Artificial intelligence to assist decision-making on pharmacotherapy: A feasibility study

Michael Bücker [a], Kreshnik Hoti [b], Olaf Rose [c,*]

[a] *Münster School of Business -FH Münster - University of Applied Sciences, Münster, Germany*
[b] *Faculty of Medicine, University of Pristina, Prishtina, Kosovo*
[c] *Institute of Pharmacy, Pharmaceutical Biology and Clinical Pharmacy, Paracelsus Medical University, Salzburg, Austria*

ARTICLE INFO

ABSTRACT

Background: Artificial intelligence (AI) has the capability to analyze vast amounts of data and has been applied in various healthcare sectors. However, its effectiveness in aiding pharmacotherapy decision-making remains uncertain due to the intricate, patient-specific, and dynamic nature of this field.

Objective: This study sought to investigate the potential of AI in guiding pharmacotherapy decisions using clinical data such as diagnoses, laboratory results, and vital signs obtained from routine patient care.

Methods: Data of a previous study on medication therapy optimization was updated and adapted for the purpose of this study. Analysis was conducted using R software along with the tidymodels extension packages. The dataset was split into 74% for training and 26% for testing. Decision trees were selected as the primary model due to their simplicity, transparency, and interpretability. To prevent overfitting, bootstrapping techniques were employed, and hyperparameters were fine-tuned. Performance metrics such as areas under the curve and accuracies were computed.

Results: The study cohort comprised 101 elderly patients with multiple diagnoses and complex medication regimens. The AI model demonstrated prediction accuracies ranging from 38% to 100% for various cardiovascular drug classes. Laboratory data and vital signs could not be interpreted, as the effect and dependence were unclear for the model. The study revealed that the issue of AI lag time in responding to sudden changes could be addressed by manually adjusting decision trees, a task not feasible with neural networks.

Conclusion: In conclusion, the AI model exhibited promise in recommending appropriate medications for individual patients. While the study identified several obstacles during model development, most were successfully resolved. Future AI studies need to include the drug effect, not only the drug, if laboratory data is part of the decision. This could assist with interpreting their potential relationship. Human oversight and intervention remain essential for an AI-driven pharmacotherapy decision support system to ensure safe and effective patient care.

1. Introduction

Contemporary pharmacotherapy practices are primarily guided by randomized controlled trials and therapy guidelines, with general practitioners and pharmacists often facing challenges in keeping up with the rapidly evolving therapeutic standards. This can lead to limited adherence to guidelines, particularly in prevalent conditions like hypertension, dyslipidemia, coronary heart disease, and heart failure.[1–3] Medication errors, especially prescribing errors, are common in clinical settings.[4] Clinical decision-support systems are software designed to be a direct aid to clinicians. They are widely accepted and integrated into current medical and pharmaceutical practice.[5,6] Artificial intelligence (AI) in the medical field is engaged predominantly in diagnosis or whenever a large amount of data needs to be analyzed.[7–9] In areas more directly pertaining to medication use, AI-based solutions have been proposed as means of addressing issues around medication self-administration with focus on guarding the correct use and application of devices, such as insulin pens and inhalers.[10] In contrast to its application in diagnosis and application, AI is rarely employed in pharmacotherapy decision-making. There are valid justifications for the limited

integration of AI in therapy decisions. Pharmacotherapy is a highly delicate process, where errors can lead to significant and immediate repercussions. Additionally, complete patient information may not always be accessible in a digital form. Factors such as patients' emotions, concerns, preferences, and expectations must be considered, yet translating these emotional aspects into quantifiable data poses challenges.[11] Establishing a therapeutic relationship is deemed to be an integral part of therapy.[12,13] CDSS in pharmacotherapy can be designed as a simple input-output software, like a drug-drug interaction checker. Algorithm-based software with interactive feedback loops is much more helpful, as it can respond to preset personal valuations. Unfortunately, software which is just reconciling medication lists may face restrictions in multimorbid patients with contradicting therapies.[14] A simple example is the prescription of two blood-pressure lowering drugs. While this combination therapy is clearly indicated in most cases, it might be a drug-related problem in certain other patients. More information is needed here to draw the right conclusion. Currently there is a lack of literature and digital solutions exploring AI use in the area of pharmacy practice and more specifically in pharmacotherapy.[15]

1.1. Objectives

The aim of this study was to investigate the feasibility of AI in optimizing complex patients' pharmacotherapy.

