eHealth2014 – Health Informatics Meets eHealth
A. Hörbst et al. (Eds.)

doi:10.3233/978-1-61499-397-1-25

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, medical 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

1. 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 [1]. 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 pharmacogenomic test results is hindered by difficulties in the storage, exchange, and interpretation of pharmacogenomic data [2]. In particular, interoperable electronic health record systems that are 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 the world. Finally,

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

most institutions lack easily accessible computer-based clinical decision support (CDS) 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 [3].

To make patient genotype data on essential pharmacogenes globally available to routine medical care, simple and intuitive systems are required in order 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.

2. Methods

The Medicine Safety Code (MSC) was implemented as a standardized 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, the MSC data can also be embedded in electronic health records. 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/QXGqrLF2h8xuqzIyCGJE2hzPzVzrND... (truncated).

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

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 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, [4]), while phenotype translations and clinical guidelines were taken from the Clinical Pharmacogenomics Implementation consortium (CPIC, [5]), the Dutch Pharmacogenomics Working Group [6] or the American College of Rheumatology [7]. The knowledge base was implemented as an OWL 2 DL ontology called *Genomic CDS*. We used TrOWL [8], 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.

3. 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 [9].

3.1. Coverage of Pharmacogenomics Domain

The current version of the MSC (version 0.2) is capable of representing data on 385 genetic markers in 58 important pharmacogenes (Table 1).

Table 1. The key pharmacogenes which can be represented with the current version of the Medicine Safety Code.

ABCB1, ABCC2, ABCG2, ACE, ADRB1, ADRB2, AHR, ALOX5, BRCA1, COMT, CYP1A1, CYP1A2, CYP2A6, CYP2B6, CYP2C19, CYP2C8, CYP2C9, CYP2D6, CYP2J2, CYP3A4, CYP3A5, DPYD, DRD2, F5, G6PD, GSTM1, GSTP1, HLA-B*1502, HLA-B*5701, HMGCR, IL28B, KCNH2, KCNJ11, MTHFR, NAT1, NAT2, NQO1, NR1I2, P2RY1, P2RY12, PTGIS, PTGS2, SCN5A, SLC15A2, SLC19A1, SLC22A1, SLC22A2, SLC22A6, SLCO1B1, SLCO1B3, SLCO2B1, SULT1A1, TPMT, TYMS, UGT1A1, UGT2B15, UGT2B7, VKORC1

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 2).

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 2. Based on data from general practitioners and medical specialists in private practice in Germany in the year 2011, these compounds were associated with 123,12 million drug prescriptions per year; 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 covered by the MSC might cover an even larger fraction of prescribed medications in the future.

Table 2. Overview of pharmaceutical compounds for which guidelines have been curated and added to the Genomic CDS knowledge base.

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

3.2. Performance

We found that the TrOWL reasoner offered excellent reasoning performance with the Genomic CDS ontology, returning results within 3-4 seconds. 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.

4. 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 MSC captures the result of genetic testing independently from the platform that was used to produce the results. Therefore, a wide variety of genotyping or sequencing platforms can be used to generate standardized representations of the most important test results on essential pharmacogenes.

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. It can be applied both in pre-clinical scenarios (e.g., as a reference taxonomy for pharmacogenomic research), as well as for clinical applications (pharmacogenomic decision support, patient stratification in clinical trials). Since it leverages OWL 2 and semantic web technologies, it can be more easily connected to a vast collection of biomedical information resources, and used with a wide variety of tools.Last but not least, a broad adoption of the MSC will also provide benefits for drug development. When data on a minimal set of essential pharmacogenes are made cheaply available for large patient populations, these pharmacogenes could be used for patient stratification during clinical trials ('tailored therapeutics'). This could help to bring new therapies to market faster and with fewer losses caused by concerns about safety and efficacy, which are often encountered in the development of non-tailored therapeutics.

4.1. Technical Considerations

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.

4.2. 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 [10]. Pulley *et al.* reported encouraging results about using pharmacogenomic CDS for optimizing anticoagulant therapy at the Vanderbilt University Medical Center [11]. Lærum *et al.* recently reported good feedback from clinicians when testing a prototype of a pharmacogenomic decision support application for immunosuppressant dosing [12].

There exists some previous work on CDS systems outside the pharmacogenomics domain where decision support logic is partly or fully based on OWL reasoning. One of the earliest examples is a system developed by Bouamrane *et al.*[13], which employs OWL reasoning together with other rule systems for preoperative assessment in order to identify potential risks and complications. A more recent example is the Lung Cancer Assistant system, which employs OWL reasoning for lung cancer treatment selection [14].

4.3. Future Work

As an important next step towards practical application, we are preparing for user tests with medical professionals and pharmacists in order to optimize the system for use in realistic clinical settings. We will also work on creating an Android app for the system, so that MSCs can be interpreted without accessing the MSC web service.

Another important next step towards practical application is the analysis of potential risks introduced by the system, and how these risks can be addressed. A noteworthy potential risk could be overconfidence in the recommendations made in

matching clinical guidelines, i.e., relevant information related to drug and medication safety could be neglected because that information is not a phamarcogenomic one. Furthermore, in order to make well-informed decisions, users of the system need to be made aware of which pharmacogenomic markers are covered, and which are *not* covered.

