A Treatment Recommender Clinical Decision Support System for Personalized Medicine: Method Development and Proof-of-concept for Drug Resistant Tuberculosis

Lennert Verboven (lennert.verboven@uantwerpen.be)
University of Antwerp
Toon Calders
University of Antwerp
Steven Callens
Ghent University Hospital
John Black
University of Cape Town
Gary Maartens
University of Cape Town
Kelly E. Dooley
Johns Hopkins Medicine
Samantha Potgieter
University of the Free State
Robin M. Warren
Stellenbosch University
Kris Laukens
University of Antwerp
Annelies Van Rie
University of Antwerp

Research Article

Keywords: Clinical Decision Support System, treatment individualisation, machine learning

Posted Date: February 16th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-199180/v1

Title: A treatment recommender clinical decision support system for personalized medicine: method development and proof-of-concept for drug resistant tuberculosis

Authors: Lennert Verboven[1,2], Toon Calders[2], Steven Callens[3], John Black[4], Gary Maartens[5], Kelly E. Dooley[6], Samantha Potgieter[7], Robin M. Warren[8], Kris Laukens[2], Annelies Van Rie[1]

Affiliations:

1. Torch Consortium FAMPOP Faculty of Medicine and Health Sciences, University of Antwerp, Belgium
2. ADReM Data Lab, Department of Computer Science, University of Antwerp, Belgium
3. Department of Internal Medicine & Infectious diseases, Ghent University Hospital, Ghent, Belgium
4. Department of Internal Medicine, University of Cape Town and Livingstone Hospital, Port Elizabeth, South Africa
5. Division of Clinical Pharmacology, Department of Medicine, University of Cape Town, South Africa
6. Divisions of Clinical Pharmacology and Infectious Diseases, Johns Hopkins University School of Medicine, Baltimore, MD USA
7. Division of Infectious Diseases, Department of Internal Medicine, Faculty of Health Sciences, University of the Free State, Bloemfontein, South Africa
8. DSI-NRF Centre of Excellence for Biomedical Tuberculosis Research, SAMRC Centre for Tuberculosis Research, Division of Molecular Biology and Human Genetics, Stellenbosch University, Cape Town, South Africa

Corresponding author: Lennert Verboven

ABSTRACT

Background

Individualised or precision medicine tailors care based on the patient's or pathogen's genotypic and phenotypic characteristics. An automated Clinical Decision Support System (CDSS) could help translate the genotypic and phenotypic characteristics into optimal treatment and thus facilitate implementation of individualized treatment by less experienced physicians.

Methods

We developed a hybrid knowledge- and data-driven treatment recommender CDSS. Stakeholders and experts first define the knowledge base by identifying and quantifying drug and regimen features for the prototype model input. In an iterative manner, feedback from experts is harvested to generate model training datasets, machine learning methods are applied to identify complex relations and patterns in the data, and model performance is assessed by estimating the precision at one, mean reciprocal rank and mean average precision. Once the model performance no longer iteratively increases, a validation dataset is used to assess model overfitting.

Results

We applied the novel methodology to develop a treatment recommender CDSS for individualized treatment of drug resistant tuberculosis as a proof of concept. Using input from stakeholders and three rounds of expert feedback on a dataset of 355 patients with 129 unique drug resistance profiles, the model had a 95% precision at 1 indicating that the highest

ranked treatment regimen was considered appropriate by the experts in 95% of cases. Use of a validation data set however suggested substantial model overfitting, with a reduction in precision at 1 to 78%.

Conclusion

Our novel and flexible hybrid knowledge- and data-driven treatment recommender CDSS is a first step towards the automation of individualized treatment for personalized medicine. Further research should assess its value in fields other than drug resistant tuberculosis, develop solid statistical approaches to assess model performance, and evaluate their accuracy in real-life clinical settings.

KEYWORDS

Clinical Decision Support System, treatment individualisation, machine learning

INTRODUCTION

Evidence-based medicine aims to integrate individual clinical experience with the best available external scientific evidence to develop trustworthy clinical practice guidelines and optimize clinical decision making [1, 2]. Under the evidence-based medicine paradigm, the clinician uses sound evidence to formulate the best therapeutic choice for their patient, most often through a standardized public health approach where national or international guidelines are implemented. More recently, personalised medicine has gained increasing attention. Personalized or precision medicine tailors medical decisions to the individual patient based on their predicted response to treatment in order to administer “therapy with

64 the right drug at the right dose in the right patient" [3, 4]. To implement personalized
65 medicine, diagnostic tests are performed to determine the patient's and/or the pathogen's
66 phenotypic and genetic characteristics. Integrating individual patient genomic information
67 into a clinical decision is however challenging, especially for non-experts, given the rapid
68 evolution in knowledge on the genotype-phenotype associations. The use of a clinical decision
69 support system (CDSS) could facilitate the use of personalized medicine approaches by less
70 experienced physicians and other health care workers at the time and location of patient care
71 [5, 6].

72 CDSSs for guiding treatment decisions can either be data-driven or knowledge-driven.
73 Knowledge-driven CDSSs use a rule-based system, implement guidelines developed by
74 national or international organizations such as CDC and WHO, and operate at a rather coarse
75 level and do not consider all available patient or pathogen information [7, 8]. In contrast, data
76 driven CDSSs use techniques such as machine learning and data mining aim to use all relevant
77 data to learn complex relations and dynamics from past experience and reveal patterns in the
78 data in order to assist with the complex decision making. For personalized medicine, data-
79 driven CDSSs are attractive as data is increasingly being collected and stored [9-13].

80 Recommender systems use machine learning and data mining techniques to predict the
81 preference a user would give to a specific item based on their preference history [14].
82 Recommender systems are mostly used to make personalized recommendations in e-
83 commerce (e.g. Amazon), online media (e.g. Netflix), social media, and online news feeds [15].
84 Most recommender systems either use collaborative filtering, content-based filtering, or a
85 combination of the two. Collaborative filtering recommender systems predict which items a

user might like based on other similar users that watched similar items [16] and implicit (i.e. a user watched a movie) or explicit (i.e. a user gave a 5-star rating) preferences by the user. Content-based recommender systems predict which new items the user will like by learning a classifier of the likes and dislikes of a user using the features associated with the items they like. [14, 17]. Data driven CDSSs and recommender systems are uncommon in clinical practice, mostly due the perception that they are a 'black box' tool [6] with a decision process that lacks transparency, even though transparent recommender systems exist [18, 19].

Crowdsourcing is a problem-solving model in which a large open group of actors try to collectively solve a larger problem [20] with many applications such as street mapping (OpenStreetMap; a collaborative effort of mappers contributing to create and maintain world map data), and data science (Netflix Prize; an open competition to develop the best collaborative filtering algorithm for Netflix [21]).

