

RESEARCH ARTICLE

Optimizing pharmacogenomic decision-making by data science

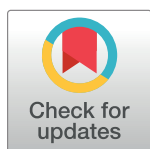
Amir M. Behdani[✉], Jessica Lai[✉], Christina Kim, Lama Basalelah, Trey Halsey, Krista L. Donohoe^{‡*}, Dayanjan Wijesinghe^{ID**}

Department of Pharmacotherapy and Outcomes Science, Virginia Commonwealth University School of Pharmacy, Richmond, Virginia, United States of America

✉ These authors contributed equally to this work.

‡ KLD and DW also contributed equally and joint senior authors on this work.

* kldonohoe@vcu.edu (KLD); wijesingheds@vcu.edu (DW)



Abstract

Healthcare systems have made rapid progress towards combining data science with precision medicine, particularly in pharmacogenomics. With the lack of predictability in medication effectiveness from patient to patient, acquiring the specifics of their genotype would be highly advantageous for patient treatment. Genotype-guided dosing adjustment improves clinical decision-making and helps optimize doses to deliver medications with greater efficacy and within safe margins. Current databases demand extensive effort to locate relevant genetic dosing information. To address this problem, Patient Optimization Pharmacogenomics (POPGx) was constructed. The objective of this paper is to describe the development of POPGx, a tool to simplify the approach for healthcare providers to determine pharmacogenomic dosing recommendations for patients taking multiple medications. Additionally, this tool educates patients on how their allele variations may impact gene function in case they need further healthcare consultations. POPGx was created on Konstanz Information Miner (KNIME). KNIME is a modular environment that allows users to conduct code-free data analysis. The POPGx workflow can access Clinical Pharmacogenomics Implementation Consortium (CPIC) guidelines and subsequently be able to present relevant dosing and counseling information. A KNIME representational state transfer (REST) application program interface (API) node was established to retrieve information from CPIC and drugs that are exclusively metabolized through CYP450, and these drugs were processed simultaneously to demonstrate competency of the workflow. The POPGx program provides a time-efficient method for users to retrieve relevant, patient-specific medication selection and dosing recommendations. Users input metabolizer gene, genetic allele data, and medication list to retrieve clear dosing information. The program is automated to display current guideline recommendations from CPIC. The integration of this program into healthcare systems has the potential to revolutionize patient care by giving healthcare practitioners an easy way to prescribe medications with greater efficacy and safety by utilizing the latest advancements in the field of pharmacogenomics.

OPEN ACCESS

Citation: Behdani AM, Lai J, Kim C, Basalelah L, Halsey T, Donohoe KL, et al. (2024) Optimizing pharmacogenomic decision-making by data science. PLOS Digit Health 3(2): e0000451. <https://doi.org/10.1371/journal.pdig.0000451>

Editor: Nino Isakadze, Johns Hopkins Medicine, UNITED STATES

Received: March 27, 2023

Accepted: January 17, 2024

Published: February 8, 2024

Copyright: © 2024 Behdani et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Author summary

Our team has developed a new program called Patient Optimization Pharmacogenomics (POPGx) to help healthcare providers efficiently determine the best pharmacogenomic medication recommendations for their patients. Pharmacogenomic medication selection and dosing recommendations improve patient safety and therapy efficacy by taking the patient's genotype into account. This process is currently time-consuming and cumbersome for healthcare providers, which can result in human-error and decreases overall pharmacogenomic utilization. POPGx was built using the Konstanz Information Miner (KNIME) which is a modular, low-code environment that allows for rapid data analysis. The program utilizes Clinical Pharmacogenomics Implementation Consortium (CPIC) guidelines to provide relevant dosing and counseling information for each metabolizer type and allows healthcare providers to input a patient's metabolizer gene, genetic allele data, and medication list. Our program is designed to streamline the pharmacogenomic therapy selection healthcare providers conduct for their patients. The goal of POPGx as demonstrated herein is to improve patient care by increasing medication safety and efficacy.