2. Material and methods

2.1. Dataset

The data used for the study was taken with permission from a previous clinical study on Medication Management, in which clinical pharmacists and general practitioners (GPs) collaboratively optimized pharmacotherapy of community patients for at least 12 months. The published study protocol and the study results provide a detailed description of the WestGEM study.[16,17] Inclusion criteria of the original study were an age $\geq$ 65 years, a minimum of 3 chronic disorders affecting two different organ systems, at least one cardiovascular disease, at least one visit to the general practitioner in each of the preceding 3-month intervals, five or more long-term drug treatments (>3 months) with systemic effects and the ability to complete questionnaires, with assistance if required. Exclusion criteria were a life expectancy of <12 months (assessed by the treating primary care physician) and participation in another clinical study.

All patients were originally recruited in 12 general practitioners' practices in the Westphalia-Lippe area, Germany. Data consisted of diagnoses, vital signs, chief complaints, symptoms, laboratory data and the previously optimized medication. The original drug therapy optimization was done by the pharmacotherapy experts of the WestGEM study.

2.2. Data preparation

Due to the limited size of the dataset, the variables pertaining to medication and diagnoses were transformed into binary form. Instead of individual drug names such as Atenolol, Bisoprolol, and Carvedilol, drug classes were utilized. For instance, the category of Beta-Blockers was employed in place of specific drug names. Additionally, drug dosage was simplified into two categories: high dose and normal dose. The aggregation of drugs into classes was carried out to increase the number of patients represented in the medication data. The study did not consider potential distinctions between specific active agents for the purpose of this research. Laboratory findings were maintained in their original numerical structure. The anonymized data was updated to conform to contemporary clinical standards and guidelines by the researchers for the purpose of AI modeling.

2.3. AI modeling

AI models can be divided into transparent models (i.e.: decision trees) and opaque models (i.e.: neural networks), where a decision is difficult or even impossible to follow. For this study, decision trees, random forests and neural networks have initially been tested. Due to the small number of patients and drugs of the sample and the desired traceability of results, a decision tree was chosen to explore feasibility. Decision trees were built on the approach of Breiman et al[18] R software, version 4.2.3 (R Foundation for Statistical Computing, Vienna, Austria) with the package tidymodels (R package version 0.1.0.) was used for calculations and modeling. All codes and data are published at https://github.com/mchlbckr/paper_ai_pharmacotherapy/tree/main and are also available in the appendix. To avoid overfitting, bootstrapping was employed. This was applied on an initial split of 74% of data with simultaneous optimization of hyperparameters (complexity parameter, tree depth, minimal number of cases per node). Based on the AUC, the best decision tree was chosen. The remaining 26% of the data was used to test predictive performance on the hold out data set.

2.4. Ethics

All patients gave informed consent. The underlying study was funded by the European Union and the state of North Rhine-Westphalia (Ziel 2, IuK & Gender Med.NRW, GW, 2076). It was entered into the controlled trials register (ISRCTN 41595373) and use for research was approved by the ethics committee of the Medical Association of Westphalia-Lippe, Germany (AKZ-2013–292-f-s). For this study, only anonymized and modified data was used, which cannot be tracked back to patients (secondary data analysis).

3. Results

3.1. Results of AI testing

Data of 101 patients with use of 929 drugs out of 76 drug classes was generated. Basic patient data is shown in Table 1. There were no missing values. According to the inclusion criteria, the population comprised of elderly patients with polymedication and multiple morbidities.

