International Journal of Epidemiology, 2020, 23–24k
doi: 10.1093/ije/dyz123
Advance Access Publication Date: 4 August 2019
Cohort Profile

Cohort Profile

Cohort Profile: The Right Drug, Right Dose, Right Time: Using Genomic Data to Individualize Treatment Protocol (RIGHT Protocol)

Suzette J Bielinski,[1]* Jennifer L St Sauver,[1,2] Janet E Olson,[1,3] Nicholas B Larson,[4] John L Black III,[3,5] Steven E Scherer,[6] Matthew E Bernard,[7] Eric Boerwinkle,[8] Bijan J Borah,[2,9] Pedro J Caraballo,[10] Timothy B Curry,[3,11] HarshaVardhan Doddapaneni,[12] Christine M Formea,[13] Robert R Freimuth,[14] Richard A Gibbs,[6] Jyothsna Giri,[3] Matthew A Hathcock,[4] Jianhong Hu,[12] Debra J Jacobson,[4] Leila A Jones,[3] Sara Kalla,[12] Tyler H Koep,[15] Viktoriya Korchina,[12] Christie L Kovar,[12] Sandra Lee,[12] Hongfang Liu,[14] Eric T Matey,[3,13] Michaela E McGree,[4] Tammy M McAllister,[3] Ann M Moyer,[5] Donna M Muzny,[12] Wayne T Nicholson,[11] Lance J Oyen,[13] Xiang Qin,[12] Ritika Raj,[12] Véronique L Roger,[1,16] Carolyn R Rohrer Vitek,[3] Jason L Ross,[15] Richard R Sharp,[3,9] Paul Y Takahashi,[17] Eric Venner,[12] Kimberly Walker,[12] Liwei Wang,[14] Qiaoyan Wang,[12] Jessica A Wright,[3,13] Tsung-Jung Wu,[12] Liewei Wang,[3,18] and Richard M Weinshilboum[3,18†]

[1]Division of Epidemiology, Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA, [2]Robert D and Patricia E Kern Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, MN, USA, [3]Center for Individualized Medicine, Mayo Clinic, Rochester, MN, USA, [4]Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA, [5]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA, [6]Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA, [7]Department of Family Medicine, Mayo Clinic, Rochester, MN, USA, [8]Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA, [9]Division of Health Care Policy and Research, Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA, [10]Division of General Internal Medicine, Department of Medicine, Mayo Clinic, Rochester, MN, USA, [11]Department of Anesthesia and Perioperative Medicine, Mayo Clinic, Rochester, MN, USA, [12]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA, [13]Department of Pharmacy, Mayo Clinic, Rochester, MN, USA, [14]Division of Digital Health Sciences, Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA, [15]OneOme, LLC, Minneapolis, MN, USA, [16]Division of Cardiovascular Diseases, Department of Internal Medicine, Mayo Clinic, Rochester, MN, USA, [17]Division of Community Internal Medicine, Department of Medicine, Mayo Clinic, Rochester, MN, USA and [18]Division of Clinical Pharmacology, Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Rochester, MN, USA

*Corresponding author. Division of Epidemiology, Department of Health Sciences Research, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA. E-mail: bielinski.suzette@mayo.edu
†The authors worked collaboratively as the Mayo-Baylor RIGHT 10K Pharmacogenomics Consortium.

Editorial decision 20 May 2019; Accepted 31 May 2019

Why was the cohort set up?

Genetic modulation of drug response can cause life-threatening adverse drug reactions, increase susceptibility to drug–drug interactions and alter therapeutic effectiveness.[1–3] Pharmacogenomics has the potential to immediately impact the care of patients and is a hallmark of genomic medicine worldwide. Ideally, pharmacogenomics-guided drug selection will produce optimal effects for specific indications, reduce delays in care to effective therapies, improve patient safety and reduce health care costs. Over the past decade, a large number of pharmacogenomic variants with demonstrated clinical utility have been identified and corresponding international guidelines for clinical implementation have been developed and published by the Clinical Pharmacogenetics Implementation Consortium (CPIC),[4] Canadian Pharmacogenomics Network for Drug Safety (CPNDS)[5] and Royal Dutch Association for the Advancement of Pharmacy Dutch Pharmacogenetics Working Group (DPWG).[6] There is now general agreement that pharmacogenomics represents an area within genomic science that could have a significant positive impact on clinical medicine and, ultimately, affect every patient.

Pre-emptive sequencing of patients interrogates large numbers of pharmacogenomic variants and integrates clinically actionable results into a patient's electronic health record (EHR) for efficient use by clinicians at the point of care. It is unknown, however, whether integration of pre-emptive pharmacogenomic data into the EHR will significantly improve patient outcomes and reduce health care costs. Therefore, the Right Drug, Right Dose, Right Time: Using Genomic Data to Individualize Treatment Protocol (RIGHT Protocol) study was designed to recruit a large group of patients for pre-emptive pharmacogenomic testing, to develop the EHR infrastructure to deliver point of care clinical decision support and to study the effects of integrating pre-emptive pharmacogenomic testing into applied clinical practice on patient outcomes. As a result of these efforts, we have recruited 11 098 patients to study pre-emptive pharmacogenomic implementation in an integrated medical system. The RIGHT cohort therefore serves as a scientific resource to broadly address the following scientific questions.

- What are the associations between genetic variants and clinical outcomes in patients prescribed drugs known to interact with these genes? The outcomes we will investigate include therapeutic response (e.g. pain control), time to optimal therapeutic goal, number of drug and/or dose changes, health care utilization (e.g. clinic visits and hospital days), adverse drug events/toxicity (e.g. gastrointestinal upset), and subsequent clinical events (e.g. stroke).
- Does pre-emptive pharmacogenomic testing improve patient outcomes and reduce health care costs through optimized treatment? By comparing patients with pharmacogenomic data to pharmacogenomics-naïve patients we will test the hypothesis that pre-emptive pharmacogenomic testing reduces adverse drug events, lowers health care utilizations and costs, and improves therapeutic response, patient safety and satisfaction.
- What clinical outcomes are associated with pharmacogenomic variants of unknown or indeterminate clinical significance or for which evidence for changes in prescribing practices is limited? A critical barrier in identifying pharmacogenomic variants that are clinically significant is the reliance on costly clinical trials that investigate a single drug or outcome. Observational studies such as the RIGHT Protocol study will provide valuable insights regarding the cost-effectiveness of identifying drug–gene interactions that result in adverse drug events and other clinically significant outcomes.
- How does pre-emptive pharmacogenomic testing impact clinician and patient perspectives and the broad spectrum of clinical practice? Understanding the potential influence of pre-emptive pharmacogenomics in prescribing practices, as well as identifying the possible impacts to clinician and patient interactions, can elucidate risks and opportunities that support safe integration. Identifying gaps and providing education is vital to support wide-scale implementation when diffusing medical innovations.

