



feature



FDA drug labeling: rich resources to facilitate precision medicine, drug safety, and regulatory science

Hong Fang¹, Stephen C. Harris¹, Zhichao Liu¹, Guangxu Zhou¹, Guoping Zhang¹, Joshua Xu¹, Lilliam Rosario², Paul C. Howard¹ and Weida Tong¹, weida.tong@fda.hhs.gov

Here, we provide a concise overview of US Food and Drug Administration (FDA) drug labeling, which details drug products, drug–drug interactions, adverse drug reactions (ADRs), and more. Labeling data have been collected over several decades by the FDA and are an important resource for regulatory research and decision making. However, navigating through this data is challenging. To aid such navigation, the FDALabel database was developed, which contains a set of approximately 80 000 labeling data. The full-text searching capability of FDALabel and querying based on any combination of specific sections, document types, market categories, market date, and other labeling information makes it a powerful and attractive tool for a variety of applications. Here, we illustrate the utility of FDALabel using case scenarios in pharmacogenomics biomarkers and ADR studies.

Introduction

FDA drug labels contain rich and comprehensive information about drug products, such as disease indications, target populations, drug–drug interactions, and ADRs. The label of a prescription drug is prepared by manufacturers and approved by the FDA and, thus, in its final form, reflects the collective input from regulators, drug manufacturers, and scientific experts. Drug-labeling data have been an important resource for diverse applications, including the support for policy development [1–4], drug discovery and development [5,6], support for pharmacogenomics applications for personalized medicine [7–9], and scientific research [2,3,7,10–13].

Drug labeling is not static and around 400–500 new or updated drug-labeling documents

are added every week to the current total of approximately 80 000 structured product labels (SPLs) (www.fda.gov/ForIndustry/DataStandards/StructuredProductLabeling/default.htm). Drug labeling has changed over time because of evolving FDA regulations and has increased in content and length, with a standard format to guide the safe and effective use of drugs [14] (www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm065010.htm). For example, each prescription drug labeling in the current format is often complex, with over 20 pages of text and tables covering a range of information about the drug product. This rapid pace of change and the complexity in content illustrate the need for an advanced bioinformatics environment with

robust and powerful data management and search capabilities to facilitate the application of drug labeling information.

Here, we describe the FDALabel database developed by the FDA as a web-based application (www.fda.gov/ScienceResearch/BioinformaticsTools/ucm289739.htm). The tool allows access to the most up-to-date drug-labeling data and facilitates their use in regulatory science, drug development, scientific research, and clinical application, such as: (i) enabling easy querying of drug information for research and monitoring of ADRs (e.g., Boxed Warning, drug-induced liver injury (DILI), pharmacogenomics biomarker) to advance pharmacovigilance; (ii) supporting research through integrating drug-labeling data with other drug databases and

disease information; (iii) allowing critical drug information to be readily accessed by physicians, healthcare professionals, and patients for enhanced healthcare; and (iv) facilitating transparent knowledge exchange among the public, pharmaceutical companies, and government regulatory agencies. To illustrate its utility, we provide examples of using FDALabel for pharmacovigilance and precision medicine. Lastly, we discuss the future direction of FDALabel development to better support research and regulatory science for improved drug safety and public health.

FDA drug labeling is a rich resource for drug information

Structured Product Labeling (SPL)

In April 2005, the FDA published a guidance document requiring the submission of labeling contents in an electronic format with standardized SPL data structures (using a machine-readable XML format and in compliance with Health Level 7) [15,16]. Drug-labeling SPLs contain diverse information about a product, such as the carton and container labels, prescribing information, patient labeling, drug advertisements, and promotional materials. The FDA archive comprises SPLs for human drugs, biologic products, animal drugs, human devices, and human vaccines. Approximately 96% (77 000 out of 80 000) SPLs are for human drugs, including human prescription drugs, biologic products, and over-the-counter drugs. In this article, we use 'drug labeling' as a broad term to define the drug information in the SPLs of the FDA for human prescription drugs. Of note, one prescription drug can have many drug products

because of the differences in regulatory application, dosage forms, routes of administration, and manufacturers. Although there are only approximately 1600 human prescription drugs, there are approximately 31 000 SPLs.