Decision trees were generated for eight frequently used cardiovascular drug classes, namely ACE-inhibitors/angiotensin receptor blockers (ACEI/ARB), thiazide diuretics, beta-blockers (BB), dihydropyridine calcium channel-blockers (DHP-CCB), mineralocorticoid receptor-blockers (MRA), loop diuretics, isosorbide mono- or dinitrate (nitrates, ISMN/ISDN) and inhaled nitroglycerin. As an illustration, a decision tree for loop diuretics is displayed in Fig. 1 and for ACEI/ARB in Fig. 2. The probability for a decision and the affected percentage of patients is shown. The probability for patients with heart failure to take a loop diuretic was 82%. The model identifies patients with renal failure as a second indication.

Table 1
Patient baselines of the study population ($n = 101$).

Age, mean (years)		78 (SD 6.6)
Female gender		51 (50.5%)
BMI, mean		27.8 (SD 4.6)
GFR, mean (ml/min)		53.2 (SD 17.5)
LDL-cholesterol, mean (mg/dl)		114.5 (SD 39.2)
Average number of diagnoses		6.5 (min. 3, max. 12)
Average number of drugs		9.2 (min. 5, max. 15)
Number of patients with:	hypertension	84 (83.2%)
	coronary artery disease	49 (48.5%)
	atrial fibrillation	31 (30.7%)
	hyperlipidemia	50 (49.5%)

BMI: body mass index; GFR: glomerular filtration rate, max.: maximum; min.: minimum; SD: standard deviation.
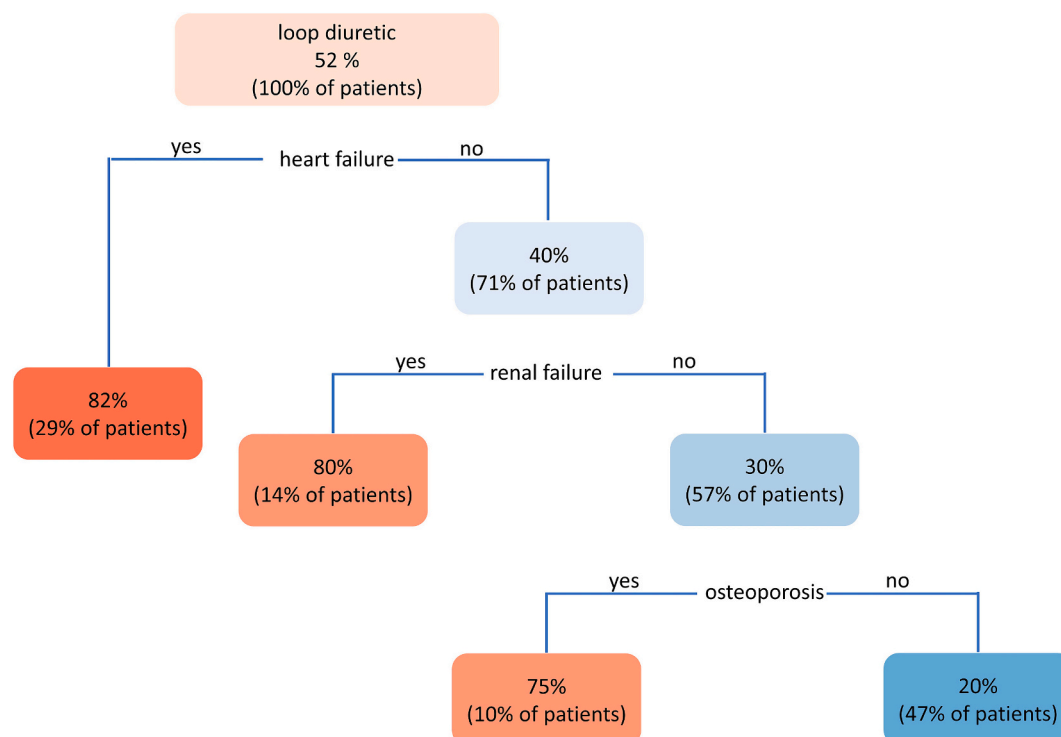
Fig. 1. Decision tree for loop diuretics. Percentage of people with this drug and a certain indication and the ratio of all patients to whom this applies is given (%).

