

# An Incremental Adoption Pathway for Developing Precision Medicine Based Healthcare Infrastructure for Underserved Settings

Suranga N. Kasthurirathne, BEng<sup>a</sup>, Paul G. Biondich, MD, MS<sup>b,c</sup>, Burke W. Mamlin, MD<sup>b,c</sup>, Theresa A. Cullen, MD, MS<sup>b,c</sup>, Shaun J. Grannis, MD, MS<sup>b,c</sup>

<sup>a</sup>School of Informatics and Computing, Indiana University, Indianapolis Indiana, USA, <sup>b</sup>Center for Biomedical Informatics, Regenstrief Institute, Indianapolis, Indiana, USA, <sup>c</sup>Indiana University School of Medicine, Indianapolis, Indiana, USA

## Abstract

*Recent focus on Precision medicine (PM) has led to a flurry of research activities across the developed world. But how can understaffed and underfunded health care systems in the US and elsewhere evolve to adapt PM to address pressing healthcare needs? We offer guidance on a wide range of sources of healthcare data / knowledge sources as well as other infrastructure / tools that could inform PM initiatives, and may serve as low hanging fruit easily adapted on the incremental pathway towards a PM based healthcare system. Using these resources and tools, we propose an incremental adoption pathway to inform implementers working in underserved communities around the world on how they should position themselves to gradually embrace the concepts of PM with minimal interruption to existing care delivery.*

## Keywords:

Precision medicine; vulnerable populations; information systems

## Introduction

The need to offer providers a more complete picture of an individuals' health has led to the advent of precision medicine (PM), which is defined as an emerging approach to optimize clinical decision-making by taking individual variability in genes, environment and lifestyle into account [8]. By harnessing measurements from multiple modalities such as clinical and genomic evaluations, environmental exposures, behavioral patterns, and many others, we can develop a far more comprehensive view of the patient's health status and its trajectory over time.

Despite the recent emergence of the PM initiative, its core underlying principals have been known for decades [20], and manifested over the years in various forms such as personalized medicine [12] and the learning healthcare system [6]. Further, PM inspired activities such as the delivery of patient specific allergy alerts [1] and tailored immunization guidance [11] have been routinely used for many decades. A notable difference between these and more recent PM initiatives is the intensified focus on so-called “-omics”, which includes the study of a body of information such as the genome, proteome, metabolome (metabolites), transcriptome (RNA transcripts), autoantibody profiles, etc., [13] made possible by the sequencing of the human genome [21] and rapid advances in high-throughput laboratory technologies and systems approaches across the fields of computer and biological sciences [7].

While -omics based therapeutics have gained widespread publicity and demonstrate significant potential [19], there remain barriers to translating and broadly implementing these

discoveries in the care delivery process. For example, -omics based PM therapeutics are expensive and resource-intensive activities that currently are ill-suited for large-scale clinical care [4; 10] because breakthroughs in many -omics based therapeutics will not be available for use until research and discovery activities progress over the next 5 - 10 years [15].

By definition, the scope of PM is much larger than -omics. PM promises to provide a more complete picture of an individuals' health by capturing actionable information detailing individuals' health status and their socioeconomic context to inform provider decision-making. Thus, the core concepts of PM are relevant to modern medical care and beneficial for both developed and underserved settings. Examples of easily implementable low cost PM-based care that does not rely on -omics based treatment include identifying high-risk patients based on socioeconomic factors and detecting adverse drug events occurring among members of genetically similar groups.

While -omics based PM initiatives are typically cost intense and thus, currently unsuitable for widespread use, the overall aims of PM are relevant and invaluable for healthcare delivery. Thus, all healthcare systems, even those of the underserved world, stand to learn from this approach. Recent focus on PM has led to a flurry of research activities across the developed world. But how can understaffed and underfunded health care systems in the US and elsewhere evolve to adapt PM to address pressing healthcare needs? Which PM based solutions are most suitable for use across underserved settings, if any? Also, which needs should be prioritized as implementers begin the thought process leading toward a PM based healthcare system?

We advocate that PM is heavily dependent on leveraging both novel and existing data and knowledge sources to ensure the accessibility and availability of information surrounding an individual and their environment. Thus, not all PM based activities may be cost and resource intensive. In this paper, we offer guidance regarding the wide range of sources of healthcare data and knowledge, as well as other infrastructure and tools that could inform PM initiatives. Such elements may serve as low hanging fruit easily and incrementally adopted on a pathway towards a PM based healthcare system. Using these resources and tools, we propose a progressive adoption pathway to inform implementers across the underserved world on how they should position themselves to gradually incorporate the principles of PM with minimal interruption to existing healthcare delivery.

