Patient Genetic Module in Pharmacology

JIŘÍ POTUČEK1, MAREK MATEJÁK[1,2], JIŘÍ DOUŠA[1], OLGA ŠtĚPÁNKOVÁ[3]
[1]Mediware a.s., Evropská 655/116, 160 00 Prague 6,
[2]First Faculty of Medicine, Charles University in Prague,
[3]Czech Institute of Informatics, Robotics and Cybernetics, ČVUT Prague
CZECH REPUBLIC
jiri.potucek@mediware.cz www.mediware.cz

Abstract Personalized medicine that is based on genetic testing opens new possibilities for better pharmacological treatments of a patient, even for physicians or pharmacists who are not familiar with the interpretation of genetic data. Genetic data can be automatic processes that use knowledge from many already accessible databases, collecting results from many scientific studies. This paper presents the idea of a system for the personalized optimization of drug dosing using the integration of pharmacogenomics, pharmacokinetics and pharmacodynamics. For such a system, it is necessary to store the genetic data of the patient as a set of the patient's alleles of selected genes.

Keywords: pharmacogenomics, personalized medicine, genetic test, personalized pharmacokinetics, personalized pharmacodynamics, genetic health records

1. Introduction

Proteins are the base molecules of the biochemical processes in the human body, processes such as enzymatic chemical reactions, membrane transports, receptor signalizations, allosteric regulations, etc. All metabolic or signalization pathways are composed from these processes. However, a selected protein in these processes should not be exactly the same for each subject. The variability and concentration of these proteins is caused by variabilities of the genes where it is encoded. These small changes between the alleles of the gene are translated into small changes in the protein structure and, therefore, also into different electro-chemical properties of the protein. In other words, "the same processes" can have different rates and different equilibria for different patients.

Modern pharmacology should already be built on the biochemistry of the complex physiological processes that are composed from the metabolic or signalization pathways. Critical processes are described using many of the proteins that are involved with only one or a few of them, which are rate-limiting. In pharmacology, the metabolism of the effective drug into the non-effective metabolite is typically the most critical process. Different drugs have different rate-limiting metabolic enzymes, which are typically coded by only one gene. This gene can have different alleles in different subjects, which means a different metabolic speed of the drug for different patients.

2. Problem formulation Structured data

HUGO Gene Nomenclature Committee (HGNC) already specifies the names of all human genes (Povey, et al., 2001). However, each gene or group of genes can have its own committee for its alleles and mutation names, e.g., The Human Cytochrome P450 (CYP) Allele Nomenclature Database (Sim and Ingelman-Sundberg, 2010). Typically, if somebody discovers a new allele of a gene, then he can push its description to the gene's committee. This is very useful for finding out more information about this mutation in the future. The identification of alleles is the first step to personalized medicines. This is because known alleles can have described properties such as Michaelis-Menten coefficient of encoded proteins.

Nowadays, the usage of allele's properties is very simplified. In practice, the most spread nomenclature for laboratory and clinical observations is Logical Observation Identifiers and Codes - LOINC (Forrey, et al., 1996), which is compatible with HL7 standards (Dolin, et al., 2006). Instead of coding chemical rate coefficients, only a few categories of enzyme speeds are used - Ultrarapid, Extensive, Intermediate or Poor Metabolizer. Recognizing that the mutation of an enzyme is faster, slower

or in normal rate is also very useful information.

Genetic testing is becoming cheaper and cheaper. However, it is still not so cheap that it can be repeated before each prescription of new drugs. Archived results of genetic testing can be reused because inherited DNA code remains almost the same throughout a subject's life. As the usage of this information can be improved to the level of complex simulation, such as Physiomodel (Mateják and Kofranek, 2015; Mateják and Kofránek, 2010; Mateják and Kofránek, 2011) or HumMod, it is much better to store the names of the identified alleles rather than only the codes of a few categories such as in the mentioned case of the enzyme's rates. Today, format and data media for storing these genetic patient's health records can almost be completely accessible such as QR codes named as "Safety-code" (Samwald and Adlassnig, 2013), mobile phones, a USB disk, a smart card, an ambulant computer, a network server and so on. Data media should be preferred as everybody can have access to this information at anytime and anywhere such as a person's QR code in his health insurance card. This can quickly give information to each physician, rescuer or pharmacist about the personal genetic prerequisites of a selected treatment.