Accuracy of predictions on the test data set was >50% for all drug classes, except for DHP-CCB, where the numbers decreased from an accuracy of 59% in training data to 38% in testing data, indicating there is still overfitting involved in the model used. Results for AUC and accuracy on both training and test data set for all eight drug classes are displayed in Table 2.

3.2. Challenges identified

While the model was developed, some barriers for AI in pharmacotherapy decision-making were realized.

3.2.1. Novel therapies and drug alerts

AI models base their decisions on probabilities derived from existing data. However, when faced with unexpected changes such as new treatment options or drug safety alerts, these models may not immediately adapt to the new circumstances until they have sufficient data to support the decision. This delay in updating recommendations could lead to inappropriate suggestions and may result in patients being excluded from potentially beneficial new treatments. For instance, the transition from a triple to a quadruple therapy for heart failure patients serves as a relevant example. In situations where swift responses are required, particularly in the case of sudden drug safety alerts, adjusting the recommendation probabilities in a neural network can be challenging. This process necessitates the modification of all historical patient data, which is not only time-consuming but may also be unfeasible. One possible solution to this issue is the use of decision trees, which allow for the straightforward adjustment of probabilities for drug recommendations in specific medical conditions. For instance, in the context of heart failure with reduced ejection fraction, a SGLT2-inhibitor could be assigned a 100% probability for eligible patients, thereby accelerating the adaptation process. A neural network would respond much slower. In this example, a neural network model would most probably only suggest a SGLT2-inhibitor when 85% of the patients with heart failure have received it, which might take months and years to occur, depending on the number of patients.

3.2.2. Laboratory data and vital signs

Laboratory results and vital signs were accessible for the majority of patients, yet proved inadequate for accurately predicting appropriate pharmacotherapy. For instance, blood pressure readings while on an ACE inhibitor treatment may vary, with some patients exhibiting normal levels and others experiencing elevated readings. This distinction is challenging for AI to discern, as the model lacks access to the patient's baseline blood pressure prior to drug intervention, a common limitation in clinical settings. Similarly, the issue extends to all laboratory data, such as uncertainty regarding the impact of statin therapy on LDL-cholesterol levels without knowledge of the patient's baseline values. Consequently, AI models must approach vital signs and laboratory data differently, recognizing the complexities involved. Addressing this challenge requires cautious integration of these parameters into AI models.

4. Discussion

This research offers insights into the utilization of artificial intelligence within an innovative clinical decision support system aimed at guiding pharmacotherapy. The study sought to investigate the feasibility of this approach and strategies for overcoming potential obstacles. Decision-making in this system was primarily based on patient diagnoses and demographic information, with laboratory data and vital signs not factored in due to a lack of clear correlations. Notably, the model achieved a 100% accuracy rate in recommending the appropriate drug for three out of eight medications, demonstrating promising potential. For drugs with lower accuracy rates, it is posited that performance could be enhanced with a larger sample size. The modeling exercise facilitated the exploration of key considerations regarding the integration of AI in the complex realm of pharmacotherapeutic decision-making, where errors could have significant health implications.

