *Addiction* 1993;88:791–804.
- Hanlon JT, Schmader KE, Samsa GP, *et al*. A method for assessing drug therapy appropriateness. *J Clin Epidemiol* 1992;45:1045–51.
- 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.
- Mukherjee B, Ou H-T, Wang F, *et al*. A new comorbidity index: the health-related quality of life comorbidity index. *J Clin Epidemiol* 2011;64:309–19.
- 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.
- Rubin DB. *Multiple imputation for nonresponse in surveys*. Hoboken, N.J.: Wiley-Interscience, 2004.
- Buuren Svan, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw* 2011;45.
- Grandt D, Lappe V, Schubert I. *Barmer Arzneimittel report 2018*. Wuppertal, 2018.
- Wood AM, White IR, Royston P. How should variable selection be performed with multiply imputed data? *Stat Med* 2008;27:3227–46.

- 50 Sauerbrei W. The use of resampling methods to simplify regression models in medical statistics. *J R Statist Soc C* 1999;48:313–29.
- 51 Kerr KF, Wang Z, Janes H, *et al.* Net reclassification indices for evaluating risk prediction instruments: a critical review. *Epidemiology* 2014;25:114–21.
- 52 James G, Witten D, Hastie T, *et al.* *An introduction to statistical learning: with applications in R*. 8th edn. New York, Heidelberg, Dordrecht, London: Springer, 2017.
- 53 Nakagawa S, Schielzeth H, O'Hara RB. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods Ecol Evol* 2013;4:133–42.
- 54 Steyerberg EW, Vickers AJ, Cook NR, *et al.* Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010;21:128–38.
- 55 Nagelkerke NJD. A note on a general definition of the coefficient of determination. *Biometrika* 1991;78:691–2.
- 56 Hastie T, Tibshirani R, Friedman JH. *The elements of statistical learning: data mining, inference, and prediction*. 12th edn. New York: Springer, 2017.
- 57 R Core Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2016.
- 58 Pinheiro J, Bates D, DebRoy S. *nlme: linear and nonlinear mixed effects models. R package. version 3*, 2016: 1–128.
- 59 Kundu S, Aulchenko YS, Janssens A, *et al.* *PredictABEL: assessment of risk prediction models. R package version. version 1.2-2*, 2014: 2.
- 60 Lumley T. *mitools: tools for multiple imputation of missing data. R package. version 2.3*, 2014.
- 61 Nakazawa M. *fmsb: functions for medical statistics book with some demographic data. R package. version 0.5.2*, 2015.
- 62 Dahl DB. *xtable: export tables to latex or HTML. R package. version 1.8-2*, 2016.
- 63 Xie Y. *knitr: a general-purpose package for dynamic report generation in R. R package. version 1.13*, 2016.
- 64 Collins GS, Reitsma JB, Altman DG, *et al.* Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med* 2015;162:55–63.
- 65 O'Caomh R, Cornally N, Weathers E, *et al.* Risk prediction in the community: a systematic review of case-finding instruments that predict adverse healthcare outcomes in community-dwelling older adults. *Maturitas* 2015;82:3–21.
- 66 Wallace E, Stuart E, Vaughan N, *et al.* Risk prediction models to predict emergency hospital admission in community-dwelling adults: a systematic review. *Med Care* 2014;52:751–65.
- 67 Coleman EA, Min S-joon, Chomiak A, *et al.* Posthospital care transitions: patterns, complications, and risk identification. *Health Serv Res* 2004;39:1449–66.
- 68 Campbell SE, Seymour DG, Primrose WR, *et al.* A systematic literature review of factors affecting outcome in older medical patients admitted to hospital. *Age Ageing* 2004;33:110–5.
- 69 García-Pérez L, Linertová R, Lorenzo-Riera A, *et al.* Risk factors for hospital readmissions in elderly patients: a systematic review. *QJM* 2011;104:639–51.
- 70 Alonso-Morán E, Nuño-Solinis R, Onder G, *et al.* Multimorbidity in risk stratification tools to predict negative outcomes in adult population. *Eur J Intern Med* 2015;26:182–9.
- 71 Wallace E, Hinchey T, Dimitrov BD, *et al.* A systematic review of the probability of repeated admission score in community-dwelling adults. *J Am Geriatr Soc* 2013;61:357–64.
- 72 Kansagara D, Englander H, Salanitro A, *et al.* Risk prediction models for hospital readmission: a systematic review. *JAMA* 2011;306:1688–98.
- 73 Vest JR, Gamm LD, Oxford BA, *et al.* Determinants of preventable readmissions in the United States: a systematic review. *Implement Sci* 2010;5:88.
- 74 Stevenson JM, Williams JL, Burnham TG, *et al.* Predicting adverse drug reactions in older adults; a systematic review of the risk prediction models. *Clin Interv Aging* 2014;9:1581–93.
- 75 Glynn RJ, Rosner B. Methods to evaluate risks for composite end points and their individual components. *J Clin Epidemiol* 2004;57:113–22.
- 76 Steyerberg EW, Eijkemans MJ, Harrell FE, *et al.* Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. *Med Decis Making* 2001;21:45–56.
- 77 van Calster B, van Smeden M, de CB, *et al.* Regression shrinkage methods for clinical prediction models do not guarantee improved performance: simulation study. *Stat Methods Med Res* 2020;962280220921415.
- 78 Swart E, Gothe H, Geyer S, *et al.* Gute praxis Sekundärdatenanalyse (GPs): Leitlinien und Empfehlungen. *Gesundheitswesen* 2015;77:120–6.