SCHOLARONE™
Manuscripts

Author Manuscript

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of record. Please cite this article as doi:10.1002/hep.29802.

**Candidate biomarkers for the diagnosis and prognosis of drug-induced liver injury: an international collaborative effort**

Rachel J. Church[1,2*], Gerd A. Kullak-Ublick[3,4*], Jiri Aubrecht[5], Herbert L. Bonkovsky[6], Naga Chalasani[7], Robert J. Fontana[8], Jens C. Goepfert[9], Frances Hackman[10], Nicholas M. P. King[11], Simon Kirby[10], Patrick Kirby[12,13], John Marcinak[12], Sif Ormarsdottir[14,15], Shelli J. Schomaker[5], Ina Schuppe-Koistinen[16], Francis Wolenski[12], Nadir Arber[17], Michael Merz[3,18], John-Michael Sauer[11], Raul J. Andrade[19], Florian van Bömmel[20], Thierry Poynard[21**], and Paul B. Watkins[1,2**]


**Author Affiliations**

[1]Institute for Drug Safety Sciences, University of North Carolina at Chapel Hill, Research Triangle Park, NC

[2]Division of Pharmacotherapy and Experimental Therapeutics, UNC Eshelman School of Pharmacy, Chapel Hill, North Carolina

[3]Department of Clinical Pharmacology and Toxicology, University Hospital Zurich, University of Zurich, Switzerland

[4]Mechanistic Safety, Novartis Global Drug Development, Basel, Switzerland

[5]Drug Safety R&D, Pfizer Inc,.Groton, CT

[6]Wake Forest University School of Medicine, Winston-Salem, NC

[7]School of Medicine, Indiana University, Indianapolis, IN

[8]University of Michigan, Ann Arbor, MI

[9]Natural and Medical Sciences Institute, Reutlingen, Germany

[10]Pfizer Inc., New York, USA

[11]Predictive Safety Testing Consortium, Tucson AZ, USA

1

[12]Takeda Pharmaceuticals, Deerfield, IL

[13]Current Affiliation: Sanofi, Framingham, MA

[14]Global Patient Safety, AstraZeneca R&D, Gothenburg, Sweden

[15]Current affiliation: Department of Gastroenterology and Hepatology, University Hospital, Iceland

[16]Science for Life Laboratory, Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden

[17]Tel Aviv Sourasky Medical Center, Tel Aviv University, Israel

[18]Discovery and Investigative Safety, Novartis Institutes for Biomedical Research, Basel, Switzerland

[19]Unidad de Gestión Clínica de Aparato Digestivo, Instituto de Investigación Biomédica de Málaga-IBIMA, Hospital Universitario Virgen de la Victoria, Universidad de Málaga, Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd), Málaga, Spain

[20] Section of Hepatology, Clinic of Gastroenterology and Hepatology, University Hospital Leipzig, Leipzig, Germany

[21]Department of Hepatology, Groupe Hospitalier Pitié Salpêtrière, University Pierre et Marie Curie, INSERM UMR 938, Paris, France

*Authors share first authorship

** Authors share senior authorship

**Author Email Addresses**
Rachel J. Church: rchurch@unc.edu
Gerd A. Kullak-Ublick: gerd.kullak@usz.ch

Jiri Aubrecht: jiri.aubrecht@pfizer.com

Herbert L. Bonkovsky: hbonkovs@wakehealth.edu

Naga Chalasani: nchalasa@iu.edu

Robert J. Fontana: rfontana@med.umich.edu

Jens Goepfert: jens.goepfert@nmi.de

Frances Hackman: frances.hackman@pfizer.com

Nicholas M. P. King: nking@c-path.org

Simon Kirby: simon.kirby@pfizer.com

Patrick Kirby: patrick.kirby@sanofi.com

John Marcinak: john.marcinak@takeda.com

Sif Ormarsdottir: sifor@landspitali.is

Shelli J. Schomaker: shelli.j.schomaker@pfizer.com

Ina Schuppe-Koistinen: ina.schuppe-koistinen@scilifelab.se

Francis Wolenski:francis.wolenski@takeda.com

Nadir Arber: narber@post.tau.ac.il

Michael Merz: michael.merz@t-online.de

John-Michael Sauer: jsauer@c-path.org

Raul J. Andrade: andrade@uma.es

Florian van Bömmel: florian.vanboemmel@medizin.uni-leipzig.de

Thierry Poynard: thierry@poynard.com

Paul B. Watkins: pwatkins@ad.unc.edu

**Key Words:** Predictive Safety Testing Consortium (PSTC), Safer and Faster Evidence-based

Translation (SAFE-T), Drug-Induced Liver Injury Network (DILIN), hepatotoxicity, reference

interval

3

**<u>FOOTNOTE PAGE</u>**

**Corresponding Author Contact Information**

Rachel J. Church

Institute for Drug Safety Sciences

6 Davis Drive

Research Triangle Park, NC 27709

Email: rchurch@unc.edu

Tel: (919) 226-3145

**Abbreviations**

DILI: drug-induced liver injury

ALT: alanine aminotransferase

AST: aspartate aminotransferase

ALP: alkaline phosphatase

TBIL: total bilirubin

PSTC: Predictive Safety Testing Consortium

SAFE-T: Safer and Faster Evidence-based Translation Consortium

IMI: Innovative Medicine Initiative

DILIN: Drug-Induced Liver Injury Network

AFP: alpha fetoprotein

ARG1: arginase-1

4

CDH5: cadherin-5

FABP1: fatty acid binding protein 1

GSTα: glutathione S-transferase alpha

K18: total cytokeratin 18

ccK18: caspase cleaved cytokeratin 18

MCSFR: macrophage colony stimulating factor receptor

OPN: osteopontin

GLDH: glutamate dehydrogenase

LECT2: leucocyte cell-derived chemotaxin-2

PON1: paraoxonase 1 normalized to prothrombin protein

SDH: sorbitol dehydrogenase

miR-122: microRNA-122

ULN: upper limit of normal

BMI: body mass index

NMI: Natural and Medical Sciences Institute

IQR: interquartile range

AI: apoptotic index

LLN: lower limit of normal

LLoQ: lower limit of quantification

5

ROC: receiver operator characteristic

CI: confidence interval

CV: coefficient of variation

AUC: area under the curve

APAP: acetaminophen

NSAID: non-steroidal anti-inflammatory drug

INR: International Normalized Ratio

MELD: Model for End stage Liver Disease

6

University of Nottingham.

7

## ABSTRACT

Current blood biomarkers are suboptimal in detecting drug-induced liver injury (DILI) and predicting its outcome. We sought to characterize the natural variabilty and performance characteristics of fourteen promising DILI biomarker candidates. Serum or plasma from multiple cohorts of healthy volunteers (n=192 and =81), subjects who safely took potentially hepatotoxic drugs without adverse effects (n=55 and =92) and DILI patients (n=98, =28, and =143) were assayed for microRNA-122 (miR-122), glutamate dehydrogenase (GLDH), total keratin 18 (K18), caspase cleaved K18 (ccK18), glutathione S-transferase alpha (GSTα), alpha fetoprotein (AFP), arginase-1 (ARG1), osteopontin (OPN), sorbitol dehydrogenase (SDH), fatty acid binding protein (FABP1), cadherin-5 (CDH5), macrophage colony stimulating factor receptor (MCSFR), paraoxonase 1 (PON1, normalized to prothrombin protein), and leucocyte cell-derived chemotaxin-2 (LECT2). Most candidate biomarkers were significantly altered in DILI cases compared to healthy volunteers. GLDH correlated more closely with gold standard alanine aminotransferase (ALT) than miR-122 and there was a surprisingly wide inter- and intra-individual variability of miR-122 levels among the healthy volunteers. Serum K18, OPN, and MCSFR levels were most strongly associated with liver-related death or transplant within 6 months of DILI-onset. Prediction of prognosis among DILI patients using Model for End-stage Liver Disease (MELD) was improved by incorporation of K18 and MCSFR levels. *Conclusion*: GLDH appears to be more useful than miR-122 in identifying DILI patients. K18, OPN and MCSFR are promising candidates for prediction of prognosis during an acute DILI event. Serial assessment of these biomarkers in large prospective studies will help further delineate their role in DILI diagnosis and management.

Drug-induced liver injury (DILI) is a serious concern for patients, clinicians and pharmaceutical companies, accounting for over half of the acute liver failure cases observed in Western countries [1, 2]. Current detection and assessment of DILI relies on measurement of analytes that have been utilized for decades. Serum enzyme activities of aminotransferases, [alanine aminotransferase (ALT) and aspartate aminotransferase (AST)] and alkaline phosphatase (ALP) are quantified as measures of hepatocellular or cholestatic injury, respectively, whilst serum total bilirubin (TBIL) concentration is frequently used to assess global liver function. However, alterations in these biomarkers are not mechanistically informative, can occur for a variety of reasons unrelated to hepatic injury [3-5], and can be observed with drugs that do not have the potential to cause clinically significant DILI [6, 7]. They also are not specific or selective for DILI versus other causes of liver injury. Moreover, it is not currently possible to distinguish benign liver chemistry elevations from those that could lead to liver failure. Hence, even mild aminotransferase elevations can increase liver safety concern, especially in early clinical trials. A combination of an injury marker (aminotransferases) and a functional marker (TBIL) in the absence of cholestasis and when other causes have been excluded (i.e., Hy's Law case criteria) is a widely accepted prognostic model for DILI outcome [8]. However, most of the patients that meet Hy's Law case criteria will not require a liver transplant or die. Ongoing efforts are exploring methods to improve severity prediction utilizing traditional tests [9]. Nevertheless, there is a clear unmet need for biomarkers that are mechanistically informative and sensitive, as well as specific for prediction of DILI progression or resolution.

Substantial resources have been committed to understanding DILI and advancing candidate biomarkers that add value to traditional liver tests. Particularly, the Critical Path Institute's Predictive Safety Testing Consortium (PSTC) in the United States and the Safer and Faster

9

Evidence based Translation (SAFE-T) consortium within the Innovative Medicines Initiative (IMI) in Europe were leading major efforts to validate and qualify novel DILI biomarkers [10]. The Drug Induced Liver Injury Network (DILIN) is a multi-center network in the United States created to prospectively biobank blood and tissue specimens from patients who have experienced DILI [11]. Each subject in this registry has undergone an unprecedented degree of phenotyping and most have at least six-month follow-up data aiding assessments of long-term outcomes and prognosis. Given the overlapping goals of these three organizations, a cross functional collaboration was established to study performance characteristics of candidate DILI biomarkers.

Candidate DILI biomarkers have been identified in preliminary evaluations. However, clinical application of these candidate markers requires robust performance for DILI detection and prognosis compared to current standards. Further, it is critical that normal reference intervals be established for these biomarkers against which data from patients can be measured. Herein the results from an international collaborative effort among PSTC, SAFE-T, and DILIN are presented. Alpha fetoprotein (AFP), arginase 1 (ARG1), cadherin 5 (CDH5), fatty acid binding protein 1 (FABP1), glutathione S transferase alpha (GSTα), total keratin 18 (K18), caspase cleaved (cc)K18, macrophage colony stimulating factor receptor (MCSFR), osteopontin (OPN), glutamate dehydrogenase (GLDH), leucocyte cell-derived chemotaxin-2 (LECT2), paraoxonase 1 (PON1, normalized to prothrombin protein), sorbitol dehydrogenase (SDH), and microRNA-122 (miR-122) were assayed in serum or plasma from two cohorts of normal healthy subjects, two cohorts of patients that safely took potentially hepatotoxic drugs, and three cohorts of DILI patients (**Table 1**). DILI cohorts included patients who at 6 months had recovered completely, had persistent DILI or who died or required a liver transplant due to the DILI event within 6 months of DILI-onset.

10

## MATERIALS AND METHODS

### Human Subject Sample Collection

Demographic data for the respective population cohorts can be found in **Supplemental Tables 1-3**. All studies were conducted in accordance with protocols approved by local Institutional Review Boards and written consent was received from all participants. Inclusion and exclusion criteria for each cohort is described in more detail in **Supplemental Methods.** The primary causative drug implicated for all DILI patients is listed in **Supplemental Table 4.** Following isolation of serum or plasma, samples were biobanked at -80°C until analysis and may have been frozen for multiple years.

### Biomarker Analysis

The 14 candidate biomarkers that were quantified in this study are listed in **Table 1**. Detailed methods for biomarker quantification can be found in **Supplemental Methods**. Briefly, serum or plasma traditional biomarker levels were quantified by clinical chemistry at the local institutions. Serum or plasma candidate biomarker levels were quantified in assays designed by or optimized by Natural and Medical Sciences Institute (NMI, Reutlingen, Germany) or at contract laboratories. Detailed information regarding validation parameters for all assays used in this study can be found in **Supplemental Table 5**.

### Statistical Analysis

Descriptive statistics, median with interquartile range (IQR), were used to describe continuous variables, and frequency and percent were used to describe categorical variables. All statistical analyses were performed using JMP Genomics v8.1 or SAS software (SAS, Cary, NC) or GraphPad Prism 7.01 (La Jolla, CA). Biomarker distribution was visualized and the majority of

11

displayed a log normal distribution. For consistency, the absolute value of all biomarkers, with the exception of the Apoptotic Index (AI), were log transformed for statistical analyses. Statistical significance was considered $p<0.05$.

**Reference Interval Determination**

PSTC healthy volunteer data and SAFE-T Tel Aviv healthy volunteer data were analyzed for the determination of reference intervals. The reference interval lower limit of normal (LLN) and ULN was defined by the 5th and 95th percentile of the population, respectively, using a mixed model approach and fitted a random subject term. For two of the markers (ccK18 and GSTα) a substantial number of values were below the lower limit of quantification (LLoQ), so a maximum likelihood estimate for a truncated log-normal distribution was used to estimate the ULN. Because ~90% of the data for K18 was below the LLoQ, only the ULN was calculated by a nonparametric method. PSTC collected three serial biomarker measurements in all subjects and a mixed model was used to obtain the variance components for inter-and intra-individual variation assuming log-normal distribution. The reference interval was obtained using the estimated mean and standard deviation for the log-normal distribution.