An interesting direction of future research is the integration of drug-drug interaction knowledge with drug-gene interaction knowledge. This approach could also potentially help to reduce the problem of 'alert fatigue', i.e., the phenomenon that health care professionals are burdened by an excessive number of drug-drug interaction alerts that are not relevant to individual patients. In the near term, however, we expect that pharmacogenomic decision support systems will leave the integration of drug-drug and drug-gene interaction knowledge to health care professionals.

4.4. Concluding Remarks

The MSC is intended to complement local pharmacogenomics initiatives by providing a simple method for making pharmacogenomic data more portable across geographic regions and health care networks. Costs associated with pharmacogenomic testing could be reduced by facilitating the re-use of genetic test results and a reduction in redundant testing.

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.

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] R. S. Epstein, T. P. Moyer, R. E. Aubert, D. J. OKane, F. Xia, R. R. Verbrugge, B. F. Gage, and J. R. Teagarden, "Warfarin Genotyping Reduces Hospitalization Rates: Results From the MM-WES (Medco-Mayo Warfarin Effectiveness Study)," *J. Am. Coll. Cardiol.*, vol. 55, no. 25, pp. 2804–2812, Jun. 2010.

[2] C. L. Overby, P. Tarczy-Hornoch, J. I. Hoath, I. J. Kalet, and D. L. Veenstra, "Feasibility of incorporating genomic knowledge into electronic medical records for pharmacogenomic clinical decision support," *BMC Bioinformatics*, vol. 11, no. Suppl 9, pp. S10–S10.

[3] K. Kawamoto, D. Lobach, H. Willard, and G. Ginsburg, "A national clinical decision support infrastructure to enable the widespread and consistent practice of genomic and personalized medicine," *BMC Med. Inform. Decis. Mak.*, vol. 9, no. 1, p. 17, 2009.

[4] M. Hewett, D. E. Oliver, D. L. Rubin, K. L. Easton, J. M. Stuart, R. B. Altman, and T. E. Klein, "PharmGKB: the Pharmacogenetics Knowledge Base," *Nucleic Acids Res.*, vol. 30, no. 1, pp. 163–165, Jan. 2002.

[5] "CPIC: Clinical Pharmacogenetics Implementation Consortium [PharmGKB]." [Online]. Available: http://www.pharmgkb.org/page/cpic. [Accessed: 26-Jul-2013]
[6] "DPWG: Dutch Pharmacogenetics Working Group." [Online]. Available: https://www.pharmgkb.org/page/dpwg. [Accessed: 27-Jan-2014]
[7] "American College of Rheumatology." [Online]. Available: http://www.rheumatology.org/. [Accessed: 27-Jan-2014]
[8] E. Thomas, J. Z. Pan, and Y. Ren, "TrOWL: Tractable OWL 2 Reasoning Infrastructure," in *the Proc. of the Extended Semantic Web Conference (ESWC2010)*, 2010.
[9] "genomic-cds - Knowledge Bases and Clinical Decision Support Systems for Realizing Personalized Medicine - Google Project Hosting." [Online]. Available: https://code.google.com/p/genomic-cds/. [Accessed: 27-Jan-2014]
[10] J. J. Swen, M. Nijenhuis, A. de Boer, L. Grandia, A. H. M. der Zee, H. Mulder, G. a P. J. M. Rongen, R. H. N. van Schaik, T. Schalekamp, D. J. Touw, J. van der Weide, B. Wilffert, V. H. M. Deneer, and H.-J. Guchelaar, "Pharmacogenetics: From Bench to Byte|[mdash]| An Update of Guidelines," *Clin. Pharmacol. Ther.*, vol. 89, no. 5, pp. 662–673, Mar. 2011.
[11] J. M. Pulley, J. C. Denny, J. F. Peterson, G. R. Bernard, C. L. Vnencak-Jones, A. H. Ramirez, J. T. Delaney, E. Bowton, K. Brothers, K. Johnson, D. C. Crawford, J. Schildcrout, D. R. Masys, H. H. Dilks, R. A. Wilke, E. W. Clayton, E. Shultz, M. Laposata, J. McPherson, J. N. Jirjis, and D. M. Roden, "Operational Implementation of Prospective Genotyping for Personalized Medicine: The Design of the Vanderbilt PREDICT Project," *Clin. Pharmacol. Ther.*, vol. 92, no. 1, pp. 87–95, 2012.
[12] H. Lærum, S. Bremer, S. Bergan, and T. Grünfeld, "A taste of individualized medicine: physicians' reactions to automated genetic interpretations," *J. Am. Med. Inform. Assoc.*, pp. amiajnl–2012–001587, Sep. 2013.
[13] M.-M. Bouamrane, A. Rector, and M. Hurrell, "Using OWL ontologies for adaptive patient information modelling and preoperative clinical decision support," *Knowl. Inf. Syst.*, vol. 29, no. 2, pp. 405–418, Nov. 2011.
[14] M. B. Sesen, E. Jiménez-Ruiz, R. Banares-Alcántara, and J. Brady, "Evaluating OWL 2 Reasoners in the context of Clinical Decision Support in Lung Cancer Treatment Selection," in *OWL Reasoner Evaluation (ORE) Workshop*, Ulm, Germany, 2013.