

Title Page

A Novel Computer-Assisted Drug-Induced Liver Injury Causality Assessment Tool (DILI-CAT)

Hans L. Tillmann, MD^{1,2}, Ayako Suzuki, MD^{3,4}, Michael Merz, MD⁵, Richard Hermann,
MD⁶, Don C. Rockey, MD⁷

¹ Division of Gastroenterology, Hepatology & Nutrition, East Carolina University, 600 Moyer Blvd., Greenville, NC 27834, USA, tillmannh14@ecu.edu

² Greenville VA Health Care Center, 401 Moyer Blvd., Greenville, NC 27834, USA

³ Duke University Medical Center, 40 Duke Medicine Circle, Durham, NC 27710, USA, Ayako.suzuki@duke.edu

⁴ Durham VA Medical Center, 508 Fulton Street, Durham, NC 27705, USA

⁵ AstraZeneca, independent consultant, Zum Roten Stein 12, 79108 Freiburg, Germany, michael.merz@t-online.de

⁶ AstraZeneca, 1 MedImmune Way, Gaithersburg, MD 20878, USA, richard.hermann@astrazeneca.com

⁷ Digestive Disease Research Center, Medical University South Carolina, 96 Jonathan Lucas, CSB, Suite 908, Charleston, SC 29425, USA, Rockey@MUSC.edu

Corresponding author:

Hans L. Tillmann, MD

Division of Gastroenterology, Hepatology & Nutrition

East Carolina University

600 Moye Blvd.

Greenville, NC 27834, USA

Telephone: 252-830-2149, ext 143162

E-mail: tillmannh14@ecu.edu

Manuscript word count (max 3500): 2420

Key words: hepatotoxicity, hepatitis, cyproterone, amoxicillin/clavulanate, augmentin, cefazolin, isoniazid

Abstract

Objective: We hypothesized that a drug-characteristic DILI-phenotype could be defined and be used to develop a computer-assisted DILI causality assessment-tool (DILI-CAT).

Design: A drug-specific DILI-phenotype was developed for amoxicillin-clavulanate (AMX/CLA), cefazolin, cyproterone, and polygonum multiflorum using data from published case series, and subsequently a DILI-CAT Score (DILI-CAT-S) was created for each drug. The phenotype was made up of the following three parameters: (1) latency, (2) R-value, and (3) AST/ALT ratio (also de Ritis ratio). A point allocation system was developed with points allocated depending on the degree of deviation from the core of published data for the three phenotypic parameters.

Results: The four drugs had a significantly different phenotype based on the three parameters utilized. For example, the median cyproterone latency was 150 days versus less than 43 days for the other three drugs (median: 26 for AMX/CLA, 20 for cefazolin, and 20 days for polygonum multiflorum; $p < 0.001$). The R-value for the four drugs was also significantly different (median: cyproterone [12.4] and polygonum multiflorum [10.9]) from AMX/CLA (1.44) and cefazolin (1.57; $p < 0.001$). The resulting DILI-CAT-S effectively separated cyproterone and polygonum multiflorum from AMX/CLA and cefazolin, respectively ($p < 0.001$). Notably, because of overlap in phenotype AMX/CLA and cefazolin could not be differentiated by DILI-CAT-S.

Conclusion: DILI-CAT is a data-driven, diagnostic tool built to define drug-specific phenotypes. Data presented here provide proof of principle that a drug-specific, data-driven causality assessment tool can be developed for different drugs and raise the possibility that such a process could improve and standardize causality assessment methods.

Funding: DCR was supported by the NIH, grant P30 DK 123704

Introduction

Drug-induced liver injury (DILI) is an important cause of acute liver injury and liver-related morbidity and mortality.¹⁻⁴ DILI is also a major concern in drug development and post-marketing surveillance, as evidenced by hepatotoxicity being a leading cause for market withdrawal of licensed drugs.⁵ Moreover, DILI diagnosis is extremely challenging, as liver biochemistry abnormalities may be caused by a variety of different causes of liver injury.⁶⁻⁸