Who is in the cohort?

Setting

Mayo Clinic is a non-profit academic medical institution with major campuses in Rochester, MN; Scottsdale and Phoenix, AZ; and Jacksonville, FL. The Mayo Clinic Health Systems has dozens of locations in MN, IA and WI. Together, these sites provide comprehensive patient care, education in clinical medicine and medical sciences, and extensive programmes in research. In 2018, Mayo Clinic

provided direct care for 1.3 million people from all 50 states and 136 countries.

The Mayo Clinic Biobank is an institutional resource comprised of adult community volunteers who donated biological specimens, provided risk factor data, allowed access to EHR data for research and consented to participate in additional studies.[7] To date, over 57 000 subjects have consented to participate. Since its inception, this resource has been used for more than 250 ancillary studies and has distributed over 125 000 samples. The Mayo Clinic Biobank was the source cohort for the RIGHT Protocol study, which allowed the timely recruitment of 11 098 participants with previously stored DNA samples.

Mayo Clinic EHR data are available for all Biobank participants. However, patients use different health care providers for different reasons and may change providers and health care systems over time. Therefore, the Mayo Clinic EHR may not contain complete data for all Biobank participants. However, most of the Biobank participants (82%) reside within the region captured by the Rochester Epidemiology Project (REP) medical records-linkage system. The REP is an established research infrastructure that links the health care records of persons residing in a 27-county region of southern MN and western WI (approximately 1.1 million persons in 2010).[8–10] Briefly, the REP links the health care records of Mayo Clinic, Olmsted Medical Center and its satellites, Olmsted County Public Health Services, Zumbro Valley Health Center and some smaller health care providers in this region. As residents of this region visit one of the REP collaborators for health care, the REP captures the electronic data related to the health care visits. In addition, the full texts of health care records are available for abstraction for all persons in the REP system since 1966. Overall, 99% of the Biobank participants have at least one non-Mayo Clinic medical record captured by the REP. Therefore, linking the Biobank participants to the REP infrastructure provides access to more complete coverage and follow-up of the Biobank population.

Recruitment

Beginning in 2012, Biobank participants were recruited for the RIGHT pilot study, which has been previously described.[11] In brief, 1013 patients consented to pharmacogenomic testing and return of pharmacogenomic results to their EHR for use in their clinical care. This pilot resulted in the infrastructure to support the expansion of the cohort. Beginning in May 2016, we began a targeted effort to further enroll an additional 10 000 persons. Patients with prior allogeneic haematopoietic stem cell transplants and/or haematologic malignancies were not eligible to participate due to the fact that circulating lymphocytes were used as the DNA source. All non-White Biobank participants ($n = 1442$) were preferentially invited to increase the diversity of the cohort. Participants were then prioritized by those that receive the majority of their care at Mayo Clinic based on EHR length and a history of clinic visits in a primary care clinic (e.g. Internal Medicine, Family Medicine). All eligible Biobank participants were mailed an invitation to participate in the RIGHT Protocol study and a consent form. The invitation letter gave recipients three options: (i) participate in the study, (ii) decline to participate, or (iii) declare themselves ineligible because they did not intend to utilize Mayo Clinic for their future health care. Consenting RIGHT participants agreed to allow use of their stored Biobank samples for clinical pharmacogenomic testing, deposit of pharmacogenomic results into their EHR for clinical use, and use of EHR and pharmacogenomic data for research. If no response was received after a period of 4 weeks, one additional attempt was made with a second mailing. This expanded recruitment effort was completed in 16 months. Overall, 18 199 persons were approached, 56% consented to participate, 29% did not respond and 15% self-reported ineligibility or refused (Table 1).

Cohort characteristics

Table 1 summarizes the characteristics of the RIGHT pilot participants, recruitment results for the expansion of the cohort and the characteristics of the full RIGHT cohort. Table 2 compares characteristics of the RIGHT cohort, Mayo Clinic Biobank and REP 27-county populations.

How often have they been followed up?

Prospective and retrospective studies

A primary goal of the RIGHT Protocol study was to implement pre-emptive pharmacogenomics for Mayo Clinic patients. Therefore, by design, we anticipate the majority of our follow-up will be conducted passively, via the EHR, in order to estimate patient outcomes after implementation of pre-emptive pharmacogenomics. Prospective studies, defined as occurring after pharmacogenomic data were placed in the EHR, permit comparisons of the RIGHT cohort to a pharmacogenomics-naïve group to determine the effects of genotype-guided prescribing. As of January 1, 2018, 99% of the cohort was living and 85% had a clinic visit in 2017 (Table 3). The RIGHT Protocol study also enables retrospective studies by leveraging decades of drug exposures to compare drug response across genotype prior to pharmacogenomic results being placed in the EHR (pharmacogenomics-naïve time period). The median length of the medical record is 22 years for the cohort (Table 3). For either study type, outcomes can include

Table 1. Characteristics of 1013 RIGHT pilot participants and the 18 199 patients invited to participate in the expansion of the RIGHT protocol