**Biomarkers of DILI Detection**

Receiver operator characteristic (ROC) curve analysis was utilized to determine the performance of traditional and candidate biomarkers for detection of DILI patients. Biomarkers were considered predictive of DILI if both the ROC AUC and the lower end of the 95% confidence interval (CI) >0.5. All healthy volunteer and patient datasets (PSTC, SAFE-T, and DILIN) were used for this analysis. The relationship of liver-specific GLDH and miR-122 to ALT was examined. Correlation of GLDH and miR-122 with ALT was determined using Pearson's r.

12

**Biomarkers of DILI Prognosis**

Accurate outcome assessments were available only for the DILIN subjects. ROC curve analysis was utilized to determine which biomarkers measured in the initial DILIN sera could significantly predict which patients died/required a liver transplant or developed unresolved DILI. For a detailed description of this analysis, refer to **Supplemental Methods**.

13

## RESULTS

### Biomarker Levels in Healthy Volunteers

The natural variation of candidate biomakers was explored in two cohorts of healthy volunteers (see **Supplemental Tables 1** and **2)**. Multiple samples returned values below the LLoQ for AFP, ccK18, GSTα, K18, SDH, and miR-122. The geometric mean, inter-subject variation and reference interval (5th and 95th percentiles) for each biomarker are presented in **Table 2**. Due to the large number of samples with K18 values that fell below the LLoQ, only the estimated 95th percentile was calculated for this biomarker. In both healthy volunteer cohorts, inter-subject variability of miR-122 was also high [% coefficient of variation (CV) of 90.89 and 213.51 in the PSTC and SAFE-T cohorts, respectively]. Further, miR-122 also showed substantial intra-subject variability in the PSTC cohort (intrasubject %CV of 93.56) and this appeared to be most prominent among black individuals (**Figure 1**).

The biomarker reference intervals between PSTC and SAFE-T showed substantial overlap, although the geometric mean tended to be slightly higher in the SAFE-T cohort. ARG1 levels, however, were considerably increased in the SAFE-T cohort, compared to the PSTC cohort.

### Biomarker Performance for Detecting DILI

All candidate DILI biomarkers significantly identified patients with DILI with the exception of LECT2 (**Table 3**). The lower CI limit of LECT2 was only 0.45, indicating that the data cannot rule out random agreement between predictions and outcome. K18, ccK18, FABP1, and GLDH had AUCs > 0.9, indicating that these biomarkers are the most accurate candidate biomarkers for the detection of DILI.

14

GLDH levels showed a very strong correlation with ALT levels (r=0.88, $p<0.0001$; **Figure 2A**). miR-122 levels were also significantly correlated with levels of ALT, although the strength of the correlation was reduced compared to that for GLDH (r=0.66; $p<0.0001$; **Figure 2B**).

**Biomarker Alterations by Drug Class**

The SAFE-T DILI cohorts contain data from patients with acetaminophen (APAP)-related liver injury, as well as from patients who experienced idiosyncratic DILI related to various compounds (refer to **Supplemental Table 4**). To determine if one or more DILI compounds/classes produces signature biomarker changes that are unique compared to APAP-related hepatotoxicity, SAFE-T DILI patient data were divided into broad drug classes (see **Supplemental Methods**). In general, biomarkers, including ALT, tended to be the most altered from other drug classes in APAP-related hepatotoxicity, emphasizing the acute and severe injury this compound causes (**Supplemental Figure 1A**). However, several biomarkers were the most elevated in flupirtine DILI. Specifically, TBIL ($p<0.05$ compared to APAP, antibiotics, chemo, NSAID, and other; **Supplemental Figure 1B**), CDH5 ($p<0.05$ compared to APAP, antibiotics, chemo, and other; **Supplemental Figure 1C**), and MCSFR ($p<0.05$ compared to APAP, antibiotics, chemo, NSAID, other; **Supplemental Figure 1D**) were significantly elevated in flupirtine-related liver injury. Flupirtine is an aminopyridine used as a non-opioid analgesic that is well recognized to cause serious DILI in Europe (not available in the US); all SAFE-T patients with flupirtine-related liver injury met Hy's Law case criteria. When DILIN data were divided into drug categories, no significant differences were observed; however, DILIN data does not have patients with either APAP- or flupirtine-related hepatotoxicity (data not shown).

Differences between cohorts in amoxicillin with clavulanic acid (Augmentin) DILI were explored. We found that DILIN patients with Augmentin-induced hepatotoxicity had

15

significantly elevated levels of ALT, ARG1, FABP1, GST-α, K18, and ccK18 ($p<0.05$ for all) compared to SAFE-T DILI patients (**Supplemental Table 6**).

**Biomarker Performance as Prognostic Markers**

*Death/Transplant*

The DILIN samples utilized in the present study were limited to those collected within 2 weeks of DILI onset and, per protocol, patients were tracked for at least 6 months. A small subsets of DILIN patients died/required a liver transplant (n=15) within 6 months and it was determined that the DILIN event was the cause [12]. ROC curve analysis demonstrated that traditional biomarkers including the international normalized ratio (INR), AST, and TBIL were predictive of death/liver tranplant (lower 95% CI limit >0.5; **Table 4**). Of the subset of candidate biomarkers measured in the DILIN dataset (**Table 1**), elevated levels of OPN, K18, MCSFR, ccK18, FABP1, and AFP significantly predicted death/transplant. INR was found to have the strongest association with death/transplant (AUC=0.920) closely followed by OPN (AUC=0.858). The predictive biomarkers had >2X and >7X fold changes over DILI patients that did not experience liver failure and healthy volunteers, respectively (**Supplemental Table 7**).

The values of K18 and ccK18 measured in 98 DILIN patients enabled the calculation of an Apoptotic Index (AI, see **Supplemental Methods** for details) and AI (ccK18:K18 ratio) was also explored as a prognostic biomarker of death/liver tranplant (**Table 4**). Although both K18 and ccK18 levels were elevated in patients that experienced death/liver transplant, compared to patients that did not, the AI was significantly reduced in patients who died/required a liver transplant (**Figure 3A-C**). An AI was additionally calculated in 64 SAFE-T patients with DILI. Patients with flupirtine-related DILI did not have a significantly different mean AI, compared to patients with APAP-induced liver injury or DILI associated with other compounds (**Figure 3D**).

16

Current prognostic models for liver transplant and death were explored to identify whether incorporation of biomarkers that passed filtering criteria (OPN, K18, MCSFR, and AFP; see **Supplemental Methods** for details) could improve prediction. Model of End-stage Liver Disease (MELD) score $\geq$ 20 was highly sensitive and MELD score $\geq$ 30 was highly specific for prognosis of death/transplant in the DILIN population. Incorporating K18 and MCSFR levels with MELD score (when MELD values were from 20-29; **Figure 4**) improved the specificity of using MELD score $\geq$ 20 (specificity of 0.889 when incorporating K18 and MCSFR with MELD score $\geq$ 20 vs. 0.738 with MELD score $\geq$ 20 alone; **Supplemental Table 8**) without reducing the sensitivity of using MELD $\geq$ 20 alone (sensitivity of 0.933 for both). Hy's Law showed moderate performance for prediction of death or liver transplant in this DILIN cohort (sensitivity of 0.8 and specificity of 0.634).

*Unresolved DILI*

Nineteen patients in the DILIN cohort had unresolved DILI at their six month follow-up visit (persistently elevated ALT, AST, ALP, or TBIL over ULN with no competing etiology). Consistent with previous data, only elevated levels of ALP predicted the outcome of these patients (lower CI limit >0.5; **Supplemental Table 9**) [13]. Within this subset, 6 of the 19 patients with unresolved DILI had ALP levels that were elevated at their six month follow-up visit. When data was reanalyzed to determine if candidate biomarkers could predict this subset of patients, it would found that GST-$\alpha$ (measured within two weeks of DILI-onset) was significantly lower in patients with prolonged ALP elevation (ROC AUC = 0.760, 95% CI = 0.509-1.0).

## DISCUSSION

Traditional biomarkers of DILI may not be adequately liver-specific, offer little mechanistic insight into mode of injury, and are not sufficiently prognostic for injury outcome/resolution. In the current study, the performance of 14 candidate DILI biomarkers was explored in cohorts of healthy volunteers, patients who received known DILI-eliciting compounds without developing liver injury, and in patients who experienced DILI.

Most of the biomarker reference intervals showed sizeable overlap between the SAFE-T and PSTC healthy volunteer cohorts. ARG1 levels, however, were substantially higher in the SAFE-T compared to PSTC volunteers (**Table 2**). This difference could not be accounted for by differences in racial diversity or age. Additionally. each biomarker was quantified at the same facility, with the same validated assays, making technical variability an unlikely explanation to account for this difference. However, biomarker stability at -80° should be explored as a possible explanation for  ARG1 population differences, given that samples were stored from 3 months to 3 years. The influence of race (primarily whites vs. African Americans) on biomarker reference ranges was explored in the PSTC cohort. Although small differences were observed (such as for ARG1), there was considerable overlap between biomarker ranges and unique reference intervals were deemed to be unnecessary in this small population.

The biomarkers examined in this study were selected based on preliminary performance data generated by SAFE-T in a small pilot cohort of DILI patients and healthy volunteers (data not shown). Therefore, it is not surprising that most of the candidate biomarkers were significantly elevated in  DILI. In particular, K18, FABP1, GLDH, and ccK18 had ROC AUCs >0.90, suggesting that these biomarkers in particular may be useful when screening for DILI. The sensitivity of these biomarkers for the detection of DILI or hepatotoxicity, in general, is in

18

agreement with previously reported data [14-20]. Biomarker differences in patients with Augmentin-induced heptotoxicity were explored between the DILIN and SAFE-T cohorts and several candidate biomarkers were found to be elevated in the DILIN patients. DILIN patients also had higher mean elevations in serum ALT, therefore DILIN patients may in general have been sicker than SAFE-T patients. Indeed, while SAFE-T enrolled patients with ALT >3X ULN, DILIN inclusion criteria specifies that patients must have an ALT >5X ULN.

Although very large elevations in serum ALT are generally believed to be liver-specific, ALT elevations are also noted following strenuous exercise and in patients with muscle diseases such as muscular dystrophy [3, 21]. A biomarker with greater liver-specificity could be valuable in the clinic when the source of an ALT elevation is uncertain. GLDH and miR-122 are highly liver-specific and are not altered in response to muscle injury [22, 23]. GLDH is a large protein found within the matrix of mitochondria, enriched in the liver [24]. The inter-individual variability of GLDH ranged from ~53% in the PSTC cohort to ~80% in the SAFE-T cohort and intra-individual variability was minimal (35%). In contrast, there was significant inter-individual variability of miR-122, a liver-specific miRNA that makes up as much as 70% of hepatic miRNA content [25], in both cohorts (~91 and 213% in the PSTC and SAFE-T cohorts, respectively) and remarkable intra-individual variability, most evident among African Americans, was also observed (~94%).

MiR-122, unlike other candidate biomarkers, may not simply be leaked passively from injured cells, although this is believed to be the primary mechanism following injury. Instead, evidence suggests that miR-122 can be released actively from the liver, at least in part within extracellular vesicles, in response to stress [26]. For example, it has recently been shown that in response to stimuli and in the absence of overt hepatocyte death, miR-122 can be released and can modulate

19

activation of innate immune cells or directly regulate kidney release of erythropoetin [27, 28]. The variability that is inherent to miR-122 levels not only between individuals but also within the same individual therefore likely results from physiologic processes unrelated to damage to the liver. The complexity of the assay (relative to other biomarker assays), and the lack of a universally accepted method for data normalization, may also contribute to the variability observed. For these reasons, the PSTC has recently deprioritized pursuit of miR-122 as a liver-specific biomarker, in favor of GLDH which is measured by routine clinical chemistry.

GLDH demonstrated enhanced correlation with ALT levels and improved performance for detection of DILI, as compared to miR-122. Previous research explored GLDH as a biomarker of hepatocellular necrosis in patients with liver impairment and found it to be superior to other candidate biomarkers (miR-122 was not evaluated in that study) [18]. Other studies have also demonstrated that GLDH is elevated in patients with APAP-related toxicity [29, 30]. Furthermore, GLDH has been proposed as a potentially early indicator of recovery from DILI due to the fast elimination of GLDH observed in subjects recovering from accidental APAP overdose with persisting high levels of ALT [14].

It should be noted that in spite of the large inter- and intra-individual variations in serum levels of miR-122 that we report here, miR-122 appears to be useful in predicting liver injury after APAP overdose. A recent study found that miR-122 demonstrated the highest performance for prediction of APAP-induced acute liver injury in a large cohort of overdose patients with normal ALT levels at presentation confirming results of an earlier study [14, 31]. In contrast, GLDH was not useful in this context. Because the method for quantitation of miR-122 in these studies differed from ours, it is unclear if the levels of miR-122 in patients susceptible to APAP injury were simply above the range of inter-and intra-individual variation meaured in the healthy

20

volunteers in our study. While the assumption is that elevated serum levels of miR-122 reflect early hepatocyte stress or injury due to APAP, given the increasing appreciation of the physiological roles of miR-122 it is possible that individuals with high baseline serum levels of miR-122 are more susceptible to APAP injury.

The prognostic performance of GLDH and miR-122 was not determined in the current study because these biomarkers were not measured in the DILIN patient cohort (due to sample volume limitatinos) where outcome data was systematically collected. However, semi-quantitative measurements of miR-122 have previously been conducted in a subset of DILIN patients [32]. In contrast to the data observed in APAP-induced hepatotoxicity, reduced serum levels of miR-122 and albumin were observed in patients that died within six months of DILI onset, compared to patients that recovered. Collectively, these data suggest that both miR-122 and GLDH likely have utility in predicting and managing DILI and factors related to extent of injury at serum collection time and biomarker half life may influence the interpretation of biomarker alterations. Research exploring the kinetics of these biomarkers in DILI may aid in interpretation of these biomarkers in the clinic.