The utilization of an AI-driven model in decision-making processes may result in delays in implementing sudden therapy changes for patients. For instance, if a registered drug is rejected due to safety concerns, the AI model may not reflect this change in recommendations
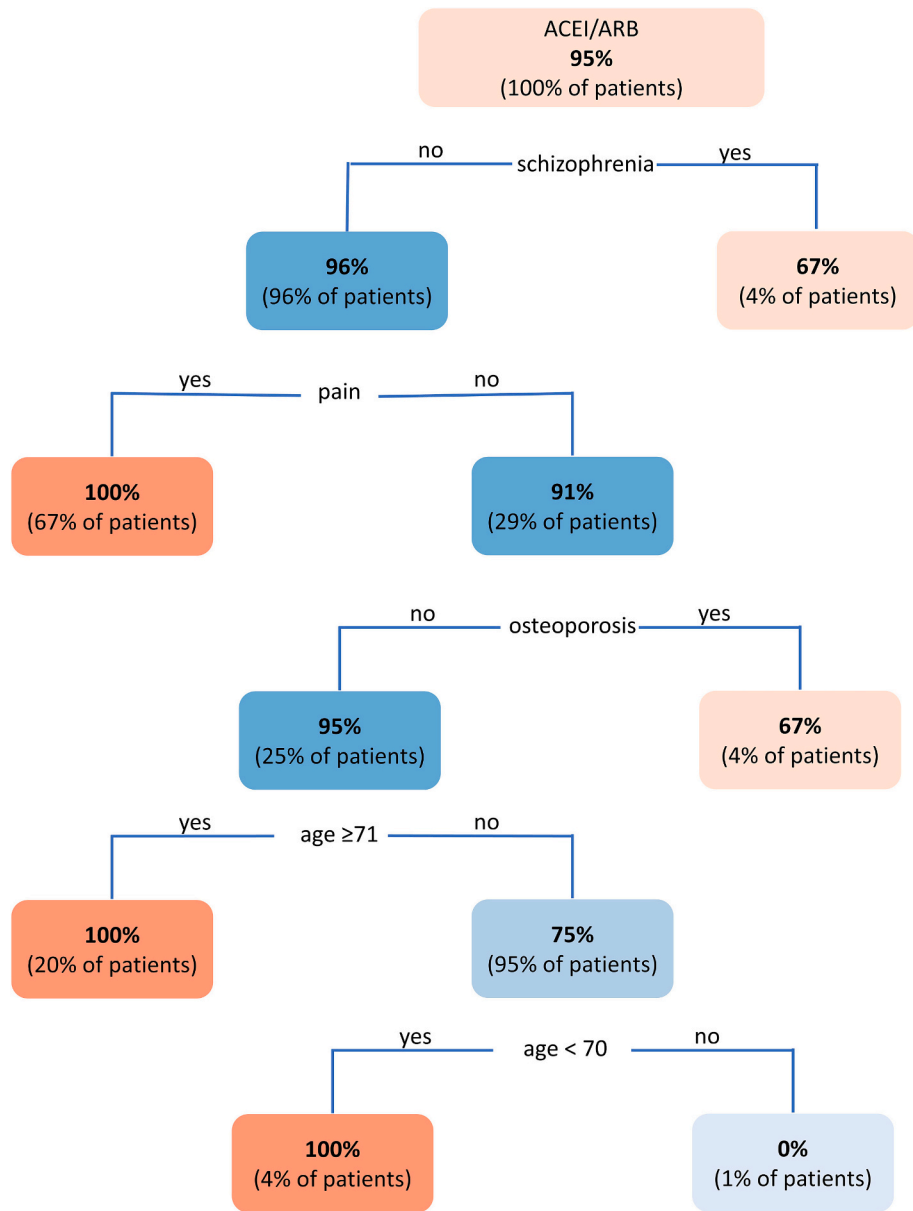
Fig. 2. Decision tree for ACEI/ARB. Percentage of people with this drug and a certain indication and the ratio of all patients to whom this applies is given (%).

Table 2
AI results on accuracy for eight specific drugs.