In this study we report the development of a fully automated hybrid knowledge- and data-driven CDSS to identify the optimal treatment regimen for individual patients and present the personalized treatment of drug resistant tuberculosis as a use case to explain the methodology and its potential application in global health.

METHODS

To ensure a transparent and standardized process, we adapted the multi-step approach for the development of a decision aid [22] (Fig 1). In this methods section, we describe the complete process of developing the treatment recommender CDSS. In the results section, we

present the results of the application of this developed treatment recommender CDSS for the individualized treatment of drug resistant tuberculosis.

Defining the scope and assembling the expert panel

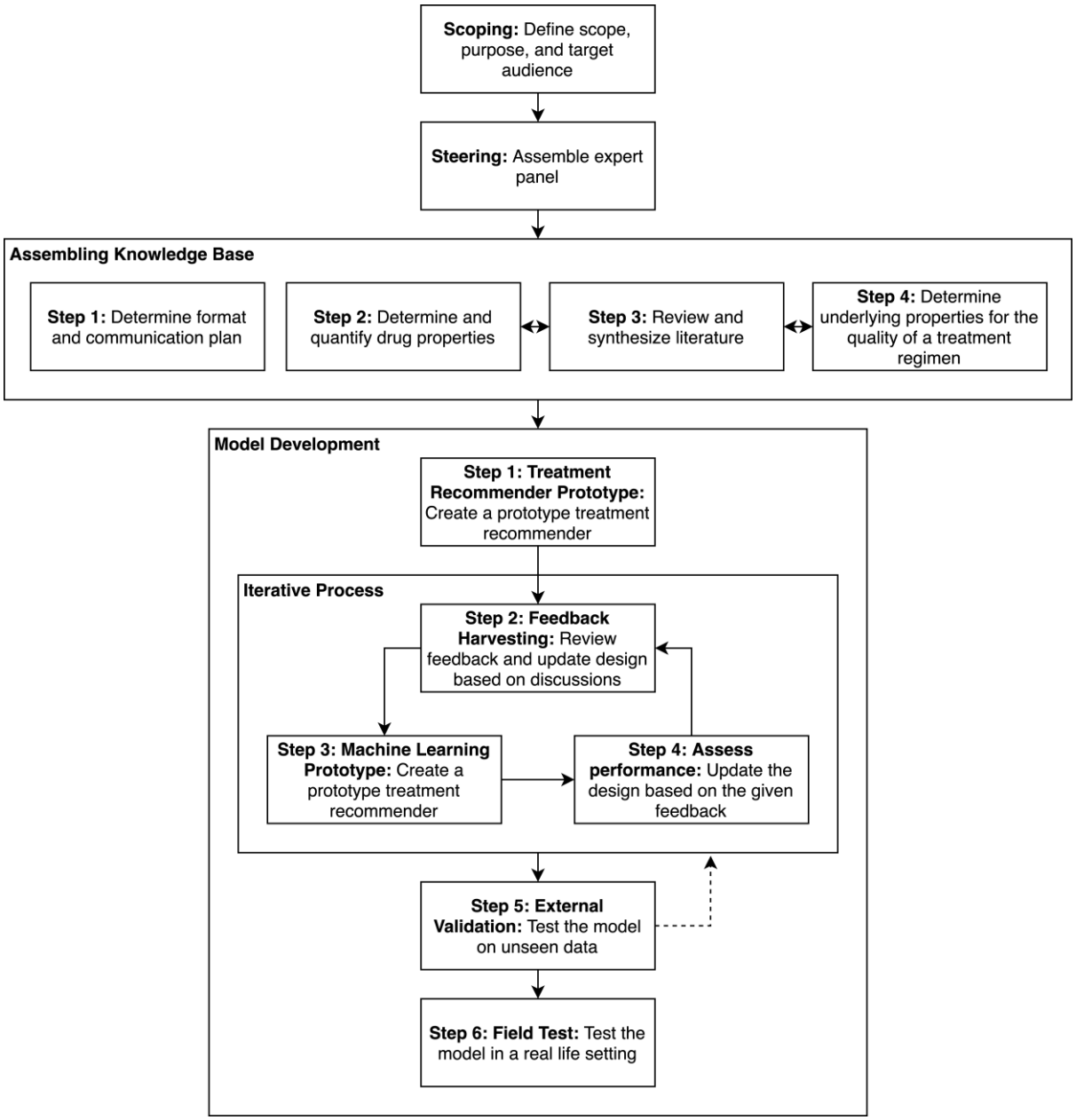
After defining the purpose and target audience of the treatment recommender CDSS (scoping step), an expert panel was assembled (steering step). To ensure a multidisciplinary and holistic approach, the panel represented expertise in pathogen genomics (knowledge of genotype-phenotype associations), pharmacology (characteristics of the drugs, drug-drug interactions, and synergy and antagonism between drugs), clinical experts, as well as health system and health economic experts. In addition, patients were consulted to provide their perspective (e.g. the impact of side effects on their quality of life).

Assembling the knowledge base

The knowledge base was developed in an iterative manner by combining review of published literature, non-published data when gaps in the literature were identified, and consensus building between experts using a standardized format for efficient expert feedback. First, the key features of relevant individual drugs were determined. Next, the features of the treatment regimen were established, hereafter referred to as regimen features, including the number of effective drugs required and how drug features are aggregated into regimen features. Third, data input requirements and decisions on user-friendly design of the treatment recommender and communication of the recommendation were made.

Model development

126 *Figure 1:teps in the development process of an automated clinical treatment recommendation system*



127

128 Development of the model consists of 6 steps (Figure 1). First, a prototype was developed to
129 rank all possibly valid treatment regimens for an individual patient. Second, expert feedback
130 was harvested on a sample of the top scoring regimens for patients that are representative
131 of the target population. Third, the expert feedback was used to develop a training dataset.

Fourth, the model performance in recommending the optimal individual treatment regimen was assessed. When the model performance was shown to be suboptimal, the rankings for the patient-regimen pairs obtained by the random forest classifier were used in determining the sampling for the next round of expert feedback harvesting. This process of machine learning, expert data harvesting and assessment of model performance was repeated until the model no longer substantially improved. Fifth, the final model was tested on a different real-life clinical dataset to assess the degree of overfitting to the training data and verify that the model is transferable to new data. Overfitting occurs when a model corresponds too closely or exactly to the data on which it was trained, and may therefore underperform on new and unseen data. The final step in the development process was a field test to assess the effectiveness of the treatment recommender CDSS model for individualized treatment in clinical trial participants. In the intervention arm, the minimum required patient information and data on genomics information (treatment recommender CDSS input) was used by the treatment recommender CDSS to proposes the optimal individualized treatment for that patient (treatment recommender CDSS output). In the section below, we describe the first three steps in greater detail.