Unlike diseases such as viral hepatitis, where diagnostic testing may confirm or exclude the diagnosis with high sensitivity and specificity, DILI is a diagnosis based on specific clinical features associated with a particular drug and requires exclusion of other causes of liver diseases. A variety of causality assessment methods (CAMs) have been developed and often use point-scoring systems (ie, Roussel Uclaf Causality Assessment Method [RUCAM], “clinical diagnostic scale” [CDS]).⁹⁻¹¹ While these scoring systems have attempted to quantify the causality likelihood for DILI, a structured expert opinion process (SEOP), such as described by the Drug Induced Liver Injury Network (DILIN), has been shown to be superior to RUCAM.¹² We have previously shown that different drugs have different clinical DILI characteristics or phenotypes.¹³ We believe that one of the reasons why SEOP is superior to RUCAM is that experts recognize specific clinical phenotypes (ie, its “signature” or typical characteristics) for different drugs that cause DILI. Unfortunately, a major limitation of the expert opinion approach is that it is not widely available in clinical practice and is thus not generalizable. It is likely that one

reason why the expert opinion approach is successful is that experts have great experience with DILI and are often familiar with drug-specific DILI phenotypes.

Here, we hypothesized that not only are there clinical DILI features that are typical for each drug and make up a typical phenotype, but also that such drug-specific DILI phenotypes could be used to develop a novel DILI causality assessment tool (DILI-CAT), incorporating data-driven drug-specific DILI phenotypes. We aimed to create a quantitative data-driven algorithm (DILI-CAT) to define drug-specific DILI phenotypes using characteristic DILI features. These include (1) latency, (2) R-value (the ratio of alanine aminotransferase [ALT] to the upper limit of normal for [ALT] / alkaline phosphatase [ALP] to the upper limit of normal for ALP) and (3) the aspartate aminotransferase [AST] / ALT ratio [de Ritis ratio] derived from available literature that included patient-level data. We also sought to develop drug-specific scoring systems to quantify the resemblance of acute liver injury events to the defined DILI phenotypes.

Methods

We performed a literature search to identify published case-series studies reporting clinical features in patients with DILI caused by single specific drugs. Although most studies that were identified failed to report patient-level data necessary to develop a robust DILI phenotype, we identified four case series that fulfilled the requirement of having detailed patient-level data for latency, R-value, and AST/ALT ratio (de Ritis) at onset. These included one study for each of the following four drugs: Cyproterone (n=22),¹⁴ amoxicillin-clavulanate (n=35),¹⁵ cefazolin (n=19),¹⁶ and polygonum multiflorum (n=18).¹⁷

Design

We considered the drug-specific DILI phenotype to be made up primarily of its latency (in days), absolute R-value, and AST/ALT ratio. A quantitative “drug-specific” scoring system was then developed that allocates points based on these three specific DILI features. For DILI-CAT scoring, a drug-specific DILI-CAT Score (DILI-CAT-S) would be developed based on the distribution of the respective latencies, R-values, and AST/ALT ratios in identified case series. A separate DILI-CAT-S was developed for each drug. Furthermore, depending on the difference in phenotype being more marked in latency, R-value, or AST/ALT ratio, the respective DILI-CAT-S was weighted in the respective category, resulting in latency-weighted DILI-CAT-S, R-value-weighted DILI-CAT-S, or AST/ALT ratio-weighted DILI-CAT-S. The category of the three (latency, R-value, and AST/ALT ratio) in which the greatest differences between two

drugs in non-parametric comparison were detected received two-fold greater weight compared to the remaining two categories.

We hypothesized that the closer a drug's DILI features (ie, latency, R-value, and AST/ALT ratio) are to the published clinical characteristics, the more likely the case is to be a bona fide DILI case. In other words, the closer a case's values are to the interquartile range (IQR) of values in published DILI cases for that drug, the more likely that injury is related to the drug in question.

In the model, points were allocated based on the closeness of the parameter in question (latency, R-value, AST/ALT ratio) for each specific drug to the IQR, or 50% core interval (Figure 1, Supplemental Figure 1a & 1b) as derived from known cases (patient/case level data).¹⁴⁻¹⁷ Proportionally fewer points were allocated when values for the parameter in question fell outside the IQR. By definition, there was variability in points allocated for each drug at different latency, R-value, and AST/ALT ratio because allocated points were dependent on the respective 50% core interval (IQR) and percentiles for each drug (Table 1 and Supplemental Table 1). Any parameter falling within the core interval was allocated 20 points. Values outside the 50% core interval (IQR) were given fewer points (Table 1). Deductions were given for values outside of the range of the values for respective drug's phenotype range (Figure 1, Table 2, Supplemental Table 1). Additional deductions were also given when values were far outside the IQR; these were defined as "outliers" (Supplemental Material Appendix 1).