Introduction

Pharmacogenomics refers to the application of patient genomic information to optimize medication selection and dosing and is a form of precision medicine that has grown exponentially in the past 50 years.[1] A relationship between genetic polymorphism and medication metabolism has been well established by multiple studies, such as in the case of the cytochrome P450 (CYP) enzymes family. DNA sequencing elucidates likely genotypes of many enzymes, including CYPs. Enzymes play a decisive role in drug metabolism, influencing both treatment efficacy and the potential for adverse drug reactions. Variations in alleles of enzyme genes can alter enzyme expression levels in vivo, ultimately affecting drug metabolism in the body. By utilizing pharmacogenetics, healthcare providers can predict and optimize treatment outcomes and patient safety [1].

Data science is a broad interdisciplinary field that involves the use of scientific methods, systems, and algorithms to guide the extraction of information and insights from data. It involves the use of various statistical and computational techniques to analyze and interpret data, and to predict and extrapolate decisions based on that data [2]. It also involves the creation of visualization and dashboards to communicate insights to stakeholders. Furthermore, the availability of health information has increased exponentially over the decades. Digital health programs can simplify access and understanding of the expansive amount of health information available in databases to bridge the understanding of complex topics between provider and patient [3,4]. The ultimate goal of data science in healthcare is to identify patterns and trends in medical data and turn them into actionable insights that can guide decision-making processes, ultimately leading to improved patient care. Patient care can be improved by earlier detection of diseases and optimal treatment with minimal delay to enhance life quality and expectancy. Additionally, data science can help make treatments more affordable by identifying more efficient and cost-effective methods [2].

Data science and pharmacogenomics are two fields that intersect in many ways. Genomic counseling is a process of communication between a healthcare provider and a patient that involves evaluating the risks, benefits, and limitations of genetic testing and other genomic information to help make informed decisions about health outcomes. The goal of this project is to utilize data science to facilitate personalized medicine by identifying the genetic factors

that contribute to a patient's specific condition. This information can be used to develop targeted treatments that are more likely to be effective for a particular patient, based on their genetic makeup. Using data science techniques, such as statistical analysis and machine learning, a data scientist can analyze patient data to identify genetic markers that can be used to predict a patient's risk of developing certain side effects or drug interactions, as well as how a patient will respond to a particular medication. This information can then be used by health-care providers to adjust the dosage of a medication to optimize its effectiveness and reduce the risk of side effects [5].

Pharmacogenetics can significantly affect an individual's metabolic response to a medication which in turn will affect how the drug works in the body. An individual's metabolic response is typically classified into four primary categories: ultrarapid metabolizers, extensive metabolizers, intermediate metabolizers, and poor metabolizers. These categories denote how efficiently a drug is metabolized by enzymes in the body. Ultrarapid metabolizers represent patients who metabolize drugs at a faster rate compared to the general population. For active drugs, this will lead to diminished therapeutic effects at standard dosages. For prodrugs, this will lead to increased drug levels in the bloodstream and a higher risk of adverse reactions. Extensive metabolizers form the reference group; people who fall into this category metabolize medications at a predictable rate and experience predictable therapeutic responses. Intermediate metabolizers metabolize drugs at a moderately reduced rate compared to extensive metabolizers. This means that for active drugs, this will lead to higher levels of the drug in the bloodstream with greater risk of side effects. For prodrugs, this will lead to suboptimal levels of the drug in the bloodstream. Lastly, poor metabolizers have significantly reduced or even absent enzyme activity and metabolize drugs much slower than the reference population. For active drugs, this can lead to very heightened drug levels in the bloodstream with a heightened risk of adverse reactions. For prodrugs, this can lead to little to no bioactivation, and thus very suboptimal levels of the drug in the bloodstream [6].

A significant challenge in incorporating pharmacogenetic data into standard clinical practice has been the absence of clear guidelines on adjusting medication use or dosage based on genetic test outcomes [7]. To address this, the Pharmacogenomics Knowledge Base (PharmGKB) collaborated with the National Institutes of Health (NIH) to establish the Clinical Pharmacogenetics Implementation Consortium (CPIC). Notably, CPIC's guidelines emphasize the utilization of existing genetic test results rather than recommending which tests to order. This focus becomes more pertinent given the rising accessibility, decreasing costs of genetic tests, and the emergence of direct-to-consumer testing options.

The current approach to making pharmacogenetics-guided medication selection and dosing is through CPIC. Healthcare providers can use CPIC to optimize treatment selection and dosing. CPIC guidelines are standardized, peer-reviewed, evidence-based, and are freely available to the public [8]. PharmGKB is a comprehensive online database detailing gene-drug associations and genotype-phenotype relationships. In collaboration with PharmGKB, CPIC publishes guidelines based on these associations [9].

