

Towards a global IT system for personalized medicine: the Medicine Safety Code initiative

Matthias SAMWALD ^{a,1}, José Antonio MINARRO-GIMÉNEZ ^a, Kathrin BLAGEC ^a and Klaus-Peter ADLASSNIG ^{a,b}

^a*Section for Medical Expert and Knowledge-Based Systems, Medical University of Vienna, Vienna, Austria*

^b*Medexter Healthcare GmbH, Vienna, Austria*

Abstract. The availability of pharmacogenomic data of individual patients can significantly improve physicians' prescribing behavior, lead to a reduced incidence of adverse drug events and an improvement of effectiveness of treatment. The Medicine Safety Code (MSC) initiative is an effort to improve the ability of clinicians and patients to share pharmacogenomic data and to use it at the point of care. The MSC is a standardized two-dimensional barcode that captures individual pharmacogenomic data. The system is backed by a web service that allows the decoding and interpretation of anonymous MSCs without requiring the installation of dedicated software. The system is based on a curated, ontology-based knowledge base representing pharmacogenomic definitions and clinical guidelines. The MSC system performed well in preliminary tests. To evaluate the system in realistic health care settings and to translate it into practical applications, the future participation of stakeholders in clinical institutions, researchers, pharmaceutical companies, genetic testing providers, health IT companies and health insurance organizations will be essential.

Keywords. Personalized medicine, pharmacogenetics, biological ontologies, medical informatics, clinical decision support systems

Introduction

The availability of pharmacogenomic data of individual patients can significantly improve physicians' prescribing behavior, lead to a reduced incidence of adverse drug events and an improvement of effectiveness of treatment. Unfortunately, major barriers to the wide-spread implementation of pharmacogenomics in clinical practice remain. Even though genetic testing is constantly becoming cheaper and faster, the financial cost and turnaround time remain a hindrance to wide-spread implementation. Furthermore, the practical application of insights from pharmacogenomic tests is hindered by difficulties in the storage, exchange, and interpretation of pharmacogenomic data. In particular, interoperable electronic health record systems capable of storing structured genetic test results and transmitting those results to other care providers are needed, but such infrastructures are still unavailable in most parts of

¹ Corresponding Author: <matthias.samwald@meduniwien.ac.at>

the world. Finally, most institutions lack easily accessible computer-based clinical decision support systems that can assist clinicians with the interpretation of pharmacogenomic data – a vital component to the use of these data at the point of care.

To make patient's genotype data on essential pharmacogenes globally available to routine medical care, simple and intuitive systems are required to minimize the need for specialized infrastructure, software, and knowledge. The Medicine Safety Code initiative is an effort to improve the ability of clinicians and patients to share pharmacogenomic data and to use it at the point of care.

1. Methods

The Medicine Safety Code (MSC) was implemented as a two-dimensional (2D) barcode that captures data about genetic variants in essential pharmacogenes of an individual patient (Figure 1). The 2D barcode is based on the Quick Response (QR) code standard, which has become widely popular in recent years because it can be decoded quickly and reliably, has relatively high information density and can encode hyperlinks to pages on the World Wide Web. In addition to the QR code representation, MSC data can also be embedded into electronic health records.



Figure 1. A Medicine Safety Code printed on a card carried by the patient is decoded with a smartphone to display drug dosing recommendations.

The data contained in each MSC is a standardized string consisting of a base-URL, a version number and a compressed representation of the pharmacogenomic markers of the patient, e.g.:

<http://safety-code.org/v0.2/QXGqrLF2h8xuzIyCGJE2hzPzVzrND...> (truncated)

This string does not only contain anonymous, compressed genetic data, but also double-acts as a URL that leads to a web page for decoding and interpreting these data. This means that common mobile devices can be used for interpreting the pharmacogenomic data contained in MSCs without requiring the installation of dedicated software. The system presents recommendations for all medications where pharmacogenomic guidelines for the specific patient apply, and medical professionals can quickly select information about a specific medication (e.g., a medication that is considered for prescription).