Several biomarkers showed promise as prognostic biomarkers for death/transplant in DILI. In particular OPN, K18, and MCSFR performed well as predictors of death/transplant. Increased levels of each of these biomarkers were observed in DILIN death/transplant patients compared to all others. OPN showed the best performance for prognosis of all candidate biomarkers in DILIN patients. OPN is associated with liver regeneration due to activation of hepatic stem cells [33]. While elevated levels of AFP, which is also released from hepatic stem cells, were prognostic for death/transplant, the performance of this biomarker was reduced compared to OPN (AUC= 0.687 vs. 0.858 for AFP and OPN, respectively).

21

The increase in OPN observed in this study is in contrast to recent data demonstrating that patients with acute liver failure (from various etiologies) who died/received transplant had reduced levels of plasma OPN compared to those that recovered [34]. The difference between these studies may be related to the timing of sample collection. Entry into the previous study required acute liver failure to be occurring at enrollment (INR $\geq$ 1.5 and encephalopathy), suggesting more advanced injury than present in the current cohort.

MCSFR, another marker of inflammation, is amongst the most promising prognostic candidate biomarkers (AUC=0.775) in the DILIN data for death/transplant. MCSFR, the receptor for CSF-1, is thought to be shed from activated macrophages during DILI [35]. Interestingly, reduced levels of CSF-1 were associated with poor outcome in patients experiencing APAP-induced liver injury and this was thought to suggest that macrophages and an innate immune response are necessary for regeneration following liver injury [36]. The cause of this discrepancy is unclear, but may be related to the type of DILI examined in the current study (idiosyncratic vs intrinsic). MCSFR levels were considerably higher in SAFE-T patients experiencing flupirtine-induced hepatotoxicity compared to the 19 cases of APAP-induced hepatotoxicity, despite ALT levels being markedly higher in APAP-induced liver injury (**Figure 2**). Although no SAFE-T DILI patients died or required a liver transplant, all patients that experienced flupirtine-induced DILI met Hy's Law case criteria, suggesting that increased MCSFR may be indicative of severe idiosyncratic DILI (vs. APAP-induced liver injury).

K18 also showed value as a prognostic biomarker in DILIN patients; K18 levels were elevated in patients that died or needed a liver transplant, compared to patients that survived. We found that incorporating K18 (threshold: log normalized value of 7.98) into a model that stratified risk based on MELD score cutoffs of $\geq$ 20 and $\geq$ 30 improved specificity of using MELD $\geq$ 20, alone,

22

without decreasing sensitivity. The addition of MCSFR further improved the specificity, slightly. The value of K18 for prediction models of death following hepatotoxicity is in agreement with previously published literature [15, 37].

K18 is an intermediate filament found in epithelial cells, including hepatocytes. Necrosis results in passive leakage of full length K18 into circulation while cleavage of K18 by caspases results in leakage of ccK18 into circulation following apoptotis [38]. Apoptosis is thought to be a more benign form of injury because apoptosis is not believed to result in the release of damage associated molecular pattens and subsequent activation of the innate immune system [39, 40]. Determination of the ccK18:K18 ratio, the AI, is believed to reflect the proportion of cell death that can be attributed to apoptosis. DILIN patients that died or needed a liver transplant had lower AIs than patients who recovered from DILI (i.e. consistent with necrosis as the predominate form of cell death in patients that died or needed a liver transplant). Because biopsies are not routinely conducted in the clinic, validation of the AI in humans is challenging. Nevertheless, pilot data in DILIN patients suggests that the AI may be useful in predicting the degree of apoptosis vs necrosis in liver tissue [41].

It should be noted that in the DILIN cohort two weeks was the maximum time between DILI-onset (the time at which a patient's serum liver chemistries first qualified for DILIN entry) and research blood collection, but the time between symptom onset (when known) and research blood collection varied to a larger degree, ranging from 2 to 90 days. We were unable to detect significant correlations between biomarker levels and days between symptom onset and research blood collection (r= -0.014 to 0.222; **Supplemental Table 10**). For example, the interval between symptom onset and determination of serum levels of K18 and MCSFR had Pearson r's of 0.142 and 0.163, respectively. Because variation in biomarker release and clearance kinetics

23

may significantly impact interpretation of biomarker levels, future studies should include serial samples collected over a broad range of intervals from symptom onset.

Resulting from the data presented here, both the Food and Drug Administration and the European Medicines Agency issued Letters of Support that explicitly encourage the exploratory use of selected biomarkers in drug registration trials and further development of K18, OPN, and MCSFR as potential diagnostic or prognostic DILI biomarkers [42, 43]. Further exploration of both miR-122 and GLDH as liver-specific alternatives to ALT was also encouraged by these regulatory agencies.

In summary, the large inter- and intra-subject variation in miR-122 and the recent recognition of its regulated release from the liver without hepatocyte death may complicate its interpretation in the clinic but it is likely still valuable in certain contexts such as in the setting of APAP-induced hepatotoxicity. Alternatively, MCSFR may be elevated to a greater degree in severe idiosyncratic DILI. GLDH was a sensitive biomarker for detection of DILI and should be useful in certain clinical contexts to exclude muscle injury as a source of serum biomarkers [22]. K18, FABP1, and ccK18 were also highly sensitive for DILI detection. OPN, K18, and MCSFR show promise as biomarkers that can identify those DILI patients who will succumb to the DILI event unless transplanted. The combined use of K18, MCSFR and the MELD score improved specificity without reducing the sensitivity compared to use of a MELD score of ≥20 alone.

Based on the data reported here, follow-up initiatives should include (i) further exploration of the prognostic value of the biomarkers endorsed by regulatory agencies via broad application in clinical trials with serial sample collection (ii) correlation of the mechanism of DILI with the performance of the biomarkers (e.g. intrinsic DILI versus immune activation), and (iii)

24

assessment of the performance of biomarkers in drug-induced versus other causes of liver injury (e.g. viral hepatitis or autoimmune hepatitis). These efforts will allow the most promising biomarkers to be validated and qualified for routine use in clinical DILI assessment.

25

## REFERENCES

1.      Ostapowicz G, Fontana RJ, Schiodt FV, Larson A, Davern TJ, Han SH, *et al.* Results of a prospective study of acute liver failure at 17 tertiary care centers in the United States. Ann Intern Med 2002;137:947-54.

2.      Wei G, Bergquist A, Broome U, Lindgren S, Wallerstedt S, Almer S, *et al.* Acute liver failure in Sweden: etiology and outcome. J Intern Med 2007;262:393-401.

3.      Pettersson J, Hindorf U, Persson P, Bengtsson T, Malmqvist U, Werkstrom V, *et al.* Muscular exercise can cause highly pathological liver function tests in healthy men. Br J Clin Pharmacol 2008;65:253-9.

4.      Purkins L, Love ER, Eve MD, Wooldridge CL, Cowan C, Smart TS, *et al.* The influence of diet upon liver function tests and serum lipids in healthy male volunteers resident in a Phase I unit. Br J Clin Pharmacol 2004;57:199-208.

5.      Ruhl CE and Everhart JE. Determinants of the association of overweight with elevated serum alanine aminotransferase activity in the United States. Gastroenterology 2003;124:71-9.

6.      Aithal GP, Watkins PB, Andrade RJ, Larrey D, Molokhia M, Takikawa H, *et al.* Case definition and phenotype standardization in drug-induced liver injury. Clin Pharmacol Ther 2011;89:806-15.

7.      Harrill AH, Roach J, Fier I, Eaddy JS, Kurtz CL, Antoine DJ, *et al.* The effects of heparins on the liver: application of mechanistic serum biomarkers in a randomized study in healthy volunteers. Clin Pharmacol Ther 2012;92:214-20.

8.      Senior JR. Alanine aminotransferase: a clinical and regulatory tool for detecting liver injury-past, present, and future. Clin Pharmacol Ther 2012;92:332-9.

26

9.      Robles-Diaz M, Lucena MI, Kaplowitz N, Stephens C, Medina-Caliz I, Gonzalez-

Jimenez A, *et al.* Use of Hy's law and a new composite algorithm to predict acute liver

failure in patients with drug-induced liver injury. Gastroenterology 2014;147:109-118 e5.

10.     Matheis K, Laurie D, Andriamandroso C, Arber N, Badimon L, Benain X, *et al.* A

generic operational strategy to qualify translational safety biomarkers. Drug Discov

Today 2011;16:600-8.

11.     Fontana RJ, Watkins PB, Bonkovsky HL, Chalasani N, Davern T, Serrano J, *et al.* Drug-

Induced Liver Injury Network (DILIN) prospective study: rationale, design and conduct.

Drug Saf 2009;32:55-68.

12.     Hayashi PH, Rockey DC, Fontana RJ, Tillmann HL, Kaplowitz N, Barnhart HX, *et al.*

Death and liver transplantation within 2 years of onset of drug-induced liver injury.

Hepatology 2017;66:1275-1285.

13.     Fontana RJ, Hayashi PH, Barnhart H, Kleiner DE, Reddy KR, Chalasani N, *et al.*

Persistent liver biochemistry abnormalities are more common in older patients and those

with cholestatic drug induced liver injury. Am J Gastroenterol 2015;110:1450-9.

14.     Antoine DJ, Dear JW, Lewis PS, Platt V, Coyle J, Masson M, *et al.* Mechanistic

biomarkers provide early and sensitive detection of acetaminophen-induced acute liver

injury at first presentation to hospital. Hepatology 2013;58:777-87.

15.     Antoine DJ, Jenkins RE, Dear JW, Williams DP, McGill MR, Sharpe MR, *et al.*

Molecular forms of HMGB1 and keratin-18 as mechanistic biomarkers for mode of cell

death and prognosis during clinical acetaminophen hepatotoxicity. J Hepatol

2012;56:1070-9.

27

6. Dear JW, Antoine DJ, Starkey-Lewis P, Goldring CE, and Park BK. Early detection of paracetamol toxicity using circulating liver microRNA and markers of cell necrosis. Br J Clin Pharmacol 2014;77:904-5.

7. Thulin P, Nordahl G, Gry M, Yimer G, Aklillu E, Makonnen E, *et al.* Keratin-18 and microRNA-122 complement alanine aminotransferase as novel safety biomarkers for drug-induced liver injury in two human cohorts. Liver Int 2014;34:367-78.

8. Schomaker S, Warner R, Bock J, Johnson K, Potter D, Van Winkle J, *et al.* Assessment of emerging biomarkers of liver injury in human subjects. Toxicol Sci 2013;132:276-83.

9. Pelsers MM, Morovat A, Alexander GJ, Hermens WT, Trull AK, and Glatz JF. Liver fatty acid-binding protein as a sensitive serum marker of acute hepatocellular damage in liver transplant recipients. Clin Chem 2002;48:2055-7.

0. Mikus M, Drobin K, Gry M, Bachmann J, Lindberg J, Yimer G, *et al.* Elevated levels of circulating CDH5 and FABP1 in association with human drug-induced liver injury. Liver Int 2016;

1. McMillan HJ, Gregas M, Darras BT, and Kang PB. Serum transaminase levels in boys with Duchenne and Becker muscular dystrophy. Pediatrics 2011;127:e132-6.

2. Flanigan KM, Voit T, Rosales XQ, Servais L, Kraus JE, Wardell C, *et al.* Pharmacokinetics and safety of single doses of drisapersen in non-ambulant subjects with Duchenne muscular dystrophy: results of a double-blind randomized clinical trial. Neuromuscul Disord 2014;24:16-24.

3. Zhang Y, Jia Y, Zheng R, Guo Y, Wang Y, Guo H, *et al.* Plasma microRNA-122 as a biomarker for viral-, alcohol-, and chemical-related hepatic diseases. Clin Chem 2010;56:1830-8.

28

4.  Schmidt ES and Schmidt FW. Glutamate dehydrogenase: biochemical and clinical aspects of an interesting enzyme. Clin Chim Acta 1988;173:43-55.

5.  Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, and Tuschl T. Identification of tissue-specific microRNAs from mouse. Curr Biol 2002;12:735-9.

6.  Holman NS, Mosedale M, Wolf KK, LeCluyse EL, and Watkins PB. Subtoxic Alterations in Hepatocyte-Derived Exosomes: An Early Step in Drug-Induced Liver Injury? Toxicol Sci 2016;151:365-75.

7.  Momen-Heravi F, Bala S, Kodys K, and Szabo G. Exosomes derived from alcohol-treated hepatocytes horizontally transfer liver specific miRNA-122 and sensitize monocytes to LPS. Sci Rep 2015;5:9991.

8.  Rivkin M, Simerzin A, Zorde-Khvalevsky E, Chai C, Yuval JB, Rosenberg N, *et al.* Inflammation-Induced Expression and Secretion of MicroRNA 122 Leads to Reduced Blood Levels of Kidney-Derived Erythropoietin and Anemia. Gastroenterology 2016;151:999-1010 e3.

9.  McGill MR, Sharpe MR, Williams CD, Taha M, Curry SC, and Jaeschke H. The mechanism underlying acetaminophen-induced hepatotoxicity in humans and mice involves mitochondrial damage and nuclear DNA fragmentation. J Clin Invest 2012;122:1574-83.

0.  McGill MR, Staggs VS, Sharpe MR, Lee WM, and Jaeschke H. Serum mitochondrial biomarkers and damage-associated molecular patterns are higher in acetaminophen overdose patients with poor outcome. Hepatology 2014;60:1336-45.

29

1.    Dear JW, Clarke JI, Francis B, Allen L, Wraight J, Shen J, *et al.* Risk stratification after paracetamol overdose using mechanistic biomarkers: results from two prospective cohort studies. Lancet Gastroenterol Hepatol 2017;

2.    Russo MW, Steuerwald N, Norton HJ, Anderson WE, Foureau D, Chalasani N, *et al.* Profiles of miRNAs in serum in severe acute drug induced liver injury and their prognostic significance. Liver Int 2017;37:757-764.

3.    Arai M, Yokosuka O, Fukai K, Imazeki F, Chiba T, Sumi H, *et al.* Gene expression profiles in liver regeneration with oval cell induction. Biochem Biophys Res Commun 2004;317:370-6.

4.    Srungaram P, Rule JA, Yuan HJ, Reimold A, Dahl B, Sanders C, *et al.* Plasma osteopontin in acute liver failure. Cytokine 2015;73:270-6.