CPIC stands as an invaluable resource that aids in interpreting results and tailoring medication regimens. However, its user interface proves to be a challenge for many. Navigating the platform often directs users to dense scientific articles that require time-consuming analysis and interpretation. This not only is time-intensive but also increases the potential for error, as users may inadvertently misinterpret or overlook crucial information, especially when sifting through details for multiple drugs and gene abnormalities. Despite these challenges, many professionals still resort to manually searching CPIC, given the lack of a more efficient alternative. Our tool aims to simplify this process, extracting and presenting only the relevant information, thus streamlining the workflow and minimizing the risk of errors.

To address concerns with the current approach of using gene/drug guideline consortiums, Patient Optimization Pharmacogenomics (POPGx) was constructed. POPGx is an automated program that provides a straightforward user-interface for providers to input patient-specific genetic allele data, enzymes of interest, and medications to retrieve clear dosing recommendations. The objective of this paper is to describe the development of POPGx and its goal to create an easier to use approach for healthcare providers to determine pharmacogenomic dosing recommendations for patients taking multiple medications.

Materials and Methods

POPGx, unlike the current gene/drug guideline consortium searching method, allows for multiple drugs in a patient's regimen to be inputted into the system at one time. It retrieves and displays pertinent dosing recommendations in a centralized manner. The goal is to decrease the number of errors that can occur from the current tedious method of searching consortiums one medication at a time and having to navigate multiple guidelines. Healthcare providers can then easily interpret patients' pharmacogenetic testing in relation to their patient's medication to determine the correct dosing for efficacy and patient safety with POPGx.

POPGx was developed on the Konstanz Information Miner (KNIME) interface. The KNIME interface is an open-source modular environment which allows code-free data analysis to users by visual assembly and interactive execution of a data pipeline [10]. In this platform, individual tasks are represented by nodes that can be used to access, merge, analyze, and visualize data. Nodes can be used for integration of data from the outside sources if the data is publicly available. Since CPIC has its database and application program interface (API) available to the public, all the data can be accessed through a KNIME rest API node.

CPIC has a publicly available Representational State Transfer Application Programming Interface (REST API) [9]. This allows users to connect applications to the CPIC database and retrieve information. The GET Request node in KNIME analytical platforms acts as a bridge connecting CPIC and the KNIME user interface. After establishing a connection between KNIME and CPIC, comprehensive data cleaning and aggregation took place to meet the application's goals. Subsequently, a user-friendly application user interface was developed to meet the needs of providers and patients with minimal training. The workflow was later deployed on KNIME Server to allow users to use this application remotely on any device with internet access. See Fig 1.

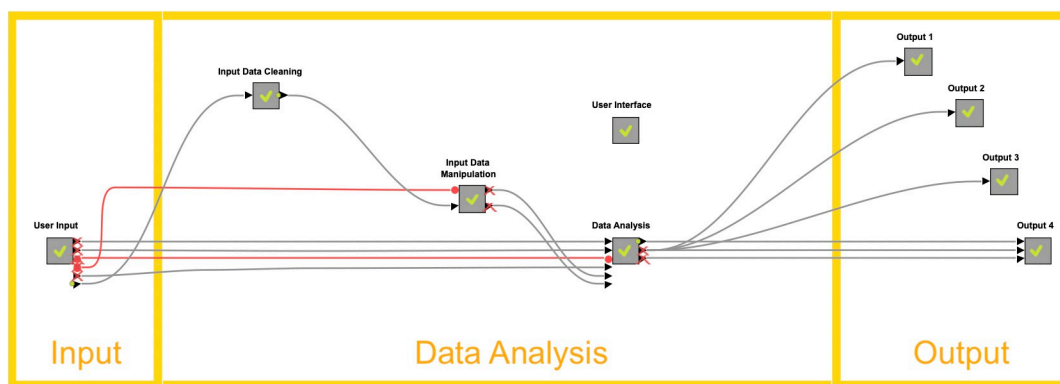


Fig 1. An illustration of the overall setup of CPIC in the KNIME analytics platform.