The system is backed by a curated knowledge base that contains information about essential pharmacogenes, including haplotype definitions, phenotype translations and clinical guidelines. Haplotype definitions were adapted from the Pharmacogenomics Knowledge Base (PharmGKB) [1], while phenotype translations and clinical guidelines were taken from the Clinical Pharmacogenomics Implementation consortium (CPIC) [2], the Dutch Pharmacogenomics Working Group [3] or the American College of Rheumatology [4]. The knowledge base was implemented as an OWL 2 DL ontology called *Genomic CDS*. We used TrOWL [5], a highly scalable OWL 2 DL reasoner, for all reasoning tasks, such as providing the decision support service, analyzing aggregated data and checking the consistency and coherence of the knowledge base.

The MSC system was tested by printing, scanning and interpreting exemplary MSC QR codes.

2. Results

The project website and demos of the decision support service are available at <http://safety-code.org/>. The source code of the software and the Genomic CDS ontology are available from the Google Code project website at [6].

The current version of the MSC (version 0.2) is capable of representing data on 385 genetic markers in 58 important pharmacogenes. The curated knowledge base contains 299 pharmacogenomic decision support rules curated from the Clinical Pharmacogenomics Implementation consortium, the Dutch Pharmacogenomics Working Group or American College of Rheumatology, covering 62 pharmaceutical compounds (Table 1).

Table 1. Overview of pharmaceutical compounds for which guidelines have been curated and added to the Genomic CDS knowledge base (ordered by number of prescriptions per year in Germany).

Pantoprazole, simvastatin, metoprolol, omeprazole, phenprocoumon, allopurinol, glimepiride, citalopram, estrogen-containing oral contraceptives, clomipramine, clopidogrel, mirtazapine, venlafaxine, carvedilol, amitriptyline, tramadol, esomeprazole, glibenclamide, sertraline, doxepin, carbamazepine, duloxetine, paroxetine, escitalopram, risperidone, trimipramine, olanzapine, oxycodone, tamoxifen, lansoprazole, azathioprine, codeine, haloperidol, clozapine, flecainide, flupenthixol, aripiprazole, phenytoin, propafenone, tacrolimus, zuclopenthixol, warfarin, rabeprazole, nortriptyline, imipramine, atomoxetine, capecitabine, abacavir, moclobemide, peginterferon alfa-2a, ribavirin, mercaptopurine, acenocoumarol, desipramine, fluorouracil, gliclazide, irinotecan, peginterferon alfa-2b, tegafur, thioguanine, tolbutamide, voriconazole

To yield an overview of the potential influence that the MSC could have on clinical care, we analyzed prescription statistics of drugs containing substances listed in Table 1. Based on data from general practitioners and medical specialists in private practice in Germany [7], these compounds were associated with 123,12 million drug prescriptions in the year 2011; amounting to costs of 3,76 billion Euros (drug dispensation inside

hospitals is not included in these statistics). New guidelines for other drugs are constantly issued, so the set of pharmacogenes represented by the MSC might cover an even larger fraction of prescribed medications in the future.

We found that the TrOWL reasoner [5] offered excellent reasoning performance with the Genomic CDS ontology, returning results within 3-4 seconds (system specifications: Intel Xeon E3-1230, 3,30 GHz, 16 GB RAM). Even though the response time of TrOWL is vastly superior to other reasoning systems, it is still not optimal for use in busy medical routine. This issue is mitigated by the fact that the MSC system includes a caching functionality, so that results are available without delay for MSCs that have already been processed and cached by the system once. With cached results, the entire process of using an MSC (i.e., picking up a card with an MSC, scanning the QR code, selecting a medication, reading the associated decision support message) can be accomplished within 15 seconds.

3. Discussion

The MSC could act as an enabling technology for the widespread dissemination and clinical implementation of pharmacogenomic data and decision support for several reasons. It is easily implementable, since the MSC is based on technology that is low-cost and widely available and requires minimal integration with existing clinical infrastructure. The 2D barcodes used by the MSC can be printed on personalized cards that patients can carry in their wallets, or they can be incorporated in paper-based lab reports. The 2D barcodes can be quickly decoded using common smartphone devices to yield clinical decision support messages that are pertinent to the individual's genetic profile. If a customized implementation is desired, the MSC can be deployed locally. The system specifications and source code are made openly available.