The strategy that was ultimately utilized to generate a scoring system required several assumptions. First, we postulated that each of the four drugs chosen (or any other drug, for that matter) would exhibit differences in one or more of the three clinical categories (latency, R-value, AST/ALT ratio) compared to at least some other drugs. We also postulated that the clinical categories that exhibited the greatest differences would be the most valuable in differentiating clinical phenotypes between drugs (ie, latency for one drug would be much longer than the others, and therefore was highly valuable in defining the clinical phenotype). For the specific category with the greatest discriminating potential (latency, R-value, AST/ALT ratio), that category's value was doubled. For example, if for a specific drug, latency exhibits the greatest statistical difference compared to R-value and AST/ALT ratio differences, then latency points were doubled.

Statistics

Each drug's phenotype was informed by the IQR, percentiles, maximum and minimum values, and definition of outlier values for each of the three clinical variables (latency, R-value, AST/ALT ratio).

A non-parametric Mann-Whitney rank test was used to compare drug phenotypes (for each of the three clinical variables latency, R-value, AST/ALT ratio) to each other. Differences in variables were defined statistically (Table 3). The smaller the Mann-Whitney value, the greater the difference, and a Mann-Whitney U value of "0" reflects complete separation of parameters between groups (ie, the Mann-Whitney number

comparing latency for cyproterone and cefazolin was zero, reflecting that all latencies for cyproterone were longer than any cefazolin latencies).

For each drug, a drug-specific DILI-CAT-S was developed using a scoring algorithm. To compare drug-specific DILI-CAT-S' performance, each drug was evaluated using its respective DILI-CAT-S against the three other drugs, where the significance of difference was assessed using the Mantel-Haenszel test for trend considering five-point incremental scores as ordinal categories. Data handling was done using Microsoft[®] Excel[®], and IBM[®] SPSS[®] version 25 was used for statistical analysis.

Role of the funding source

There was no funding for this study. The corresponding author of this manuscript certifies that the contributors' and conflicts of interest statements included in this paper are correct and have been approved by all co-authors.

Results

The typical latency for cyproterone was considerably longer (median 150 days, IQR 114-240 days) than that for the other three drugs (which ranged from median 20 to median 26 days; IQR 16,25-24 days to 7-44 days; Table 2, Figure 1, Supplemental Figure 1). The R-values also varied but at the same time were similar for some of the drugs, with median values of 10.9 (IQR 7.6-13.5) and 12.4 (IQR 9.83-17.84) for polygonum multiflorum and cyproterone, respectively, and 1.4 (IQR 0.74-2.92) and 1.6 (IQR 1.07-2.87) for AMX-

CLA and cefazolin, respectively (Table 2). The AST/ALT ratios largely overlapped among all these four drugs (from 0.52 to 0.78).

Phenotypic differences among drugs

The DILI-CAT-S utilizes weighting of individual phenotypic features (latency, R-value, and AST/ALT ratio, as described in the Methods).

Cyproterone showed the greatest difference in latency compared to the other three drugs ($p < 0.001$, Table 3). While polygonum multiflorum differed from cyproterone most strongly in terms of latency ($p < 0.001$, Table 3) and to lesser extent in AST/ALT ratio ($p = 0.008$, Table 3), polygonum multiflorum differed from AMX-CLA and cefazolin significantly only in R-value ($p < 0.001$, Table 3). AMX-CLA and cefazolin differed only mildly in the AST/ALT ratio from each other ($p = 0.038$, Table 3). Based on the respective greatest difference, as defined by lowest U-value (Table 3), the following weighting was applied:

- For cyproterone a latency-weighted (thus latency valued double) DILI-CAT was applied for comparison against all other three drugs.
- For polygonum multiflorum, a latency-weighted DILI-CAT was applied for comparison against cyproterone, but an R-value-weighted DILI-CAT was applied for comparison against both AMX-CLA and cefazolin.
- For both AMX-CLA and cefazolin, a latency-weighted DILI-CAT was applied against cyproterone, an R-value-weighted DILI-CAT against MP, and, finally,

AMX-CLA and cefazoline were compared using an AST/ALT ratio-weighted DILI-CAT.

Cyproterone-derived DILI-CAT

In order to create a DILI-CAT score, as outlined in the Methods, points were allocated based on latency, R-value, and AST/ALT ratios. As predetermined, a latency-weighted DILI-CAT would be applied that shows significant difference against all three other drugs (Table 4, Supplemental Figure 2a).