<https://doi.org/10.1371/journal.pdig.0000451.g001>

Results

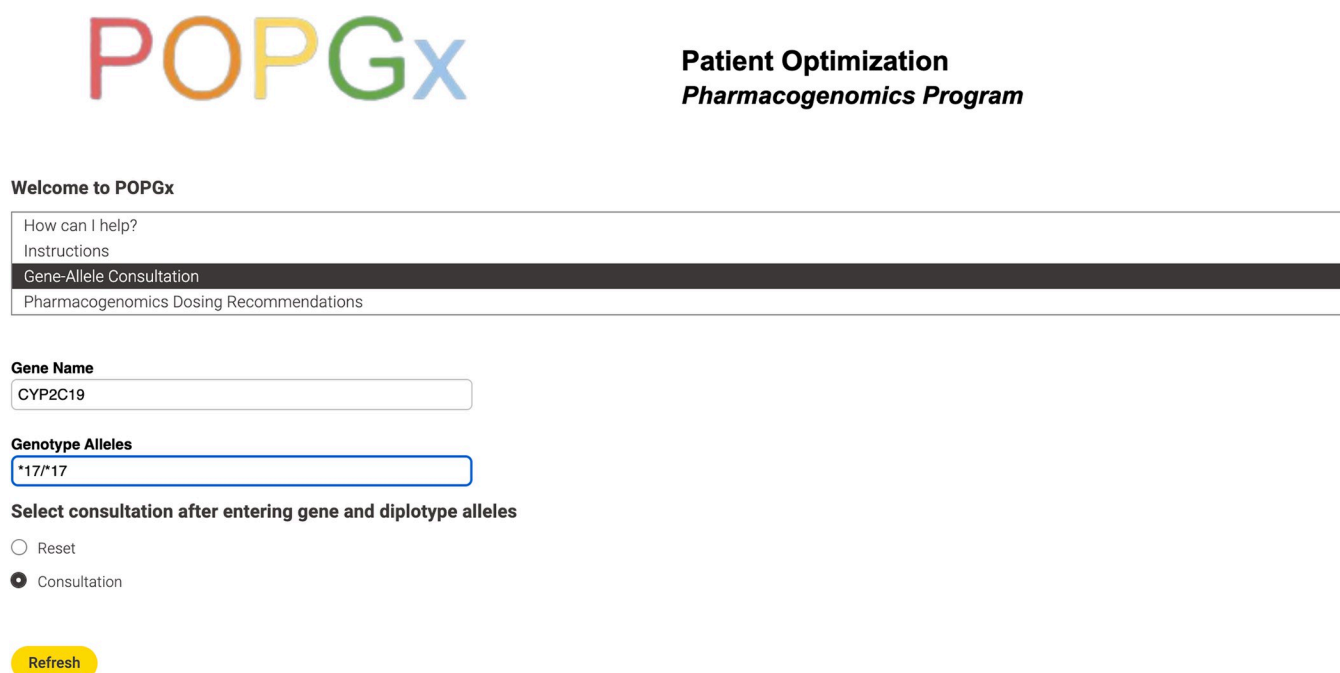
POPGx features a straightforward user interface that adapts to the needs and medical expertise of its users. The program has an “Instructions” option to allow first-time users to quickly learn how to use POPGx. There are two main services POPGx provides: 1. Gene-Allele Consultation, and 2. Pharmacogenomics Dosing Recommendations.

Gene-Allele Consultation

The “Gene-Allele Consultation” service is mainly tailored for patient use. This service calls for the user to input the gene name and the patient’s phenotype alleles to view the gene result and its corresponding consultation See Fig 2.

The gene result reports the patient’s likely metabolizer type based on their phenotype alleles. Gene results for metabolizer expression include: ultrarapid, extensive, intermediate, or poor. The corresponding consultation for that patient’s likely metabolizer gene result defines their metabolizer’s function status, associated risks and outcomes, and points the patient to consult a healthcare provider if needed. This consultation retrieval service increases patient accessibility to understanding the link between metabolizer type and medication therapy outcomes. See Fig 3.

Additionally, the “Gene-Allele Consultation” provides the user with a list of possible impacted medications related to the gene name and phenotype allele inputs. See Fig 4. This list of medications is related to suboptimal patient outcomes resulting from phenotypes diverging from the normal metabolizer. This curated list allows patients to quickly skim through and recognize any medication they may be taking. This at-a-glance list provides another layer for patients and providers to detect suboptimal therapy and enhance patient safety accordingly.