The MSC system makes it possible to provide access to up-to-date decision support, anywhere. This includes outpatient and emergency situations, regardless of whether the care facility has implemented a genome-enabled electronic medical record and local clinical decision support algorithms. The infrastructure also allows patients to opt-in or opt-out of the system at any time, since they can always choose whether they want to make their MSCs available to healthcare providers.

The Genomic CDS ontology we developed as a backend for the MSC system can be used to represent, organize and reason over the growing wealth of pharmacogenomic knowledge, as well as to identify errors, inconsistencies and lacking definitions in source data sets or individual patient data.

To succeed, the establishment of a network of stakeholders in clinical institutions, researchers, pharmaceutical companies, genetic testing providers, health IT companies and health insurance organizations is essential. Persons interested in participating in the Medicine Safety Code initiative are welcome to visit <http://safety-code.org/> and to subscribe to the mailing list of the initiative.

3.1. Related work

While the MSC system is unique in its focus on providing pharmacogenomic data and decision support in a decentralized setting, some systems have been described that implement pharmacogenomic CDS in local institutions or regional infrastructures.

Swen *et al.* reported on the implementation of pharmacogenomic decision support rules in a computerized drug prescription system in the Netherlands [8]. Pulley *et al.* reported encouraging results about using pharmacogenomic CDS for optimizing anticoagulant therapy at the Vanderbilt University Medical Center [9]. Lærum *et al.* recently reported good feedback from clinicians when testing a prototype of a pharmacogenomic decision support application for immunosuppressant dosing [10].

3.2. Future work

As an important next step towards practical application, we are preparing for user tests with medical professionals to optimize the system for use in realistic clinical settings.

Acknowledgements

We want to thank other participants and supporters of the Medicine Safety Code initiative, including Robert R. Freimuth, Richard D. Boyce, Michel Dumontier and Simon Lin. The research leading to these results has received funding from the Austrian Science Fund (FWF): [PP 25608-N15].

References

- [1] Hewett M, Oliver DE, Rubin DL, Easton KL, Stuart JM, Altman RB, et al. PharmGKB: the Pharmacogenetics Knowledge Base. *Nucleic Acids Res.* 2002 Jan 1;30(1):163–5.
- [2] CPIC: Clinical Pharmacogenetics Implementation Consortium [PharmGKB] [Internet]. [cited 2013 Jul 26]. Available from: <http://www.pharmgkb.org/page/cpic>
- [3] DPWG: Dutch Pharmacogenetics Working Group [Internet]. [cited 2014 Jan 27]. Available from: <https://www.pharmgkb.org/page/dpwg>
- [4] American College of Rheumatology [Internet]. [cited 2014 Jan 27]. Available from: <http://www.rheumatology.org/>
- [5] Thomas E, Pan JZ, Ren Y. TrOWL: Tractable OWL 2 Reasoning Infrastructure. the Proc of the Extended Semantic Web Conference (ESWC2010). 2010.
- [6] genomic-cds - Knowledge Bases and Clinical Decision Support Systems for Realizing Personalized Medicine - Google Project Hosting [Internet]. [cited 2014 Jan 27]. Available from: <https://code.google.com/p/genomic-cds/>
- [7] Schwabe U, Paffrath D. *Arzneiverordnungs-Report 2012 aktuelle Daten, Kosten, Trends und Kommentare.* Berlin: Springer; 2012.
- [8] Swen JJ, Nijenhuis M, Boer A de, Grandia L, Zee AHM der, Mulder H, et al. Pharmacogenetics: From Bench to Byte - An Update of Guidelines. *Clin Pharmacol Ther.* 2011 Mar 16;89(5):662–73.
- [9] Pulley JM, Denny JC, Peterson JF, Bernard GR, Vnencak-Jones CL, Ramirez AH, et al. Operational Implementation of Prospective Genotyping for Personalized Medicine: The Design of the Vanderbilt PREDICT Project. *Clin Pharmacol Ther.* 2012;92(1):87–95.
- [10] Lærum H, Bremer S, Bergan S, Grünfeld T. A taste of individualized medicine: physicians' reactions to automated genetic interpretations. *J Am Med Inform Assoc.* 2013 Sep 3;amiajnl-2012-001587.