Physician Labeling Rule (PLR)

In 2006, FDA issued the final regulation 'Requirements on Content and Format of Labeling for Human Prescription Drug and Biological Products' (www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/LawsActsandRules/ucm084159.htm), which is also known as the 'Physician Labeling Rule (PLR)' as described in 21 CFR 201.56 and 201.57. Its central goal is to provide structured labeling information that is easy to access, read, and use by both the FDA and the public. Specific requirements were added, such as 'Highlights of Prescribing Information', 'Table of Contents', and 'Reorders and Reorganizes' sections. The revisions were mandatory for new drug labeling as well as for drugs that have been on the market up to 5 years before the effective date of the PLR (June 30, 2006). Consequently, the PLR enhances the consistency in drug labeling and makes it a valuable resource for drug review, development, and research.

As summarized in Table 1, 'Highlights of Prescribing Information' is a half-page summary of the essential safety and efficacy information for approved human prescription drug and biological products. By contrast, full prescribing information has a total of 17 labeling sections with more detailed content compared with the 'Highlights'. The additional sections include Drug Abuse and Dependence (Section 9), Over Dosage

(Section 10), Clinical Pharmacology (Section 12), Nonclinical Toxicology (Section 13), Clinical Studies (Section 14), and Patient Counseling Information (Section 17). The information in each section is structured and specified for certain drug information. Of note, patients receive the prescription drug information with limited information from the Full Prescribing Information (FPI) specified for patient (Medication Guides).

Pharmacogenomics data

In January 2013, the guidance for industry on 'Clinical Pharmacogenomics' information preparation for labeling was released [14] (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM337169.pdf>). A new subsection, Pharmacogenomics, was added under Clinical Pharmacology (Section 12). The guidance document indicates that, 'If applicable, a 'Pharmacogenomics' subsection should be included in the CLINICAL PHARMACOLOGY section 12 (e.g., as '12.5 Pharmacogenomics') of the prescribing information (PI) and should include clinically relevant data or information on the effect of genetic variations affecting drug therapy'. Recently, many pharmacogenomics biomarkers have been included in drug labeling (e.g., the drugs warfarin, boceprevir, Nuedexta®, Brovana®, and pantoprazole) [8], which allows clinicians to apply these drugs to the specific populations who are most likely to benefit from precision medicine [8,9]. The pharmacogenomics biomarkers found in drug labeling can be categorized as: (i) involved in drug metabolism variability (e.g., CYP enzymes) among individuals; (ii) associated with increased risk for adverse

TABLE 1

Summary of 'Highlights of Prescribing Information': the essential safety and efficacy information for approved human prescription drug and biological products.

Title	Notes
Limitations Statement	For example, These highlights do not include all the information needed to use ZIAGEN safely and effectively. See full prescribing information for ZIAGEN
Product Names	For example, ZIAGEN (abacavir) tablets, for oral use
Initial U.S. Approval	For example, Initial US approval: 1998
Boxed Warning	Commonly known as a black box warning
Recent Major Changes	The month and year the change was made
Indications and Usage	Concise Summary of Labeling Section 1
Dosage and Administration	Concise Summary of Labeling Section 2
Dosage Forms and Strengths	Concise Summary of Labeling Section 3
Contraindications	Concise Summary of Labeling Section 4
Warnings and Precautions	Concise Summary of Labeling Section 5
Adverse Reactions	Concise Summary of Labeling Section 6
Drug Interactions	Concise Summary of Labeling Section 7
Use in Specific Populations	Concise Summary of Labeling Section 8

BOX 1

Example applications using drug labeling data**Drug interactions**

Drug labeling has a specific section to summarize the findings relating to drug–drug interactions and their associated adverse events in drug application. Some specific questions can be queried against the labeling data, such as which drugs used to treat HIV (in the Drug Indications and Usage Section) are known to interact with methadone and which drugs will interact with disulfiram (in the Drug Interactions Section).