Variable	AUC training	ACC training	AUC testing	ACC testing
Nitroglycerin	99%	99%	100%	100%
MRA	86%	88%	100%	100%
ISMN/ISDN	68%	85%	50%	81%
Loop diuretics	66%	64%	70%	65%
Beta-blocker	64%	80%	50%	65%
Thiazide diuretics	57%	76%	56%	69%
ACEI/ARB	55%	97%	NA	100%
DHP-CCB	52%	59%	31%	38%

AUC: area under the curve, ACC: accuracy, ACEI: ACE-inhibitor, ARB: angiotensin receptor blocker, DHP-CCB: dihydropyridin calcium chanel blocker, MRA: mineralocorticoid receptor antagonists, NA: not available (too many patients with ACEI/ARB did not allow for an AUC reading).

until a significant portion, such as 85%, of patients cease using the drug. Consequently, manual intervention would be required to update or remove all previous cases in the system. Similarly, there would be a time lag in incorporating new guideline recommendations into the AI model. This issue can be effectively addressed by employing decision tree-based models, as the initial decision-making parameters can be easily adjusted to zero or 100 to mitigate such delays.

The model that was developed did not incorporate patient preferences and chief complaints, despite their significant role in formulating a pharmacotherapy plan. However, it is conceivable that these factors could be digitized and integrated into the model. Recent research indicates that the inclusion of patients' preferences and shared decision-making is not consistently practiced in current clinical settings, despite clear recommendations to do so.[19] Due to time constraints in consulting with healthcare professionals, patients may find it beneficial to interact with an avatar to discuss their medical condition without feeling pressured.[20] This differs from the expectations of most physicians and pharmacists, who believe that only human interactions can effectively handle such sensitive situations. However, it still is questionable whether AI can translate these emotions into suitable categories for decision-making in pharmacotherapy. Additionally, challenges related to upholding the principle of person-centered care should also be taken

into consideration. Nevertheless, the outcomes of an AI-driven conversation must be incorporated into pharmacotherapy decision support systems if a comprehensive digitalization of prescribing practices is desired. However, interpreting laboratory data and vital signs poses challenges for the model, as they can be influenced by the patient's current medication regimen.

The acceptable level of accuracy in predicting the appropriate drug for a specific patient remains uncertain. With accuracies ranging from 38% to 100% achieved by this model, a threshold must be established before implementing an AI-based clinical decision support system. Any accuracy exceeding 50% would surpass the current therapeutic standard. Nonetheless, given the potential risks associated with an incorrect therapeutic decision, the responsibility currently lies with physicians and pharmacists, with AI-based recommendations serving as a clinical tool that carries a residual risk of inaccuracy, akin to drug-drug interaction checkers. This underscores the notion that an AI-based clinical decision support system for pharmacotherapy should serve to assist rather than replace healthcare professionals. Even if the AI system makes superior decisions compared to the average healthcare provider, it remains uncertain whether patients would prefer AI-generated treatment plans over human decisions. A key takeaway from this feasibility study is that certain data, such as laboratory results and vital signs, are challenging to interpret solely from a theoretical standpoint without considering the patient's individual state, symptoms, or complaints. For example, if antibiotics are seen only with elevated leucocytes or CRP, our AI model would rather suggest to discontinue antibiotic therapies with such laboratory data. As a consequence, future data sets on pharmacotherapy can interpret laboratory and vital data only, if the effect of the drug is provided. In this example, this would be the CRP lowering effect of an antibiotic. A single point retrospective data set with only laboratory data and drugs hence is not helpful for an AI model with a similar focus to what we have explored. The system would require ongoing maintenance and updates by human operators with current data, who must determine whether new information supersedes prior knowledge. The model exhibits inertia, as it takes time for a significant portion of patients to change therapy before the new recommendation can be based on this updated majority.