5.    Andersson U, Lindberg J, Wang S, Balasubramanian R, Marcusson-Stahl M, Hannula M, *et al.* A systems biology approach to understanding elevated serum alanine transaminase levels in a clinical trial with ximelagatran. Biomarkers 2009;14:572-86.

6.    Stutchfield BM, Antoine DJ, Mackinnon AC, Gow DJ, Bain CC, Hawley CA, *et al.* CSF1 Restores Innate Immunity After Liver Injury in Mice and Serum Levels Indicate Outcomes of Patients With Acute Liver Failure. Gastroenterology 2015;149:1896-1909 e14.

7.    Bechmann LP, Jochum C, Kocabayoglu P, Sowa JP, Kassalik M, Gieseler RK, *et al.* Cytokeratin 18-based modification of the MELD score improves prediction of spontaneous survival after acute liver injury. J Hepatol 2010;53:639-47.

30

Manuscript: HEP-17-1863

8.    Caulin C, Salvesen GS, and Oshima RG. Caspase cleavage of keratin 18 and reorganization of intermediate filaments during epithelial cell apoptosis. J Cell Biol 1997;138:1379-94.

9.    Sauter B, Albert ML, Francisco L, Larsson M, Somersan S, and Bhardwaj N. Consequences of cell death: exposure to necrotic tumor cells, but not primary tissue cells or apoptotic cells, induces the maturation of immunostimulatory dendritic cells. J Exp Med 2000;191:423-34.

0.    Basu S, Binder RJ, Suto R, Anderson KM, and Srivastava PK. Necrotic but not apoptotic cell death releases heat shock proteins, which deliver a partial maturation signal to dendritic cells and activate the NF-kappa B pathway. Int Immunol 2000;12:1539-46.

1.    Church RJ, Kullak-Ublick GA, Kleiner DE, Bonkovsky HL, Chalasani N, Fontana RJ, *et al.* Candidate liver safety biomarkers provide prognostic and mechanistic insights in patients with drug-induced liver injury. Hepatology 2017;66:24A-25A.

2.    FDA. *U.S. Food & Drug Administration Letter of Support Initiative*.   6/15/17]; Available from:
https://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/UCM517355.pdf

3.    EMA. *EMA Letter of support for drug-induced liver injury (DILI) biomarker*.   6/15/17]; Available from:
http://www.ema.europa.eu/docs/en_GB/document_library/Other/2016/09/WC500213479.pdf.

31

**<u>FIGURE LEGENDS</u>**

**Figure 1. Intra-Individual Variabilty in PSTC Cohort Observed in microR-122 (miR-122) Quantifications.** Three fasting blood samples were collected from volunteers over the course of 21 days. Greater intra-subject variability was observed in miR-122 levels amongst black subjects in this study compared to white subjects. Each bar represents an individual subject while circles represent data for miR-122 measurements 1 (white), 2 (black), and 3 (gray).

**Figure 2. Correlation Between Levels of Alanine Aminotransferase (ALT) and Liver-Specific Biomarkers.** Correlation of ALT with glutamate dehydrogenase (GLDH; **A**) and microRNA-122 (miR-122; **B**). Data points are individual PSTC (Predictive Safety Testing Consortium) and SAFE-T (Safer and Faster Evidence-based Translation) subject samples and represent individuals that did not have drug-induced liver injury (DILI; gray) and individuals that did have DILI (black). Values are log normalized. Pearson r is shown.

**Figure 3. Assessment of Keratin 18 (K18) Measurements.** Differences in serum K18 (**A**), caspase cleaved K18 (ccK18; **B**), and Apoptotic Index (AI; **C**) between Drug-Induced Liver Injury Network patients that did not die/require a liver transplant by 6 months post-DILI-onset and those that did. Differences in AI (**D**) between SAFE-T patients experiencing DILI associated with flupirtine utilization and patients experiencing DILI unrelated to flupirtine. Data points represent individual patients. A dotted line for AI is drawn at 0.5, representing a score that suggests an equal contribution of apoptosis and necrosis. Values for K18 and ccK18 are log normalized. Significance is \*\*$p<0.01$ and \*\*\*$p<0.001$.

**Figure 4. Incorporation of Candidate Biomarkers Into Model of End-Stage Liver Disease (MELD) Score Prognostic Model.** Prognostic model optimized for prediction of

32

death/transplant in Drug-Induced Liver Injury (DILIN) patients (n=141) using the MELD score, total keratin 18 (K18), and macrophage colony stimulating factor receptor (MCSFR) measurements (sensitivity=0.933, specificity=0.889). White boxes represent branching points, light grey boxes represent patients not predicted to have an adverse outcome, and dark grey boxes represent patients predicted to die/require transplant. Numbers in italics represent false results (*i.e.* 18 false positives and 1 false negative).

33

## COLLABORATORS

Guruprasad P. Aithal
Daniel J. Antoine
Uwe Bamberger
Thomas Berg
Patricia Chandler
Heidrun Ellinger Ziegelbauer
Neil S. French
Balasz Fülöp
Thomas Knorpp
M Isabel Lucena
Thomas O. Joos
Douglas A. Keller
Johan Lindberg
Eric Maller
Eric Massana
Jessica Mwinyi
An Ngo
Yen Ngo
Kevin B. Park
Hugo Perazzo
Hannes Planatscher
Mercedes Robles-Díaz
Camilla Stephens
Maximilian Schmeding
Nicole Schneiderhan-Marra
Eckart Schott
Joachim Tillner
Christoph Wandel
Stefan Weiler
K. Rajender Reddy
Huiman Barnhart
Jiezhun Gu
Paul H. Hayashi
Jawad Ahmad
Andrew Stolz
Victor Navarro
Jay Hoofnagle

34

## ACKNOWLEDGEMENTS

35

## Table 1. Candidate Biomarkers

| Candidate Biomarker | Physiological Function | Tissue Localization | Potential Utility in DILI | Cohorts Analyzed |
|---|---|---|---|---|
| AFP | plasma protein thought to be the fetal form of albumin | high levels in liver progenitor cells | regeneration (progenitor cells) | DILIN, PSTC, SAFE-T |
| ARG1 | hydrolase enzyme that catalyzes the hydrolysis of arginine to urea and orthinine | high levels in liver; lower levels in erythrocytes, kidney and brain | cell injury/death | DILIN, PSTC, SAFE-T |
| CDH5 | calcium-dependent transmembrane adherens junction protein important for endothelial cell integrity and cell-cell adhesion | broad localization including liver | susceptibility | DILIN, PSTC, SAFE-T |
| FABP1 | protein involved in binding and transport of fatty acid | high levels in the liver; lower levels in kidney and gastrointestinal tract | cell injury/death | DILIN, PSTC, SAFE-T |
| GSTα | phase II detoxification enzyme that catalyzes the conjugation of glutathione with various electrophiles | high levels in the liver and multiple tissues | cell injury/death | DILIN, PSTC, SAFE-T |
| K18 /ccK18 | type I intermediate protein expressed in epithelial cells responsible for cell structure and integrity. Caspase cleavage results in a fragmented form of protein (ccK18). | broad localization including liver | cell injury/death, mechanism | DILIN, PSTC, SAFE-T |
| MCSFR | receptor on macrophages/monocytes for CSF, a cytokine that controls the proliferation, differentiation, and function of macrophages | broad localization including liver | inflammation | DILIN, PSTC, SAFE-T |
| OPN | phosphoprotein involved in migration/infiltration of inflammatory and cancer cells | broad localization including liver | inflammation, regeneration (progenitor cells) | DILIN, PSTC, SAFE-T |
| GLDH | mitochondrial matrix protein that catalyzes the conversion of 2-oxoglutarate to L-glutamate | high levels in the liver; lower levels in kidney and brain | cell death, mechanism | PSTC, SAFE-T |
| LECT2 | protein involved in the recruitment of neutrophils | high expression in the liver; lower expression in the testis | regeneration (hepatocytes) | PSTC, SAFE-T |
| PON1 | HDL-associated enzyme that participates in paraoxonase, arylesterase, and dyazoxonase activites. Useful in diagnosis of NAFLD and NASH when normalized to prothrombin. | produced in liver, released constitutively into circulation | function | PSTC, SAFE-T |
| SDH | enzyme involved in carbohydrate metabolism that converts sorbitol into fructose | high levels in the liver, kidney, and testis; lower levels in multiple tissues | cell injury/death | PSTC, SAFE-T |
| miR-122 | liver-specific miRNA that post-transcriptionally regulates mRNA involved in processes including hepatocyte differentiation and lipid/cholesterol metabolism | high levels in the liver | cell injury/death | PSTC, SAFE-T |

Abbreviations: AFP, alpha fetoprotein; ARG1, arginase 1; CDH5, cadherin 5; FABP1, fatty acid binding protein 1; GST-α, glutathione S transferase alpha; K18, total keratin 18; ccK18, caspase cleaved K18; MCSFR, macrophage colony stimulating factor receptor; OPN, osteopontin; GLDH, glutamate dehydrogenase; LECT2 leukocyte cell derived chemotaxin 2; PON1, paroxonase 1 normalized to prothrombin protein; SDH, sorbitol dehydrogenase; miR-122, microRNA-122; HDL, high-density lipoprotein; NAFLD, non-alcoholic fatty liver disease; NASH, nonalcoholic steatohepatitis; DILIN, Drug-Induced Liver Injury Network; PSTC, Predictive Safety Testing Consortium; SAFE-T, Safer and Faster Evidence-based Translation

**Table 2: Biomarker Reference Intervals in Healthy Volunteers**

| Biomarker | Unit | Matrix | PSTC (n=81) * | | | | | SAFE-T (n=192) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Est. Geometric Mean | LLN<br>Est. 5th Percentile | ULN<br>Est. 95th Percentile | Inter-Subj %CV | Intra-Subj %CV | Est. Geometric Mean | LLN<br>Est. 5th Percentile | ULN<br>Est. 95th Percentile | Inter-Subj %CV |
| AFP | ng/ml | serum | 0.68 | 0.24 | 1.98 | 61.53 | 31.93 | 0.99 | 0.28 | 3.54 | 90.21 |
| ARG1 | ng/ml | serum | 7.63 | 3.00 | 19.46 | 46.03 | 37.46 | 35.97 | 18.38 | 70.38 | 42.57 |
| ccK18 | U/L | serum | 90.65 | 31.59 | 260.16 | 70.97 | 34.76 | 139.99 | 52.46 | 373.55 | 65.39 |
| CDH5 | ng/ml | serum | 2798.89 | 1853.77 | 4225.87 | 18.00 | 17.69 | 2287.79 | 1222.52 | 4281.33 | 39.52 |
| FABP1 | ng/ml | serum | 6.91 | 3.29 | 14.55 | 32.75 | 32.86 | 9.21 | 4.57 | 18.54 | 44.55 |
| GLDH | U/L | serum | 2.71 | 1.01 | 7.24 | 52.74 | 34.53 | 3.00 | 0.95 | 9.51 | 79.68 |
| GSTα | ng/ml | serum | 6.31 | 0.68 | 60.00 | 119.54 | 71.86 | 6.61 | 0.71 | 64.11 | 172.57 |
| LECT2 | ng/ml | plasma | 252.27 | 142.07 | 447.96 | 28.64 | 20.97 | 177.96 | 84.74 | 373.74 | 47.50 |
| miR-122 | copies/μl | serum | 2152.98 | 347.05 | 13356.52 | 90.89 | 93.56 | 3173.64 | 368.02 | 27367.61 | 213.51 |
| MCSFR | ng/ml | plasma | 334.81 | 196.1 | 571.64 | 30.08 | 13.89 | 306.67 | 175.98 | 534.39 | 34.75 |
| OPN | ng/ml | serum | 4.13 | 1.66 | 10.31 | 52.15 | 26.61 | 6.54 | 2.68 | 15.99 | 58.56 |
| PON1 | ng/μg | plasma | 4.91 | 2.02 | 11.93 | 44.16 | 34.57 | 9.44 | 4.18 | 21.35 | 52.81 |
| SDH | U/L | serum | 3.02 | 1.18 | 7.75 | 43.43 | 41.01 | 1.79 | 0.79 | 7.17 | 101.57 |
| K18** | U/L | serum | | | 121.35 | | | | | 151.14 | |

*Three serial collections were collected for each individual. The mean value for each individual was used for all statistical analyses with the exception of intra-individual %CV

**90% of K18 data was below the lower limit of quantification, therefore only an upper reference interval was determined

Abbreviations: PSTC, Predictive Safety Testing Consortium; SAFE-T, Safer and Faster Evidence-based Translation; Est, established; LLN, lower limit of normal; ULN, upper limit of normal; CV, coefficient of variation; CI, confidence interval AFP, alpha fetoprotein; ARG1, arginase 1; ccK18, caspase cleaved K18; CDH5, cadherin 5; FABP1, fatty acid binding protein 1;  GLDH, glutamate dehydrogenase; GSTα, glutathione S transferase alpha; LECT2, leukocyte cell derived chemotaxin 2; miR-122, microRNA-122; MCSFR, macrophage colony stimulating factor receptor; OPN, osteopontin; PON1, paroxonase 1 normalized to prothrombin protein; SDH, sorbitol dehydrogenase; K18, keratin 18

**Table 3: Biomarker Performance in DILI Identification\***

| Category | Biomarker | AUC | 95% CI |
|---|---|---|---|
| Traditional | ALT | 0.990 | 0.984 – 0.996 |
| Traditional | AST | 0.975 | 0.963 - 0.987 |
| Traditional | ALP | 0.902 | 0.873 -0.930 |
| Traditional | TBIL | 0.857 | 0.821 - 0.892 |
| Candidate | K18 | 0.947 | 0.928 - 0.966 |
| Candidate | FABP1 | 0.916 | 0.890 - 0.941 |
| Candidate | ccK18 | 0.911 | 0.887 - 0.935 |
| Candidate | GLDH | 0.907 | 0.870 - 0.945 |
| Candidate | MCSFR\*\* | 0.854 | 0.822 - 0.887 |
| Candidate | miR-122 | 0.831 | 0.779 - 0.883 |
| Candidate | AFP | 0.826 | 0.793 - 0.859 |
| Candidate | GSTα | 0.827 | 0.792 - 0.862 |
| Candidate | SDH | 0.819 | 0.763- 0.876 |
| Candidate | OPN | 0.758 | 0.718- 0.799 |
| Candidate | CDH5 | 0.658 | 0.614 - 0.701 |
| Candidate | PON1 | 0.612 | 0.542 - 0.682 |
| Candidate | ARG1 | 0.564 | 0.519 - 0.609 |
| Candidate | LECT2 | 0.519 | 0.450- 0.588 |

Abbreviations: AUC, area under the curve; CI, confidence interval; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase; TBIL, total bilirubin; K18, keratin 18; FABP1, fatty acid binding protein 1; GLDH, glutamate dehydrogenase; ccK18, caspase cleaved K18; MCSFR, macrophage colony stimulating factor receptor; SDH, sorbitol dehydrogenase; miR-122, microRNA-122; AFP, alpha fetoprotein; GST-α, glutathione S transferase alpha; OPN, osteopontin; PON1, paroxonase 1; CDH5, cadherin 5; ARG1, arginase 1; LECT2, leukocyte cell derived chemotaxin 2; PSTC, Predictive Safety Testing Consortium; SAFE-T, Safer and Faster Evidence-based Translation; DILI, Drug-Induced Liver Injury; DILIN, DILI Network.