## Methods

Based on findings of the Precision Medicine initiative cohort program [14] and seminal publications [3; 8; 18] we sought to

identify data and knowledge sources that could be leveraged to support PM-based care delivery. However, these cannot operationalize PM based care without appropriate supporting systems and processes. Thus, we also identified a list of key infrastructural components that would enable the use of data and knowledge sources for delivering appropriate PM based care.

Next, we sought to describe an incremental adoption pathway for underserved settings that seek to adopt PM initiatives by proposing when each of the aforementioned data, knowledge an infrastructure components should be integrated into exiting healthcare systems. We considered five criteria to evaluate where each component fit into our proposed adoption pathway:

- (1) The availability of robust technical tools / platforms for ready use (Availability): *Are mature toolsets with demonstrated evidence of use available for adoption?*
- (2) The level of provider and user education / training required for adoption (Workforce capacity): *Does the tool/approach require specialist technical skills? How difficult is it to develop local capacity to manage and operate these systems?*
- (3) The practicality of integrating it into the existing health ecosystem and workflows (Integration to healthcare): *Can the component be feasibly integrated into existing/emerging infrastructure? Will it lead to disruptions, and does it require significant behavior changes for use?*
- (4) Cost burden/sustainability: *How sustainable is care delivery using these new components?*
- (5) Clinical impact: *Does robust evidence indicate this element meaningfully improves clinical outcomes?*

We adopted the ThoughtWorks Technology Radar (TTR) [26] to present our incremental PM adoption pathway. The TTR is

a living document prepared by the ThoughtWorks Corporation, a leading software development company, to assess the risks and rewards of existing and nascent technologies, and the strategic importance of each of these for their organization at a specific point in time. TTR presents a concise overview of techniques, tools, platforms/languages and frameworks (four quadrants of a circle), and recommend their strategic importance to technology organizations by allocating each of these to one of four bins - adopt, trial, asses and hold (four concentric rings of a circle). In adopting TTR for our own needs, we modified it as follows:

- (a) Rather than the four quadrants presented in the TTR, we propose three slices (data, knowledge and infrastructure)
- (b) Rather than the four rings presented in the TTR, we propose three: **Adopt** (highly significant components that must be adopted immediately), **Aspire** (components that should be focused on only after the adopt phase is competed) and **Hold** (components that should be held off for a future date where further advances are made).

As our model is intended to represent a static snapshot the current state of PM-based research, and not serve as a 'living document', we will forgo TTR's 'new', 'moved' 'or no change' icons to indicate how the influence of each component has shifted over time.

## Results

### Various components and their adaption

We identified the following data components (Table 1), knowledge components (Table 2) and infrastructure (Table 3) as nessesary for enabling PM based care systems.

Table 1. Data types, examples and sources

Id	Data type	Examples	'Exemplar' or 'Typical' Sources
Source 1: Clinical, demographic and behavioral data			
1	Patient demographics and contact information	Birthdate, gender, race, sexuality, address, educational status, occupation	(1A) EHR systems; (1B) Patient registries (PR)
2	Behavioral and lifestyle measures	Physical activity levels, smoking/tobacco/alcohol use, assessment of other risk factors	(2A) EHR systems; (2B) Smart phones, wearable and home based devices
3	Sensor-based data	Cardiac rate and rhythm monitoring, respiratory rate, physical activity etc.	(3) Smart phones, wearable and home based devices
4	Structured clinical data	Medication, problem and diagnosis lists, vital signs, lab results, family health history measures etc.	(4A) EHR systems; (4B) Non -omics based Laboratory Information Systems (LIS)
5	Unstructured clinical data	Narrative free text data/reports, EKG/EEG waveform data, radiology/medical imaging	(5A) EHR systems; (5B) Picture Archiving and Communication Systems (PACS)
6	-Omics data derived from bio-specimens	Genomics, proteomics, metabolites, cell-free DNA, single cell studies, infectious exposures, standard clinical chemistries, histopathology etc.	(6) External/ancillary Laboratory Information Systems that support -omics testing
Source 2: Geospatial socio-economic, environmental and public health data			
7	Socioeconomic data	Education, unemployment and crime rates, access to transportation, social services, health resources etc.	(7) Various state, private/not-for-profit, and advocacy organizations
8	Public health data	Health insurance coverage, disease rates, life expectancy, obesity rates etc.	(8) Various state and private/not-for-profit monitoring organizations
9	Healthcare resource data	Data on healthcare providers and facilities	(9A) Provider registries; (9B) Health facility registries
Source 3: Other			
10	Data obtained from social media	Behavioral data, over the counter medication purchases, food / drug consumption etc.	(10A) Facebook, twitter, other social media; (10B) Pharmacy billing data