For example, it is known that a dosing of Warfarin is dependent on the mutations of the genes VKORC1, CYP4F2, GGCX, CYP2C9, CYP3A4, CYP2C18, CYP1A1, CYP1A2 and CYP2C19 (Johnson, et al., 2011; Jonas and McLeod, 2009; Lenzini, et al., 2010). With enough information about their alleles, it is possible to estimate the optimal dosing or whether it is better to use other anticoagulant drugs. The logic of decision is based on the metabolic pathways of the pharmacokinetics and pharmacodynamics of the selected drug. Even these data are already accessible as an output of The Pharmacogenomics Knowledgebase – PharmGKB (Klein, et al., 2001), where much research is collected together for these purposes. For the calculation of optimal drug dosing, specialized tools can be used such as MW-Pharm (Potucek, et al., 2013). For this purpose, there are also non-commercial alternative academic simulation tools such as Chemical library (Mateják, et al., 2015), Physiolibrary (Mateják, 2014; Mateják, 2015; Mateják, et al., 2015; Mateják, et al., 2014), etc.

The base of pharmacogenomics is the selection of proteins, which are rate-limiting in pharmacokinetics and pharmacodynamics pathways. Examples of the base set of genes of these proteins are in Tables 1-4. Most drugs are metabolized to a non-effective metabolite using Cytochrome P450. These enzymes bind the oxygen molecule to the drug molecule using a porphyrin ring with an iron atom in the middle, which is attracted to the oxygen molecule. Binding the oxygen molecule to the iron slightly changes the shape of the ring such as in other hemoproteins (Mateják, 2015; Mateják, et al., 2015). If the drug molecule is inside the enzyme cavity, then this option will insert the oxygen to the specific place in the drug molecule. A created metabolite is non-effective in contrast with the original effective drug.

Table 1, Rate-limiting enzymes

Gene	#alleles	Protein family description
ACE	1	Angiotensin converting enzyme
ADH1B	1	Alcohol dehydrogenase
ADH7	1	
ALOX5	1	Leukotrienes synthesis (AA)
COMT	4	Catecholamine catabolism
CYP1A1	4	Cytochrome P450 (monooxygenase)
CYP1A2	20	
CYP2A6	12	
CYP2B6	36	
CYP2C19	37	
CYP2C8	4	
CYP2C9	20	
CYP2D6	16	
CYP2J2	1	
CYP3A4	13	
CYP3A5	15	
CYP4F2	1	
PTGIS	1	
DPYD	15	Pyrimidine catabolism
G6PD	6	Glucose-6-P metabolism
GSTM1	1	GSH S-transferase
GSTP1	2	
HMGCR	11	Mevalonate/cholesterol pathway
MTHFR	2	Thymidine/methionine synthesis
NAT1	7	N-acetyltransferase
NAT2	8	
NQO1	2	NAD(P)H dehydrogenase
PTGS2	3	Prostaglandin G/H synthase (AA)
SULT1A1	3	Sulfotransferase
TPMT	5	Thiopurine methyltransferase
TYMS	2	Thymidylate synthase
UGT1A1	11	UDP Glucuronosyl transferase
UGT2B15	1	
UGT2B7	1	
VKORC1	10	Vitamin K activator

Table 2, Receptors

Gene	#alleles	Protein family description
ADRB1	2	β receptors
ADRB2	3	
AHR	1	Aryl hydrocarbon receptor
DRD2	4	Dopamin receptor
NR1I2	1	Nuclear receptor
P2RY1	2	G-protein coupled receptor
P2RY12	4	