\*Statistical data for all biomarkers was calculated using patient data from PSTC healthy volunteers (n=81), SAFE-T healthy volunteers (n=192), SAFE-T subjects that safely received DILI-eliciting compounds (n=147), SAFE-T DILI patients (n=126). DILIN patient data (n=143) was also used for all biomarkers with the exception of GLDH, miR-122, SDH, PON1, and LECT2.

\*\*DILIN measurements collected in serum. All other measurements collected in plasma.

**Table 4. Prognostic DILI Biomarkers for Death/Liver Transplant\***

| Category | Biomarker | AUC | 95% CI | Value at Youden's J* |
|---|---|---|---|---|
| Traditional | INR | 0.920 | 0.864-0.977 | 0.47 |
| Traditional | TBIL | 0.821 | 0.733-0.909 | 5.57 |
| Traditional | AST | 0.7 | 0.587-0.814 | 5.05 |
| Traditional | ALT | 0.606 | 0.433-0.78 | 6.68 |
| Traditional | ALP | 0.597 | 0.433-0.76 | 5.01 |
| Candidate | OPN | 0.858 | 0.759-0.957 | 3.38 |
| Candidate | K18 | 0.832 | 0.737-0.927 | 7.98 |
| Candidate | MCSFR | 0.775 | 0.654-0.896 | 6.94 |
| Candidate | ccK18 | 0.778 | 0.676-0.881 | 6.96 |
| Candidate | FABP1 | 0.721 | 0.608-0.833 | 4.21 |
| Candidate | AFP | 0.687 | 0.566-0.809 | 1.57 |
| Candidate | CDH5 | 0.623 | 0.498-0.748 | 8.01 |
| Candidate | ARG1 | 0.588 | 0.436-0.741 | 3.47 |
| Candidate | GST-α | 0.536 | 0.359-0.713 | 6.88 |
| Candidate | AI | 0.761 | 0.627-0.895 | 0.37 |

\*All values with the exception of AI are log normalized. Youdin's J is a statistic that estimates the probability of an informed decision.

Abbreviations: AUC, area under the curve; CI, confidence interval; INR, international normalized ratio; TBIL, total bilirubin; AST, aspartate aminotransferase; ALT, alanine aminotransferase; ALP, alkaline phosphatase; OPN, osteopontin; K18, cytokeratin 18; MCSFR, macrophage colony stimulating factor receptor; ccK18, caspase cleaved K18; FABP1, fatty acid binding protein 1; AFP, alpha fetoprotein; ARG1, arginase 1; CDH5, cadherin 5; GST-α, glutathione S transferase alpha; AI, apoptotic index
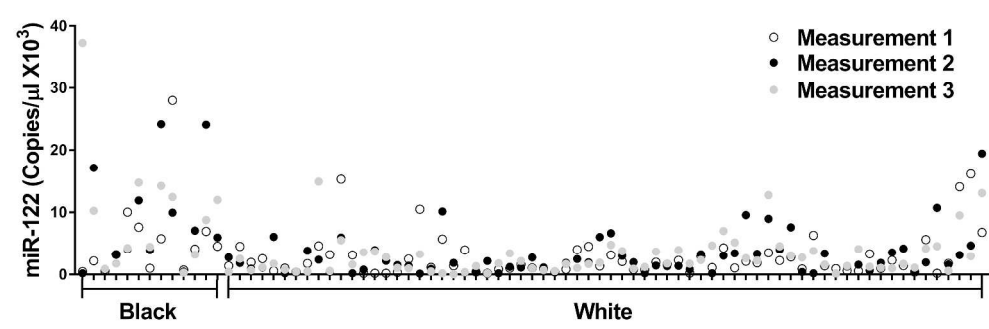
499x170mm (300 x 300 DPI)

**A.**

**B.**
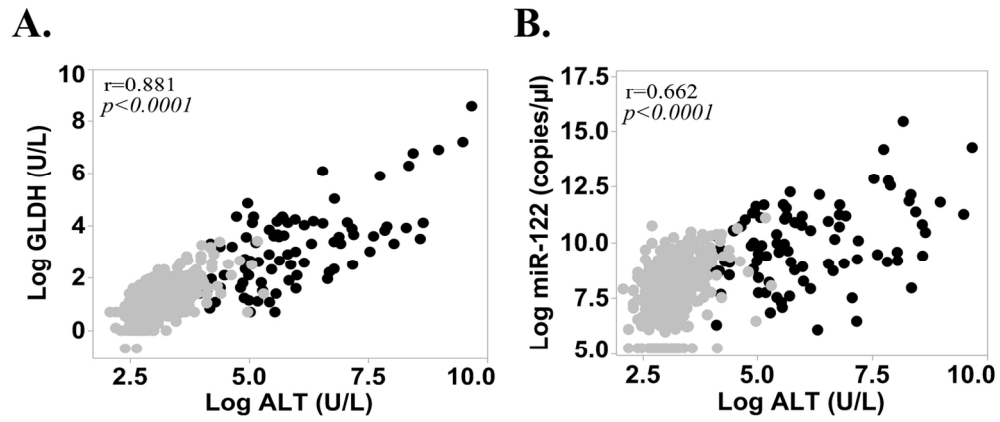


65x28mm (600 x 600 DPI)

**A.**



**B.**



**C.**



**D.**



123x108mm (600 x 600 DPI)

DILIN cases with MELD, K18, and MCSFR
data

N=141

Death/transplant: 15

All others: 126

MELD ≥ 30

N=10

Death/transplant: 9

All others: *1*

MELD 20-29

N=37

Death/transplant: 5

All others: 32

MELD < 20

N=94

Death/transplant: *1*

All others: 93

K18 ≥ 7.98 **AND** MCSFR ≥ 6.94

N=22

Death/transplant: 5

All others*: 17*

K18 < 7.98 **OR** MCSFR < 6.94

N=15

Death/transplant: 0

All others: 15

**Sensitivity: 0.933; Specificity: 0.889**
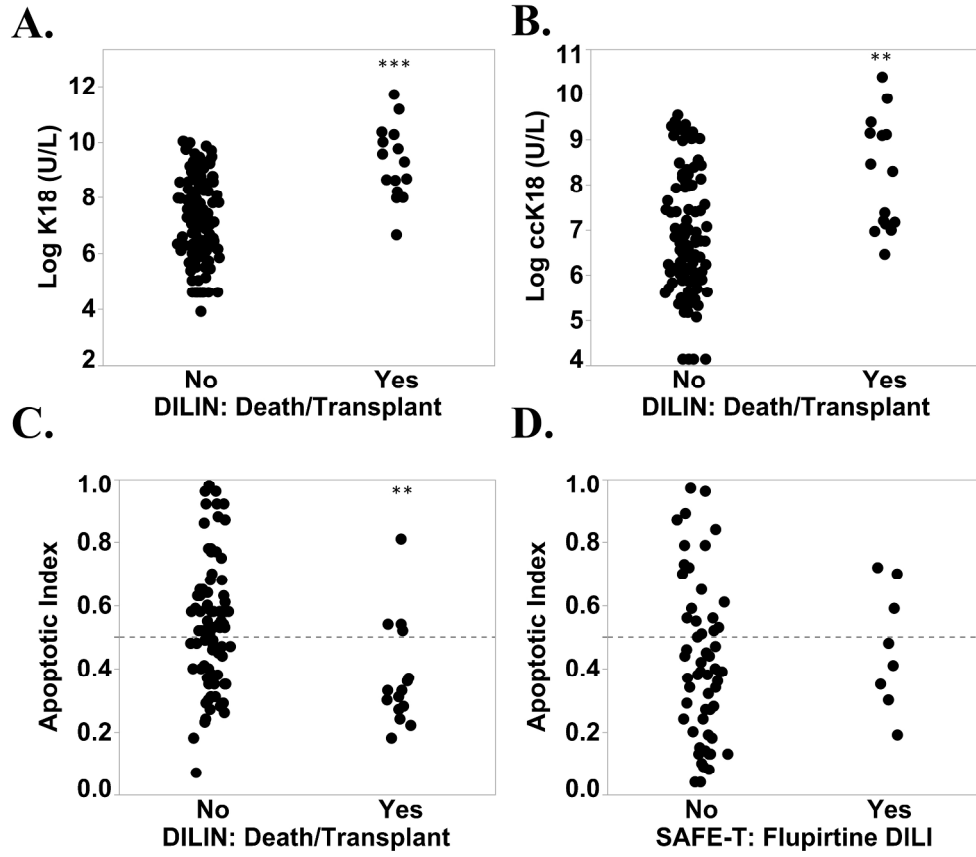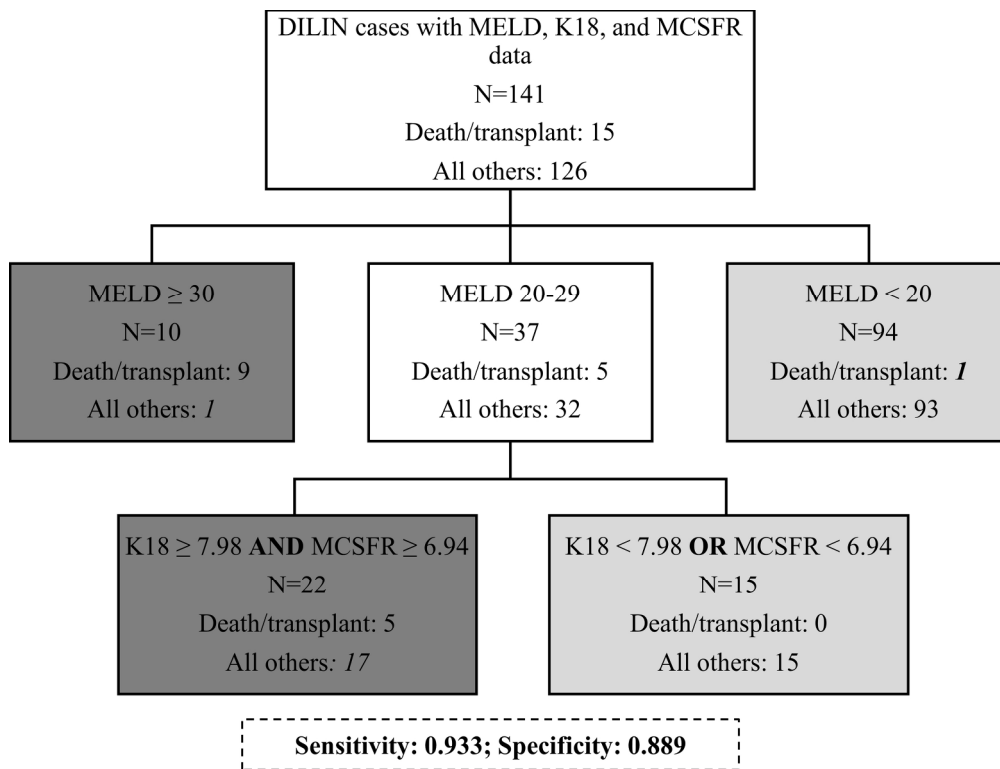
103x78mm (600 x 600 DPI)