Characteristic		RIGHT expansion results				
	RIGHT pilot participants[a] $n = 1013$	Invited $n = 18\,199$	Consented $n = 10\,085$	Did not respond $n = 5331$	Ineligible/ refused $n = 2783$	Complete RIGHT cohort $n = 11\,098$
Sex, *n* (%)						
Female	538 (53)	11 043 (61)	6150 (61)	3192 (60)	1701 (61)	6688 (60)
Male	475 (47)	7156 (39)	3935 (39)	2139 (40)	1082 (39)	4410 (40)
Age on January 1, 2016, years, *n* (%)						
18–24	0 (0)	151 (1)	58 (1)	76 (1)	17 (1)	58 (1)
25–34	0 (0)	1701 (9)	647 (6)	784 (15)	270 (10)	647 (6)
35–44	2 (<1)	1722 (9)	825 (8)	744 (14)	153 (6)	827 (7)
45–54	178 (18)	2390 (13)	1299 (13)	831 (16)	260 (9)	1477 (13)
55–64	658 (65)	3769 (21)	2069 (21)	1138 (21)	562 (20)	2727 (25)
65–74	175 (17)	5084 (28)	3217 (32)	1056 (20)	811 (29)	3392 (31)
75+	0 (0)	3382 (19)	1970 (20)	702 (13)	710 (26)	1970 (18)
Race, *n* (%)						
White	966 (95)	16 759 (92)	9482 (94)	4689 (88)	2588 (93)	10 448 (94)
Non-White	39 (4)	1177 (6)	524 (5)	492 (9)	161 (6)	563 (5)
Black	7 (1)	124 (1)	50 (<1)	64 (1)	10 (<1)	57 (1)
Asian	9 (1)	276 (2)	91 (1)	139 (3)	46 (2)	100 (1)
AIAN	0 (0)	30 (<1)	16 (<1)	8 (<1)	6 (<1)	16 (<1)
NHPI	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Other and mixed	23 (2)	747 (4)	367 (4)	281 (5)	99 (4)	390 (3)
Unknown	8 (1)	263 (1)	79 (1)	150 (3)	34 (1)	87 (1)
Ethnicity, *n* (%)						
Non-Hispanic	1006 (99)	17 906 (98)	9967 (99)	5205 (98)	2734 (98)	10 973 (99)
Hispanic	7 (1)	275 (2)	112 (1)	115 (2)	48 (2)	119 (1)
Unknown	0 (0)	18 (<1)	6 (<1)	11 (<1)	1 (<1)	6 (<1)
Self-reported education at time of Biobank consent (2009–2017), *n* (%)						
High school graduate or General Educational Development (GED) or less	87 (9)	2754 (15)	1262 (13)	897 (17)	595 (21)	1349 (12)
Some college or Associates degree (including community college)	326 (32)	5614 (31)	2938 (29)	1817 (34)	859 (31)	3264 (29)
Four year college graduate (Bachelor's degree)	238 (24)	3580 (20)	1992 (20)	1079 (20)	509 (18)	2230 (20)
Graduate or professional school	362 (36)	6116 (34)	3848 (38)	1476 (28)	792 (28)	4210 (38)
Unknown	0 (0)	135 (1)	45 (<1)	62 (1)	28 (1)	45 (<1)

[a]Previously recruited in 2012–2013.
AIAN, American Indian or Alaska Native; NHPI, Native Hawaiian or Pacific Islander.

health care utilization, drug type and dose changes, time to optimal therapy, and adverse drug events including subsequent clinical events (e.g. stent thrombosis). Finally, the availability of genetic sequence data allows for discovery of additional drug–gene interactions for which clinical evidence is limited or non-existent.

Active follow-up

Participants in the cohort may, with appropriate approvals, be contacted directly to assess pharmacogenomics-related outcomes. To date, we have conducted three surveys of RIGHT participants. The RIGHT pilot study participants were surveyed in 2014 to assess their understanding of their *CYP2D6* results.[12] Of those invited, 86% completed the survey.[13] In 2017, 5000 subjects from the expanded cohort were surveyed after enrollment to collect information on prior testing, expectations and concerns about pharmacogenomic testing, as well as their views about their current medications and medications in general. Response to this survey was exceptionally high (94%) and reflects a very high level of interest in

Table 2. Characteristics as of January 1, 2016 of the RIGHT cohort, Mayo Clinic Biobank and Rochester Epidemiology Project

Characteristic	RIGHT cohort $n = 11\,098$	Mayo Clinic Biobank $n = 56\,988$	Rochester Epidemiology Project 27-counties $n = 582\,466$
Sex, *n* (%)			
Female	6688 (60)	33 478 (59)	309 707 (53)
Male	4410 (40)	23 510 (41)	272 759 (47)
Age on January 1, 2016, years, *n* (%)			
18–24	58 (1)	536 (1)	69 237 (12)
25–34	647 (6)	3227 (6)	98 800 (17)
35–44	827 (7)	4026 (7)	85 385 (15)
45–54	1477 (13)	7238 (13)	94 059 (16)
55–64	2727 (25)	13 020 (23)	101 644 (17)
65–74	3392 (31)	15 115 (27)	71 216 (12)
75+	1970 (18)	13 826 (24)	62 125 (11)
Race, *n* (%)			
White	10 448 (94)	54 316 (95)	520 716 (89)
Non-White	563 (5)	1967 (3)	41 434 (7)
Black	57 (1)	702 (1)	14 248 (2)
Asian	100 (1)	572 (1)	11 423 (2)
AIAN	16 (<1)	117 (<1)	1491 (<1)
NHPI	0 (0)	22 (<1)	842 (<1)
Other and mixed	390 (3)	554 (1)	13 430 (2)
Unknown race	87 (1)	705 (1)	20 316 (3)
Ethnicity, *n* (%)			
Non-Hispanic	10 973 (99)	5370 (94)	505 374 (87)
Hispanic	119 (1)	783 (1)	22 428 (4)
Unknown	6 (<1)	2504 (4)	54 664 (9)
Socio-economic characteristics			
≥ High school (%)	99	97	41*
≥ Bachelor's degree (%)	58	50	15*

*Education data are available for 329 589 (57%) persons; percentages are based on this denominator.
AIAN, American Indian or Alaska Native; NHPI, Native Hawaiian, or Pacific Islander.

pharmacogenomics in our population. A follow-up survey is currently in the field to assess responses following receipt of pharmacogenomic results. Additionally, there are opportunities to study clinician perspectives, pharmacogenomic testing behaviours and clinical actions following deposition of pharmacogenomic results into the EHR.

Loss to follow-up

The recruitment design of the study prioritized patients for whom Mayo Clinic is their primary care clinic and therefore loss to follow-up is mitigated. Likewise, use of the REP allows for data capture within the surrounding community practices. However, data capture decreases as the distance of the patient's residence from Rochester, MN increases. Loss to follow-up will occur for those patients moving or choosing a provider out of our catchment area. However, as indicated in Table 3, the RIGHT cohort reflects a stable patient population with 91% receiving at least one prescription in 2017.

What has been measured?

DNA sequencing and pharmacogenomic phenotyping

For the 1013 pilot-study participants, lymphocyte-derived DNA sequencing of 84 pharmacogenes was completed at Mayo Clinic using Version 1 (v.1) of the PGRN-Seq assay.[14] For the expanded RIGHT cohort, the Clinical Laboratory Improvement Amendments (CLIA)-certified and College of American Pathologists (CAP)-accredited Baylor College of Medicine's Human Genome Sequencing Center Clinical Laboratory sequenced 77 genes using version 3 (v.3.) of the PGRN-Seq assay (now termed PGx-seq). This panel includes structurally complex genomic regions important to optimal drug treatment including *CYP2D6* and *HLA* regions. *CYP2D6* is involved in the metabolism of 25% of the drugs currently on the market[15] and *HLA* regions are associated with severe drug hypersensitivities (e.g. Stevens-Johnson Syndrome). In v.1, all gene coverage included all gene model exons plus both 2 kb upstream and 1 kb downstream. In v.1, targeted exons and