Table 2. Knowledge types, description and sources

Id	Knowledge type	Description	'Exemplar' or 'Typical' Sources
11	Drug-drug interactions	Information on how a drug affects the activity of another when administered together.	(11) Drug knowledge bases, such as Medscape and Drugbank.

12	Drug-gene interactions	Information on how a patient's genetic metabolic affects their ability to clear a drug.	(12) The Drug Gene Interaction Database (DGIdb) [27]; PharmGKB [25]
13	Allergen databases	Information on protein allergens and a patient's susceptibility	(13) Online databases such as allergome.org, allergen.org etc. [17]
14	Genomics knowledge bases	Information on human genes, genetic phenotypes and disorders, and relationships between phenotype and genotype.	(14) Tools such as Online Mendelian Inheritance in Man® (OMIM) [2], HGVSbaseG2P [24]
15	Public Health Genomics	The use of genomic discoveries to improve healthcare and disease prevention	(15) Public Health Genomics Knowledge Base [28]

Table 3. Infrastructure resources

Id	Infrastructure component	Description
16	EHR systems	(16) An electronic version of patients' current and past medical condition that may include all key administrative clinical data relevant to a patients care under a particular provider, including demographics, progress notes, problems, medications, diagnosis etc.
17	Picture Archiving and Communication Systems (PACS)	(17) A medical imaging tool that provides storage and access to images from medical imaging technologies such as ultrasound (US), magnetic resonance (MR), Nuclear Medicine imaging, positron emission tomography (PET), endoscopy (ES), digital radiography (DR), radiography (CR) etc.
18	Laboratory Information Systems (LIS) and tools	(18A) Traditional LIS (18B) LIS that support -omics based testing (18C) -Omics based testing equipment
19	Patient, provider and facility registries	(19A) Patient registries (19B) Provider / health worker registries (19C) Facility registries
20	Disease registries	(20) Registries used to track clinical care and outcomes of a defined patient population suffering from one or more chronic diseases such as diabetes, coronary artery disease, cancer or asthma.
21	Medical terminologies	(21) Various clinical terminologies for reporting diseases (ICD), SNOMED, laboratory data (LOINC) etc.
22	Messaging standards	(22) Various standards introduced to exchange electronic health data across multiple systems. Examples: the Health Level 7 (HL7) family of standards and the DICOM standard.
23	Health Information Exchange (HIE)	(23) A well rounded HIE would consist of many components listed above, such as point of care applications (EHR systems), various registries and a Shared Health Record (SHR) etc.
24	Biobank systems	(24) A biorepository that stores biological samples for use in laboratory based (genomics and traditional) testing.

The information presented in Tables 1, 2 and 3 were fit into a PM technology radar (Figure 1) based on the five evaluation criteria presented above. The PM technology radar is comprised of three slices (data, knowledge and infrastructure) and three concentric circles (adopt, aspire and hold).

#### Description of the PM Technology radar

The Data slice: As step 1 we recommend collecting various unstructured and structured clinical data from EHR, PACS and LIS systems that may already co-exist within a healthcare facility. As step 2, we recommend investigating and integrating other healthcare data such as various registries and other geospatial information sources that may already exist outside the borders of the immediate healthcare facility. This facilitates the use of non-clinical socioeconomic and public health data that could be used to paint a better picture of patients' overall health. Step 3 consists of -omics data and personal data sources such as social networks and smart phones, wearable and home based devices. We also note that some data types such as patient demographics and contact information may be obtained from multiple sources that belong to different stages. In such an event, we recommend that data sources are adopted incrementally over each stage, with patient demographics captured from EHR systems in stage 1, and gradually supplemented with registry data during stage 2.

The knowledge slice: We recommend adopting drug-drug interaction information as step 1. All other knowledge sources are based on -omics, and thus, should be held off for step 3.