**Supplemental Figure 1**
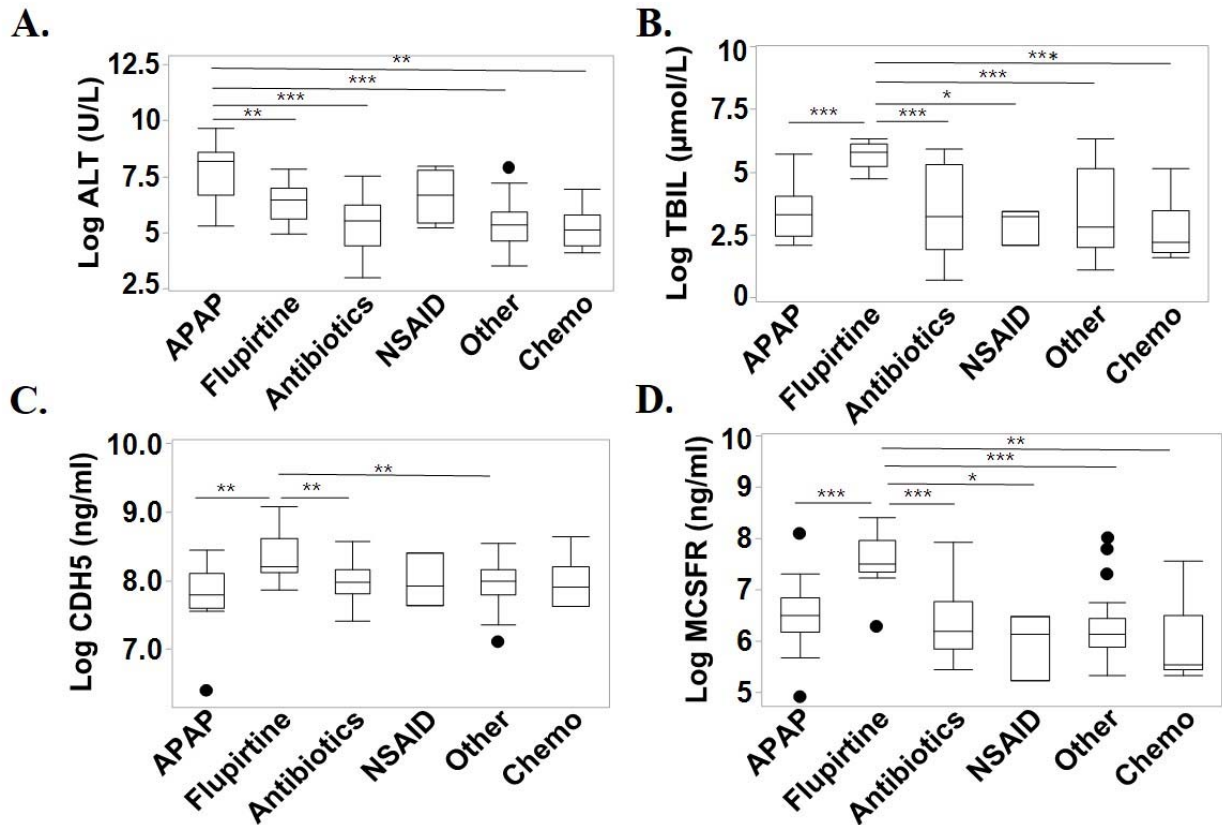


**Supplemental Figure 1: Biomarker Differences by Drug Class in Safer and Faster Evidence-based Translation (SAFE-T) DILI Patients.** Differences in mean alanine aminotransferase (ALT; **A**), total bilirubin (TBIL; **B**), cadherin 5 (CDH5; **C**), and macrophage colony stimulating factor receptor (MCSFR; **D**) between SAFE-T drug-induced liver injury (DILI) patients based on drug classes. Drug classes are acetaminophen (APAP; n=19), flupirtine (n=14), antibiotics (n=35), chemotherapeutics (n=7), non-steroidal anti-inflammatory drugs (NSAIDs; n=4), and others (n=45). The box in each box plot extends from the 25th percentile to the 75th percentile of data values; whiskers extend to minimum and maximum data with data outliers represented by circles. TBIL and CDH5 measurements were collected in serum while MCSFR measurements were collected in plasma. Significance is *$p<0.05$, **$p<0.01$, and ***$p<0.001$.

Manuscript: HEP-17-1863

## SUPPLEMENTAL TABLES

### Supplemental Table 1: PSTC Demographics

| PSTC Healthy Volunteers (N=81) | |
| --- | --- |
| **Age, y, median (IQR)** | 39 (29.5-50.5) |
| **Sex, n (%)** | |
| *Male* | 40 (49.4) |
| *Female* | 41 (50.6) |
| **Race, n (%)** | |
| *White* | 68 (84) |
| *Black* | 13 (16) |
| **BMI (kg/m$^2$), median (IQR)** | 27.8 (23.7-31.35) |
| **Liver biochemistries, median (IQR)** | |
| *ALT (U/L)* | 20 (15.5-28) |
| *AST (U/L)* | 22 (19-25) |
| *ALP (U/L)* | 65 (54.5-76.5) |
| *TBIL (µmol/L)* | 8.55 (6.84-11.97) |

Abbreviations: PSTC, Predictive Safety Testing Consortium; IQR, interquatile range; BMI, body mass index; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase; TBIL, total bilirubin

## Supplemental Table 2: SAFE-T Demographics

| | SAFE-T Study | | | | |
| | Healthy Volunteers | No DILI | | DILI | |
| | Tel Aviv | Protocol 4 | Protocol 5 | Swiss DILI | Protocol 3A |
| | N=192 | N=55 | N=92 | N=28 | N=98 |
|---|---|---|---|---|---|
| **Age, y, median (IQR)** | 52 (42-62) | 29 (24-39) | 52.5 (43.3-61) | 56 (42-66.8) | 53 (38-66.3) |
| **Sex, n (%)** | | | | | |
| *Male* | 103 (53.7) | 32 (58.2) | 31 (33.7) | 15 (53.6) | 41 (41.8) |
| *Female* | 88 (45.8) | 23 (41.8) | 61 (66.3) | 13 (46.4) | 9 (47.4) |
| *Missing* | 1 (0.5) | | | | |
| **Race, n (%)** | | | | | |
| *White* | | 33 (60) | 68 (73.9) | 25 (89.3) | 90 (91.8) |
| *Black* | | 20 (36.4) | 11 (12) | 2 (7.1) | 1 (1.02) |
| *Asian* | | 1 (1.8) | 1 (1.1) | 0 (0) | 5 (5.1) |
| *Other* | | 1 (1.8) | 10 (10.8) | 1 (3.6) | 1 (5.3) |
| *Missing* | | | 2 (2.2) | | |
| **BMI (kg/m²), median (IQR)** | 25.9 (23.1-29.2) | 21.6 (19.3-24.8) | 25.2 (22-29.4) | 24.6 (22-27.4) | 25.7 (23.4- 29) |
| **Liver biochemistries, median (IQR)** | | | | | |
| *ALT (U/L)* | 22 (18-29) | 21.5 (18-35) | 25 (18-32) | 278 (144-1877) | 322 (137.8-884) |
| *AST (U/L)* | 23 (20-26) | 26 (20.8-33) | 26.5 (22-30) | 152 (64-728) | 138.6 (66.5-349) |
| *ALP (U/L)* | 66 (54-81) | 71 (57-9.8) | 62 (46-73.5) | 84.5 (65-246.8) | 181 (101- 254) |
| *TBIL (μmol/L)* | 10.26 (8.6-13.7) | 6.8 (5.1-10.3) | 7 (5-9.8) | 8.5 (5.3-32) | 42 (11.5-247) |
| *INR* | | | | 1.1 (1-1.3) | 1.3 (1-1.6) |
| **Hy's Law, n (%)** | | | | | |
| *No* | | | | 20 (71.4) | 53 (54.1) |
| *Yes* | | | | 4 (14.3) | 35 (35.7) |
| *Missing* | | | | 4 (14.3) | 10 (10.2) |
| **Pattern of Injury, n (%)** | | | | | |
| *Cholestatic* | | | | 6 (21.4) | 5 (5.1) |
| *Mixed* | | | | 1 (3.6) | 24 (24.5) |
| *Hepatocellular* | | | | 21 (75) | 69 (70.4) |
| **R Value, median (IQR)** | | | | 7.6 (1.3-75) | 5.7 (2.3- 27.9) |

Abbreviations: SAFE-T, Safer and Faster Evidence-based Translation; DILI, drug-induced liver injury; IQR, interquartile range; BMI, body mass index; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase; TBIL, total bilirubin; INR, international normalized ratio; Hy's Law (ALT>3X Upper Limit of Normal, ULN, TBIL>2X ULN, ALP<2X ULN)

Manuscript: HEP-17-1863

## Supplemental Table 3: DILIN Demographics

| | DILI Outcome | | | | |
|---|---|---|---|---|---|
| | Recovered N=89 | Unresolved N=19 | Death/Transplant N=15 | Unknown N=20 | *p* |
| Age, y, mean ± S.D | 46.2 ± 16.9 | 43.1 ± 16.1 | 52.6 ± 20.8 | 39.9 ± 15.9 | NS |
| Sex, n (%) | | | | | NS |
| *Male* | 32 (36) | 10 (52.6) | 9 (60) | 11 (55) | |
| *Female* | 57 (64) | 9 (47.4) | 6 (40) | 9 (45) | |
| Race, n (%) | | | | | NS |
| *White* | 65 (73) | 12 (63.2) | 10 (66.7) | 16 (80) | |
| *Black* | 13 (14.6) | 6 (31.5) | 2 (13.3) | 3 (15) | |
| *Asian* | 5 (5.6) | 0 (0) | 2 (13.3) | 0 (0) | |
| *Other* | 6 (6.8) | 1 (5.3) | 1 (6.7) | 1 (5) | |
| Ethnicity, n (%) | | | | | NS |
| *Hispanic* | 13 (14.6) | 2 (10.5) | 0 (0) | 3 (15) | |
| *Non-Hispanic* | 76 (85.4) | 17 (89.5) | 15 (100) | 17 (85) | |
| BMI (kg/m$^2$), mean ± S.D. | 28.5 ± 7.3 | 28.1 ± 9.6 | 27.4 ± 6.6 | 26.5 ± 5.0 | NS |
| Liver biochemistries, median (IQR) | | | | | |
| *ALT (U/L)* | 527 (228.8-1258.5) | 357 (128-1106) | 907 (152-1536) | 247 (106-458.3) | NS |
| *AST (U/L)* | 306 (126.3-755.3) | 290 (71-664) | 865 (220-987) | 130 (63.25-612.3) | 0.01 |
| *ALP (U/L)* | 165 (127.3-323.5) | 216 (173- 327) | 146 (120-297) | 229.5 (152.3-356.8) | NS |
| *TBIL (μmol/L)* | 93.2 (26.5-221.9) | 165.9 (73.5-311.2) | 311.2 (261.6-434.3) | 177.8 (47.9-262.5) | <0.0001 |
| *INR* | 1.1 (1-1.3) | 1.1 (1-1.4) | 3 (1.7-4.4) | 1.0 (0.9-1.2) | <0.0001 |
| Hy's Law, n (%) | | | | | 0.007 |
| *No* | 54 (61.4) | 12 (63.2) | 3 (20) | 15 (75) | |
| *Yes* | 34 (38.6) | 7 (36.8) | 12 (80) | 5 (25) | |
| Pattern of Injury, n (%) | | | | | NS |
| *Cholestatic* | 17 (19.3) | 8 (42.1) | 3 (20) | 10 (50) | |
| *Mixed* | 16(18.2) | 2 (10.5) | 3 (20) | 4 (20) | |
| *Hepatocellular* | 55 (62.5) | 9 (47.4) | 9 (60) | 6 (30) | |
| R Value, median (IQR) | 8.2 (2.3-19.9) | 3.8 (1- 14.7) | 13.7 (3.2-36.6) | 2.2 (0.9-9.3) | NS |
| MELD Score, median (IQR) | 16.1 (103-21.7) | 16.7 (12.2-19.1) | 33.2 (28.9-40) | 17.4 (12.9-20.2) | <0.0001 |

Abbreviations: DILI, drug-induced liver injury; BMI, body mass index; IQR, interquartile range; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALP, alkaline phosphatase; TBIL, total bilirubin; INR, international normalized ratio; Hy's Law (ALT>3X Upper Limit of Normal, ULN, TBIL>2X ULN, ALP<2X ULN); MELD, Model of End-stage Liver Disease

**Supplemental Table 4: Primary Causative Drugs in DILI Patients**

| Primary Causative Drug | Patients (n) | | Drug Class |
|---|---|---|---|
| | SAFE-T | DILIN | |
| Acetaminophen | 19 | | APAP |
| Acetaminophen and Others | 3 | | APAP + Others |
| Acetazolamide | 1 | | Others |
| Allopurinol | | 1 | Others |
| Althiazide | 1 | | Others |
| Amino Acids Nos | | 1 | Others |
| Amiodarone | | 1 | Others |
| Amoxicillin | 1 | 1 | Antibiotics |
| Amoxicillin W/Clavulanic Acid | 9 | 11 | Antibiotics |
| Amphetamines | 1 | | Others |
| Anabolic Agents For Systemic Use | 3 | 4 | Others |
| Anakinra | | 1 | Others |
| Antiinflammatory And Antirheumatic Products, | | 1 | Others |
| Antithymocyte Immunoglobulin | | 1 | Others |
| Asparaginase | | 1 | Chemotherapy |
| Atorvastatin | 5 | | Others |
| Azathioprine | 1 | 1 | Others |
| Azithromycin | 2 | 3 | Antibiotics |
| Baclofen | 1 | | Others |
| Beta-Interferon | 2 | | Others |
| Bortezomib | 1 | | Others |
| Bupropion | | 1 | Others |
| Camellia Sinensis | | 1 | Others |
| Carbamazepine | 1 | 2 | Others |
| Carbohydrates/Proteins/Minerals/Vitamins, Com | | 1 | Others |
| Cefalexin | | 1 | Antibiotics |
| Cefazolin | | 2 | Antibiotics |
| Cefotaxime | | 1 | Antibiotics |
| Ceftriaxone | 2 | 1 | Antibiotics |
| Celecoxib | 1 | | NSAID |
| Centrally Acting Sympathomimetics | | 1 | Others |
| Ciprofloxacin | | 4 | Antibiotics |
| Clarithromycin | | 1 | Antibiotics |
| Cyclophosphamide | 1 | | Chemotherapy |
| Cyclosporine A | 1 | | Others |
| Dantrolene | | 1 | Others |
| Dapsone | | 1 | Antibiotics |
| Daptomycin | | 1 | Antibiotics |
| Darunavir | | 2 | Others |
| Diclofenac | | 2 | NSAID |
| Disulfiram | 1 | 1 | Others |
| Doxepin | 1 | | Others |
| Doxycycline | 1 | 1 | Antibiotics |
| Erythromycin W/Sulfisoxazole | | 1 | Antibiotics |
| Escitalopram | | 2 | Others |