Table 3. Medical record data for the RIGHT cohort

Characteristics	RIGHT cohort $n = 11\,098$
Alive as of January 1, 2018, *n* (%)	11 021 (99)
Clinical contact in 2017, % yes	85
Length of medical record prior to 2017, median years interquartile range (IQR)	22 (19, 27)
Clinic visits per year from 2004 to 2017, median visits (IQR)	27 (16, 44)
Prescribed medications in 2017, *n* (%)	
None	1010 (9)
1–2	1725 (16)
3–5	2732 (25)
6–10	2992 (27)
11–20	2124 (19)
≥ 21	515 (5)
Prescribed a drug with therapeutic recommendations[a] 2004–2017, *n* (%)	
abacavir	2 (<1)
allopurinol	494 (4)
amitriptyline	664 (6)
atazanavir	1 (<1)
azathioprine	134 (1)
capecitabine	20 (<1)
carbamazepine	102 (1)
citalopram	1531 (14)
escitalopram	524 (5)
clomipramine	8 (<1)
clopidogrel	636 (6)
codeine	3704 (33)
desipramine	20 (<1)
doxepin	179 (2)
fluorouracil	625 (6)
fluvoxamine	16 (<1)
imipramine	29 (<1)
ivacaftor	0 (0)
mercaptopurine	11 (<1)
nortriptyline	723 (7)
ondansetron	1927 (17)
oxcarbazepine	29 (<1)
paroxetine	369 (3)
peginterferon alfa-2a	4 (<1)
peginterferon alfa-2b	3 (<1)
ribavirin	8 (<1)
phenytoin	41 (<1)
rasburicase	0 (0)
sertraline	1268 (11)
simvastatin	3562 (32)
tacrolimus	449 (4)
tamoxifen	180 (2)
tegafur	0 (0)
thioguanine	0 (0)
trimipramine	3 (<1)
tropisetron	0 (0)
voriconazole	53 (<1)

[a]Therapeutic recommendations are available at https://www.pharmgkb.org/guidelines

2 kb upstream/1 kb downstream for *CYP2A6* were included. For PGx-seq, these upstream and downstream targets were dropped. In both v.1 and PGx-seq, intron 6 of *CYP3A4* was targeted. In PGx-seq, the entire ~31 kb region for *CYP2D6* was targeted including both of the nearby pseudogenes. For both versions, all of the Illumina ADME and Affymetrix DMET single nucleotide polymorphism (SNP) sites not already covered by gene targets were included. Over 250 genes have either DNA sequencing or SNP genotyping information available (Table 4).

For the RIGHT pilot, *CYP2C19*, *CYP2C9*, *VKORC1*, *SLCO1B1* and *CYP2D6*, sequencing results were interpreted at the Mayo Clinic Personalized Genomics Laboratory, a CLIA-certified and CAP-accredited laboratory. These genes were chosen because they have strong clinical practice recommendations when patients are prescribed warfarin, clopidogrel, simvastatin, tamoxifen, opioid analgesics and selective serotonin reuptake inhibitors.[16] For the expanded RIGHT cohort, *CYP1A2, CYP2C9, CYP2C19, CYP2D6, CYP3A4, CYP3A5, DPYD, SLCO1B1, TPMT, UGT1A1, VKORC1, HLA-A* and *HLA-B* are interpreted and reported in the EHR by the Personalized Genomics Laboratory. Point of care drug–gene alerts access these reports and present a description of the variants identified for each gene to the prescriber. In addition, the clinician receives an interpretation of the impact of the findings on gene function as well as pharmacogenomic implications. Table 5 includes the drug–gene alerts active in the Mayo Clinic EHR as of December 31, 2018.

A pharmacogenomic interpretive report (OneOme LLC, Minneapolis, MN) provides information to clinicians on how an individual patient's genes may affect drug response. Proprietary algorithms and curated clinical data are used to generate a personalized report based on a patient's genomic results. The report lists drugs categorized into three different bins according to genomic impact on different metabolic or physiologic effects: 'Major gene–drug interaction' (red), 'Moderate gene–drug interaction' (yellow) and 'Minimal gene–drug interaction' (green). The drugs are further classified by primary therapeutic area. For each drug, the report provides the patient's results of the interpreted genes along with additional clinical annotation and interpretations in both textual and graphical form to provide clinicians with a better understanding of the complex information. These static reports flow through the Personalized Genomics Laboratory and are placed into the EHR for each patient. These reports may be viewed by health care providers, but do not interact with the EHR.

Drug prescription information

Drug prescription information is available via the REP. Complete outpatient electronic prescription data are

Table 4. Genetic data available in the RIGHT study

Gene	Sequence data		Genotype data
	Version 1	Version 3	
ABCA1	Yes	No	No
ABCB1	Yes	Yes	Yes
ABCB11	Yes	No	Yes
ABCB4	No	No	Yes
ABCB7	No	No	Yes
ABCC1	No	No	Yes
ABCC2	Yes	No	Yes
ABCC3	No	No	Yes
ABCC4	No	Yes	Yes
ABCC5	No	No	Yes
ABCC6	No	No	Yes
ABCC8	No	No	Yes
ABCC9	No	No	Yes
ABCG1	Yes	No	Yes
ABCG2	Yes	Yes	Yes
ABP1	No	No	Yes
ACAD10	No	No	Yes
ACE	Yes	No	No
ADH1A	No	No	Yes
ADH1B	No	No	Yes
ADH1C	No	No	Yes
ADH4	No	No	Yes
ADH5	No	No	Yes
ADH6	No	No	Yes
ADH7	No	No	Yes
ADRB1	Yes	No	No
ADRB2	Yes	Yes	No
AGR3	No	No	Yes
AHR	Yes	No	Yes
AKAP9	No	No	Yes
ALB	No	No	Yes
ALDH1A1	No	No	Yes
ALDH2	No	No	Yes
ALDH3A1	No	No	Yes
ALDH3A2	No	No	Yes
ALOX5	Yes	No	No
ANKK1	No	Yes	No
ANKRD17	No	No	Yes
ANXA1	No	No	Yes
AOX1	No	No	Yes
AOX2P	No	No	Yes
APOA1	Yes	No	No
APOA2	No	No	Yes
ARID5B	Yes	No	No
ARNT	No	No	Yes
ARSA	No	No	Yes
ASB4	No	No	Yes
ATP7A	No	No	Yes
ATP7B	No	No	Yes
ATXN7L2	No	No	Yes
BCKDK	No	No	Yes
BDNF	Yes	No	No
C14orf164	No	No	Yes
C16orf54	No	No	Yes
C18orf16	No	No	Yes
C1orf87	No	No	Yes
C3orf22	No	No	Yes
C4orf17	No	No	Yes
CACNA1C	Yes	No	No
CACNA1S	Yes	No	No
CACNB2	Yes	No	No
CBR1	No	No	Yes
CBR3	No	No	Yes
CCDC101	No	No	Yes
CDA	No	No	Yes
CES1	Yes	Yes	No
CES2	Yes	No	Yes
CFTR	No	Yes	No
CHST1	No	No	Yes
CHST10	No	No	Yes
CHST11	No	No	Yes
CHST13	No	No	Yes
CHST2	No	No	Yes
CHST3	No	No	Yes
CHST4	No	No	Yes
CHST5	No	No	Yes
CHST6	No	No	Yes
CHST7	No	No	Yes
CHST8	No	No	Yes
CHST9	No	No	Yes
COMT	Yes	Yes	Yes
CRHR1	Yes	No	No
CRIP3	No	No	Yes
CROT	No	No	Yes
CYP11A1	No	No	Yes
CYP11B1	No	No	Yes
CYP11B2	No	No	Yes
CYP17A1	No	No	Yes
CYP19A1	No	No	Yes
CYP1A1	No	No	Yes
CYP1A2	Yes	Yes	Yes
CYP1B1	No	No	Yes
CYP20A1	No	No	Yes
CYP21A2	No	No	Yes
CYP24A1	No	No	Yes
CYP26A1	No	No	Yes
CYP26C1	No	No	Yes
CYP27A1	No	No	Yes
CYP27B1	No	No	Yes
CYP2A13	No	No	Yes
CYP2A6	Yes	Yes	Yes
CYP2A7	No	No	Yes
CYP2B6	Yes	Yes	Yes
CYP2B7P1	No	No	Yes