Step 1: Developing the treatment recommender prototype

Based on the knowledge base assembled, the prototype computes a quality score for every valid patient-regimen pair (Fig 2). A valid regimen is defined as a regimen that only contains valid drugs, and valid drugs are defined as drugs that are effective (no resistance detected) and can be included in the individualized treatment regimen because of absence of clinical contraindications, drug stock outs, or country-specific drug licensing issues.

154 The number of possible regimens r for a given patient is:

155 $$r = \binom{n}{k} = \frac{n!}{k!\,(n-k)!} \quad 1)$$

156

157 where n is the number of drugs available for a specific patient and k is the number of effective
158 drugs required in a treatment regimen. If the number of drugs to be included in the
159 individualized treatment regimen can vary, then the total number of possible regimens for
160 each patient equals the sum of formula 1 for all possible values of k.

161 The total number of unique resistance profiles rp for a disease is:

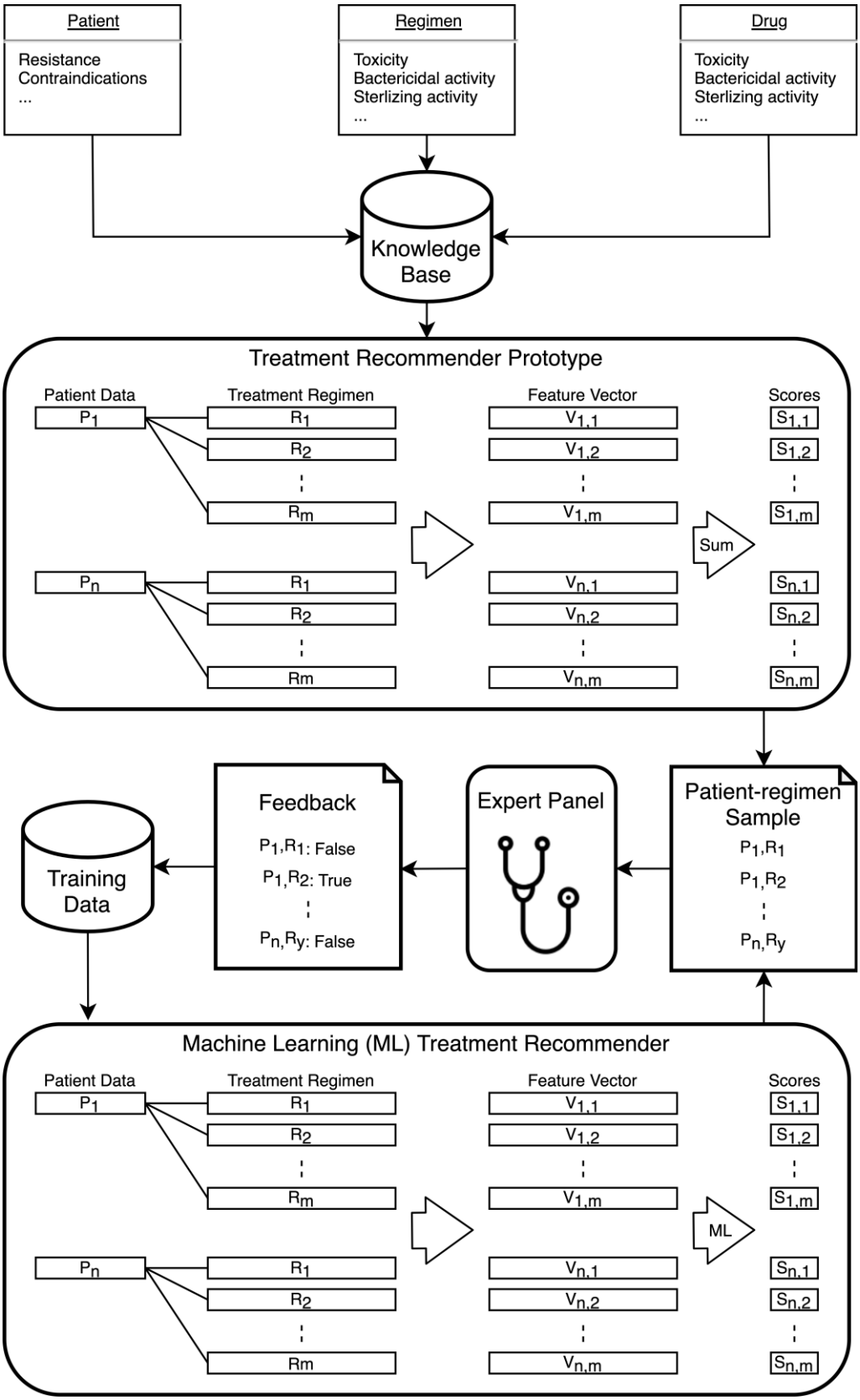
162 $$rp = \sum_{k=0}^{d} \binom{d}{k} = 2^d \quad 2)$$

163 where d is the total number of available drugs for the disease of interest.

164 In the prototype, all regimen features were normalised from 0 (bad) to 1 (good) on the patient
165 level, meaning that the entire 0 to 1 interval is used to represent each regimen feature even
166 in patients with few available drugs. A higher score can be interpreted as better for the patient
167 and the highest scoring regimen is assumed to be the best regimen for that patient. After
168 normalising all regimen features and inverting negative features, such that for all regimen
169 features a lower score means worse for the patient, the sum of all features equates to the
170 quality score for that regimen, which is then used to rank the regimens for individual patients.