4.1. Limitations

This research was conducted with a limited sample size of 101 patients, who had been optimized. It is widely recognized that AI models require extensive data to achieve accuracy. Despite the meticulous optimization of pharmacotherapy for the patients by multiple pharmacists and physicians, incorporating up-to-date standards, there were gaps in the available information, particularly concerning laboratory data and vital signs. One proposed method to enhance data quality is the utilization of expert panels. However, it should be noted that expert panels are resource-intensive and may not yield large quantities of data either.[21]

Patients' preferences and unique needs should be taken into account alongside factors such as pharmacogenomics, costs, and individual circumstances, as there is no universally perfect pharmacotherapy regimen. It can be challenging to retrieve a large number of patients who have received specialized optimization of their pharmacotherapy. Additionally, the sample population should be representative of the specific setting in which the treatment is intended to be utilized. For instance, when considering the typical cardiovascular patient in a pharmacy or clinical setting, it is important to include both healthy individuals and those with common indications in the database. However, the sample group in this particular study consisted mainly of a homogeneous cohort of elderly cardiovascular patients residing in the community.

5. Conclusions

Based on the diagnoses of the patients, the current decision tree-based AI methodology demonstrated a notably high level of accuracy in recommending appropriate medications such as ACEI/ARB, MRA, and nitroglycerin to the patients. Challenges with recommending other medications are anticipated to be addressable through the utilization of larger and more diverse patient datasets. The decision tree approach offers advantages over neural networks by allowing for the modification of data in response to unforeseen circumstances, such as drug safety alerts, contraindications, or the introduction of new drugs or therapies. The development of an improved model necessitates a larger cohort of well-optimized patients, which may be challenging to acquire. While further research is clearly needed, this study has the potential to address issues and propose solutions to enhance the feasibility of AI in pharmacotherapy decision-making.

Ethical approvement

The basic study was approved by the ethics committee of the Medical Association of Westphalia-Lippe, Germany (AKZ-2013–292-f-s).

CRediT authorship contribution statement

Michael Bücker: Writing – review & editing, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Kreshnik Hoti:** Writing – review & editing, Software, Resources, Methodology, Investigation, Conceptualization. **Olaf Rose:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

Michael Bücker declares no conflict of interest. Kreshnik Hoti declares no conflict of interest. Olaf Rose declares no conflict of interest.

References

1. Hobbs FDR, Erhardt L. Acceptance of guideline recommendations and perceived implementation of coronary heart disease prevention among primary care physicians in five European countries: the Reassessing European Attitudes about Cardiovascular Treatment (REACT) survey. *Fam Pract*. 2002;19(6):596–604. https://doi.org/10.1093/fampra/19.6.596.
2. Komajda M, Follath F, Swedberg K, et al. The EuroHeart Failure Survey programme–a survey on the quality of care among patients with heart failure in Europe. Part 2: treatment. *Eur Heart J*. 2003;24(5):464–474. https://doi.org/10.1016/s0195-668x(02)00700-5.
3. Vonbank A, Saely CH, Rein P, Sturn D, Drexel H. Current cholesterol guidelines and clinical reality: a comparison of two cohorts of coronary artery disease patients. *Swiss Med Wkly*. 2013;143, w13828. https://doi.org/10.4414/smw.2013.13828.
4. Naseralallah L, Stewart D, Price M, Paudyal V. Prevalence, contributing factors, and interventions to reduce medication errors in outpatient and ambulatory settings: a

systematic review. *Int J Clin Pharmacol.* 2023;45(6):1359–1377. https://doi.org/10.1007/s11096-023-01626-5.

5. Jaspers MWM, Smeulers M, Vermeulen H, Peute LW. Effects of clinical decision-support systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic review findings. *J Am Med Inform Assoc.* 2011;18(3):327–334. https://doi.org/10.1136/amiajnl-2011-000094.

6. Sim I, Gorman P, Greenes RA, et al. Clinical decision support systems for the practice of evidence-based medicine. *J Am Med Inform Assoc.* 2001;8(6):527–534. https://doi.org/10.1136/jamia.2001.0080527.