Table 3, Membrane channel transporters

Gene	#alleles	Protein family description
ABCB1	7	ABC transporter
ABCC2	6	
KCNH2	5	K+ voltage-driven channel
KCNJ11	1	
SCN5A	3	Na+ voltage-driven channel
SLC15A2	4	Oligopeptide transporter
SLC19A1	5	Folate transporter
SLC22A1	9	Organic cation transporter
SLC22A2	5	
SLC22A6	1	Organic anion transporter
SLCO1B1	22	
SLCO1B3	2	
SLCO2B1	1	

Table 4, Signalling proteins

BRCA1	16	Breast cancer 1
F5	1	Coagulation factor V
IL28B	1	Interferon-λ 3
HLA-B*1502	2	Human leukocyte antigen complex (MHC)
HLA-B*5701	1	

3. Problem solution Pharmacogenomics integration

Integration of the pharmacogenomics data described in the previous section leads to a complex system that automatically predicts the patient's status during selected treatments. Thus, a better estimation of drug dosing can be done using a patient's genetic data. This is a result of the integration of drug pharmacokinetic and pharmacodynamics pathways with the knowledge of specific alleles of genes of rate-limiting proteins.

As usual, the physician must conduct the diagnosis and carry out the initial selection of the drugs to treat the patient. If the patient does not have sufficient genetic records for the selected drugs, then the genetic testing can provide the patient's alleles. As in Fig 1, the pharmacogenomics module represents the knowledgebase from which the set of genes that are necessary for genetic testing can be automatically expressed. When some alleles for the drugs are known, the pharmacogenomics module can give the inputs for the personalized optimization of the drug dosing. These inputs can be as simple as only categorizing the enzymes into a fast, slow or normal rate.

The recommendation to drug dosing is provided using the pharmacogenomics module outputs and data of previous patient treatment observations. Other result of personalised drug dosing can be also recommendation not to use the selected drugs or frequency of next patient observations during treatments.

4. Conclusion Discussion

Pharmacogenomics is a new progressive discipline that integrates a huge amount of physiological approaches together (Evans and Relling, 1999). It will soon be included in the medical practice of personalized medicine because each patient is genetically different. These differences must be taken into the account, especially in communities of different populations.

Additionally, simple markers can be very useful because they can predict overdosing, effectivity or insufficient dosing, even for minor communities of patients. They can also predict some side effects of drugs, etc. As a result, the physician will know if there are any necessary additional observations.

In future, it could be that this approach will give a sufficient amount of information to the more complex simulation of a patient's status. This can be used to cure complex symptoms, where such treatments are almost impossible today (Mesko, 2014).

However, the knowledge of pharmacogenomics is still insufficient and most of the approaches are not exact enough to be accepted even by the scientific community. The development of such a system to use known information can generate new questions and answers because there is a huge space for improvements.

Extensions of this pharmacogenomics system can be for statistical purposes, where genetic and observation data from many patients can validate or improve the knowledge of pharmacogenomics.

The system that uses the last versions of many scientific databases closes the loop between

development and clinical practice. Each currently published study that is exact enough to be stored on any of these knowledge databases can be immediately used to provide new recommendations of optimal drug dosing for a patient's genotype.

5. Conflicts of interest

The authors are not in any financial and personal relationships with other people or organizations that could inappropriately influence (bias) their actions.

6. Acknowledgments

The authors appreciate the funding of this work by project "Centrum of pharmacogenomics and informatics" 4.2PT03/1211, PHJ s.r.o. Praha 6.

References

Dolin, R.H., *et al.* HL7 clinical document architecture, release 2. *Journal of the American Medical Informatics Association* 2006;13(1):30-39.