AMX/CLA-derived DILI-CAT

AMX/CLA had a more complex clinical pattern; the most significant differences found in the three clinical categories depended on the comparator drug. For example, latency was most distinct from cyproterone (Table 3), R-value most different from polygonum multiflorum (Table 3), and AST/ALT ratio most distinct from cefazolin (Table 3). Latency weighting provided the greatest discrimination between AMX-CLA and cyproterone, while an R-value-weighted DILI-CAT provided the greatest discrimination between AMX-CLA and polygonum multiflorum (Table 4). In contrast, AST/ALT ratio weighting did not provide added discriminatory value (Table 4, Supplemental Figure 2b).

Cefazolin-derived DILI-CAT performance

Cefazolin was similar to AMX/CLA in all three clinical categories. It had the greatest difference in latency compared to cyproterone and was most distinct in R-value compared

to polygonum multiflorum. Its AST/ALT ratio was most distinct from AMX/CLA (though minimally different [$p = 0.038$; Table 3]).

While the AMX/CLA-derived DILI CAT did not separate AMX/CLA from cefazolin, a cefazolin-derived DILI-CAT was able to better separate cefazolin from AMX/CLA ($p=0.008$; Table 4, Supplemental Figure 2c). This is likely because the cefazolin phenotype shows less variation in latency, R-value and AST/ALT ratio compared to the AMX/CLA phenotype, where more AMX/CLA cases overlap with cefazolin's phenotype than vice versa (Table 2).

Polygonum multiflorum-derived DILI-CAT performance

Polygonum multiflorum was most different from cyproterone in the latency category, while the latency of polygonum multiflorum and either AMX/CLA or cefazolin was similar (Table 3). Polygonum multiflorum's R-value was distinct from AMX/CLA and cefazolin (Table 3). Therefore, a latency-weighted polygonum multiflorum-derived DILI-CAT shows the greatest separation from cyproterone (Table 4), while an R-value-weighted DILI-CAT shows the greatest separation against AMX/CLA as well as cefazolin (Table 4, Supplemental Figure 2d).

Discussion

Here, we have developed a data-driven approach (DILI-CAT) that can be used to create a drug-specific DILI phenotype and enhance DILI causality assessment. We demonstrate

that drug phenotypes differ significantly and that our algorithmic approach allows for differentiation of DILI caused by different drugs.

RUCAM, the commonly used causality assessment method (tool) developed almost three decades ago,⁹ is often considered the most reliable approach to DILI causality assessment when an expert opinion assessment is not available.¹² However, neither RUCAM nor any of the other currently available causality assessment tools uses a drug-specific approach.¹⁸ An SEOP is considered superior to RUCAM,^{1,2} which is likely due to the fact that experts probably consider a drug's phenotype even when there is no mechanism to quantitate the influence of a drug-specific phenotype. Implicit in the findings presented here is that allowing in a formal process for inclusion of a drug phenotype enhances the DILI adjudication process by including phenotypic characteristics of drug-specific DILI. We speculate that this will be helpful to experts and, perhaps to an even greater degree, to nonexperts.¹⁹

An algorithmic data-driven and drug-specific diagnostic tool such as DILI-CAT has a number of advantages. DILI-CAT's data-driven approach is objective. It can be optimized via weighting of specific variables, which will allow for better discrimination between different drugs. The intrinsic propensity for hepatotoxicity of a drug (ie, the likelihood or probability that a specific drug would cause liver injury) could be included in the mathematical algorithm (Supplement Appendix 2, 3, and 4; Supplement Table 2), and could be readily derived from the published literature.^{20,21} Scoring for competing causes in DILI-CATs could also be included, allowing for grading of individual drugs

along a causality scale (Supplemental Appendix 2). Further, the flexible format of DILI-CAT allows it to be programmed for use by any drug, as long as the DILI phenotype of a drug can be characterized (eg, with a sufficient number of known DILI cases to estimate percentiles of the drug-specific features). Finally, the approach should be considered to be a “living” process, meaning that additional cases could be added as more published cases become available so as to create a more robust DILI signature.