(continued)

(continued)

Table 4. Continued

Gene	Sequence data		Genotype data
	Version 1	Version 3	
CYP2C18	No	No	Yes
CYP2C19	Yes	Yes	Yes
CYP2C8	No	Yes	Yes
CYP2C9	Yes	Yes	Yes
CYP2D6	Yes	Yes	Yes
CYP2D7P1	No	No	Yes
CYP2E1	No	Yes	Yes
CYP2F1	No	No	Yes
CYP2J2	No	Yes	Yes
CYP2R1	Yes	No	No
CYP2S1	No	No	Yes
CYP39A1	No	No	Yes
CYP3A4	Yes	Yes	Yes
CYP3A43	No	No	Yes
CYP3A5	Yes	Yes	Yes
CYP3A7	No	No	Yes
CYP46A1	No	No	Yes
CYP4A11	No	No	Yes
CYP4A22	No	No	Yes
CYP4B1	No	No	Yes
CYP4F11	No	No	Yes
CYP4F12	No	No	Yes
CYP4F2	No	Yes	Yes
CYP4F3	No	No	Yes
CYP4F8	No	No	Yes
CYP4X1	No	No	Yes
CYP4Z1	No	No	Yes
CYP4Z2P	No	No	Yes
CYP51A1	No	No	Yes
CYP7A1	No	No	Yes
CYP7B1	No	No	Yes
CYP8B1	No	No	Yes
DBH	Yes	No	No
DCK	No	No	Yes
DKFZp779M0652	No	No	Yes
DPYD	Yes	Yes	Yes
DRD1	Yes	No	No
DRD2	Yes	Yes	No
DRD3	No	Yes	No
DRD4	No	Yes	No
EGFR	Yes	Yes	No
EPHX1	No	No	Yes
EPHX2	No	No	Yes
ESR1	Yes	No	No
F2	No	No	No
F5	No	Yes	No
FAAH	No	Yes	Yes
FAM19A3	No	No	Yes
FKBP5	Yes	No	No
FMO1	No	No	Yes
FMO2	No	No	Yes
FMO3	No	No	Yes
FMO4	No	No	Yes

(continued)

Table 4. Continued

Gene	Sequence data		Genotype data
	Version 1	Version 3	
FMO5	No	No	Yes
FMO6P	No	No	Yes
G6PD	Yes	Yes	Yes
GGCX	No	Yes	No
GLCCI1	Yes	No	No
GRIK4	No	Yes	No
GRIP2	No	No	Yes
GRK4	Yes	No	No
GRK5	Yes	No	No
GSTA1	No	No	Yes
GSTA2	No	No	Yes
GSTA3	No	No	Yes
GSTA4	No	No	Yes
GSTA5	No	No	Yes
GSTM1	No	No	Yes
GSTM2	No	No	Yes
GSTM2P1	No	No	Yes
GSTM3	No	No	Yes
GSTM4	No	No	Yes
GSTM5	No	No	Yes
GSTO1	No	No	Yes
GSTP1	No	Yes	Yes
GSTT1	No	No	Yes
GSTT2	No	No	Yes
GSTT2B	No	No	Yes
GSTZ1	No	No	Yes
HLA-A	No	Yes	No
HLA-B	No	Yes	No
HMGCR	Yes	Yes	Yes
HNF1A	No	Yes	No
HNF4A	No	Yes	No
HNMT	No	No	Yes
HOMEZ	No	No	Yes
HSD11B2	Yes	No	No
HTR1A	Yes	No	No
HTR2A	Yes	Yes	No
HTR2C	No	Yes	No
ICK	No	No	Yes
IFNL3	No	Yes	No
IGFBP7	No	Yes	No
IL28B	No	No	No
KCNH2	Yes	Yes	No
KIF6	No	Yes	No
LDLR	Yes	Yes	No
LEP	No	Yes	No
LEPR	No	Yes	No
LOC26172	No	No	Yes
LOC286186	No	No	Yes
LOC619207	No	No	Yes
LOC651536	No	No	Yes
LOC96610	No	No	Yes
LST-3TM12	No	No	Yes
MAOA	Yes	No	Yes

(continued)