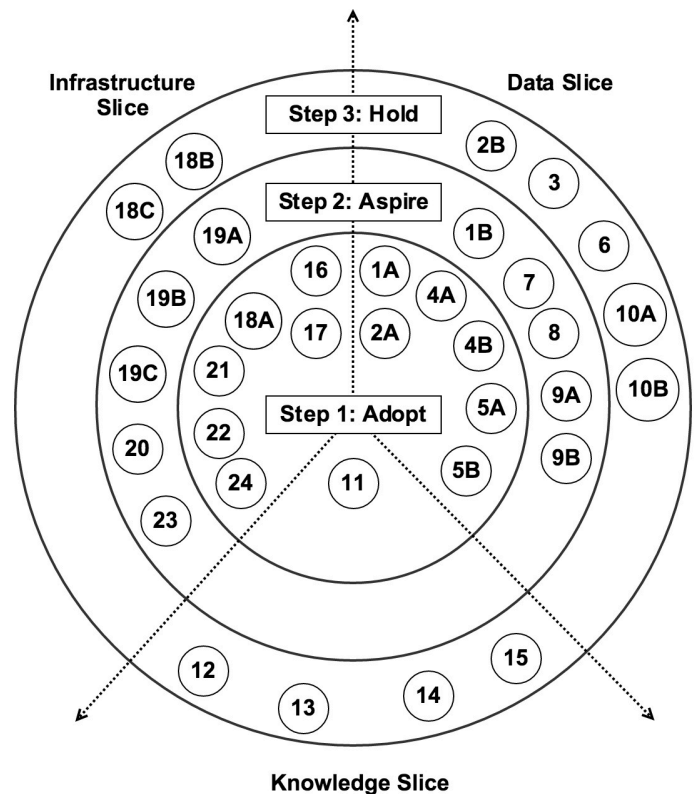


Figure 1. PM technology radar

The infrastructure slice: As step 1, we advocate adopting core health IT components such as EHR, PACS and LIS systems, together with medical terminology/vocabularies and

messaging standards for collecting clinical data. Step 2 consists of incorporating different health infrastructure components that together, form a robust health ecosystem; ideally a fully functional health information exchange. We recommend that -omics based infrastructure components be placed on hold, to be enabled together with genomic based data and knowledge sources after appropriate advanced are made in the field.

## Discussion

Enabling PM based care systems can be seen as two distinct but related challenges: (a) optimizing the use of existing sources of readily available data and knowledge spread across clinical, socioeconomic, and public health spheres, and (b) the integration of -omics based therapeutics and other data obtained from smart devices and additional sources outside of the traditional healthcare delivery domain.

A fundamental requirement for all PM-based systems is the efficient management and use of structured information to inform providers' view of patients' health status. Thus, basic EHR systems that are widely used within underserved settings can serve as an initial stepping-stone on the incremental pathway toward building a more complete PM based healthcare system. To ensure comprehensive adoption of PM-based systems, readily obtainable clinical and behavioral data must also be sustainably captured and appropriate infrastructure components must be built to ensure that this data, together with additional sources of knowledge, are available for actionable use.

In evaluating the PM technology radar, it is evident that our selection of priorities for each step matches a specific pattern; the collection of data, knowledge, and infrastructure defined as step 1 (adopt) implies that the first step in enabling PM based care delivery lies in consolidating standardized data collection within a healthcare facility. As per figure 1, it is also evident that a significant portion of infrastructure resources should be implemented during the early stages of the process, as terminology and messaging standards are crucial to ensure standardized data collection.

The components assigned to step 2 (aspire) characterize expansion of data collection to resources that may lie beyond traditional disconnected healthcare facilities and towards a greater Health Information Exchange (HIE) spreading across a larger demographic area and resources. Adopting an HIE represents significant investment. However, many tools and platforms that enable HIE-based infrastructure have been in existence for years, and are available free of charge [9]. There is also significant value in shifting towards an HIE as these components can contribute to, and inform many other needs beyond PM based care delivery. In comparison, resources identified as step 3 (hold) represent a consolidation of -omics and other resources that have the potential to significantly impact healthcare delivery, but are too immature and/or expensive for adoption at current time.

The PM technology radar also indicates that for underserved settings, early efforts should focus on using PM based initiatives to improve more common care delivery needs such as identifying patients at higher risk, better clinical decision support and medication adherence, as opposed to introducing new treatment processes for specific and less common diseases.