Manuscript: HEP-17-1863

| Primary Causative Drug | Patients (n) | Drug Class | Primary Causative Drug |
|---|---|---|---|
| Etoricoxib | 1 | | NSAID |
| Exemestane | 1 | 1 | Others |
| Fenofibrate | 1 | 2 | Others |
| Fingolimod | 1 | | Others |
| Flavocoxid | | 1 | Others |
| Flucloxacillin | 6 | | Antibiotics |
| Flupirtine | 14 | | Flupirtine |
| Fluvastatin | 1 | | Others |
| Gabapentin | 1 | | Others |
| Herbals | | 7 | Others |
| Hydralazine | | 2 | Others |
| Hydroxycut - Ephedra Free | | 2 | Others |
| Ibuprofen | 2 | | NSAID |
| Imetelstat | | 1 | Chemotherapy |
| Infliximab | | 2 | Others |
| Ipilimumab | | 1 | Chemotherapy |
| Isoniazid | | 12 | Anti-TB |
| Isoniazid/pyrazinamide/rifampin | 1 | | Anti-TB |
| Isoniazid/pyrazinamide/rifampin/ ethambutol | 1 | | Anti-TB |
| Leflunomide | 2 | | Others |
| Letrozole | 1 | | Others |
| Levofloxacin | 1 | 3 | Antibiotics |
| Lisinopril | | 1 | Others |
| Metamizole | 3 | | Others |
| Mercaptopurine | | 2 | Chemotherapy |
| Meropenem | 1 | | Antibiotics |
| Methyldopa | 1 | 2 | Others |
| Micafungin | | 1 | Others |
| Minocycline | | 5 | Antibiotics |
| Montelukast | | 1 | Others |
| Moxifloxacin | | 1 | Antibiotics |
| Mushrooms | 1 | | Others |
| Nefazodone | | 1 | Others |
| Nicotinic Acid | | 3 | Others |
| Nitrofurantoin | 1 | 2 | Antibiotics |
| Octreotide | | 1 | Others |
| Olanzapine | 1 | | Others |
| Oxaliplatin | | 2 | Chemotherapy |
| Oxymethalone | 1 | | Others |
| Pantaprazole | 1 | | Others |
| Pentamidine | 1 | | Others |
| Pegaspargase | | 2 | Chemotherapy |
| Phenprocoumon | 1 | | Others |
| Phenylpropanolamine | | 1 | Others |
| Phenytoin | | 3 | Others |
| Piperacillin Sodium W/Tazobactam | 2 | 1 | Antibiotics |
| Pravastatin | | 1 | Others |

| Primary Causative Drug | Patients (n) | Drug Class | Primary Causative Drug |
|---|---|---|---|
| Prednisolone | 1 | | Others |
| Pregabalin | 1 | 1 | Others |
| Propylthiouracil | | 2 | Others |
| Quetiapine | | 2 | Others |
| Rifampin | 1 | | Anti-TB |
| Several Antibiotics | 5 | | Antibiotics |
| Several Chemoterapeutics | 5 | | Chemotherapy |
| Simvastatin | 1 | 1 | Others |
| Sulfamethoxazole W/Trimethoprim | 1 | 11 | Antibiotics |
| Sulfasalazine | | 1 | Others |
| Tacrolimus | 1 | | Others |
| Temozolomide | 1 | | Chemotherapy |
| Terbinafine | 1 | | Others |
| Thiamazole | 1 | | Others |
| Valaciclovir | | 1 | Others |
| Valproic Acid | | 1 | Others |
| Other | | 3 | Others |
| Query Outstanding | | 1 | |

Abbreviations: SAFE-T, Safer and Faster Evidence-based Translation; DILIN. Drug-Induced Liver Injury Network; APAP, acetaminopen

## Supplemental Table 5: Biomarker Validation Data

| Analyte | Type of Assay | Sample Matrix analyzed | unit | LOD | LLoQ | ULoQ | intra-assay precision (% CV) | inter-assay precision (% CV) | dilutional linearity of high conc sample | Spike-in recovery (%) | short term stability (24h at RT and 4°C) | F/T stability, 3 cycles |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ccK18 | ELISA | Serum | U/L | 16.2 | 62.5 | 1000 | 2.2 | 5.7 - 7.9 | up to 1:16 | 112 - 118 | yes | yes |
| K18 | ELISA | Serum | U/L | 20 | 100 | 5000 | 3.7 | 6.1 - 9.4 | up to 1:32 | 83 - 107 | yes | yes |
| GLDH | Activity Assay | Serum | U/L | 0.3 | 1 | 80 | 0.4 - 7.7 | 1.5 - 6.4 | 1:4 - 1:256 | ND | yes, > 6h | yes |
| GSTα | Immunoassay | Serum | ng/mL | 1.79 | 1.82 | 373 | 1 - 14 | 11-Sep | 1:5 - 1:10 | 77 - 94 | yes | yes |
| AFP | Immunoassay | Serum | ng/mL | 0.367 | 0.367 | 584 | 16-Feb | 13-Jul | 1:5 - 1:40 | 99 - 106 | yes | yes |
| ARG1 | Immunoassay | Serum | ng/mL | 1.6 | 7.4 | 800 | 6.4 - 11.9 | 4.3 - 15.7 | 1:4 - 1:256 | 84 - 88 | yes | yes |
| OPN | Immunoassay | Serum | ng/mL | 1.25 | 1.25 | 1149 | 5-Jan | 6 - 11 | 1:5 - 1:10 | 81 - 85 | yes | yes |
| SDH | Activity Assay | Serum | U/L | 0.3 | 0.5 | 50 | 0.6 - 10.6 | 1.7 - 13.4 | up to 1:32 | ND | yes, > 6h | yes |
| miR-122 | RT-qPCR | Serum | copies/µL | ND | 384 | 5089837 | 1.3 - 12.1 | 0.5 - 25.4 | ND | ND | 2h RT, 5h 4°C | yes |
| FABP1 | Immunoassay | Serum | pg/mL | 3.1 | 15.6 | 16000 | 5.6 | 6.7 - 18.1 | 1:2 - 1:2048 | 110 - 115 | yes | yes |
| CDH5 | ELISA | Serum | ng/mL | 0.36 | 3.13 | 100 | 6 | 4.7 - 7.2 | 1:40 - 1:640 | 50 - 83 | yes | yes |
| MCSFR | Immunoassay | EDTA-Plasma | pg/mL | 170 | 600 | 10000 | 1.1 - 13.9 | 8.0 - 28.0 | up to 1:3,200 | 71 -79 | yes | yes |
| PON1 | Immunoassay | EDTA Plasma | ng/mL | 0.06 | 0.35 | 600 | 5.9 | 8.3 - 12.3 | 1:20 - 1:160 | 64 - 82 | 4h RT, 24h 4°C | yes |
| Prothrombin | Immunoassay | EDTA Plasma | µg/mL | 0.8 | 1.92 | 200 | 4.7 | 1.7 - 4.5 | 1:40 - 1:320 | 79 - 108 | yes | yes |
| LECT2 | Immunoassay | EDTA Plasma | ng/mL | 2 | 5.56 | 300 | 7.8 | 11.7 - 12.6 | 1:40 - 1:1.280 | 94 - 118 | yes | yes |

Abbreviations: limit of detection (LOD) , lower limit of quantification (LLoQ),upper limit of quantification (ULoQ), coefficient of variability (CV), concentration (conc), hours (h), room termperature (RT), freeze/thaw (F/T), total cytokeratin 18 (K18), caspase cleaved cytokeratin 18 (ccK18), glutamate dehydrogenase (GLDH), not determined (ND), glutathione-S-transferase α (GSTα), alpha fetoprotein (AFP), arginase 1 (ARG1), osteopontin (OPN), sorbitol dehydrogenase (SDH), microRNA-122 (miR-122), reverse transcription quantitative real-time PCR (RT-qPCR), liver fatty acid binding protein (L-FABP), cadherin 5 (CDH5), macrophage colony stimulating factor receptor (M-CSF-R), paroxonase 1 (PON1, normalized to prothrombin protein), leukocyte cell-derived chemotaxin 2 (LECT2)

| Supplemental Table 6: Biomarker Alterations in Augmentin-related DILI | | | |
|---|---|---|---|
| **Biomarker** | **Mean Biomarker Value (Ln)** | | *p* |
| | **SAFE-T** | **DILIN** | |
| ALT (U/L) | 4.67 | 5.46 | 0.048 |
| ARG1 (ng/ml) | 2.83 | 3.52 | 0.033 |
| FABP1 (ng/ml) | 2.82 | 3.82 | 0.04 |
| ccK18 (U/L) | 5.75 | 6.29 | 0.028 |

Abbreviations: DILI, drug-induced liver injury; SAFE-T, safer and faster evidence-based translation; DILIN, DILI network; ARG1, arginase 1; FABP1, fatty acid binding protein 1; ccK18, caspase cleaved keratin 18

**Supplemental Table 7: Biomarker Geometric Means of Healthy Volunteers and DILIN Patients**

| Biomarker | Geometric Mean | | | Fold Change | |
|---|---|---|---|---|---|
| | HV | No Death/Trans | Death/Trans | HV vs. Death/Trans | No Death/Trans vs. Death/Trans |
| OPN (ng/ml) | 5.75 | 14.17 | 41.01 | 7.13 | 2.89 |
| K18 (U/L) | 68.44 | 1358.73 | 10481.29 | 153.15 | 7.71 |
| MCSFR (ng/ml) | 315.4 | 883.93 | 2240.95 | 7.11 | 2.54 |
| ccK18 (U/L) | 121.83 | 978.14 | 3636.49 | 29.85 | 3.72 |
| FABP1 (ng/ml) | 8.54 | 50.14 | 133.7 | 15.66 | 2.67 |
| AFP (ng/ml) | 0.9 | 4.47 | 10.32 | 11.47 | 2.31 |

Abbreviations: DILIN, drug-induced liver injury network; HV, healthy volunteer; trans, transplant; OPN, osteopontin; K18, total keratin 18; MCSFR, macrophage colony stimulating factor receptor; ccK18, caspase cleaved keratin 18; FABP1, fatty acid binding protein 1; AFP, alpha fetoprotein

**Supplemental Table 8: Comparison of Prediction Models for Death/Liver Transplant**

| Model | Sensitivity | Specificity | PPV | NPV | *p* |
|---|---|---|---|---|---|
| Hy's Law | 0.8 | 0.634 | 0.207 | 0.964 | 0.0054 |
| MELD score ≥ 20 | 0.933 | 0.738 | 0.298 | 0.989 | <0.0001 |
| MELD score ≥ 30 | 0.6 | 0.992 | 0.9 | 0.954 | <0.0001 |
| Modified Hy's Law* | 0.733 | 0.611 | 0.183 | 0.951 | 0.0303 |
| ALF Algorithm* | 0.533 | 0.817 | 0.258 | 0.936 | 0.0075 |
| MELD + K18/MCSFR | 0.933 | 0.889 | 0.5 | 0.991 | <0.0001 |

Abbreviations: PPV, positive predictive value; NPV, negative predictive value; Hy's Law (ALT>3X Upper Limit of Normal, ULN, TBIL>2X ULN, ALP<2X ULN); MELD, Model of End-stage Liver Disease; ALF, acute liver failure; K18, total keratin 18; MCSFR, macrophage colony stimulating factor receptor.

*Robles-Diaz M, Lucena MI, Kaplowitz N, Stephens C, Medina-Caliz I, Gonzalez-Jimenez A, *et al.* Use of Hy's law and a new composite algorithm to predict acute liver failure in patients with drug-induced liver injury. Gastroenterology 2014;147:109-118 e5.

**Supplemental Table 9: Prognostic Biomarkers for Unresolved DILI at 6 Months Post-Onset**

| Category | Biomarker | AUC | 95% CI |
| --- | --- | --- | --- |
| Traditional | ALP | 0.67 | 0.562-0.777 |
| Traditional | TBIL | 0.629 | 0.497-0.761 |
| Traditional | ALT | 0.544 | 0.39-0.7 |
| Traditional | INR | 0.528 | 0.376-0.679 |
| Traditional | AST | 0.516 | 0.369-0.664 |
| Candidate | GST-α | 0.633 | 0.485-0.78 |
| Candidate | ARG1 | 0.614 | 0.48-0.747 |
| Candidate | ccK18 | 0.58 | 0.442-0.719 |
| Candidate | OPN | 0.562 | 0.436-0.688 |
| Candidate | FABP1 | 0.562 | 0.418-0.706 |
| Candidate | CDH5 | 0.539 | 0.406-0.673 |
| Candidate | AFP | 0.538 | 0.397-0.679 |
| Candidate | K18 | 0.519 | 0.374-0.664 |
| Candidate | MCSFR | 0.516 | 0.385-0.647 |
| Candidate | AI | 0.53 | 0.359-0.702 |

*All values with the exception of AI are log normalized

Abbreviations: AUC, area under the curve; CI, confidence interval; INR, international normalized ratio; TBIL, total bilirubin; AST, aspartate aminotransferase; ALT, alanine aminotransferase; ALP, alkaline phosphatase; OPN, osteopontin; K18, cytokeratin 18; MCSFR, macrophage colony stimulating factor receptor; ccK18, caspase cleaved K18; FABP1, fatty acid binding protein 1; AFP, alpha fetoprotein; ARG1, arginase 1; CDH5, cadherin 5; GST-α, glutathione S transferase alpha; AI, apoptotic index

**Supplemental Table 10: Biomarker Correlation with Days Between Symptom Onset and Biospecimen Collection in DILIN Patients**

| Biomarker | Pearson's r |
|---|---|
| AFP | 0.077 |
| ARG1 | -0.014 |
| CDH5 | 0.222 |
| K18 | 0.142 |
| ccK18 | 0.212 |
| FABP1 | -0.002 |
| GST-α | -0.048 |
| MCSFR | 0.163 |
| OPN | 0.158 |

Abbreviations: DILIN, drug-induced liver injury network AFP, alpha fetoprotein; ARG1, arginase 1; CDH5, cadherin 5; K18, cytokeratin 18; ccK18, caspase cleaved K18  FABP1, fatty acid binding protein 1; GST-α, glutathione S transferase alpha; MCSFR, macrophage colony stimulating factor receptor; OPN, osteopontin;

## SUPPLEMENTAL METHODS

### PSTC Healthy Volunteers

All subjects in this cohort (n=81; **Supplemental Table 1**) were recruited at the Jasper Clinic, Inc., Kalamazoo, MI, USA. Three fasting blood samples (n=243 total samples) were collected from 81 subjects over 21 days. Inclusion criteria included age between 18 and 70 years, no underlying medical conditions or use of chronic medications, and a body mass index (BMI) < 35 (kg/m$^2$) (two exceptions with BMIs of 35.3 and 37.6 kg/m$^2$). Exclusion criteria included a positive test for human immunodeficiency virus, and/or active hepatitis B or hepatitis C viral infections, a medical intervention performed within three months of study enrollment, a positive pregnancy test, or unwillingness to refrain from illicit drug/alcohol/tobacco use or strenuous exercise during the study.