Table 4. Continued

Gene	Sequence data		Genotype data
	Version 1	Version 3	
MAOB	No	No	Yes
MAT1A	No	No	Yes
METTL1	No	No	Yes
MKL1	No	No	Yes
MTHFR	No	Yes	No
MYOF	No	No	Yes
NAT1	No	No	Yes
NAT2	Yes	Yes	Yes
NNMT	No	No	Yes
NPPB	Yes	No	No
NPR1	Yes	No	No
NQO1	No	No	Yes
NR1I2	No	No	Yes
NR1I3	No	No	Yes
NR3C1	Yes	No	Yes
NR3C2	Yes	No	No
NTRK2	Yes	No	No
NTSR1	No	No	Yes
NUDT15	No	No	No
OPA3	No	No	Yes
OPRM1	No	Yes	No
ORM1	No	No	Yes
ORM2	No	No	Yes
PARL	No	No	Yes
PCP4L1	No	No	Yes
PEAR1	Yes	No	No
PGAP3	No	No	Yes
PNMT	No	No	Yes
PON1	No	Yes	Yes
PON2	No	No	Yes
PON3	No	No	Yes
POR	Yes	No	Yes
PPARD	No	No	Yes
PPARG	No	No	Yes
PPP1R9A	No	No	Yes
PRSS53	No	No	Yes
PTGIS	Yes	No	Yes
PTGS1	Yes	No	No
QPRT	No	No	Yes
RALBP1	No	No	Yes
RPL13	No	No	Yes
RXRA	No	No	Yes
RYR1	Yes	Yes	No
RYR2	Yes	Yes	No
SCN1A	No	Yes	No
SCN5A	Yes	Yes	No
SDCBP	No	No	Yes
SERPINA7	No	No	Yes
SETD4	No	No	Yes
SHANK3	No	No	Yes
SLC10A1	No	No	Yes
SLC10A2	No	No	Yes

(continued)

Table 4. Continued

Gene	Sequence data		Genotype data
	Version 1	Version 3	
SLC13A1	No	No	Yes
SLC15A1	No	No	Yes
SLC15A2	Yes	No	Yes
SLC16A1	No	No	Yes
SLC19A1	No	Yes	Yes
SLC22A1	Yes	Yes	Yes
SLC22A11	No	No	Yes
SLC22A12	No	No	Yes
SLC22A13	No	No	Yes
SLC22A14	No	No	Yes
SLC22A2	Yes	Yes	Yes
SLC22A3	Yes	No	Yes
SLC22A4	No	No	Yes
SLC22A5	No	No	Yes
SLC22A6	Yes	No	Yes
SLC22A7	No	No	Yes
SLC22A8	No	No	Yes
SLC25A27	No	No	Yes
SLC28A1	No	No	Yes
SLC28A2	No	No	Yes
SLC28A3	No	No	Yes
SLC29A1	No	No	Yes
SLC29A2	No	No	Yes
SLC41A2	No	No	Yes
SLC47A1	Yes	No	No
SLC47A2	Yes	No	No
SLC5A6	No	No	Yes
SLC6A3	Yes	No	No
SLC6A4	Yes	Yes	No
SLC6A6	No	No	Yes
SLC7A5	No	No	Yes
SLC7A5P1	No	No	Yes
SLC7A7	No	No	Yes
SLC7A8	No	No	Yes
SLC9A7	No	No	Yes
SLCO1A2	Yes	No	Yes
SLCO1B1	Yes	Yes	Yes
SLCO1B3	Yes	No	Yes
SLCO2B1	Yes	Yes	Yes
SLCO3A1	No	No	Yes
SLCO4A1	No	No	Yes
SLCO5A1	No	No	Yes
SOD2	No	Yes	No
SPG7	No	No	Yes
SPN	No	No	Yes
SR140	No	No	Yes
SULT1A1	No	Yes	Yes
SULT1A2	No	No	Yes
SULT1A3	No	No	Yes
SULT1A4	No	No	Yes
SULT1B1	No	No	Yes
SULT1C2	No	No	Yes

(continued)

Table 4. Continued

Gene	Sequence data		Genotype data
	Version 1	Version 3	
SULT1C4	No	No	Yes
SULT1E1	No	No	Yes
SULT2A1	No	No	Yes
SULT2B1	No	No	Yes
SULT4A1	No	No	Yes
TBXAS1	Yes	No	Yes
TCL1A	Yes	No	No
TFF3	No	No	Yes
TMEM120A	No	No	Yes
TMEM176A	No	No	Yes
TMPRSS11B	No	No	Yes
TOP1P1	No	No	Yes
TP53	No	Yes	No
TPMT	Yes	Yes	Yes
TPSG1	No	No	Yes
TSEN2	No	No	Yes
TTBK1	No	No	Yes
TYMS	No	Yes	Yes
UGT1A1	Yes	Yes	Yes
UGT1A10	No	No	Yes
UGT1A3	No	No	Yes
UGT1A3-10 exon1	No	Yes	No
UGT1A4	Yes	Yes	Yes
UGT1A5	No	No	Yes
UGT1A6	No	No	Yes
UGT1A7	No	No	Yes
UGT1A8	No	No	Yes
UGT1A9	No	No	Yes
UGT2A1	No	No	Yes
UGT2A2	No	No	Yes
UGT2B10	No	No	Yes
UGT2B11	No	No	Yes
UGT2B15	No	Yes	Yes
UGT2B17	No	No	Yes
UGT2B28	No	No	Yes
UGT2B4	No	No	Yes
UGT2B7	No	Yes	Yes
UGT8	No	No	Yes
VDR	Yes	No	No
VEGFA	No	Yes	No
VKORC1	Yes	Yes	Yes
XDH	No	No	Yes
XYLB	No	No	Yes
ZNF280B	No	No	Yes
ZNF423	Yes	No	No
ZNF662	No	No	Yes

available from January 1, 2004 through to the present. Filled drug prescription data are not available. However, medication reconciliation occurs when patients visit a health care provider. Patients are asked to list all of the medications they are currently taking, including prescribed medications, over the counter medications and supplements. All medication information is coded using the RxNorm and National Drug File-Reference Terminology (NDF-RT) coding systems.[17,18] Combination drugs with multiple ingredients are counted under the category of the main ingredient, or under the combination drug category when applicable. The REP electronic indexes may be searched to identify all persons with either a prescription or report of medication use in the desired time frame. Prescriptions for drugs with pharmacogenomic guidelines for the RIGHT cohort are summarized in Table 3.

Table 5. Mayo Clinic clinical decision support interventions

Gene	Drug	Year went live
CYP2C19	Clopidogrel	2014
	Citalopram	2015
	Escitalopram	2015
CYP2D6	Codeine	2013
	Tramadol	2013
	Tamoxifen	2013
	Paroxetine	2015
	Fluoxetine	2015
	Fluvoxamine	2015
	Venlafaxine	2015
SLCO1B1	Simvastatin	2014
CYP2C9/VKORC1	Warfarin	2014
TPMT	Mercaptopurine	2013
Genotype or phenotype	Thioguanine	2013
	Azathioprine	2013
*HLA-B*1502*	Carbamazepine	2013
*HLA-B*5701*	Abacavir	2013
*HLA-B*5801*	Allopurinol	2014
CYP3A5	Tacrolimus	2016
DPYD	5-FU	2017
NUDT15	Mercaptopurine	2018
	Thioguanine	2018
	Azathioprine	2018
*HLA-A*3101*	Carbamazepine	2018

Outcomes

Relevant pharmacogenomic outcomes (e.g. response to treatment or adverse events) may be assessed electronically using coded data from the available EHRs and billing systems. Such data include all International Classification of Diseases (ICD-9 and ICD-10) and Current Procedural Terminology (CPT) coded diagnoses and procedures from these EHRs, as well as health care utilizations including hospitalizations, office visits and emergency room visits. We also maintain access to the full text of the EHRs in the REP, and this text is available for chart abstraction or for

natural language processing (NLP) techniques. We have used such techniques to identify prescription drug allergies from the linked EHRs (see below).