171 Step 2: Harvesting expert feedback

172 *Figure 2:The model development flow*



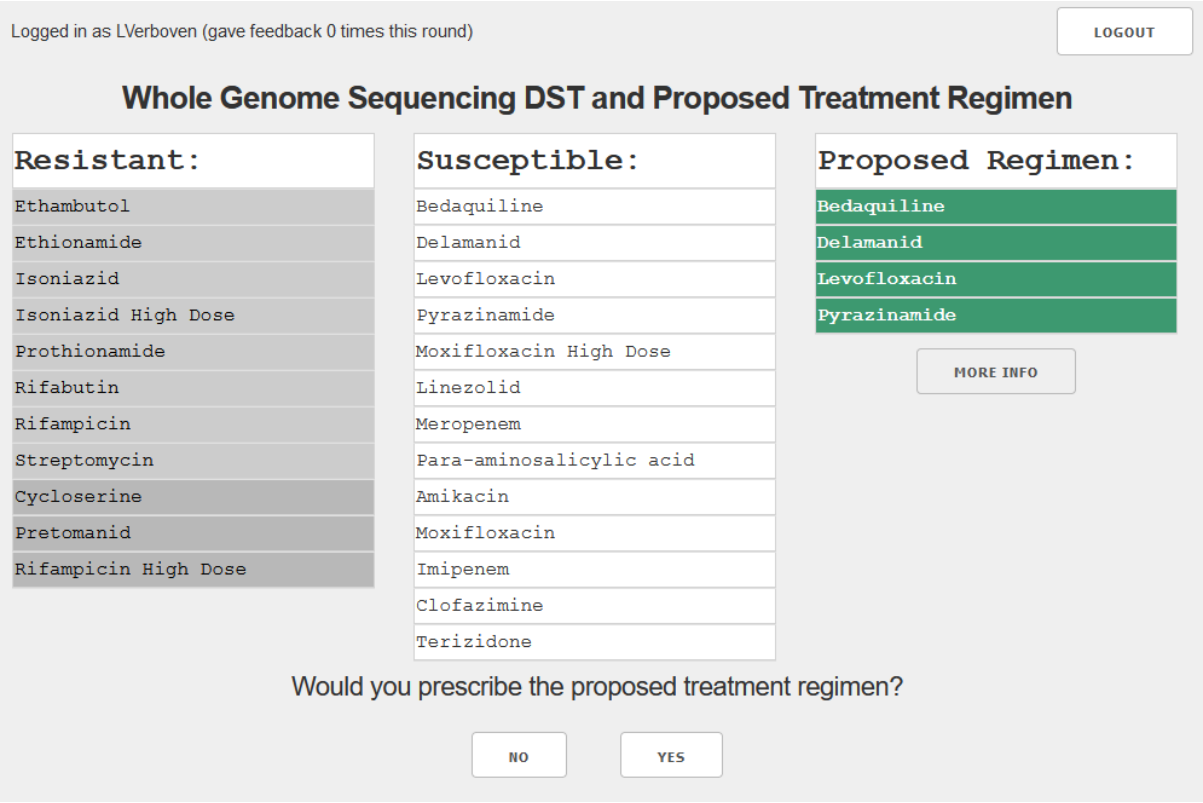
173

For each of the patients selected to represent the target patient population, a sample of the top scoring regimens was reviewed by clinicians experienced in treating the condition of interest (Figure 2). The number of cases reviewed was fixed and set to be large enough to generate sufficient data to train the machine learning model but small enough so that the experts were able and willing to carefully review every case presented. We sampled with replacement to allow that multiple experts provide feedback on the same patient-regimen pair, to allow that experts can provide feedback on the same patient-regimen pair multiple times, and to ensure that a single expert cannot veto a top scoring regimen. The sampling function randomly samples 1 regimen from the top 3 regimens (based on the ranking) for that patient. A regimen i that has been sample before however has a probability p_i

$$p_i = 1 - {p_{initial}}^{f_i} \quad 3)$$

where $p_{initial}$ is the initial probability and f_i is the amount of feedback already received for regimen i, to be removed from the top 3 making room for the next regimen in the ranking. Parameter $p_{initial}$ can be tuned to make regimens more or less likely to be removed from the top 3. After tuning, we found $p_{initial} = 0.2$ to give a good balance between resampling a regimen and allowing new, lower ranked regimens to be sampled.

Figure 3:Llayout of the web interface



To harvest the experts' feedback on the regimen-patient pairs, we developed a secure web interface (Figure 3) that captured the feedback in a structured format. The web interface contained four components: list of drugs that are not valid for this patient (due to resistance, contraindications or not being available), list of valid drugs for the patient, and the recommended treatment regimen. Upon review of this information, the expert is asked whether he/she would prescribe this regimen for the individual patient. If the expert responded they would not prescribe the recommended regimen to the individual patient, they were asked why (open field) and were asked to list the regimen they would prescribe. The alternative regimen increased the number of patient-regimen pairs with positive feedback. The reasons for rejection were used to identify recurring topics that were then

discussed with the experts. The outcome of these discussions led to additional literature review and possible addition, modification, or removal of drug or regimen features.

Step 3: Data driven machine learning model development

Our recommender model is different from standard recommender systems in that it does not use a patient's treatment history to propose a new treatment, instead our recommender model works by learning the underlying qualities of good treatments from other patients. The feedback given by the experts was stored in a database and served as training data for the machine learning model. We used a random forest classifier to learn the importance of features and interactions between the features of a treatment regimen because random forest classifiers are robust against overfitting and easy to develop. An additional advantage of random forests is that they are constructed by having many decision trees vote and taking the consensus of these trees. The concept of a decision tree and the way they are used in random forests is very intuitive, making these more likely to be accepted by clinicians. Using the normalised features, the model learns which regimen is good for a given patient, accounting for the fact that not all options are available. The model is trained on the patient-regimen-feedback pairs obtained through the expert feedback harvesting step and tries to predict the probability that a patient-regimen pair was deemed good by the experts (Figure 2). These probabilities are then again used to rank all patient-regimen pairs for a given patient, with the highest ranked regimen being the optimal treatment for that patient.

We used a patient level leave-one-out cross validation strategy to predict the new ranking of regimens for a given patient. In other words, when using the model to predict the ranking for

patient p, we used the entire training data excluding all training data on p to train the random forest classifier.

Step 4: Assessing model performance

We used three patient level performance measures: Precision at 1 (P@1)which assesses the highest ranked regimen, Mean Reciprocal Rank (MRR) which represents the average of how high the first regimen is ranked over all patients by the model, and Mean Average Precision (MAP) which takes the position of all appropriate regimens into account.

Precision at 1 is a performance parameter where precision at N is defined by equation 4.

$$P@N = \frac{\#good\ treatments\ in\ top\ N}{\#total\ treatments\ in\ top\ N} \quad 4)$$

P@1 is the fraction of patients for which the top ranked regimen is classified by the experts as an appropriate regimen, with appropriate regimen defined as regimen the expert would be willing to prescribe for that patient. When experts disagreed on the highest ranked regimen, a majority voting was used to determine whether the regimen is appropriate.

The mean reciprocal rank is defined by equation 5,

$$MRR = \frac{1}{|P|}\sum_{p}^{P}\frac{1}{rank_p} \quad 5)$$

where P is the set of patients and $rank_p$ the position of the first regimen classified by the experts as appropriate for patient p.

The mean average precision is defined by equation 6,

$$MAP = \frac{1}{|P|}\sum_{p}^{P} AvgP(p) \quad 6)$$

where $Avg\ (p)$ is defined in equation 7

$$Avg(p) = \frac{1}{|GT_p|}\sum_{i}^{TX_p}(P@\,i\ \times\ rel(i)) \quad 7)$$

where GT_p is the set of appropriate treatments for patient p, TX_p is the set of all valid treatment regimens for patient p, and $rel(i)$ indicates whether treatment i is an appropriate treatment for patient p ($rel(i) = 1$), or an inappropriate treatment ($rel(i) = 0$).

RESULTS

In this section, we present how the newly developed method was as a proof-of-concept, applied to develop a treatment recommender CDSS to guide the individualization of treatment for Rifampicin Resistant Tuberculosis (RR-TB).

The scope and purpose of the development of a rifampicin resistant (RR)-TB treatment recommender was defined as "improving RR-TB treatment outcomes by optimizing the individualized treatment regimen in high TB burden resource limited settings". The primary target audience were the clinicians treating patients with RR-TB in such setting.