While strategies for adopting resources defined under step 1 are increasingly well understood, step 2 may not be as simple, and would require significant buy in and policy changes. Our approach includes significant emphasis on realizing a

concordant HIE. However, this may be restricted by national or regional scale policies or law. For example, U.S. legislation bans creating a unique national patient identifier that could be used to effectively identify patients across the healthcare ecosystem.

Our reasons to delegate -omics based therapeutics as 'hold' for underserved settings are manifold, and based on practical, financial and policy based limitations rather than disagreement on the benefits of PM. Currently, genomic testing is available for approximately 2,000 clinical conditions in the U.S., and the number of available diagnostic tests is increasing exponentially [18]. However, genetics are only responsible for 20% of an individual's overall health status. Healthy behaviors (50%), environment (20%) and access to care (10%) pose significant impact on health, and are also relatively cheaper and easier to adopt [5]. Further, existing health information systems may not be geared to manage -omics based care. Research suggests that EHRs have poor support for online test ordering and provide limited decision support for genetic testing, interpretation of test results, and potential impact of results on patients and their families [22]. There is also significant need to develop, deploy, and adopt data standards to ensure data privacy, security, and integrity in managing PM based care delivery across these systems [23].

-Omics based treatment also raises significant questions regarding patient confidentiality and health payer rights and policies. In the U.S., these concerns led to the signing of the Genetic Information Nondiscrimination Act (GINA) [16]. However, other countries have been slow to follow. Given these considerations, and that projection that -omics based PM initiatives focused on discovery of disease risk factors, pharmacogenomics, disease biomarker discovery, loss-of-function mutations, new classification of diseases and clinical trials of targeted therapies are not expected to be realized over the next 5 to 10 years [15], we recommend that these components are placed on hold.

The use of smart devices for fitness and monitoring purposes is enjoying significantly greater interest and adoption in comparison to -omics based resources. However, we have assigned it to step 3 (hold) as; (a) smart devices do not contribute as significantly as other components and resources. These data often may also be collected via EHR systems or questionnaires, thus reducing their value; (b) despite tremendous interest, smart devices have not matured adequately to be robustly integrated with healthcare infrastructure, and may lead to security vulnerabilities. However, we acknowledge that smart devices present considerable potential, and may be moved to step 2 shortly.

## Conclusions

The advent of PM represents the continuation of research principals that have manifested in many forms over recent decades. Despite overemphasis on -omics based therapeutics, effective PM based care delivery also involves the efficient use of existing sources of clinical and public health data that are less expensive and easily integrated into existing healthcare infrastructure, and thus serve as 'low hanging fruits' for managing PM based care delivery. We present a PM technology radar that informs implementers on how they could gradually shift towards enabling better PM based care delivery.