We recognize limitations of the current version of DILI-CAT. One limitation is that some drugs will have overlapping phenotypes, such as was the case with cefazolin and AMX/CLA; in this situation, DILI-CAT-s will be unable to provide a clear distinction between a drug in question and other competing drugs. Another limitation is that DILI-CAT depends on having available cases with which to develop specific drug signatures. In this regard, we speculate that the number of cases needed to develop a robust signature will depend on the consistency of the drug’s phenotype.¹⁹ The more variable the phenotype, the more cases that are likely to be required to generate a precise picture of a drug’s signature.

In the future, we envision a staggered approach to DILI causality assessment. First, the likelihood of DILI could be assessed using an algorithmic methodology such as that presented here, and secondly laboratory testing could be used for confirmation. While lymphocyte transformation test (LTT) is recommended in the Japanese DDW-J scoring for DILI,²² it is unclear whether this assay is reproducible enough to be used.²³ A novel promising approach is based on assay of blood derived monocytes that are transformed

into hepatocyte like cells.²⁴ In a number of studies, this test has shown promise as a confirmatory assay.²⁵⁻²⁷

In summary, we have presented an objective and data-driven drug-specific tool (DILI-CAT) that represents a novel and substantial step forward in DILI causality assessment. This approach is likely to be extremely useful for clinicians who are not experts in DILI causality assessment, and it also has the potential to improve expert adjudication of DILI.

Contributors

All authors - study concept and design; acquisition of data; interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content.

Hans L. Tillmann and Richard Hermann have verified the underlying data.

Hans L. Tillmann - data analysis

Richard Hermann - project administration, supervision

Don C. Rockey - literature review figures, data analysis

Declaration of interests

Hans L. Tillmann is a stockholder of Abbott, AbbVie, and Gilead outside the submitted work. He reports that his wife is a full-time employee of AbbVie.

Ayaka Suzuki has nothing to disclose.

Michael Merz has nothing to disclose.

Richard Hermann is a full-time employee and stockholder at AstraZeneca.

Don C. Rockey has nothing to disclose.

Acknowledgments

Dr. Huiman Barnhart was instrumental in developing earlier versions of this approach.

Editorial assistance, funded by AstraZeneca, was provided by Peloton Advantage, LLC, an OPEN Health company. Additional statistical support was provided by Daniel Chima, an employee of AstraZeneca.

Data Statement

The data used here from published case series with data derived from the published articles, as the patient level data were derived from published cases series. Data will also be available with publication by outreach to the corresponding author and will be shared for analyses to replicate findings, after approval of a proposal and with a signed data access agreement, at a minimum.

Patient and Public Involvement statement

Due to the nature of this research that did not involve direct patient contact, not patient and public involvement was used for this study.

References

1. Wilke RA, Lin DW, Roden DM et al. Identifying genetic risk factors for serious adverse drug reactions: current progress and challenges. *Nat Rev Drug Discov* 2007; 6: 904-16.
2. Robles-Diaz M, Lucena MI, Kaplowitz N et al. Spanish DILI Registry; SLatinDILI Network; Safer and Faster Evidence-based Translation Consortium. Use of Hy's law and a new composite algorithm to predict acute liver failure in patients with drug-induced liver injury. *Gastroenterology* 2014; 147: 109-18.e5.
3. Chalasani N, Bonkovsky HL, Fontana R et al. United States Drug Induced Liver Injury Network. Features and outcomes of 899 patients with Drug-Induced Liver Injury: The DILIN prospective study. *Gastroenterology* 2015; 148: 1340-52.e7.
4. Fontana RJ, Hayashi PH, Gu J et al. DILI. Idiosyncratic drug-induced liver injury is associated with substantial morbidity and mortality within 6 months from onset. *Gastroenterology* 2014; 147: 96-108.e4.
5. Lee WM. Drug-induced acute liver failure. *Clin Liver Dis* 2013;17:575-86.
6. Narjes H, Nehmiz G. Effect of hospitalisation on liver enzymes in healthy subjects. *Eur J Clin Pharmacol* 2000; 56: 329–33.
7. Rosenzweig P, Miget N, Brohier S. Transaminase elevation on placebo during phase I trials: prevalence and significance. *Br J Clin Pharmacol* 1999; 48: 19–23.

