DILIrank: the largest reference drug list ranked by the risk for developing drug-induced liver injury in humans

Minjun Chen¹, Ayako Suzuki², Shraddha Thakkar¹, Ke Yu¹, Chuchu Hu¹ and Weida Tong¹

¹Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, US Food and Drug Administration, Jefferson, AR, USA
²Department of Medicine, University of Arkansas for Medical Sciences, Little Rock, AR, USA

Introduction

Rapid advancement in high-throughput technologies, such as high-content assay/high-throughput screening methodologies, enables the assessment of cellular responses to hundreds of drugs in a single experiment. In addition, new approaches utilizing in vitro models (e.g., 3D cell culture), ‘omics’, organ-on-a-chip, and pluripotent stem cells (iPSC) have been introduced to evaluate drug safety and efficacy through anchoring phenotypes observed in humans [1,2]. In this era of data-driven science, analysis anchored in phenotype depends on accurate and consistent annotation for a large number of drugs in a comparative analysis to derive reliable and robust emerging biomarkers.

Drug-induced liver injury (DILI) is one of the major safety concerns for drug developers, regulators, and clinicians [3,4]. Current in vivo toxicological studies are not sufficient in assessing the hepatotoxic potential of a drug in humans, as suggested by a large survey conducted by collaborative efforts [5]. Thus, active research is conducted for developing new tools and approaches to better predict DILI risk in humans [6]. A reference drug list with a sufficient number of drugs that are well annotated based on their DILI risk in humans [7] is required for enhanced methodological developments in DILI risk assessment.

The DILI annotation addressed here refers to the classification of drugs based on DILI risk for humans treated for various diseases. Establishing a DILI annotation, which reflects the frequency, causality and severity of DILI [8] for each drug, is challenging. Given the diversity of clinical manifestations of DILI, the uncertainty in causality assessment, severe underreporting of DILI cases, and the uncommon occurrence of DILI, it is difficult to identify a single resource that could provide all the information required for an accurate DILI annotation [8]. The approaches to annotate DILI risk are categorized as case report based, drug compendium based, or mixed [7,9–11]. Case reports can be retrieved from literature [12–14] or the FDA adverse event reporting system (FAERS) [15–17]. Information summarized in drug compendiums, such as Physicians’ Desk Reference [18], can also be utilized to categorize the DILI risk.

Given the lack of a ‘gold standard’ that defines DILI risk, the accuracy of the DILI annotation is difficult to evaluate, and the variability of the published annotations in terms of different schema and data sources used is of concern [7]. The FDA-approved drug labeling is the authoritative document that summarizes drug safety information used here to define DILI risk based on a systematic assessment of data from preclinical toxicological data, clinical trials, postmarketing surveillance, and the literature. Information gathered from the drug labeling might not be the perfect data source for DILI definitions, but perhaps it is ‘the closest one that can get to the truth’ [19]. We previously developed a DILI annotation schema based on the information gathered from the FDA-approved drug labeling and generated a benchmark drug list that contained 287 drugs and were grouped into three levels of DILI severity [11]. Our annotations and data set were recommended as the standardized list for model validation [20] and have been widely adopted by the research community to develop DILI predictive methodologies based on in silico [21–23], in vitro [13,24–30], and in vivo assays [31–34]. However, the previous DILI annotated drug list does not contain a sufficient number of drugs. A few models have been developed based on an expanded DILI annotation by applying our schema to a larger number of drugs [13,24].

There are reservations over the usage of drug-labeling information for defining the degree of DILI risk in humans [8]. One of the major concerns is its weakness in causality assessment. By law, regulators can issue a warning if a clinically significant hazard is identified for a drug with reasonable evidence of causality; however, a definite causal relation is not mandatorily required for drug labeling (http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfCFR/CFRSearch.cfm?fr=201.57). Moreover, the lengthy process
required for drug-labeling modification likely causes a time lag in updating the safety information in drug labeling and the most updated findings from the clinicians’ view point might not be incorporated in drug labeling [35]. Including the information derived from the up-to-date literature and publicly available safety knowledges is necessary to further improve the drug labeling-based DILI annotation.

In this study, we developed a refined annotation schema by weighing evidence of causality to overcome inherent deficits in drug labeling and improve the accuracy of DILI annotation. More specifically, the refined annotation schema was built upon the collection of adjudicated cases [verified using the standardized clinical causality assessment system; that is, Roussel Uclaf Causality Assessment Method (RUCAM)] and well-vetted cases (verified via thorough case evaluation by DILI experts) [36]. We considered both ‘adjudicated’ and ‘well-vetted’ as ‘verified’ with causality assessment, and use the unified terms ‘verified’ and/or ‘verification’ here to describe the verification processes. We developed a large database of drugs, namely a DILIrank data set, including 1036 marketed drugs approved by the FDA before 2010, which were annotated with the ranked DILI risk in humans by using the newly developed schema. Using this new reference drug list, we analyzed the landscape of DILI risk in therapeutic classes associated with FDA-approved drugs. Of note, this work solely focused on drug-induced hepatotoxicity and the toxicities to other organ systems were not considered in the schema, although they can be developed in a similar fashion.