## References

- [1] S.A. Abookire, J.M. Teich, H. Sandige, M.D. Paterno, M.T. Martin, G.J. Kuperman, and D.W. Bates, Improving allergy alerting in a computerized physician order entry system, in: *Proceedings of the AMIA Symposium*, American Medical Informatics Association, 2000, p. 2.
- [2] J.S. Amberger, C.A. Bocchini, F. Schiettecatte, A.F. Scott, and A. Hamosh, OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders, *Nucleic acids research* **43** (2015), D789-D798.
- [3] S.J. Aronson and H.L. Rehm, Building the foundation for genomics in precision medicine, *Nature* **526** (2015), 336-342.
- [4] A.S. Behl, K.A. Goddard, T.J. Flottesch, D. Veenstra, R.T. Meenan, J.S. Lin, and M.V. Maciosek, Cost-effectiveness analysis of screening for KRAS and BRAF mutations in metastatic colorectal cancer, *Journal of the National Cancer Institute* **104** (2012), 1785-1795.
- [5] Bipartisan Policy Center, What Makes Us Healthy vs. What We Spend on Being Healthy, in, 2012.
- [6] A. Budrionis and J.G. Bellika, The Learning Healthcare System: Where are we now? A systematic review, *Journal of Biomedical Informatics* **64** (2016), 87-92.
- [7] R. Chen and M. Snyder, Promise of personalized omics to precision medicine, *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* **5** (2013), 73-82.
- [8] F.S. Collins and H. Varmus, A new initiative on precision medicine, *New England Journal of Medicine* **372** (2015), 793-795.
- [9] R. Crichton, D. Moodley, A. Pillay, and C.J. Seebregts, An interoperability architecture for the health information exchange in Rwanda, (2012).
- [10] J.E. Dancey, P.L. Bedard, N. Onetto, and T.J. Hudson, The genetic basis for cancer treatment decisions, *Cell* **148** (2012), 409-420.
- [11] P.R. Dexter, S. Perkins, J.M. Overhage, K. Maharry, R.B. Kohler, and C.J. McDonald, A computerized reminder system to increase the use of preventive care for hospitalized patients, *New England Journal of Medicine* **345** (2001), 965-970.
- [12] M.A. Hamburg and F.S. Collins, The path to personalized medicine, *New England Journal of Medicine* **363** (2010), 301-304.
- [13] R.P. Horgan and L.C. Kenny, 'Omic' technologies: genomics, transcriptomics, proteomics and metabolomics, *The Obstetrician & Gynaecologist* **13** (2011), 189-195.
- [14] K. Hudson, R. Lifton, and B. Patrick-Lake, The precision medicine initiative cohort program—building a research foundation for 21st century medicine, *Precision Medicine Initiative (PMI) Working Group Report to the Advisory Committee to the Director*, ed (2015).
- [15] K. Hudson, R. Lifton, and B. Patrick-Lake, The precision medicine initiative cohort program—Building a Research Foundation for 21st Century Medicine, *Precision Medicine Initiative (PMI) Working Group Report to the Advisory Committee to the Director*, ed (2015).
- [16] K.L. Hudson, M. Holohan, and F.S. Collins, Keeping pace with the times—the Genetic Information Nondiscrimination Act of 2008, *New England Journal of Medicine* **358** (2008), 2661-2663.
- [17] A. Mari, C. Rasi, P. Palazzo, and E. Scala, Allergen databases: current status and perspectives, *Current allergy and asthma reports* **9** (2009), 376-383.
- [18] R. Mirnezami, J. Nicholson, and A. Darzi, Preparing for precision medicine, *New England Journal of Medicine* **366** (2012), 489-491.
- [19] S. Morgan, J. Hurley, F. Miller, and M. Giacomini, Predictive genetic tests and health system costs, *Canadian Medical Association Journal* **168** (2003), 989-991.
- [20] National Research Council Committee on A Framework for Developing a New Taxonomy of Disease, *Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease*, National Academies Press (US), 2011.
- [21] L. Nilsson, M. Hofflander, S. Eriksen, and C. Borg, The importance of interaction in the implementation of information technology in health care: a symbolic interactionism study on the meaning of accessibility, *Informatics for health & social care* **37** (2012), 277-290.
- [22] M.T. Scheuner, H. de Vries, B. Kim, R.C. Meili, S.H. Olmstead, and S. Teleki, Are electronic health records ready for genomic medicine?, *Genetics in Medicine* **11** (2009), 510-517.
- [23] J.D. Tenenbaum, P. Avillach, M. Benham-Hutchins, M.K. Breitenstein, E.L. Crowgey, M.A. Hoffman, X. Jiang, S. Madhavan, J.E. Mattison, and R. Nagarajan, An informatics research agenda to support precision medicine: seven key areas, *Journal of the American Medical Informatics Association* (2016), ocv213.
- [24] G.A. Thorisson, O. Lancaster, R.C. Free, R.K. Hastings, P. Sarmah, D. Dash, S.K. Brahmachari, and A.J. Brookes, HGVBbaseG2P: a central genetic association database, *Nucleic acids research* **37** (2009), D797-D802.
- [25] C.F. Thorn, T.E. Klein, and R.B. Altman, Pharmacogenomics and bioinformatics: PharmGKB, *Pharmacogenomics* **11** (2010), 501-505.
- [26] ThoughtWorks Inc., TECHNOLOGY RADAR, in, 2016.
- [27] A.H. Wagner, A.C. Coffman, B.J. Ainscough, N.C. Spies, Z.L. Skidmore, K.M. Campbell, K. Krysiak, D. Pan, J.F. McMichael, and J.M. Eldred, DGIdb 2.0: mining clinically relevant drug-gene interactions, *Nucleic acids research* (2015), gkv1165.
- [28] W. Yu, M. Gwinn, W.D. Dotson, R.F. Green, M. Clyne, A. Wulf, S. Bowen, K. Kolor, and M.J. Khoury, A knowledge base for tracking the impact of genomics on population health, *Genetics in Medicine* (2016).

### Address for correspondence

Suranga N. Kasthurirathne  
School of Informatics and Computing (SOIC) Indiana  
University, Indianapolis, IN, USA  
snkasthu@iupui.edu