8. Douglas IJ, Langham J, Bhaskaran K, et al. Orlistat and the risk of acute liver injury: self controlled case series study in UK Clinical Practice Research Datalink. *BMJ* 2013; 346: f1936.
9. Danan G, Benichou C. Causality assessment of adverse reactions to drugs. I. A novel method based on the conclusions of international consensus meetings: application to drug-induced liver injuries. *J Clin Epidemiol* 1993; 46: 1323-30.
10. Maria VAJ, Victorino RMM. Development and validation of a clinical scale for the diagnosis of drug-induced hepatitis. *Hepatology* 1997; 26: 664-9.
11. Takikawa H, Takamori Y, Kumagi T et al., “Assessment of 287 Japanese cases of drug induced liver injury by the diagnostic scale of the International Consensus Meeting,” *Hepatol Res* 2003; 27: 192–5.
12. Rockey DC, Seeff LB, Rochon J et al.. US Drug-Induced Liver Injury Network. Causality assessment in drug-induced liver injury using a structured expert opinion process: comparison to the Roussel-Uclaf causality assessment method. *Hepatology* 2010; 51: 2117-26.
13. Tillmann HL, Rockey DC. Signatures in drug-induced liver injury. *Curr Opin Gastroenterol* 2020; 36: 199-205.
14. Bessone F, Lucena MI, Roma MG, et al. Cyproterone acetate induces a wide spectrum of acute liver damage including corticosteroid-responsive hepatitis: report of 22 cases. *Liver Int* 2016; 36: 302-10.

15. Hautekeete ML, Horsmans Y, Van Waeyenberge C, et al. HLA association of amoxicillin-clavulanate--induced hepatitis. *Gastroenterology* 1999; 117: 1181-6.
16. Alqahtani SA, Kleiner DE, Ghabril M, et al. Identification and characterization of cefazolin-induced liver injury. *Clin Gastroenterol Hepatol* 2015; 13: 1328-36.e2.
17. Dong H, Slain D, Cheng J, et al. Eighteen cases of liver injury following ingestion of Polygonum multiflorum. *Complement Ther Med* 2014; 22: 70-4.
18. Tillmann HL, Suzuki A, Barnhart HX, et al. Tools for Causality Assessment in Drug Induced Liver Disease. *Curr Opin Gastroenterol* 2019; 35: 183-90.
19. Talaat N, Tillmann HL. Injury pattern recognition to discriminate competing causes of liver injury. *Liver Int* 2019; 39: 821-5.
20. Chen M, Suzuki A, Thakkar S, et al. DILrank: the largest reference drug list ranked by the risk for developing drug-induced liver injury in humans. *Drug Discov Today* 2016; 21: 648-53.
21. Björnsson ES, Hoofnagle JH. Categorization of drugs implicated in causing liver injury: Critical assessment based on published case reports. *Hepatology* 2016; 63: 590-603.
22. Takikawa H, Takamori Y, Kumagi T et al., "Assessment of 287 Japanese cases of drug induced liver injury by the diagnostic scale of the International Consensus Meeting, *Hepatology Research* 2003; 27: 192-5.

23. Whritenour J, Ko M, Zong Q, et al. Development of a modified lymphocyte transformation test for diagnosing drug-induced liver injury associated with an adaptive immune response. *J Immunotoxicol* 2017 14: 31-8.
24. Benesic A, Leitl A, Gerbes AL. Monocyte-derived hepatocyte-like cells for causality assessment of idiosyncratic drug-induced liver injury. *Gut* 2016; 65: 1555-63.
25. Benesic A, Rotter I, Dragoi D, et al. Development and Validation of a Test to Identify Drugs That Cause Idiosyncratic Drug-Induced Liver Injury. *Clin Gastroenterol Hepatol* 2018; 169: 1488-94.e5.
26. Gerhardt F, Benesic A, Tillmann HL, et al. Iberogast-Induced Acute Liver Failure-Reexposure and In Vitro Assay Support Causality. *Am J Gastroenterol* 2019; 114: 1358-9.
27. Weber S, Wong GLH, Wong VWS, et al. Monocyte-Derived Hepatocyte-Like Cell Test: A Novel Tool for in vitro Identification of Drug-Induced Liver Injury in Patients with Herbal or Dietary Supplements. *Digestion* 2020; 20: 1-4.

Figure 1. Frequency distribution of latency of cases compared to a normal distribution.

Normal distribution (A) is shown compared to distribution of non-normal distribution of latency among DILI cases due to cyproterone (B), AMX/CLA (C), cefazoline (D), and polygonum multiflorum (E). In panels B to E, frequency of cases is given on the Y axis and latency in days from drug start in the X axis; the vertical lines in panels B-E represent the interquartile range or 25th and 75th percentile.