The screenshot displays the POPGx Patient Optimization Pharmacogenomics Program interface. At the top, the logo 'POPGx' is shown in large, colorful letters (P: red, O: orange, P: yellow, G: green, x: blue). To the right, the text 'Patient Optimization Pharmacogenomics Program' is displayed. Below the logo, a 'Welcome to POPGx' section contains a navigation menu with four options: 'How can I help?', 'Instructions', 'Gene-Allele Consultation' (which is highlighted with a dark background), and 'Pharmacogenomics Dosing Recommendations'. The main form area is titled 'Gene Name' and contains a text input field with 'CYP2C19'. Below this is a 'Genotype Alleles' section with a text input field containing '*17/*17'. A label 'Select consultation after entering gene and diplotype alleles' is positioned above two radio button options: 'Reset' and 'Consultation' (which is selected). At the bottom of the form is a yellow 'Refresh' button.

Fig 2. Gene-allele consultation inputs.

<https://doi.org/10.1371/journal.pdig.0000451.g002>

Gene-Allele Expression

CYP2C19**Gene Result:** Ultrarapid Metabolizer

Consultation: This result signifies that the patient has two copies of an increased function allele. Based on the genotype result this patient is predicted to be an ultrarapid metabolizer of CYP2C19 substrates. This patient may be at risk for an adverse or poor response to medications that are metabolized by CYP2C19. To avoid an untoward drug response, dose adjustments or alternative therapeutic agents may be necessary for medications metabolized by CYP2C19. Please consult a clinical pharmacist for more information about how CYP2C19 metabolic status influences drug selection and dosing.

Fig 3. Gene-allele expression consultation based on inputs.

<https://doi.org/10.1371/journal.pdig.0000451.g003>

Pharmacogenomics Dosing Recommendations

The “Pharmacogenomics Dosing Recommendation” service is customized for healthcare provider use. This service calls for users to input gene name, diplotype alleles, and drug name to retrieve corresponding dosing recommendations. See Fig 5.

The “Pharmacogenomics Dosing Recommendation” service defines the patient’s gene status based on the above inputs. See Fig 6. The gene status describes the clinical outcome if the patient was continued on normal dosing without regard to mutant metabolizer expression. Subsequently, it provides a clear dosing recommendation with regards to the patient’s

Possible Impacted Medications

Suboptimal patient outcomes resulting from phenotypes diverging from normal metabolizer

Drug
amitriptyline (Elavil)
citalopram (Celexa)
clopidogrel (Plavix)
escitalopram (Lexapro)
lansoprazole (Prevacid)
omeprazole (PriLOSEC)
pantoprazole (Protonix)
voriconazole (Vfend)
brivaracetam (Briviact)
clomipramine (Anafranil)
dexlansoprazole (Dexilant)
doxepin (Silenor)
imipramine (Tofranil)
sertraline (Zoloft)
trimipramine (Surmontil)
carisoprodol (Soma)
clobazam (Onfi)
diazepam (Valium)
esomeprazole (Nexium)
flibanserin (Addyi)

Fig 4. Possible impacted medications list.

<https://doi.org/10.1371/journal.pdig.0000451.g004>

POPGx

**Patient Optimization
Pharmacogenomics Program**

Welcome to POPGx

How can I help?
Instructions
Gene-Allele Consultation
Pharmacogenomics Dosing Recommendations

Gene Name
CYP2C19

Diplotype Alleles
*17/*17

Drug Name
escitalopram, clopidogrel, omeprazole, voriconazole

Select consultation after entering gene, diplotype alleles, and drug

☐ Reset
☒ Consultation

Fig 5. Pharmacogenomics dosing recommendations inputs.

<https://doi.org/10.1371/journal.pdig.0000451.g005>

metabolizer type and corresponding medication. The service also allows a provider to input multiple medications into a patient's therapy regimen at one time. This unique option—inputting multiple medications instead of looking up one medication at a time—streamlines the process providers must go through to retrieve dosing recommendations, which subsequently may lead to less medication errors and boost patient safety.