Drug classification

There are several ways to classify drugs; each one has its intended application (e.g., clinical application, mechanistic study, chemical structure, etc.). For example, which drugs share the same pharmacologic class, such as kinase inhibitors, HIV protease inhibitors, or beta-adrenergic blockers, and so on. The drug-labeling indexing provides classification based on Established Pharmacologic Class (EPC), Mode-of-Action (MoA), Physiologic Effect (PE), and Chemical Ingredient by structure (CI). These classification schemes facilitate the study of drug class effect and evidence-based justification for making a labeling change to a drug class during the review process.

Adverse events

Three sections (Boxed Warning, Warnings and Precautions, and Adverse Events) summarize drug-related adverse events, which have been widely applied in pharmacovigilance and drug safety research. While the standard vocabulary to describe adverse events is not mandatory, most terminologies for adverse events are implemented with certain standard terminology (such as SNOMED and MedDRA), which facilitates the study of the adverse events data in the drug labeling.

Precision medicine

A large number of pharmacogenomics biomarkers are included in drug labeling. These biomarkers are likely to impact the effectiveness and adverse events for patients from specified subpopulations taking the drugs. Thus, the information facilitates the identification of new trends and frequency of genetic variability associated with increased risks to public health, which is an important goal in precision medicine.

events (e.g., G6PD, TPMT, and HLA-B); and (c) describing the mechanism of action of the drug (e.g., CD30), which likely impacts the effect of the drug on specified patients (www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenetics/ucm083378.htm).

In summary, the drug labeling contains rich information from clinical studies, nonclinical studies, and postmarketing experiences regarding ADRs for pharmacovigilance and pharmacogenomics. At the FDA, drug-labeling development has been a crucial element in the drug review process. Its content can also support diverse research needs. Box 1 summarizes some of the key applications using the drug-labeling data.

FDALabel database

The FDALabel database contains over 80 000 full-text SPLs. The source of FDALabel data is the SPLs of the FDA archived in the FDA Online Labeling Repository (<http://labels.fda.gov/>) and DailyMed (<http://dailymed.nlm.nih.gov/dailymed/index.cfm>). FDALabel was implemented as a secure three-tier platform with an Oracle database and is updated quarterly. The database can be accessed through a web-based application. The tool has a simple query that can be intuitively performed (e.g., full-text search, product or generic name search in version 1.0). Importantly, the advanced query functions are

implemented to perform a range of queries for, individually or in combination: (i) presence of, or text within, specific sections of the prescribing information; (ii) document types (e.g., Human Rx, Human OTC, Vaccine); (iii) marketing categories (e.g., NDA, ANDA, BLA); (iv) SPL identifiers (e.g., Product NDC Codes and SETIDs); and (v) Market start/end date. In addition, the search summary results can be downloaded as a spreadsheet that links to the original SPLs.

Of note, DailyMed is a widely used resource for drug labeling and provides SPL data. However, many unique functions (e.g., the full-text searching capability and querying based on any combination of specific drug fields and sections) are available in FDALabel and, thus, make it a more powerful, user friendly, and attractive tool (see Table S1 in the supplemental information online for comparisons between FDALabel and DailyMed). For example, we searched for the keywords 'acute liver failure' in 'Full Text Query', which resulted in 757 labeling hits. We also searched the same keywords within Boxed Warning using 'Section Present', which resulted in 556 labeling hits. In addition, we also added NDA as filter from 'Marketing Categories', with the same two queries resulting in 53 hits and 23 hits, respectively. Thus, a large number of duplications in drug labeling can be easily removed by using NDA as a filter in FDALabel,

which is not readily available in DailyMed. The unique functions of FDALabel database enable the drug-labeling content to be more easily accessed by researchers for ADR study, FDA medical officers for drug review, pharmaceutical companies for drug development and repositioning, and physicians and consumers for drug safety information. Google Analytics has shown that the number of users has increased greatly since the database opened for public access in 2012 (Fig. S1 in the supplemental information online).