A multidisciplinary steering group of experts was assembled by inviting clinicians with experience in treating RR-TB, pharmacologists with expertise in TB drugs, molecular epidemiologists with expertise in interpretation of the genotype-phenotype associations regarding drug resistance in *Mycobacterium tuberculosis*, health systems experts to assess integration of individualized treatment in routine care, and ex-RR-TB patients.

The design phase started with a stakeholder meeting where all the members of the steering group discussed the treatment recommender input parameters. Discussions focussed on number of drugs needed in an effective treatment regimen, health system burden of treatment monitoring, monitoring burden, drug toxicity, drug features, and clinical patient characteristics, and genomic drug resistance profile. Uncertainties regarding the input parameters identified during the meeting were resolved through literature search, identification of unpublished data and iterative discussion until a consensus was reached on all features to be included in the model. These discussions resulted in a set of 24 drugs that are licensed for treatment of tuberculosis in South Africa, 9 drugs features, 18 regimen features, and the consensus that 4 effective drugs need to be included in all individualized regimens (Table 1). Based on these decisions, up to 10626 valid treatment regimens were possible (Equation 1) for a patient without resistance.

Table 1: Knowledge base on drugs, drug features and regimen features included in the treatment recommender for drug resistant tuberculosis

Drugs	Amikacin, Bedaquiline, Clofazimine, Cycloserine, Delamanid, Ethambutol, Ethionamide, Imipenem, high or standard dose Isoniazid, Levofloxacin, Linezolid, Meropenem, high or standard dose Moxifloxacin, Para-aminosalicylic acid, Pretomanid, Prothionamide, Pyrazinamide, Rifabutin, high or standard dose Rifampicin, Streptomycin, Terizidone,
Drug features	Route of administration, toxicity, QT prolongation, cost, early bactericidal activity, bactericidal activity, sterilizing activity, mechanism of action, propensity to acquire resistance
Regimen features	Core or companion drug[a] [23], prevention of acquired resistance[a], four different mechanisms of action[a], fully oral regimen[a], cost[b], toxicity[b], QT prolongation[a], high early bactericidal activity[a], early bactericidal activity[b], high early bactericidal activity[a], bactericidal activity[b], high sterilizing activity[a], sterilizing activity[b], synergism[b], antagonism[b], contra-indications[a], same drug class[a], first line drugs included[b]

[a] Binary features, [b] Continuous features

We used a dataset that contained both clinical and whole genome sequencing data on 355 patients diagnosed with RR-TB in three provinces in South Africa. The whole genome sequencing data represented 129 different drug resistance profiles. A group of 6 clinicians experienced in treatment of drug resistant tuberculosis were then asked to provide feedback on the recommended treatment using a structured online survey (Figure 3).

Table 2: Training and external validation of the treatment recommender CDSS model

	Number of regimens presented to experts	**Total number of observations[a]**	**Number of participating experts**	**Precision at 1**	**Mean average precision**	**Mean reciprocal rank**
Training round 1	479	855	5	89%	53%	90%
Training round 2	445	719	6	95%	69%	97%
Training round 3	360	607	6	95%	72%	95%
Training round 1-3	1284	2181				
External Validation	375	592	5	78%	68%	87%

[a]Number of observations include the alternative proposed by the expert in case the recommended regimen was not considered appropriate by the expert. The performance figures for the training rounds indicate the performance when training and validation on all currently available training data.

Assessment of the performance of the prototype model showed precision at 1 of 89%, mean average precision (MAP) of 53% and mean reciprocal rank (MPR) of 90% (Table 2). The written comments on regimens judged to be inappropriate were discussed with the expert group which resulted in modifications to the treatment recommender CDSS prototype. For example, the efficacy of the different drugs, which was initially quantified using a single feature was changed to three features: early bactericidal activity, bactericidal activity and sterilizing activity. Using the feedback, a random forest classifier was used to identify complex relations

and patterns in the data. Based on these results, the updated treatment recommender CDSS reclassified the order of valid treatment regimens from which a sample was drawn for a second round of feedback harvesting from the expert clinicians. The model had improved, with an increase in all three performance parameters. Precision at one increased from 89% to 95%, MAP from 53% to 69% and MRR from 90% to 97%. After three rounds, the performance no longer improved, with precision at 1 and MRR stabilizing at 95% and MAP around 70%.

For the external validation, another dataset consisting of 64 unique resistance profiles for patients diagnosed with RR-TB in another province of was used. On this external validation set, the model performance was lower, with a precision at 1 of 78%, MAP of 68% and MRR of 87% (table 2).

DISCUSSION

We developed a novel treatment recommender CDSS that combines a knowledge-driven approach using feedback harvesting methodology with a data-driven approach using machine learning methods to automate the individualisation of treatment. The knowledge-driven component consists of input and feedback from stakeholders and experts. The data-driven component consists of the machine learning methods to identify complex relations and patterns.

Our approach is fundamentally different from standard knowledge-driven approaches to CDSSs and offers several advantages for personalized medicine. Standard knowledge-driven

systems implement clinical guidelines using if-then rules which allow limited treatment individualisation [6] and offer little flexibility as they need a complete overhaul when new drugs become available or new knowledge becomes available. In contrast, including new drugs and incremental knowledge on existing drugs in our novel methodology is possible without new model training through quantifying the features of the new drug or updating the relevant drug features.

Our approach is also fundamentally different from the data-driven methods that have been used for other CDSSs. Grasser et al. developed two therapy decision support systems that use historical treatment data to individualise psoriasis treatment based on patient attributes and past treatment attributes to predict the response to different therapies [6]. The main limitations of this methodology are that it does not learn the underlying properties of optimal treatment but bases decisions on which treatments have worked well in the past similar patients. Consequently, this method suffers from concept drift where statistical properties of data change over time due to the discovery of a new drug or new knowledge on drug features. Our model is less subject to concept drift, as it is possible to assign features to a new drug or update the features of the drugs when statistical properties of drug or regimen features change.

Machine learning and artificial intelligence methods such as neural networks have been used in the framework of personalized medicine to learn complex and nonlinear relationships between prognostic features and an individual patient's risk of treatment failure [11]. This approach requires a dataset with treatment outcomes. Because of the relative novelty of personalized medicine, such datasets with treatment outcome only allow the model to learn

337 the underlying properties of the current standard of care instead of novel individualized
338 treatment regimens.

339 The application of our newly developed treatment CDSS methodology to the individualization
340 of treatment of RR-TB provided proof-of-principle by demonstrating that the novel approach
341 is well suited to guide a personalized medicine approach when multiple combination of drugs
342 are possible in one individual patient. Using the proposed method, a treatment recommender
343 CDSS for personalized treatment can be built in a relatively short period of time using a
344 combination of stakeholder input, published and unpublished evidence, and expert feedback.