The number of genetic data points required to be entered will vary among individuals, as it greatly depends on the specific medication list of each patient. Not all medications have pharmacogenomic implications, meaning that not every drug would necessitate an associated genetic data point. Furthermore, the concept of multiplicity plays a role in this process. If a patient is prescribed several medications that interact with the same implicated gene, it could potentially reduce the number of data points needed and the time spent inputting this information.

Discussion

The current approach for making pharmacogenomic-based therapy selection and dosing modification is tedious. Online gene/drug guideline consortiums, such as CPIC, require healthcare providers to search for a patient's drug individually to retrieve the guideline associated with the drug's metabolizer of concern. From there, a provider would have to match the patient's diplotype to the likely phenotype. Then, the phenotype would be matched to a dose recommendation in a separate table further down the guideline. This entire process only accounts for one drug in a patient's medication regimen, and the provider would have to repeat this process for each drug in a patient's medication regimen to optimize medication efficacy and patient safety (see Fig 7).

Dosing recommendation based on genotype alleles

clopidogrel

Gene status: Increased clopidogrel active metabolite formation; lower on-treatment platelet reactivity; no association with higher bleeding risk

Dosing Recommendations: If considering clopidogrel, use at standard dose (75 mg/day)

escitalopram

Gene status: Increased metabolism when compared to normal metabolizers. Lower plasma concentrations will increase probability of pharmacotherapy failure.

Dosing Recommendations: Consider an alternative drug not predominantly metabolized by CYP2C19

omeprazole

Gene status: Decreased plasma concentrations of PPIs compared to CYP2C19 NMs; increased risk of therapeutic failure

Dosing Recommendations: Increase starting daily dose by 100%. Daily dose may be given in divided doses. Monitor for efficacy.

voriconazole

Gene status: In patients for whom an ultrarapid metabolizer genotype (*17/*17) is identified, the probability of attainment of therapeutic concentrations is small

Dosing Recommendations: Choose an alternative agent that is not dependent on CYP2C19 metabolism as primary therapy in lieu of voriconazole. Such agents include liposomal amphotericin B and posaconazole.

Fig 6. Dosing recommendations based on inputs.

<https://doi.org/10.1371/journal.pdig.0000451.g006>

If a patient is on multiple medications or has multiple genes that are not heterogenous metabolizers, the provider has to reiterate the process mentioned in Fig 7 multiple times. The process is time-consuming and will potentially increase the risk of medical errors. However, the Patient Optimization Pharmacogenomics (POPGx) program does this process fully automated using the most updated clinical data from CPIC. The only requirement would be entering the patient-related genetic data by the user, and the program will provide results in seconds.

Currently, it is time intensive to search guidelines to determine a patient's correct pharmacogenetic dosing if they have a variation in their alleles. The POPGx workflow was created to improve this process. POPGx provides comprehensive dosing recommendations in a time-efficient manner, streamlining the prescribing process for healthcare providers while also allowing patients to view their pharmacogenomic data in regard to their medications.

POPGx has been mainly experimented with and employed by pharmacists and pharmacy students, with its primary application focused on educational aspects. This tool is useful for

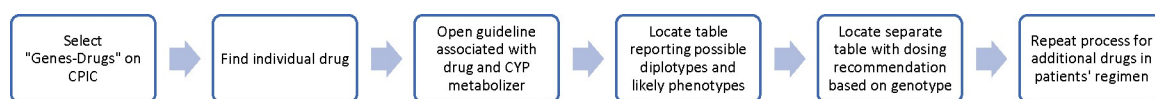


Fig 7. Gene/Drug guideline consortium process.

<https://doi.org/10.1371/journal.pdig.0000451.g007>

simulating patient scenarios for pharmacy or medical students. It helps pharmacy students become more comfortable and familiar with the concept of pharmacogenomics and how it impacts medication prescribing and dosage. Feedback from pharmacists and students indicates that the tool has been invaluable in substantially cutting down the time spent on tasks and mitigating potential errors. Although the tool's application with patients and non-pharmacist providers remains unexplored, early indications from its current users highlight its potential for streamlining operations whilst also ensuring accuracy.

The design and interface of the tool are evidently user-friendly and efficient. For a typical patient, the process is straightforward, demanding only necessary information to present desired results. The streamlined and efficient data input process enhances user engagement for both patients and for health care professionals.