**Verification process for DILI annotation**

Our previous schema to define DILI risk in humans was based on information gathered from drug labeling [8]. Here, we modified our original schema by incorporating information about whether the drugs were verified for their causality of DILI in humans, using publicly available resources as detailed below. We included the drugs approved by the FDA before the year of 2010 to expand our drug list. The developed schema is depicted in Fig. 1.

The verified drugs that cause DILI in humans were mainly collected from large DILI registries [14,37] and from an authoritative public resource (i.e., the NIH LiverTox database [38]) to warrant the data quality. In a previous international collaborative study, Suzuki et al. [14] reported 225 US marketed drugs that were verified for a cause of DILI in the accumulated cases from three major DILI registries: Spanish DILI Registry, Swedish Adverse Drug Reactions Advisory Committee Database, and the Drug-Induced Liver Injury Network (DILIN) in the USA. In this work, the RUCAM scoring system and/or expert opinions were considered as causality assessment, and only cases classified as ‘definite’, ‘highly probable’, ‘probable’, or ‘possible’ using the RUCAM score were used to

---

**FIGURE 1**

The schema to refine the previous drug-induced liver injury (DILI) annotation based on drug-labeling data by building in causality evidence. The new schema classifies drugs into four categories, with three categories verified ("Most-", "Less-", and "No-DILI concern") and leaving out one with DILI concern but without verified causality ("Ambiguous DILI concern"). * The verification for DILI cause is not required for withdrawn or boxed warning drugs.
verify the investigated drug as a cause of DILI. Chalasani et al. [37] reported 163 drugs responsible for over 800 DILI cases, all of which were verified by the DILIN causality committee using the DILIN causality categories of ‘definite’, ‘very likely’, and ‘probable’. In addition, 344 drugs were identified as the cause of DILI in the LiverTox databases in which case reports collected from literature were verified by the LiverTox experts [39]. Altogether, a total of 399 unique drugs with a single active ingredient marketed in the USA were identified from the three main data sources (i.e., Suzuki et al., DILIN, and LiverTox database) (Fig. 2). A literature search was conducted to identify additional drugs that had been verified as a cause for DILI (see Table S1 in the supplemental information online).

With the causality information mentioned above, a modified schema (Fig. 1) was used to determine the degree of DILI risk of individual drugs by complementing the drug labeling with the available evidence of verified causality. Specifically, (i) withdrawn and boxed warning warnings because DILI risk was assigned as most-DILI concern in a previous study [8]. A recent study [7] showed that withdrawn drugs and those with box warnings for severe liver injury were consistently classified as high DILI risk among several published data sets. Therefore, drugs withdrawn or with box warnings were considered as verified Most-DILI concern (’Most-DILI concern) drugs in the new schema; (ii) in the previous schema, the Most-DILI concern category included drugs that had warnings and precautions in their labels as well with descriptions suggesting severe or moderate DILI occurrence (i.e., diclofenac). As per the definition, warnings and precautions reflect safety concerns from the regulators’ viewpoints and do not necessarily require strong evidence of causality [8]. In the new schema, these drugs were further assessed using the verification process of causality; only drugs that had been verified as causal drugs were classified as the ‘Most-DILI concern. Consequently, some drugs (e.g., anidulafungin) that were previously classified as most-DILI concern but did not have the evidenced causality were reclassified as ‘Ambiguous DILI concern’; (iii) similarly, the Less-DILI concern drugs in the previous schema were reclassified as verified Less-DILI concern (‘Less-DILI concern) or ‘Ambiguous DILI concern’, depending on whether evidence of causality was available, as shown in Fig. 1; (iv) in the new schema, if a No-DILI concern drug was verified as a cause of DILI in literature, it was reclassified into the ‘Less-DILI concern group; otherwise, it was assigned as a verified No-DILI concern drug (‘No-DILI concern).

It is important to point out that the verification process is based on the reported data (i.e., existing knowledge). Given that the existing knowledge will advance over the time, we plan to continuously update the DILI annotation in the DILIRank data set and will make the updated version publicly available.