AMX/CLA, amoxicillin/clavulanate; DILI, drug-induced liver injury.

Table 1: Relative point allocation according to value relative to the distribution of values within the respective case series

	Percentage of points to be allocated	Point allocation
IQR low-IQR high	100%	20
25th to 15th percentile	50%	10
75th to 85th percentile		
15th to 10th percentile	25%	5
85th to 90th percentile		
10th percentile to minimum of range	0%	0
90th percentile to maximum of range		
Below minimum of range	-25%	-5
Above maximum of range		
Outlier*	-25%	-5

* For cases with values that are outside the range and are outliers, -50% is used.
IQR, interquartile range.

Table 2. Drug phenotypes described by Latency, R-Value & AST/ALT ratio with interquartile range and percentiles

	Allocated points	Cyproterone (n=22)			AMX-CLA (n=35)			Cefazolin (n=19)			Polygonum multiflorum (n=18)		
		Latency*	R-Value	AST/ALT ratio	Latency*	R-Value	AST/ALT ratio	Latency	R-Value	AST/ALT ratio	Latency	R-Value	AST/ALT ratio
Outlier low ¹⁻⁵		-75	-4.93	-0.17	-14.5	-2.61	-0.115	1.5	-2.47	-0.185	-49.7	-4.62	-0.045
Outlier low ⁰⁻⁷⁵	-5	19.6	1.93	0.25	1.3	-0.95	0.19	8.3	-0.71	0.09	-21.3	1.07	0.19
Below min	-5												
Incl. min	0	33.0	0.99	0.20	4.0	0.24	0.21	3.0	0.53	0.19	1.0	2.80	0.26
Incl. 10th %	5	63.3	2.19	0.39	7.6	0.55	0.35	4.0	0.82	0.28	1.0	5.23	0.32
Incl. 15th %	10	78.4	4.83	0.52	8.8	0.58	0.40	6.0	0.92	0.33	3.6	5.50	0.34
IQR low	20	114.3	8.78	0.67	17.0	0.72	0.50	15.0	1.07	0.37	7.0	6.75	0.42
Median	20	150	12.38	0.78	25.5	1.44	0.67	21	1.57	0.42	20	10.90	0.52
IQR high	20	240.5	17.92	1.23	38.0	2.94	0.91	24.0	3.43	0.74	44.8	14.33	0.73
Incl. 85th %	10	276.3	20.40	1.48	44.0	3.96	1.06	28.0	4.52	0.97	54.3	18.92	0.92
Incl. 90th %	5	305.9	25.93	1.71	50.0	6.74	1.30	29.0	6.05	0.99	72.3	21.26	1.13
Incl. max	0	425.0	29.96	2.13	63.0	13.94	1.87	29.0	10.60	1.23	120.0	26.30	2.52
Above max	-5												
Outlier high ¹⁻⁵	-5	429.9	31.63	2.08	69.5	6.27	1.53	37.5	6.98	1.31	101.4	25.69	1.20

*Latency in days from drug start to DILI onset.

Bold numbers indicate median, upper, and lower bounds of the interquartile range. Graded shading indicates further extent from the median.

AMX/CLA, amoxicillin/clavulanate; DILI, drug-induced liver injury IQR, interquartile range.

Table 3. Differences in Phenotypes Categories (Latency, R-Value & AST/ALT ratio) between drugs

	Cyproterone (n=22)			AMX-CLA (n=35)			Cefazolin (n=19)			Polygonum multiflorum (n=18)		
	Latency*	R-Value	AST/ALT ratio	Latency*	R-Value	AST/ALT ratio	Latency*	R-Value	AST/ALT ratio	Latency*	R-Value	AST/ALT ratio
Cyproterone												
Mann-Whitney U	NA	NA	NA	12.5	65.0	289.0	0	34	100	12	164	101
p-value	NA	NA	NA	<0.000	0.000	0.116	<0.000	0.000	0.004	<0.000	0.355	0.008
AMX-CLA												
Mann-Whitney U	12.5	65	289	NA	NA	NA	227.5	276	218	262.000	41	232
p-value	<0.000	<0.000	0.116	NA	NA	NA	0.057	0.306	0.038	0.319	<0.000	0.119
Cefazolin												
Mann-Whitney U	0	34	100	227	276	218	NA	NA	NA	171	17	146
p-value	<0.000	<0.000	0.004	0.057	0.306	0.038	NA	NA	NA	1.000	<0.000	0.447
Polygonum multiflorum												
Mann-Whitney U	12	164	101	262	41	232	171	17	146	NA	NA	NA
p-value	<0.000	0.355	0.008	0.319	<0.000	0.119	1.000	<0.000	0.447	NA	NA	NA

*Latency in days from drug start to DILI onset.