The advantage of using KNIME is due to modularity and the ease of use for those that do not have any data science background. Since the majority of healthcare providers do not receive comprehensive training in data science, using a platform like KNIME bridges the gaps in healthcare with the use of code-free data science. Subsequently, it allows users to do rapid prototyping with minimal data science background. For instance, this program was developed from conception to utilization in less than 15 hours for a healthcare hackathon. Moreover, different workflows built through KNIME can either be used independently or they could get connected to one another to offer a wider range of capabilities. For instance, an area of improvement could be connecting this workflow to dosing calculator programs that were previously built through KNIME to execute multiple functions simultaneously. KNIME also offers KNIME Server, an enterprise-grade solution for Advanced Analytics workloads such as sharing workflows and executing workflows [11]. This allows KNIME workflows to be used by any user without even using the KNIME software.

A current alternative for optimizing pharmacogenomic patient health outcomes is DNA tests with pharmacogenomic reports, such as GeneSight, that patients can utilize through a provider to determine how their body metabolizes their medications [12]. While services like GeneSight are individualized per patient, they are still very limited in use, difficult to order, and are inefficient in returning results; furthermore, analyses are confined to mental health disease states and often only include a patient's pharmacokinetic interactions with the drugs in question, and do not include any dosing recommendations [12]. Application of these products is also not generalizable to many patients and is expensive [12]. POPGx works around these limitations by returning information quickly without the need for a proxy, and by providing all necessary details to the provider and the patient, including a regimen and pharmacokinetics. POPGx is also applicable to hundreds of drugs and disease states, not solely restricted to psychotropics.

While POPGx can retrieve valuable pharmacogenomic information to healthcare providers and patients alike, it has some limitations. Currently, it is completely reliant on the CPIC database, which makes it vulnerable to CPIC's system interruptions or delays in upgrades. One possible solution to this would be to incorporate other pharmacogenomic databases to allow users to verify information from multiple sources. However, it is noteworthy to mention that this tool aims to refine the prevailing method of pharmacogenomic data retrieval. Currently, providers often resort to labor-intensive procedures, such as manually sifting through guideline articles to obtain pertinent data. Even renowned platforms like CPIC merely guide users to specific publications, which they then have to study to extract the necessary information. This not only consumes an exorbitant amount of time but also heightens the risk of errors. In this context, our tool presents a promising avenue to revolutionize the existing system, making it more efficient and reliable.

Additionally, POPGx does not facilitate the direct downloading of genetic and medication data. This limitation stems from the fact that the tool is not yet integrated with any Electronic Health Record (EHR) systems. However, plans for the tool's enhancement envision its linkage to EHR systems, allowing for automatic data population using patient information that will allow healthcare providers to easily access pharmacogenomic recommendations based on patient profiles. Furthermore, POPGx is currently only available on computers that meet all the application requirements, but it could become more accessible in future updates via web or mobile access.

To integrate the tool within the EHR, the plan encompasses the following steps: firstly, the specific integration needs of the tool with respect to the EHR's existing functionalities need to be identified. Prioritizing data security and patient privacy, it is essential to ensure full compliance with relevant regulatory standards, such as HIPAA. Subsequently, the most suitable data integration methodology for this purpose, considering standards like HL7 or FHIR, needs to be chosen. Given the intricacies of data flow, it's essential to establish a robust error-handling mechanism to preemptively address any discrepancies or issues. Before full-scale deployment, rigorous testing of the integrated system should be undertaken to ascertain its reliability and efficiency. Lastly, comprehensive training for end-users needs to be provided to ensure they can effectively leverage the integrated tool within the EHR environment.

Currently, POPGx has been primarily tested and utilized by pharmacists and pharmacy students for educational purposes. Its main objective during this phase has been to enhance the efficiency and effectiveness of pharmacogenomics counseling. While there is no data from a commercial website to cite, the feedback and results from these educational applications have been promising. It is important to note that direct application in real-world clinical settings, with patients directly benefiting from the tool or healthcare providers assisting patients with it, remains a future direction. We are optimistic about its potential and look forward to extending its reach to provide tangible benefits in real-world practice. Additionally, by publishing our findings and the details of our tool, we hope to inspire other innovators in the field to develop similar programs. The goal is not only to further enhance the current landscape of tools available but also to address the time-consuming processes and minimize the potential for user error inherent in existing tools.