Generation of the DILIRank data set with a verification process
Six criteria were applied to filter the FDA-approved drugs for annotations: (i) has FDA-approved labeling data; (ii) is for human use; (iii) contains a single active molecule in the dosage form; (iv) excludes food supplements, minerals, or cosmetics; (v) systemic administration through oral or parental use; (vi) has been approved before January 1, 2010. In this work, minor modifications were made to the previous criteria [8] by changing the time period in criterion (vi) from 10 years to 5 years to maximize the coverage of drugs while minimizing the changes of drug labels related to postmarketing safety experiences. Then, a total of 1036 FDA-approved unique drugs with a single active molecule for human use were collected from the DailyMed database (http://dailymed.nlm.nih.gov/, as of September 1, 2015). By using our previous schema, the 1036 FDA-approved drugs were classified into 209 Most-, 488 Less-, and 339 No-DILI concern drugs. All of these drugs and their related information (e.g., DILI classification, verified causality, and FDA approved date) are summarized in Table S1 in the supplemental material online or on our Liver Toxicity Knowledge Base website (http://www.fda.gov/ScienceResearch/BioinformaticsTools/LiverToxicityKnowledgeBase/ucm2024036.htm).

Next, we applied the verification process to the 1036 drugs classified by our previous schema, yielding 192 ‘Most-DILI concern, 278 ‘Less-DILI concern, and 312 ‘No-DILI concern drugs, all of which were verified by the evidenced causality; and leaving out 254 drugs as ‘Ambiguous DILI concern’ (Table 1).

To compare DILI risk annotation among different methodologies, four published studies were selected with the criteria of having: (i) a large number of drugs (N > 200); and (ii) drug annotation, including both DILI positives and negatives (Table 1). Among them, Xu et al. [11] developed a DILI classification system based on the combination of drug labeling and the frequency of case reports. Greene et al. [12] applied a similar schema to that of Xu et al. [11], but reclassified some drugs annotated as negatives into a new group with weak evidence of hepatotoxicity. Sakatis et al. [18] classified drugs as hepatotoxic or nonhepatotoxic based on the number of case reports documented in Physician’s Desk Reference. Zhu et al. [17] collected positives from a previous publication by Suzuki et al. [14] and defined negatives that lacked confirmative reports or references in PubMed or FDA MedWatch.

The most significant change in the new schema was the division of 488 Less-DILI concern drugs, which was an ambiguous annotation in the previous schema, into 251 DILI concern drugs with
verified causality and 237 ‘Ambiguous DILI concern’ drugs by weighing the evidenced causality. Notably, 139 of 251 ‘Less-DILI concern drugs were found labeled as DILI positives by at least one of the four published data sets, while only 39 of 237 ‘Ambiguous DILI concern’ drugs were labeled as DILI positive by any of these four data sets (Figure S1 in the supplemental information online). In other words, the likelihood of being classified as ‘hepatotoxic’ in the other four studies that applied different schema was approximately six times higher for drugs in the ‘Less-DILI concern category versus the ‘Ambiguous DILI concern’ drugs (odds ratio: 6.2, 95% confident interval: 4.0–9.7, \( P < 0.0001 \)). Given the comprehensive information collected DILI information from the drug labeling data, the verified causality, and the consistency among the annotations of DILI positives within the literature, the ‘Less-DILI concern can be considered as an intermediate category between the ‘Most-DILI concern and the ‘No-DILI concern. Therefore, DILLrank contains three-level rank in the DILI risk scale. The ‘Ambiguous DILI concern’ drugs are considered as an ambiguous annotation group, pending further characterization.

Another improvement in the new schema is the reclassification of some of the no-DILI concern drugs into ‘Less-DILI concern after the hepatotoxic potential of the drug was verified (Table S2 in the supplemental information online). Defining DILI negatives is more challenging compared with defining DILI positives; the criteria to define the true DILI negatives are not yet established. It is important that drugs with strong DILI evidence (e.g., verified as a cause of DILI in literature) are not assigned as DILI negatives, as is seen in some studies [10]. By analyzing the four selected data sets (Table 1), the percentage of drugs verified by the evidenced causality for the DILI negatives was 7.9–41.8%. Furthermore, among 50 drugs that were annotated as in a category of ‘weak evidence’ human hepatotoxicity by Greene et al. [12] but assigned as negatives by Xu et al. [11], 33 (66%) were associated with evidenced hepatotoxicity in the literature (Table S3). These high percentages of misclassification highlight the importance of modifying annotation based on updated information of human DILI risk.

The new schema also reclassified a small portion (\( N = 17 \)) of Most-DILI concern drugs annotated in the previous schema as ‘Ambiguous DILI concern’ drugs (Table S2 in the supplemental information online). Notably, ten of the 17 drugs were approved after the year 2000. Thus, evidenced causality might be not available yet because of their relatively short marketing [40]. Often, establishing evidenced hepatotoxicity takes a long time. An extreme example is propylthiouracil, a drug for which a boxed warning for hepatotoxicity was issued after over 60 years of clinical use [41]. Annotations for a few drugs with a longer marketing period were also inconsistent between the two schemas, previous versus new. For example, isocarboxizid is a monoamine oxidase inhibitor that was approved in 1959 and was previously classified as Most-DILI concern but reclassified as ‘Ambiguous DILI concern’ drugs in the new schema. This might be explained by a single case report of clinically significant DILI (accompanied by jaundice) shortly after drug approval, which was preceded by the withdrawal of iproniazid, another monoamine oxidase inhibitor [42]. Despite the long marketing period, isocarboxizid has not been acknowledged for hepatotoxicity in recent reviews [43,44] or identified in major DILI registries (e.g., DILIN [37] or Spain DILI registry [14,45]), and no convincing DILI cases were found in literature.