Evans, W.E. and Relling, M.V. Pharmacogenomics: translating functional genomics into rational therapeutics. *science* 1999;286(5439):487-491.

Forrey, A., *et al.* Logical observation identifier names and codes (LOINC) database: a public use set of codes and names for electronic reporting of clinical laboratory test results. *Clinical Chemistry* 1996;42(1):81-90.

Johnson, J.A., *et al.* Clinical Pharmacogenetics Implementation Consortium Guidelines for CYP2C9 and VKORC1 genotypes and warfarin dosing. *Clinical Pharmacology & Therapeutics* 2011;90(4):625-629.

Jonas, D.E. and McLeod, H.L. Genetic and clinical factors relating to warfarin dosing. *Trends in pharmacological sciences* 2009;30(7):375-386.

Klein, T., *et al.* Integrating genotype and phenotype information: an overview of the PharmGKB project. *Pharmacogenomics J* 2001;1(3):167-170.

Lenzini, P., *et al.* Integration of genetic, clinical, and INR data to refine warfarin dosing. *Clinical Pharmacology & Therapeutics* 2010;87(5):572-578.

Mateják, M. Physiology in Modelica. *MEFANET Journal* 2014;2(1):10-14.

Mateják, M. Adairove viazanie O2, CO2 a H+ na hemoglobín In, *Medsoft 2015*. 2015. p. 140-149.

Mateják, M.: Formalization of Integrative Physiology. Charles University in Prague. Dissertation Thesis. 2015.

Mateják, M., *et al.* Physiolibrary 2.3-An Intuitive Tool for Integrative Physiology. *IFAC-PapersOnLine* 2015;48(1):699-700.

Mateják, M. and Kofranek, J. Physiomodel-an integrative physiology in Modelica. In, *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*. IEEE; 2015. p. 1464-1467.

Mateják, M. and Kofránek, J. Rozsáhlý model fyziologických regulací v Modelice. In, *Medsoft 2010*. 2010. p. 126-146.

Mateják, M. and Kofránek, J. HumMod–Golem Edition–Rozsáhlý model fyziologických systémů. In, *Medsoft 2011*. 2011. p. 182-196.

Mateják, M., Kulhánek, T. and Matoušek, S. Adair-based hemoglobin equilibrium with oxygen, carbon dioxide and hydrogen ion activity. *Scandinavian Journal of Clinical & Laboratory Investigation* 2015:1-8.

Mateják, M., *et al.* Physiolibrary - Modelica library for Physiology. In, *10th International Modelica Conference*. Lund, Sweden; 2014.

Mateják, M., *et al.* Free Modelica Library of Chemical and Electrochemical Processes. In, *11th International Modelica Conference, Versailles, France*. Linköping University Electronic Press, Linköpings universitet; 2015. p. 359-366.

Mesko, D.B. The guide to the future of medicine: technology and the human touch. Webicina Kft.; 2014.

Potucek, J., Dousa, J. and Punt, N. The Next Generation Pharmacokinetic Software Suite, Edsim plus plus and Mwpharm plus plus, Provide Clinicians the Means to Develop and Use Pk/Pd Models in a Clinical Setting. In, *THERAPEUTIC DRUG MONITORING*. LIPPINCOTT WILLIAMS & WILKINS 530 WALNUT ST, PHILADELPHIA, PA 19106-3621 USA; 2013. p. 665-665.

Povey, S., *et al.* The HUGO gene nomenclature committee (HGNC). *Human genetics* 2001;109(6):678-680.

Samwald, M. and Adlassnig, K.-P. Pharmacogenomics in the pocket of every patient? A prototype based on quick response codes. *Journal of the American Medical Informatics Association* 2013;20(3):409-412.

Sim, S.C. and Ingelman-Sundberg, M. The Human Cytochrome P450 (CYP) Allele Nomenclature website: a peer-reviewed database of CYP variants and their associated effects. *Human genomics* 2010;4(4):278.