345 By harvesting expert feedback on patient scenario's simulated from real-life data, we created
346 a minimal dataset consisting of diverse individual treatment regimens that are representative
347 for patient care. By allowing the experts to provide an alternative regimen when disagreeing
348 with the proposed regimen, the resulting dataset contained a wide variety of accepted and
349 rejected treatment regimens.

350 While our novel method has many strengths for the field of precision medicine, including the
351 hybrid data- and knowledge-driven approach, the use of a structured 'crowdsourcing'
352 approach with predetermined experts for treatment decision making research [24], and the
353 high degree of flexibility, several limitations should be noted. First, given the novelty of our
354 approach, there is no consensus method yet to assess the performance of the model and
355 there are no clear decision boundaries on when to stop the iterative development process.
356 The external validation of our model for drug resistant TB demonstrated substantial
357 overfitting on the training data, suggesting that the current performance parameters may not
358 accurately capture the model's performance. Using a more diverse set of patient profiles and

repeating the development process on this new dataset in combination with the already collected data is likely to reduce overfitting. As such the data now used for external validation could be included as training data before bringing this model into clinical practice. Second, the model was developed using a limited group of experts. It is unknown to which degree the model development is dependent of the number and composition of the expert group. Use of other performance parameters and evaluation of the model in clinical trials in different regions of the world will be needed to determine its accuracy for real-life decision making in a personalized medicine.

In conclusion, while novel and promising, our hybrid knowledge- and data-driven treatment recommender CDSS for individualising treatment is an important first step in the development of methods aimed to facilitate the widespread implementation of personalized medicine. Further research should assess its value in fields other than drug resistant tuberculosis, develop solid statistical approaches to assess model approaches, and evaluate their accuracy in real-life clinical settings.

LIST OF ABBREVIATIONS

CDSS Clinical Decision Support System

MAP Mean Average Precision

MRR Mean Reciprocal Rank

P@1 Precision at 1

RR-TB Rifampicin Resistant Tuberculosis

DECLARATIONS

Ethics approval and consent to participate

The research was performed in accordance with the Declaration of Helsinki. Collection of the clinical strains and the determination of the genotypic drug resistance profile was approved by appropriate ethics committees. For the training dataset, ethical approval was obtained from the Human Research Ethics Committee of the University of the Witwatersrand in South Africa and the Institutional Review Board of the University of North Carolina at Chapel Hill in the United States (reference numbers M111139 and 11-2273). Approval of study activities was also obtained from relevant health authorities. Participants gave verbal informed consent by phone (recorded). The Research Ethics Committee or Institutional Review Board of Human Research Ethics Committee of the University of the Witwatersrand in South Africa (M111139) and the Institutional Review Board of the University of North Carolina at Chapel Hill in the United States (11-2273) waived the need of informed consent for patients who had died or were lost to follow-up from TB care prior to study enrolment and could not be contacted despite multiple attempts. For the validation dataset, ethical clearance was obtained from Stellenbosch University Health Research Ethics Committee. This ethics approval (ref number N09/11/269) has a waiver of informed consent. The study titled "a personalised recommender system for whole genome sequencing-based individualised treatment for drug resistant tuberculosis" reported in this manuscript was approved by the Ethics Committee of the University Hospital Antwerp and the University of Antwerp (ref number 2107101).

Consent for publication

Not applicable

Availability of data and materials

The datasets analysed are available from the corresponding author on reasonable request.

Competing interests

Not applicable

Funding

Fonds Wetenschappelijk Onderzoek, 1S39119N

Fonds Wetenschappelijk Onderzoek, G0F8316N

U.S. Food and Drug Administration, R01FD005724

National Institute of Allergy and Infectious Diseases, K24AI150349

South African Medical Research Council

Department of Science and Innovation - National Research Foundation, South Africa

Authors' contributions

LV, AVR, RW and KL set up the concept of the study. LV, KL, TC, and AVR developed the methods. AVR was principal investigator of the study that collected the data used in the training data set. RW is the principal investigator of the study that collected the data used in the validation data set. SC, GM, SP, JB, and KD contributed as experts to the data. LV analysed and interpreted the data. LV wrote the first draft; AVR and KL contributed substantially in revising the manuscript. All authors read and approved the final version of the manuscript.

Acknowledgements

We would like to thank everyone who helped make this research possible. Specifically, the people who helped organise and participated in the expert panel meetings, namely Fransceca Conradie, Heeran Makkan, Trevor Beattie, Margaretha De Vos, Anzaan Dippenaar, Marissa Klopper, Elise De Vos, Leen Rigouts, Noor Zakhura, and Makanda Goodman.

REFERENCES

1. Sackett, D.L., et al., *Evidence based medicine: what it is and what it isn't.* BMJ, 1996. **312**(7023): p. 71-2.
2. Djulbegovic, B. and G.H. Guyatt, *Progress in evidence-based medicine: a quarter century on.* Lancet, 2017. **390**(10092): p. 415-423.
3. Mancinelli, L., M. Cronin, and W. Sadee, *Pharmacogenomics: the promise of personalized medicine.* AAPS PharmSci, 2000. **2**(1): p. E4.
4. Academy of Medical Sciences, *Stratified, personalised or P4 medicine: a new direction for placing the patient at the centre of healthcare and health education.* 2015.
5. Beeler, P.E., D.W. Bates, and B.L. Hug, *Clinical decision support systems.* Swiss Med Wkly, 2014. **144**: p. w14073.
6. Grasser, F., et al., *Therapy Decision Support Based on Recommender System Methods.* J Healthc Eng, 2017. **2017**: p. 8659460.
7. Berner, E.S. and T.J. La Lande, *Overview of Clinical Decision Support Systems*, in *Clinical Decision Support Systems: Theory and Practice*, E.S. Berner, Editor. 2007, Springer New York: New York, NY. p. 3-22.
8. Zhu, N., et al., *A Decision Support System with Intelligent Recommendation for Multi-disciplinary Medical Treatment.* ACM Trans. Multimedia Comput. Commun. Appl., 2020. **16**(1s): p. Article 33.
9. Wang, M., et al., *Safe Medicine Recommendation via Medical Knowledge Graph Embedding.* arXiv 1710.05980, 2017.
10. Zhang, Y., et al., *iDoctor: Personalized and professionalized medical recommendations based on hybrid matrix factorization.* Future Generation Computer Systems, 2017. **66**: p. 30-35.
11. Katzman, J.L., et al., *DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network.* BMC Med Res Methodol, 2018. **18**(1): p. 24.