7. Koh D-M, Papanikolaou N, Bick U, et al. Artificial intelligence and machine learning in cancer imaging. *Commun Med (Lond).* 2022;2:133. https://doi.org/10.1038/s43856-022-00199-0.

8. Hollon TC, Pandian B, Adapa AR, et al. Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. *Nat Med.* 2020;26(1):52–58. https://doi.org/10.1038/s41591-019-0715-9.

9. Das N, Topalovic M, Janssens W. Artificial intelligence in diagnosis of obstructive lung disease: current status and future potential. *Curr Opin Pulm Med.* 2018;24(2):117–123. https://doi.org/10.1097/MCP.0000000000000459.

10. Zhao M, Hoti K, Wang H, Raghu A, Katabi D. Assessment of medication self-administration using artificial intelligence. *Nat Med.* 2021;27(4):727–735. https://doi.org/10.1038/s41591-021-01273-1.

11. McCahon D, Duncan P, Payne R, Horwood J. Patient perceptions and experiences of medication review: qualitative study in general practice. *BMC Prim Care.* 2022;23(1):293. https://doi.org/10.1186/s12875-022-01903-8.

12. Bolsinger J, Jaeger M, Hoff P, Theodoridou A. Challenges and opportunities in building and maintaining a good therapeutic relationship in acute psychiatric settings: a narrative review. *Front Psych.* 2019;10:965. https://doi.org/10.3389/fpsyt.2019.00965.

13. Buchegger S, Iglseder B, Alzner R, et al. Patient perspectives on, and effects of, medication management in geriatric fallers (the EMMA study): protocol for a mixed-methods pre-post study. *BMJ Open.* 2023;13(2), e066666. https://doi.org/10.1136/bmjopen-2022-066666.

14. Kerliu L, Citaku D, Rudhani I, Hughes JD, Rose O, Hoti K. Exploring instruments used to evaluate potentially inappropriate medication use in hospitalised elderly patients in Kosovo. *Eur J Hosp Pharm.* 2021;28(4):223–228. https://doi.org/10.1136/ejhpharm-2019-001904.

15. Kessler S, Desai M, McConnell W, et al. Economic and utilization outcomes of medication management at a large Medicaid plan with disease management pharmacists using a novel artificial intelligence platform from 2018 to 2019: a retrospective observational study using regression methods. *J Manag Care Spec Pharm.* 2021;27(9):1186–1196. https://doi.org/10.18553/jmcp.2021.21036.

16. Rose O, Schaffert C, Czarnecki K, et al. Effect evaluation of an interprofessional medication therapy management approach for multimorbid patients in primary care: a cluster-randomized controlled trial in community care (WestGem study protocol). *BMC Fam Pract.* 2015;16:84. https://doi.org/10.1186/s12875-015-0305-y.

17. Köberlein-Neu J, Mennemann H, Hamacher S, et al. Interprofessional medication management in patients with multiple morbidities. *Dtsch Arztebl Int.* 2016;113(44):741–748. https://doi.org/10.3238/arztebl.2016.0741.

18. Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and Regression Trees.* Routledge; 2017.

19. Fleming V, Prasad A, Ge C, et al. Prevalence and predictors of shared decision-making in goals-of-care clinician-family meetings for critically ill neurologic patients: a multi-center mixed-methods study. *Crit Care.* 2023;27(1):403. https://doi.org/10.1186/s13054-023-04693-2.

20. Hong G, Smith M, Lin S. The AI will see you now: feasibility and acceptability of a conversational AI medical interviewing system. *JMIR Form Res.* 2022;6(6), e37028. https://doi.org/10.2196/37028.

21. Meslin SMM, Zheng WY, Day RO, Tay EMY, Baysari MT. Evaluation of clinical relevance of drug-drug interaction alerts prior to implementation. *Appl Clin Inform.* 2018;9(4):849–855. https://doi.org/10.1055/s-0038-1676039.