FDALabel use cases*Pharmacogenomics biomarkers*

Some pharmacogenomics biomarkers are associated with ADRs. We queried five ADR-related biomarkers (i.e., G6PD, TPMT, DPD, HLA-B*1502, and HLA-B* 5701) in FDALabel (Table S2 in the supplemental information online) and built a network visualization of the drug–ADR relations via these biomarkers (Fig. 1). The results illustrated that patients who carry the HLA-B*5701 allele are at high risk for experiencing a hypersensitivity reaction (HR) to abacavir, while patients who carry the HLA-B*1502 allele are at a high risk of HR to carbamazepine, which could lead to Steven–Johnson syndrome, a severe ADR mentioned in the Box Warning section [17,18].

ADR study using MedDRA standard

Medical Dictionary for Regulatory Activities (MedRA; www.meddra.org/) is widely used in the USA, European Union, and Japan for ADR reporting. Extracting MedDRA standard terms from drug labeling will allow researchers, regulators, and healthcare professionals to better understand the trends and frequencies of adverse events for drugs in the current markets [6]. There are five term levels in MedDRA, from lowest to highest: LLT (Low-Level Term), PT (Preferred Term), HLT (High-Level Term), HLGT (High-Level Group Term), and SOT (System Organ Class), with PTs the most commonly used for ADR study. Our mapping showed that, out of a total of 74 229 MedDRA (version 18.0) LLTs, 11 847 LLTs have appeared in FDA-approved prescription drug labeling, which, in turn, identified 6161 PTs (out of total of 21 345 PTs in MedDRA). The top ten labeling sections that contain the most PTs are plotted by counts (Fig. 2). Not surprisingly, the Adverse Reactions section contained the most PTs (4444), followed by Precautions (2656), Warnings and Precautions (2550), Clinical Study (2475), Indications and Usage (2460), Clinic Pharmacology (2161), and Warnings (2134). Of note, the Precautions section and the Warnings section in non-PLR format