12. Chen, Y., J.D. Elenee Argentinis, and G. Weber, *IBM Watson: How Cognitive Computing Can Be Applied to Big Data Challenges in Life Sciences Research.* Clin Ther, 2016. **38**(4): p. 688-701.
13. Zauderer, M.G., et al., *Piloting IBM Watson Oncology within Memorial Sloan Kettering's regional network.* Journal of Clinical Oncology, 2014. **32**(15_suppl): p. e17653-e17653.
14. Varga, E., *Recommender Systems*, in *Practical Data Science with Python 3: Synthesizing Actionable Insights from Data*. 2019, Apress: Berkeley, CA. p. 317-339.
15. Deldjoo, Y., et al., *Recommender Systems Leveraging Multimedia Content.* ACM Computing Surveys, 2020. **53**(5): p. 1-38.
16. Breese, J.S., D. Heckerman, and C. Kadie, *Empirical analysis of predictive algorithms for collaborative filtering*, in *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*. 1998, Morgan Kaufmann Publishers Inc.: Madison, Wisconsin. p. 43–52.
17. Ricci, F., L. Rokac, and B. Shapira, *Recommender Systems Handbook*. 2009: Springer.
18. Musto, C., et al., *Linked open data-based explanations for transparent recommender systems.* International Journal of Human-Computer Studies, 2019. **121**: p. 93-107.
19. Ngo, T., J. Kunkel, and J. Ziegler, *Exploring Mental Models for Transparent and Controllable Recommender Systems: A Qualitative Study*, in *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization*. 2020, Association for Computing Machinery: Genoa, Italy. p. 183–191.
20. Howe, J., *The Rise of Crowdsourcing*, in *Wired*. 2006.
21. Netflix. *Netflix Prize*. [cited 2021 02/02/2021]; Available from: https://netflixprize.com/.
22. Coulter, A., et al., *A systematic development process for patient decision aids.* BMC Med Inform Decis Mak, 2013. **13 Suppl 2**: p. S2.
23. Van Deun, A., et al., *Principles for constructing a tuberculosis treatment regimen: the role and definition of core and companion drugs.* Int J Tuberc Lung Dis, 2018. **22**(3): p. 239-245.
24. Wazny, K., *Applications of crowdsourcing in health: an overview.* J Glob Health, 2018. **8**(1): p. 010502.

Figures

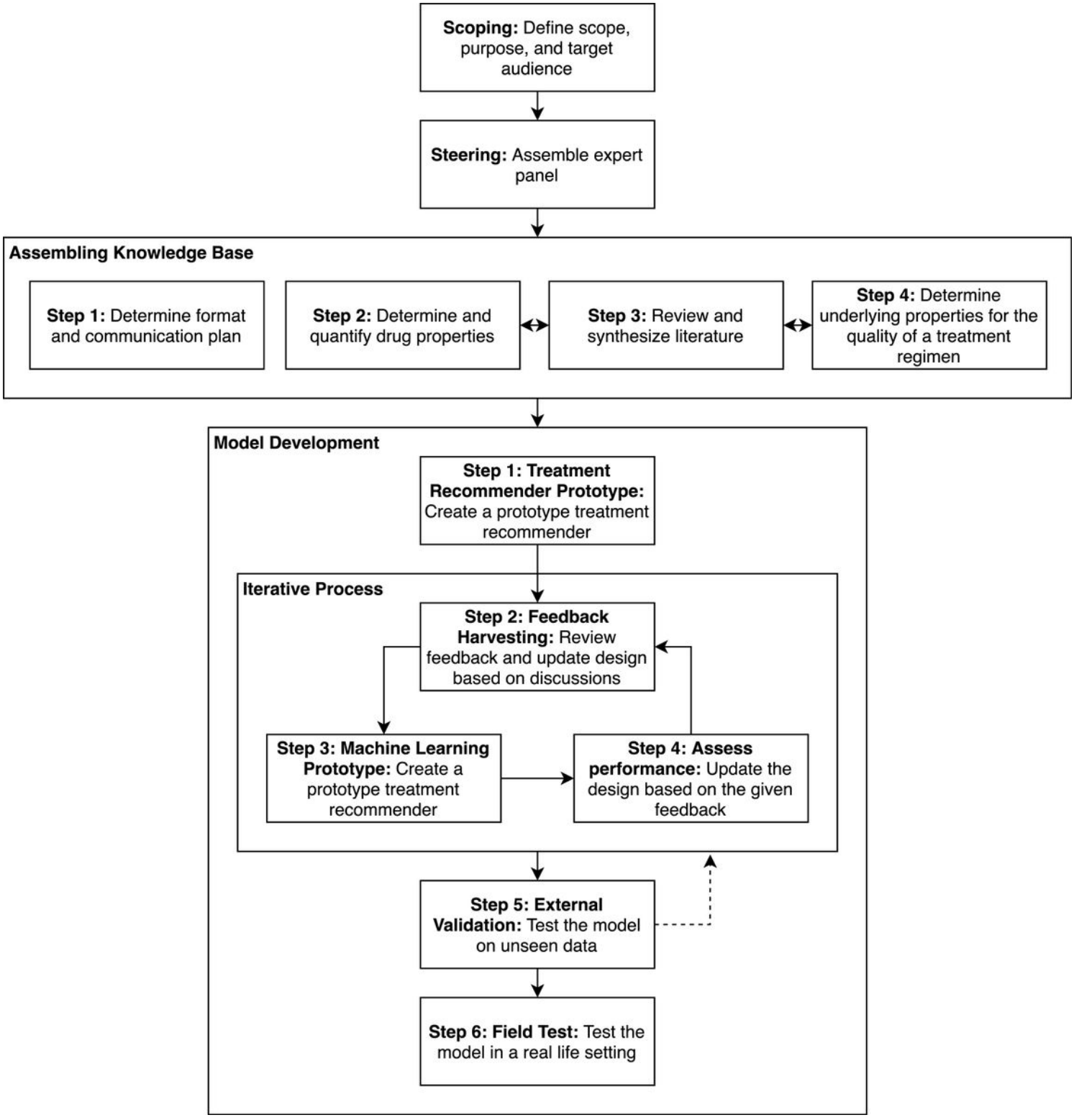


Figure 1

Steps in the development process of an automated clinical treatment recommendation system