### Analysis of the DILLrank data set

The DILI landscape using the DILLrank data set was analyzed based on the therapeutic categories (2nd level) and chemical subgroup (4th level) as defined by the World Health Organization (WHO) Anatomical Therapeutic Chemical (ATC) classification system. This analysis focused on identification of the drug subgroups enriched with a higher hepatotoxic risk (Table 2). The enrichment analysis was conducted by comparing the prevalence of ‘Most-DILI concern drugs among the investigated subgroup versus the DILLrank data set through a Fisher exact test.

At the therapeutic category level, seven therapeutic classes were found to be significantly enriched with ‘Most-DILI concern drugs, including nonsteroidal anti-inflammatory drugs (NSAIDs), antigout, antimycotics, antineoplastic, psychoanaleptics, immunostimulants, and antivirals. Notably, these therapeutic categories are well known to be associated with hepatotoxicity, and the profile noted above is in line with reports in literature [37,45]. Interes-
Intriguingly, although many of the antibiotics were assigned as high-DILI concern drugs (N = 9), as a class they do not present significantly higher hepatotoxic risk because as many as N = 67 antibiotic drugs were not classified as the high-DILI concern drugs.

We also identified the chemical subgroups enriched with higher DILI risk, including interferons, protein kinase inhibitors, acetic acid derivative NSAIDs, macrolide antibiotics, triazole antifungal, and nucleoside/nucleotide reverse transcriptase inhibitors (NRTIs), as shown in Table 2. These chemical subgroups might implicate potential DILI mechanisms involved; for example, NRTIs are well known as a drug class to cause mitochondrial toxicity, which could lead to liver failure and lactic acidosis [46].

### Discussion

In this era of data-driven science, a high-quality annotation is crucial to develop new tools and approaches for discovering DILI biomarkers by utilizing high-throughput technologies and in silico methodologies. Given that DILI is of concern for both the drug development industry and regulatory agencies, we made an effort to refine our previously reported annotation schema to improve DILI risk rank. In this new schema, the previous drug labeling-based approach was improved by weighing the verified causality. Applying this schema to the FDA-approved drugs, the DILIrank data set was generated with 1036 drugs annotated for DILI risk and classified into three verified DILI ranks (i.e., high-, low-, and no-DILI concern), leaving out drugs with DILI concern but without verified causality (ambiguous annotation). To our knowledge, it is the largest annotated DILI data set in public.

The new annotation is relevant to regulatory professionals in their DILI risk assessment. For example, a drug classified as high-DILI concern (e.g., diclofenac) was considered to have the potential to cause severe clinical outcomes, and those of low-DILI concern (e.g., heparin) can cause liver injury but rarely lead to severe outcomes. The no-DILI concern drug (e.g., phenoxybenzamine) reflects the low risk perceived and the associated liver injury must be rare or nonexistent. The DILIrank data set with improved annotation will contribute to the development of predictive models that use emerging technologies (i.e., high-throughput screening or high-content assay) and in silico methods, such as structure–activity relationships (QSARs), for the early identification of DILI risk liability during drug development. Moreover, comprehensive annotation of the FDA-approved drug list could support clinical and epidemiological investigations in DILI and aid in the clinical assessment of suspected DILI cases.

### Disclaimer

The opinions expressed by the authors do not reflect the opinions or policies of their respective institutions. Any statements in this article should not be considered a present or future policy of any regulatory agency.

### Appendix A. Supplementary data

upplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.drudis.2016.02.015.

---

**References**

10 Weng, Z. et al. (2015) A comprehensive study of the association between drug hepatotoxicity and daily dose, liver metabolism, and lipophilicity using 975 oral medications. Oncotarget 6, 17031–17038
13 Gustafsson, F. et al. (2013) A correlation between the in vitro drug toxicity of drugs to cell lines which express human P450s and their propensity to cause liver injury in humans. Toxicol. Sci. 137, 189–211
24 Aloo, M.D. et al. (2014) Human drug-induced liver injury severity is highly associated to dual inhibition of liver mitochondrial function and bile salt export pump. Hepatology 60, 1015–1022
40 Lasser, K.E. et al. (2002) Timing of new black box warnings and withdrawals for prescription medications. JAMA 287, 2215–2220