Processed allergy section clinical notes

Drug side-effects as well as drug allergies are commonly reported in the allergy section of the clinical note. As a result, the allergy section is a rich source of information for adverse drug reactions. However, descriptions of drug events are often embedded in both structured and unstructured clinical narratives, and thus processes for cleaning and standardizing the data are required. For the RIGHT Protocol study, we deployed NLP methods to process the clinical text and annotate clinical concepts to capture adverse drug events.[19–21] We used the Medical Dictionary for Regulatory Activities (MedDRA),[22,23] a medical terminology ontology adopted to describe adverse drug reactions and provide an internationally approved classification, and the Unified Medical Language System (UMLS, version 2012AB) (https://www.nlm.nih.gov/research/umls/),[24] a knowledge resource for biomedical vocabularies developed by the US National Library of Medicine. Allergy data from the EHR were pulled and recorded as unstructured text, for example: 'fluid retention, hypertension' and 'anxiety, distorted vision'. We used MedTagger,[25] an NLP pipeline to extract adverse drug events mentioned in text. The extracted descriptions were then mapped to MedDRA terms and then aggregated using UMLS. For example, in the allergy text 'fluid retention, hypertension', 'fluid retention' was first extracted and mapped to MedDRA 'body fluid retention', and mapped to 'Renal and urinary disorders' and 'Metabolism and nutrition disorders', whereas 'hypertension' was extracted and mapped to 'Vascular disorders'. The top five most common drug reactions were skin and subcutaneous tissue disorders, gastrointestinal disorders, immune system disorders, nervous system disorders, and general disorders and administration site conditions.

EHR integration and clinical decision support

Since 2013, Mayo Clinic has developed and implemented a comprehensive multidisciplinary model[26] to integrate pharmacogenomic test results and clinical decision support (CDS) interventions in the Mayo Clinic unified EHR system (Table 5). Real-time pharmacogenomic-CDS alerts advise the need for pharmacogenomic testing when it is indicated, alert for potentially actionable drug–gene interactions, explain to prescribers the nature of the drug–gene interaction, and provide patient-specific therapeutic recommendations. The rules are embedded in the clinical workflow of all prescribers and pharmacists managing prescriptions (inpatient and outpatient prescription management systems). Multiple implementation and process metrics have been developed to allow post-implementation analyses. It is possible to measure when alerts are triggered and whether the clinician changes the initial prescription following the CDS alert. Such data allow for studies of clinician behaviours related to pharmacogenomics-CDS implementation.

What has it found? Key findings and publications

In the RIGHT pilot study, we found that 99% of participants had a clinically actionable variant when considering the five genes reported in the EHR (*CYP2C19*, *CYP2C9*, *VKORC1*, *SLCO1B1* and *CYP2D6*).[27] We have also reported our experience with integrating pharmacogenomic-CDS alerts into clinical practice in recent publications.[16,26,28]

Drug-focused studies include investigating drug response and clinical outcomes associated with exposure to opioids,[29] diabetes treatment, proton pump inhibitors, statins, anaesthesia, antidepressants, anticoagulants, penicillin, beta blockers and antifungals. Other ancillary studies have assessed patient and provider perspectives on pharmacogenomics.[12,13,28] Finally, additional studies have examined associations between pharmacogenomic phenotypes and health care utilization.[30]

One of our key findings is related to our first return of results to participants. After the pilot was completed, we returned *CYP2D6* results to participants and surveyed them to determine their understanding of the data. Of the respondents, 26% said that they only somewhat understood their results and 7% did not understand them at all. Participants commonly suggested that results should be personalized, should refer to drugs they were currently taking, and should use layperson language.[12]

What are the main strengths and weaknesses?

The strengths of the RIGHT cohort derive from the large sample size, extensive EHR data, medical record access for abstraction, ability to study any drug exposure, and availability of sequence data for discovery. The RIGHT cohort are all Mayo Clinic Biobank participants who have existing, stored biospecimens (DNA, plasma, serum, white blood cells) and who can be re-contacted for future research questions or invited to participate in ancillary studies.

The ability to cost-effectively re-contact participants expands the research capability of the cohort. Despite the strengths, the RIGHT cohort has some limitations. First, the cohort was not selected from patients with any particular diagnosis or drug indication. Thus, although the RIGHT study is one of the largest studies of pre-emptive pharmacogenomics, less frequent and rare drug exposures will be difficult to study in this population. Second, because this is an observational cohort, and not a clinical trial, patients in this study have been prescribed drugs based upon prescriber practices, the dictates of insurance coverage, and not randomized methods. Thus, inferences may be limited.

Can I get hold of the data? Where can I find out more?

The RIGHT cohort is a resource for pharmacogenomic research.[11] As part of the infrastructure, a RIGHT Data Access Committee has been created to review data requests for use of RIGHT data. External access to the data is facilitated by the Mayo Clinic Biobank[31] https://www.mayo.edu/research/centers-programs/mayo-clinic-biobank/overview

Profile in a nutshell

- Pharmacogenomics is the hallmark of genomic medicine worldwide.
- 11 098 Mayo Clinic patients have consented to pre-emptive pharmacogenomic testing.
- The RIGHT study was developed to address several goals:
  - identify the associations between genetic variants and clinical outcomes in patients prescribed drugs known to interact with these genes,
  - determine if pre-emptive pharmacogenomic testing improves patient outcomes and reduces health care costs,
  - identify clinical outcomes associated with pharmacogenomic variants of unknown or indeterminate clinical significance, and
  - understand the perspectives and practices of clinicians and patients.
- DNA sequencing or SNP genotyping information is available for over 250 genes.
- Decades of prescription data are available for research via the Rochester Epidemiology Project (REP) medical records-linkage system.
- Comparisons of the RIGHT cohort to a pharmacogenomic-naïve cohort allow for the study of the effects of genotype-guided dosing.

Funding

This work was supported by the Mayo Clinic Center for Individualized Medicine, the Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, and the National Institutes of Health grants [U19 GM61388 (The Pharmacogenomics Research Network), R01 GM28157, U01 HG005137, R01 GM125633, R01 AG034676 (The Rochester Epidemiology Project)], and {U01 HG06379 and U01 HG06379 Supplement [The Electronic Medical Record and Genomics (eMERGE) Network]}.