Bold numbers in shaded areas represent the most significant differences.

AMX/CLA, amoxicillin/clavulanate; DILI, drug-induced liver injury; NA, not applicable.

Table 4: Drugs' phenotype-derived DILI-CAT scores show potential for liver injury event separation dependent on drug and applied drug-characteristic DILI-CAT derived from phenotype as outlined in Table 1

Cyproterone vs. other drugs using DILI-CAT- $S_{\text{Cyproterone}}$ median points

	Categories			DILI-CAT		
	Latency	R-Value	de Ritis	Latency weighted	R-Value weighted	de Ritis weighted
Cyproterone (n=22)	20	20	20	47.5	55	50
AMX-CLA (n=35)	-5	-5	10	5	0	20
Cefazolin (n=19)	-5	-5	10	5	5	20
PM (n=18)	-7.5	10	10	15	32.5	30
P-values for Cyproterone versus	Mantel-Haenszel test for trend					
AMX-CLA (n=35)	<0.001	<0.001	0.785	<0.001	<0.001	<0.001
Cefazolin (n=19)	<0.001	<0.001	0.595	<0.001	<0.001	<0.001
PM (n=18)	<0.001	0.984	0.352	<0.001	0.004	0.001

AMX-CLA vs. other drugs using DILI-CAT- $S_{\text{AMX/CLA}}$ median points

	Categories			DILI-CAT		
	Latency	R-Value	de Ritis	Latency weighted	R-Value weighted	de Ritis weighted
Cyproterone (n=22)	-10	-5	20	-7.5	-2.5	20
AMX-CLA (n=35)	20	20	20	55	60	60
Cefazolin (n=19)	20	20	10	70	60	60
PM (n=18)	2.5	-5	15	25	10	27.5
P-values for AMX-CLA versus	Mantel-Haenszel test for trend					
Cyproterone (n=22)	<0.001	<0.001	0.785	<0.001	<0.001	<0.001
Cefazolin (n=19)	0.357	0.379	0.356	0.453	0.496	0.972
PM (n=18)	0.003	<0.001	0.989	<0.001	<0.001	<0.001

Cefazolin vs. other drugs using DILI-CAT-S_{Cefazolin}, median points

	Categories			DILI-CAT		
	Latency	R-Value	de Ritis	Latency weighted	R-Value weighted	de Ritis weighted
Cyproterone (n=22)	-10	-10	10	-12.5	-10	0
AMX-CLA (n=35)	10	10	20	30	40	40
Cefazolin (n=19)	20	20	20	60	60	60
PM (n=18)	0	-7.5	20	17.5	10	25
P-values for Cefazolin versus	Mantel-Haenszel test for trend					
Cyproterone (n=22)	<0-001	<0-001	0-045	<0-001	<0-001	<0-001
AMX-CLA (n=35)	0-005	0-082	0-244	0-002	0-009	0-008
PM (n=18)	0-002	<0-001	0-990	<0-001	<0-001	<0-001

Polygonum Multiflorum vs. other drugs using DILI-CAT-S_{Polygonum Multiflorum}, median points

	Categories			DILI-CAT		
	Latency	R-Value	de Ritis	Latency weighted	R-value weighted	de Ritis weighted
Cyproterone (n=22)	-10	15	5	7.5	32.5	25
AMX-CLA (n=35)	20	-5	10	45	20	40
Cefazolin (n=19)	20	-5	10	45	20	35
PM (n=18)	20	20	20	57.5	60	57.5
P-values for PM versus	Mantel-Haenszel test for trend					
Cyproterone (n=22)	<0-001	0-282	0-033	<0-001	0-001	<0-001
AMX-CLA (n=35)	0-031	<0-001	0-169	0-011	<0-001	0-002
Cefazolin (n=19)	0-084	<0-001	0-07	0-016	<0-001	0-001

AMX/CLA, amoxicillin/clavulanate; DILI-CAT, drug-induced liver injury causality assessment tool; DILI-CAT-S, drug-induced liver injury causality assessment tool score; PM: polygonum multiflorum.