### SAFE-T Healthy Volunteers

Biomarker measurements from subjects in this cohort (n=192; **Supplemental Table 2**) were taken from a single fasting blood sample collected between 7 and 9 a.m. at the Tel Aviv Sourasky Medical Center, Tel Aviv Israel. All subjects were asymptomatic and in good health. They completed a detailed epidemiological questionnaire and underwent a thorough analysis of life style by a trained nutritionist. Subjects were interviewed regarding their personal and family history and underwent a comprehensive physical examination. Female subjects underwent a breast and pelvic exam by a senior surgeon and mammography was performed at age > 40 years. Heavy smokers (>20 packs/year) were offered a computed tomography scan. Men > 40 years were tested for total and free prostate-specific antigen. Further diagnostic tests were performed as needed based on the

initial screening results. Exclusion criteria included heavy alcohol intake, a history of renal or liver diseases, and a personal or family history of cancer.

**SAFE-T DILI Patients**

The clinical studies analyzed in this manuscript can be divided into (i) protocols that recruited patients diagnosed with DILI (**Supplemental Table 2**, "DILI") and (ii) protocols that recruited patients who safely took a known DILI-eliciting compound and who were prospectively monitored for several months without evidence of liver injury (**Supplemental Table 2**, "No DILI"). Fasting blood samples were collected. The SAFE-T criteria for adjudicating suspected DILI cases have been described elsewhere [1]. With few exceptions, DILI patients fulfilled the consensus criteria for DILI as previously published [2, 3].

SAFE-T DILI Patients:

*Swiss DILI study*: This study was an 8-week single-center follow-up study investigating the prognostic value of new biomarkers in patients with DILI and included 28 patients adjudicated as having DILI. None of the patients included from this protocol died/required a liver transplant during the observation period. It is unknown whether these patients developed chronic DILI.

*Protocol 3A*: This study was a 12-week multi-center follow-up study investigating the prognostic value of new biomarkers in patients with DILI and included 98 patients adjudicated as DILI. None of the patients included from this protocol died/required a liver transplant during the observation period. It is unknown whether these patients developed chronic DILI.

SAFE-T Drug-exposed No DILI Patients:

*Protocol 4:* This study was a 9-month single-center follow-up study in tuberculosis patients (n=55) starting anti-tuberculosis drug therapy. None of the patients enrolled in this protocol developed

DILI [ALT >5X upper limit of normal (ULN)] during the observation period. Biomarker measurements were made in samples collected at a time point after the patients had begun taking compound (time ranged from 1-6 months on compound).

*Protocol 5:* This study was a 3-year single-center follow-up study in rheumatoid arthritis patients and 92 patients were included in this analysis. None of the patients enrolled in this protocol developed DILI (ALT >5X ULN) during the observation period. When possible biomarker measurements were made in samples collected at time points after patients had begun taking compound (time ranged from 6-30 months on compound); however, only a baseline sample was available for some of these individuals (n=26).

**DILIN Patients**

DILIN prospectively collects clinical, laboratory, imaging, and histopathological information as well as biospecimens from patients within 6 months of suspected DILI onset at multiple centers across the United States (**Supplemental Table 3**). The criteria utilized for DILI assessment in this network has been described in detail elsewhere [4]. The current study assessed biomarkers in 143 samples and included only patients with probable, highly likely, or definite DILI and a blood sample collected within two weeks of DILI onset. Within this cohort, 15 patients died/required a liver transplant within 6 months of onset because of their DILI. Following a readjudication process, DILI was deemed to be the primary factor in all of these patients [5]. Additionally, 19 patients had unresolved DILI (persistently elevated ALT, AST, ALP, or TBIL in the absence of a competing etiology) at 6 months following onset. Of the remaining patients, 89 had recovered by their 6

months follow up visit and 20 did not return for a follow up visit, therefore it is unknown whether their liver injury had completely resolved.

### *Biomarker Quantification*

Predictive Safety Testing Consortium (PSTC) and Safer and Faster Evidence-based Translation (SAFE-T) biomarker measurements were made in either serum or EDTA-plasma (plasma) depending on which matrix was determined to be better suited for the assay. Leucocyte cell derived chemotaxin 2 (LECT2), macrophage colony stimulating factor receptor (MCSFR), and paraoxonase 1 (PON1; normalized to prothrombin protein) were quantified in plasma. All other biomarkers were quantified in serum. All Drug-Induced Liver Injury Network (DILIN) biomarker measurements were made in serum samples. Of the subset of biomarkers measured in all datasets (due to limitations on sample volume, only 9/14 biomarkers were examined in DILIN patients), the matrix for MCSFR differed between cohorts because of sample availability. For all analytes, no international reference standard was available and the measured concentrations were calculated based on individual standard proteins used for the assay calibration.

Traditional biomarkers alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), total bilirubin (TBIL), and international normalized ratio (INR) were measured at the local institutional clinical laboratories and were not obtained from stored samples. Samples utilized for candidate biomarker analyses were taken from archived samples stored at $\leq$ -70°C. Measurements were made at Natural and Medical Sciences Institute (NMI; Reutlingen, Germany) or at contract research laboratories. Briefly, ELISA assays were used to measure total keratin 18 (K18), caspase cleaved K18 (ccK18; VLVbio, Stockholm, Sweden) and cadherin 5 (CDH5; R&D Systems, Minneapolis, MN). Sandwich immunoassays were used to measure glutathione S transferase alpha (GSTα), alpha fetoprotein (AFP), osteopontin (OPN; optimized

Luminex assays from Myriad RBM, Austin, Texas) and arginase 1 (ARG1), MCSFR, PON1, prothrombin protein, fatty acid binding protein 1 (FABP1), and LECT2 (assays developed by NMI, Germany). Colorimetric applications for glutamate dehydrogenase (GLDH) activity (Roche Diagnostics, Grenzach-Wyhlen, Germany) and sorbitol dehydrogenase (SDH) activity (Sekisui Diagnostics, Lexington, MA, USA) were run on a Roche P. Modular Analyzer. PON1 was normalized to prothrombin protein because evidence suggests that this normalization method enables distinction from nonalcoholic steatohepatitis and nonalcoholic fatty liver disease [6]. Absolute quantification of microRNA-122 (miR-122) was analyzed by reverse transcription quantitative real time PCR utilizing standard reagents and real time hydrolysis probes (Life Technologies, Grand Island, New York). Differences in RNA extraction efficiency from individual serum samples were compensated for by adding a synthetic non-human miR (mmu-miR-293) to all samples prior to extraction. All PCR analyses were performed on 192.24 Dynamic Array IFC (Fluidigm). Cq values were calculated by averaging the technical triplicate Cq values, normalized by the average Cq value of the spiked mmu-miR-293 and total miR-122 copy numbers/µL were calculated.

When a biomarker value fell below the lower limit of quantification (LLoQ), that value was used as LLoQ/2.

All commercial assay kits were run according to manufacturer's recommended protocols. All non-clinical assays used for analysis of sample sets which were performed at NMI or contract research organizations were validated following a fit-for-purpose approach considering usual guidelines.

Validation of each assay was approved by a dedicated team within the SAFE-T consortium before the assays were released for sample screenings.

When permissable, an apoptotic index of injury (AI) was calculated from patient data utilizing the ratio of ccK18 to K18. Evidence has demonstrated that this ratio is only meaningful when ccK18 and K18 are above background threshhold levels [7]. In the current study, the following rules were set to establish when calculation of an AI was appropriate: a) K18 $\geq$ 500 U/L b) ccK18 $\geq$ 200 U/L c) K18 > ccK18. Using these rules, an AI was calculated for 98 DILIN patients and 64 SAFE-T DILI patients. Significance was determined by logistic regression and was considered $p < 0.05$.

### *Biomarker Differences by Drug Class*

To determine if one or more DILI compounds/classes produces signature biomarker changes that are unique compared to APAP-related DILI, SAFE-T DILI patient data were divided into broad drug classes. Data was divided as follows: APAP (n=19), flupirtine (n=14), antibiotics (n=35), chemotherapeutics (n=7), non-steroidal anti-inflammatory drugs (NSAIDs; n=4), and others (n=45). When a primary causative drug was uncertain, data were excluded (n=2). Biomarker differences in drug classes were determined in SAFE-T DILI data and DILIN patient data utilizing a one way ANOVA and Wilcoxon multiple comparison correction.

Additionally, cohort differences in patients with DILI related to amoxicillin with clavulanic acid (Augmentin) was examined between DILIN (n=11) and SAFE-T (n=9). Differences were determined using a Wilcoxon test.

### *Prognostic Model Generation*

The performance of current DILI outcome prediction models including Hy's Law, Model for End Stage Liver Disease (MELD) $\geq$ 20, MELD $\geq$ 30, along with a modified version of Hy's Law and

a novel model proposed to predict acute liver failure in DILI patients [8] were explored in the current DILIN patient cohort. Patients were assigned a binary label based on whether or not they met model criteria. Hy's Law criteria was met if patients had ALT $\geq$ 3X upper limit of normal (ULN), TBIL $\geq$ 2X ULN, and ALP < 2X ULN. A MELD score for each patient was calculated as previously described [9]. Concurrent sodium levels were not utilized in this calculation. Performance characteristics (sensitivity, specificity, positive predictive value, PPV, and negative predictive value, NPV) were determined and a contingency table and Fisher's exact test were used to establish significance.

We were interested in determining if candidate DILI biomarkers added value to predictions of death/transplant made with traditional biomarkers. The statistical literature and related data suggest that at least 10 cases are needed for every covariate in a logistic regression prediction model to avoid over-fitting; otherwise, the parameter estimators will be unstable, the covariates in the model may represent noise instead of the true effects of underlying risk factors, and precision of parameter estimators will be poor [10]. Because only n=15 patients in this cohort required a liver transplant or died as a result of DILI, construction of a predictive model using only the biomarker data from this study was not attempted. Instead, we sought to determine if incorporation of any candidate biomarkers could improve the performance of common or previously described predictive models (that use traditional biomarker data). To reduce the number of candidate biomarkers being examined in this analysis, only biomarkers considered predictive of outcome (AUC and lower tail of 95% CI both > 0.5) were carried forward. Predictive biomarkers were then used to construct a correlation matrix and Pearson's r for each biomarker combination was

determined. If any biomarkers were found to be highly correlated, only the biomarker with the greatest AUC generated in ROC curve analysis was carried forward.

Novel biomarkers were also incorporated into a model that utilized MELD score, given that MELD $\geq 20$ was the most sensitive prediction model and MELD $\geq 30$ was the most specific prediction model. Because most of the "false" tests when using MELD score were observed when a patient's MELD score was between 20 and 30, we determined if adding novel biomarker quantifications to this subset of patients would improve the MELD performance. Any patient with a MELD score <20 was considered "recovered." And patient with a MELD score $\geq$30 was considered "adverse." Using the biomarkers that passed our earlier filters (K18, OPN, MCSFR, and AFP) we first determined the single biomarker that best impoved the specifity of the MELD score model (one "adverse" patient had a MELD score <20, therefore the sensitivity of the model could not be further improved using this approach) without negatively affecting the sensitivity. Once this biomarker was identified, we determined if adding a second biomarker could improve the specificity further.

This analysis was performed using data from 141 DILIN patients (2 patient had missing laboratory values). Youden's J is a single statistic that estimates the probability of an informed decesion and captures the performance of a binary test. Therefore, The value corresponding to the best Younden's J for each biomarker was used as an unbiased cut-off threshold for calling the outcome of each patient. Receiver operating characteristic (ROC) curve performance characteristics were examined when each biomarker was added alone or when a combination of the candidate biomarkers was incorporated. The combination of biomarkers that gave the best performance (K18 and MCSFR) was reported.

## Supplemental References

1. Mikus M, Drobin K, Gry M, Bachmann J, Lindberg J, Yimer G, *et al.* Elevated levels of circulating CDH5 and FABP1 in association with human drug-induced liver injury. Liver Int 2016;

2. Aithal GP, Watkins PB, Andrade RJ, Larrey D, Molokhia M, Takikawa H, *et al.* Case definition and phenotype standardization in drug-induced liver injury. Clin Pharmacol Ther 2011;89:806-15.

3. Fontana RJ, Watkins PB, Bonkovsky HL, Chalasani N, Davern T, Serrano J, *et al.* Drug-Induced Liver Injury Network (DILIN) prospective study: rationale, design and conduct. Drug Saf 2009;32:55-68.

4. Fontana RJ, Hayashi PH, Gu J, Reddy KR, Barnhart H, Watkins PB, *et al.* Idiosyncratic drug-induced liver injury is associated with substantial morbidity and mortality within 6 months from onset. Gastroenterology 2014;147:96-108 e4.

5. Hayashi PH, Rockey DC, Fontana RJ, Tillmann HL, Kaplowitz N, Barnhart HX, *et al.* Death and liver transplantation within 2 years of onset of drug-induced liver injury. Hepatology 2017;66:1275-1285.

6. Bell LN, Theodorakis JL, Vuppalanchi R, Saxena R, Bemis KG, Wang M, *et al.* Serum proteomics and biomarker discovery across the spectrum of nonalcoholic fatty liver disease. Hepatology 2010;51:111-20.

7. Linder S, Olofsson MH, Herrmann R, and Ulukaya E. Utilization of cytokeratin-based biomarkers for pharmacodynamic studies. Expert Rev Mol Diagn 2010;10:353-9.

8. Robles-Diaz M, Lucena MI, Kaplowitz N, Stephens C, Medina-Caliz I, Gonzalez-Jimenez A, *et al.* Use of Hy's law and a new composite algorithm to predict acute liver failure in patients with drug-induced liver injury. Gastroenterology 2014;147:109-118 e5.

Manuscript: HEP-17-1863

9.    Kamath PS and Kim WR. The model for end-stage liver disease (MELD). Hepatology

       2007;45:797-805.

10.   Peduzzi P, Concato J, Kemper E, Holford TR, and Feinstein AR. A simulation study of

       the number of events per variable in logistic regression analysis. J Clin Epidemiol

       1996;49:1373-9.