Acknowledgments

We would like to thank Dr. Ravi Patel and the University of Pittsburgh School of Pharmacy for their support through the Pitt Challenge hackathon, where POPGx was developed. We would also like to thank Joshua Morris and Monther Alsultan for their support in the revision of this paper.

Author Contributions

Conceptualization: Amir M. Behdani, Jessica Lai, Christina Kim, Lama Basalelah, Trey Halsey, Krista L. Donohoe, Dayanjan Wijesinghe.

Data curation: Jessica Lai, Christina Kim.

Investigation: Amir M. Behdani.

Methodology: Amir M. Behdani, Jessica Lai.

Project administration: Amir M. Behdani, Jessica Lai, Christina Kim, Krista L. Donohoe.

Software: Jessica Lai, Christina Kim, Dayanjan Wijesinghe.

Supervision: Krista L. Donohoe, Dayanjan Wijesinghe.

Visualization: Christina Kim, Trey Halsey, Krista L. Donohoe, Dayanjan Wijesinghe.

Writing – original draft: Amir M. Behdani, Jessica Lai, Christina Kim, Lama Basalelah, Trey Halsey, Krista L. Donohoe, Dayanjan Wijesinghe.

Writing – review & editing: Amir M. Behdani, Jessica Lai, Christina Kim, Lama Basalelah, Trey Halsey, Krista L. Donohoe, Dayanjan Wijesinghe.

References

1. Weinshilboum RM, Wang L. Pharmacogenomics: Precision Medicine and Drug Response. *Mayo Clin Proc.* 2017; 92(11):1711–22. <https://doi.org/10.1016/j.mayocp.2017.09.001> PMID: 29101939
2. Sanchez-Pinto LN, Luo Y, Churpek MM. Big Data and Data Science in Critical Care. *Chest.* 2018; 154(5):1239–48. <https://doi.org/10.1016/j.chest.2018.04.037> PMID: 29752973
3. Nutbeam D. From health education to digital health literacy—building on the past to shape the future. *Glob Health Promot.* 2021; 28(4):51–5. <https://doi.org/10.1177/17579759211044079> PMID: 34719292
4. Mazanetz MP, Marmon RJ, Reisser CBT, Morao I. Drug discovery applications for KNIME: an open source data mining platform. *Curr Top Med Chem.* 2012; 12(18):1965–79. <https://doi.org/10.2174/156802612804910331> PMID: 23110532
5. Subrahmanya SVG, Shetty DK, Patil V, Hameed BMZ, Paul R, Smriti K, et al. The role of data science in healthcare advancements: applications, benefits, and future prospects. *Ir J Med Sci* 1971 -. 2022 Aug; 191(4):1473–83. <https://doi.org/10.1007/s11845-021-02730-z> PMID: 34398394
6. Zanger UM, Schwab M. Cytochrome P450 enzymes in drug metabolism: Regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol Ther.* 2013; 138(1):103–41. Available from: <https://www.sciencedirect.com/science/article/pii/S0163725813000065> <https://doi.org/10.1016/j.pharmthera.2012.12.007> PMID: 23333322
7. Relling MV, Klein TE, Gammal RS, Whirl-Carrillo M, Hoffman JM, Caudle KE. The Clinical Pharmacogenetics Implementation Consortium: 10 years later. *Clin Pharmacol Ther.* 2020; 107(1):171–5. Available from: <http://dx.doi.org/10.1002/cpt.1651>
8. Clinical Pharmacogenetics Implementation Consortium [Internet]. [Cited 2024 Jan 22]. Available from: <https://cpicpgx.org/>
9. PharmGKB [Internet]. [Cited 2024 Jan 22]. Available from: <https://www.pharmgkb.org/whatIsPharmgkb>
10. Berthold MR, Cebron N, Dill F, Gabriel TR, Kötter T, Meinl T, et al. KNIME: The Konstanz Information Miner. In: *Data Analysis, Machine Learning and Applications.* Berlin, Heidelberg: Springer Berlin Heidelberg; p. 319–26.
11. API and Database [Internet]. [Cited 2024 Jan 22]. Available from: <https://cpicpgx.org/api-and-database/>
12. Understanding the Gene-Drug Interaction Chart [Internet]. GeneSight. 2018 [Cited 2024 Jan 22]. Available from: <https://genesight.com/genetic-insights/understanding-the-gene-drug-interaction-chart/>