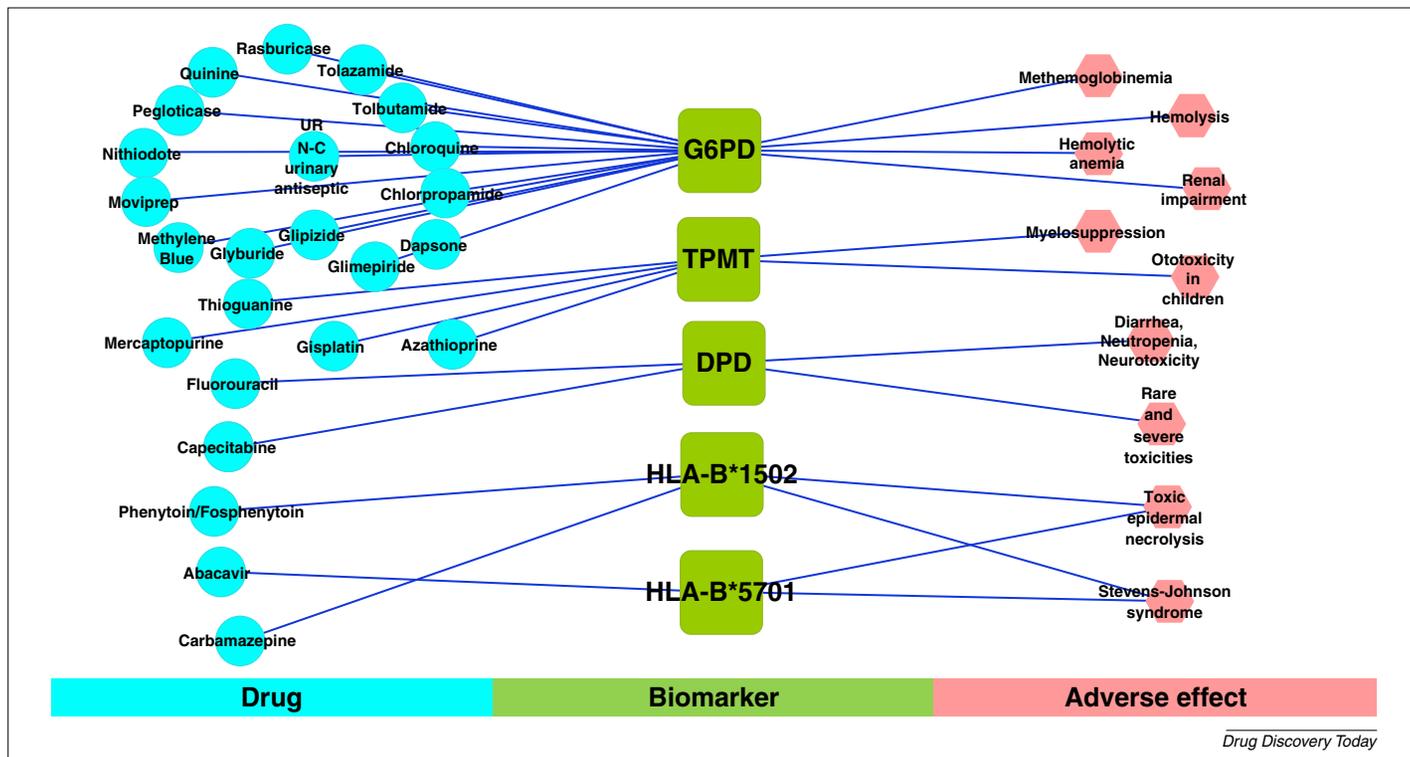


FIGURE 1

Five selected pharmacogenomics biomarkers and their associated adverse effects and drugs. A network visualization illustrates the relation among drugs (blue), biomarkers (green), and associated adverse effects (red) based on the information retrieved from FDA Label. For example, patients who carry the HLA-B*1502 allele are at a high risk for experiencing a hypersensitivity reaction (HR: such as Stevens–Johnson syndrome) to carbamazepine.

have been combined into the Warnings and Precautions section in the new drug-labeling format in PLRs.

Drug-induced liver injury study

We utilized drug-labeling data to study DILI [17], drug safety [19], and drug repositioning [5]. For example, we developed a systematic annotation method using drug-labeling information to an-

notate the potential of a drug for DILI. Specifically, a combination of keywords about DILI, which reflected not only different types, but also various severity levels of DILI, was used to search against the drug-labeling database. The study enabled the relevant DILI information to be extracted from three labeling sections (Boxed Warning, Warnings and Precautions, and Adverse Reactions) with a DILI classification scheme to

define a benchmark DILI data set that is widely used as a model in DILI study [20,21].

Concluding remarks and future perspectives

FDA drug labeling has accumulated over the past 40 years or so [since the Federal Register of June 26, 1979 (44 FR 37434)] and is an integral part of the FDA review process. In addition, many guidance documents have been issued by the FDA to facilitate its application, such as for drug discovery and development. Similar labeling resources have also been developed around the world, such as in Europe and Japan. Given the recent implementation of data standards and rapid advancement of information technology, drug-labeling data has grown tremendously, truly becoming regulatory big data for knowledge discovery and drug-centric research to improve public health. To fully utilize these regulatory big data, powerful tools and databases with flexible functions are crucial. FDALabel is one such tool, developed by the FDA to specifically support regulatory science. The tool fills the gap in the FDA where large amounts of the drug information are available, but few tools are available to take advantage of it.