Conflict of interest: J.L.B. III, L.W., R.M.W., and the Mayo Clinic have stock ownership and have licensed intellectual property to OneOme LLC.

References

1. Weinshilboum R. Inheritance and drug response. *N Engl J Med* 2003;**348**:529–37.
2. Weinshilboum RM, Wang L. Pharmacogenomics: precision medicine and drug response. *Mayo Clin Proc* 2017;**92**:1711–22.
3. Wang L, McLeod HL, Weinshilboum RM. Genomics and drug response. *N Engl J Med* 2011;**364**:1144–53.
4. Clinical Pharmacogenetics Implementation Consortium (CPIC®). https://cpicpgx.org/ (19 November 2018, date last accessed).
5. The Canadian Pharmacogenomics Network for Drug Safety (CPNDS). http://cpnds.ubc.ca/ (1 March 2019, date last accessed).
6. Dutch Pharmacogenetics Working Group (DPWG). The Dutch Pharmacogenetics Working Group (DPWG) was established in 2005 by the Royal Dutch Pharmacist's Association (KNMP). https://www.pharmgkb.org/page/dpwg (19 November 2018, date last accessed).
7. Olson JE, Ryu E, Johnson KJ *et al*. The Mayo Clinic Biobank: a building block for individualized medicine. *Mayo Clin Proc* 2013;**88**:952–62.
8. Rocca WA, Grossardt BR, Brue SM *et al*. Data Resource Profile: expansion of the Rochester Epidemiology Project medical records-linkage system (E-REP). *Int J Epidemiol* 2018;**47**:368–j.
9. St Sauver JL, Grossardt BR, Yawn BP *et al*. Data resource profile: the Rochester Epidemiology Project (REP) medical records-linkage system. *Int J Epidemiol* 2012;**41**:1614–24.
10. St Sauver JL, Grossardt BR, Yawn BP, Melton LJ 3rd, Rocca WA. Use of a medical records linkage system to enumerate a dynamic population over time: the Rochester Epidemiology Project. *Am J Epidemiol* 2011;**173**:1059–68.
11. Bielinski SJ, Olson JE, Pathak J *et al*. Preemptive genotyping for personalized medicine: design of the Right Drug, Right Dose, Right Time-using Genomic Data to Individualize Treatment Protocol. *Mayo Clin Proc* 2014;**89**:25–33.
12. Olson JE, Rohrer Vitek C, Bell EJ *et al*. Participant-perceived understanding and perspectives on pharmacogenomics: the Mayo Clinic RIGHT protocol (Right Drug, Right Dose, Right Time). *Genet Med* 2017;**19**:819–25.
13. Bielinski SJ, St Sauver JL, Olson JE *et al*. Are patients willing to incur out-of-pocket costs for pharmacogenomic testing? *Pharmacogenomics J* 2017;**17**:1–3.
14. Gordon AS, Fulton RS, Qin X, Mardis ER, Nickerson DA, Scherer S. PGRNseq: a targeted capture sequencing panel for

pharmacogenetic research and implementation. *Pharmacogenet Genom* 2016;**26**:161–68.
15. Zhou SF, Liu JP, Chowbay B. Polymorphism of human cytochrome P450 enzymes and its clinical impact. *Drug Metab Rev* 2009;**41**:89–295.
16. Caraballo PJ, Parkulo M, Blair D *et al*. Clinical decision support to implement CYP2D6 drug-gene interaction. *Stud Health Technol Inform* 2015;**216**:946.
17. Pathak J, Chute CG. Analyzing categorical information in two publicly available drug terminologies: RxNorm and NDF-RT. *J Am Med Inform Assoc* 2010;**17**:432–39.
18. Pathak J, Murphy SP, Willaert BN *et al*. Using RxNorm and NDF-RT to classify medication data extracted from electronic health records: experiences from the Rochester Epidemiology Project. *AMIA Annual Symposium Proceedings AMIA Symposium* 2011;**2011**:1089–98.
19. Wang Y, Wang L, Rastegar-Mojarad M *et al*. Clinical information extraction applications: a literature review. *J Biomed Inform* 2018;**77**:34–49.
20. Wang X, Hripcsak G, Markatou M, Friedman C. Active computerized pharmacovigilance using natural language processing, statistics, and electronic health records: a feasibility study. *J Am Med Inform Assoc* 2009;**16**:328–37.
21. Wang X, Chase H, Markatou M, Hripcsak G, Friedman C. Selecting information in electronic health records for knowledge acquisition. *J Biomed Inform* 2010;**43**:595–601.
22. Pearson RK, Hauben M, Goldsmith DI *et al*. Influence of the MedDRA hierarchy on pharmacovigilance data mining results. *Int J Med Inform* 2009;**78**:e97–e103.
23. Bousquet C, Lagier G, Lillo-Le Louët A, Le Beller C, Venot A, Jaulent MC Appraisal of the MedDRA conceptual structure for describing and grouping adverse drug reactions. *Drug Saf* 2005;**28**:19–34.
24. Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Res* 2004;**32**: D267–70.
25. Wagholikar K, Torii M, Jonnalagadda S, Liu H. Feasibility of pooling annotated corpora for clinical concept extraction. *AMIA Jt Summits Transl Sci Proc* 2012;**2012**:38.
26. Caraballo PJ, Hodge LS, Bielinski SJ *et al*. Multidisciplinary model to implement pharmacogenomics at the point of care. *Genet Med* 2017;**19**:421–9.
27. Ji Y, Skierka JM, Blommel JH *et al*. Preemptive pharmacogenomic testing for precision medicine: a comprehensive analysis of five actionable pharmacogenomic genes using next-generation DNA sequencing and a customized CYP2D6 genotyping cascade. *J Mol Diagn* 2016;**18**:438–45.
28. St Sauver JL, Bielinski SJ, Olson JE *et al*. Integrating pharmacogenomics into clinical practice: promise vs reality. *Am J Med* 2016;**129**:1093–9.e1.
29. St Sauver JL, Olson JE, Roger VL *et al*. CYP2D6 phenotypes are associated with adverse outcomes related to opioid medications. *Pharmgenomics Pers Med* 2017;**10**:217–27.
30. Takahashi PY, Ryu E, Pathak J *et al*. Increased risk of hospitalization for ultrarapid metabolizers of cytochrome P450 2D6. *Pharmgenomics Pers Med* 2017;**10**:39–47.
31. Mayo Clinic Biobank. https://www.mayo.edu/research/centers-programs/mayo-clinic-biobank/overview (19 November 2018, date last accessed).