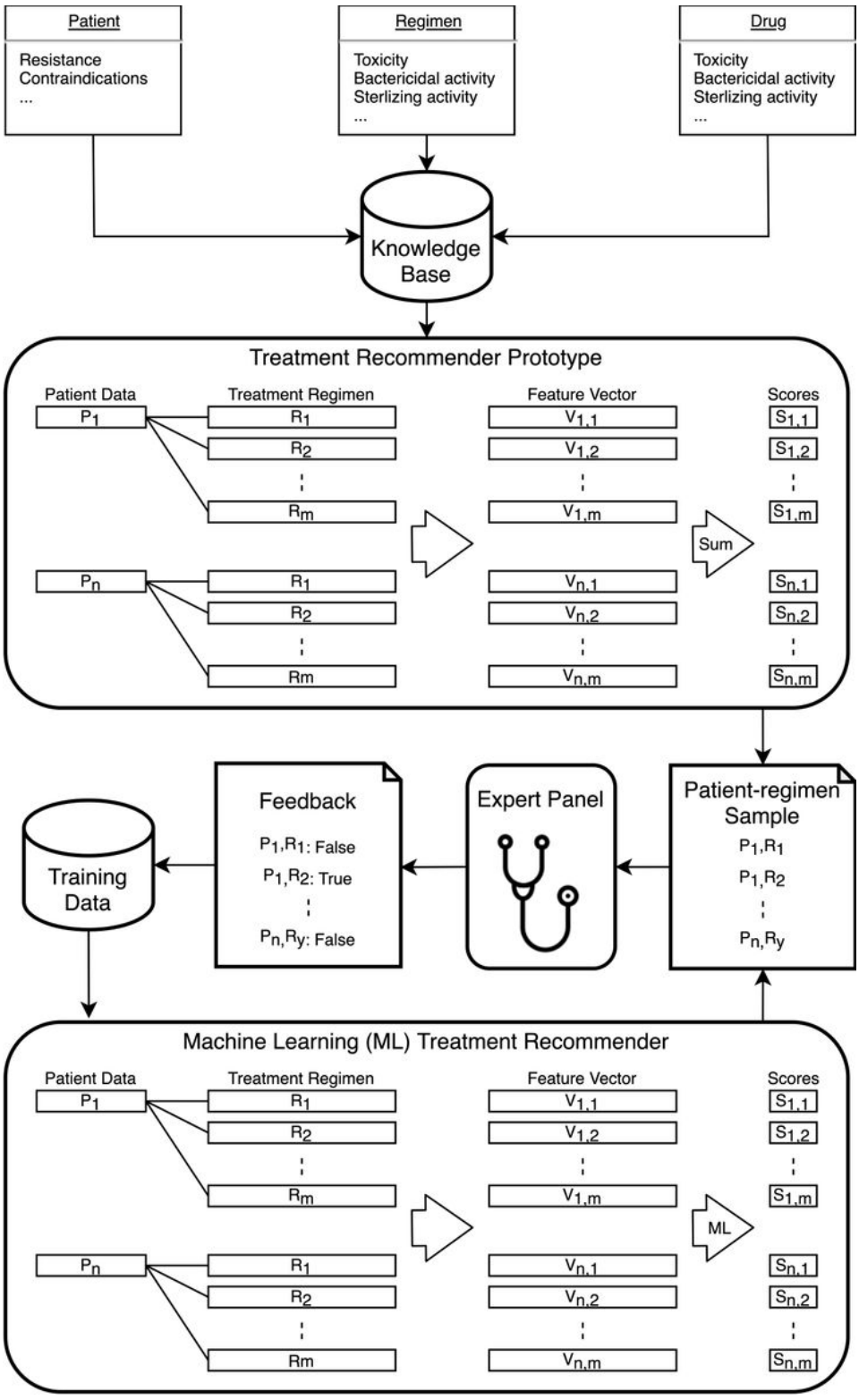


Figure 2

The model development flow

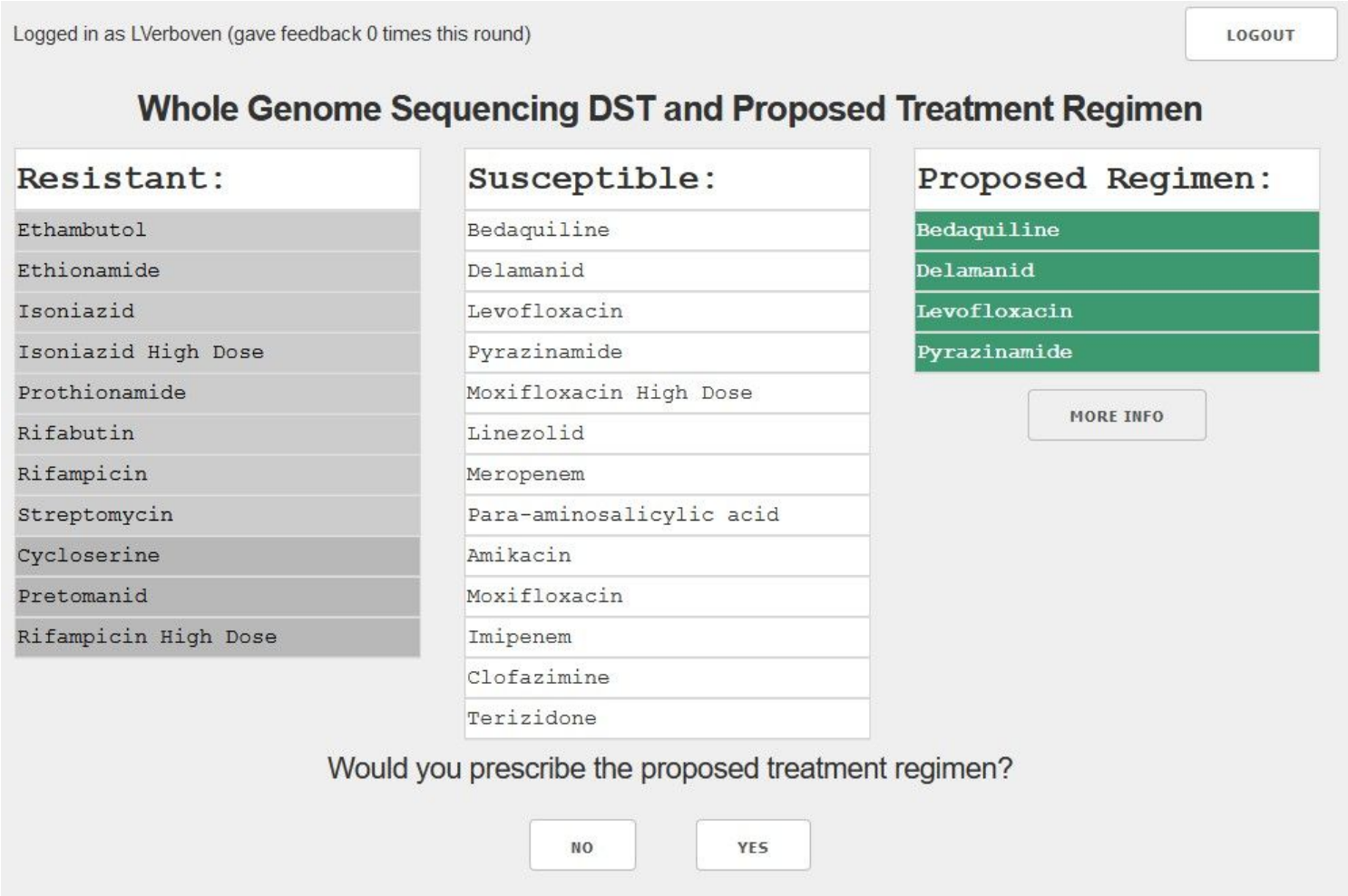


Figure 3

Layout of the web interface