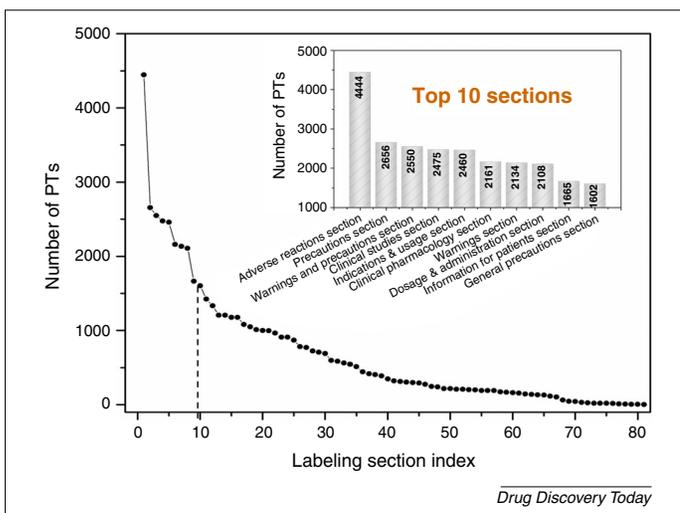


FIGURE 2

Distribution of MedDRA preferred terms (PT) in different labeling sections. The top ten sections are listed.

The drug-labeling resource is not the only one related to drugs within the FDA. Furthermore, many drug-centric databases and data standards have also been developed by the research community. However, these databases, including those developed by the FDA, are often not easily available to inform the FDA review process and drug safety research. We intend to address this challenge by expanding FDALabel by integrating it with multiple disparate database contents to provide comprehensive access to drug-related information. The expanded FDALabel will make the data accessible in a way that is useful and focused on the question asked by reviewers and researchers to discover knowledge and fill the knowledge gap. At the time of writing, the information and databases currently being evaluated for integration were: (i) Drugs@FDA, which provides drug approval history; (ii) FDA Orange Book, containing publications for approved drug products with Therapeutic Equivalence Evaluations, which provides patent related information; (iii) 'Pharmacological Class Indexing' from SPLs, which will enable searches for ADRs across drug class products. Drugs from the same pharmacological class often share similar efficacy and safety profiles; (iv) MedDRA, which provides standard terminology for international clinical results used by regulatory authorities in the pharmaceutical industry; (5) RxNorm, which provides normalized names for clinical drugs and links their names to many of the drug vocabularies and databases commonly used; (vi) FAERS, which the Adverse Event Reporting System of the FDA for drug products; and (vii) Substance Registration System-Unique Ingredient Identifier (SRS-UNII), which provides unique chemical substance information and structure for drugs. Furthermore, we will implement more options for flexible access to this integrated information, such as searches for drug class, chemical structure, topic, and so on. Our ultimate goal is to provide publicly available, rich, accurate, and complete information that facilitates transparent knowledge exchange among the public, pharmaceutical companies, and government regulatory agencies.

Disclaimer

FDALabel database is not compatible with Internet Explorer. We suggest that users use Firefox or Google Chrome as Internet browsers. FDA-Label is not a diagnostic tool and is not intended

to inform regarding choice of medicines or therapies for medical conditions. The views presented in this article do not necessarily reflect current or future opinion or policy of the U.S. Food and Drug Administration. Any mention of commercial products is for clarification and not intended as endorsement.

Acknowledgments

The project was supported by FDA internal funding. G.Z. is grateful to the National Center for Toxicological Research of the FDA for post-doctoral support through the Oak Ridge Institute for Science and Education. The authors would like to thank many FDA colleagues for their support and assistance. Particularly, we thank Crystal Allard, Helena Sviglin, and Eileen Navarro Almarino for collecting user feedback; Farrokh Sohrabi and Eric Brodsky for comprehensive comments to improve the FDALabel home webpage and tool; and Lonnie Smith for providing valuable information regarding SPL data content. We thank Qiang Shi, Minjun Chen, and Robin Stingley for their constructive review of the manuscript and helpful discussion to improve the FDALabel. We appreciate the support from the NCTR IT Team, FDA IT Scientific Enclave Team, and NCTR Web Development Team.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.drudis.2016.06.006>.

References

- 1 FDA (2006) *Promoting Safe and Effective Drugs for 100 Years*. FDA
- 2 Lesko, L.J. and Atkinson, A.J. (2001) Use of biomarkers and surrogate endpoints in drug development and regulatory decision making: criteria, validation, strategies. *Annu. Rev. Pharmacol. Toxicol.* 41, 347–366
- 3 Roberts, R. et al. (2003) Pediatric drug labeling: improving the safety and efficacy of pediatric therapies. *JAMA* 290, 905–911
- 4 Woosley, R.L. (2013) One hundred years of drug regulation: where do we go from here? *Annu. Rev. Pharmacol. Toxicol.* 53, 255–273
- 5 Bisgin, H. et al. (2012) Investigating drug repositioning opportunities in FDA drug labels through topic modeling. *BMC Bioinformatics* 13, 1–9
- 6 Kuhn, M. et al. (2010) A side effect resource to capture phenotypic effects of drugs. *Mol. Syst. Biol.* 6, 343
- 7 Ehmann, F. et al. (2015) Pharmacogenomic information in drug labels: European Medicines Agency perspective. *Pharmacogenomics J.* 15, 201–210
- 8 Frueh, F.W. et al. (2008) Pharmacogenomic biomarker information in drug labels approved by the United

- States Food and Drug Administration: prevalence of related drug use. *Pharmacotherapy* 28, 992–998
- 9 Wei, C.-Y. et al. (2012) Pharmacogenomics of adverse drug reactions: implementing personalized medicine. *Human Mol. Genet.* 21, R58–R65
- 10 Basch, E. (2010) The missing voice of patients in drug-safety reporting. *N. Engl. J. Med.* 362, 865–869
- 11 Mullard, A. (2013) 2012 FDA drug approvals. *Nat. Rev. Drug Discov.* 12, 87–90
- 12 Seminerio, M.J. and Ratain, M.J. (2013) Are drug labels static or dynamic? *Clin. Pharmacol. Ther.* 94, 302–304
- 13 Wang, B. et al. (2014) Clinical evidence supporting pharmacogenomic biomarker testing provided in US Food and Drug Administration drug labels. *JAMA Intern. Med.* 174, 1938–1944
- 14 FDA (2013) *Guidance for Industry on Clinical Pharmacogenomics Information Preparation for Labeling*. FDA
- 15 Schadow, G. (2005) HL7 structured product labeling-electronic prescribing information for provider order entry decision support. *AMIA Annu. Symp. Proc.* 2005, 110
- 16 Schadow, G. (2007) Assessing the impact of HL7/FDA Structured Product Label (SPL) content for medication knowledge management. *AMIA Annu. Symp. Proc.* 2007, 646
- 17 Chen, M. et al. (2011) FDA-approved drug labeling for the study of drug-induced liver injury. *Drug Discov. Today* 16, 697–703
- 18 Mallal, S. et al. (2008) HLA-B*5701 screening for hypersensitivity to abacavir. *N. Engl. J. Med.* 358, 568–579
- 19 Lasser, K.E. et al. (2002) Timing of new black box warnings and withdrawals for prescription medications. *JAMA* 287, 2215–2220
- 20 Chen, M. et al. (2013) Quantitative structure–activity relationship models for predicting drug-induced liver injury based on FDA-approved drug labeling annotation and using a large collection of drugs. *Toxicol. Sci.* 136, 242–249
- 21 Liu, Z. et al. (2011) Translating clinical findings into knowledge in drug safety evaluation – Drug Induced Liver Injury Prediction System (DILLps). *PLoS Comput. Biol.* 7, e1002310

Hong Fang¹
Stephen C. Harris¹
Zhichao Liu¹
Guangxu Zhou¹
Guoping Zhang¹
Joshua Xu¹
Lilliam Rosario²
Paul C. Howard¹
Weida Tong^{1,*}

¹National Center for Toxicological Research, US Food and Drug Administration, 3900 NCTR Road, Jefferson, AR 72079, USA

²Office of Computational Science, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD 20903, USA

*Corresponding author.