SAFE-T consortium

Summary Data Package
Novel clinical biomarkers of Drug-Induced Liver Injury

The Drug induced liver injury work package of Innovative Medicines Initiative SAFE-T Consortium and The Hepatotoxicity Working Group of Critical Path Institutes PSTC

Document type: DILI BM Summary Data Package

Document status: Final

Release date: 30-Sept-2016

Table of contents

List of tables

List of figures

List of abbreviations

ALP	Alkaline phosphatase
ALT	Alanine transaminase
ALF	Acute Liver Failure
APAP	Acetaminophen
AST	Aspartate transaminase
AUROC	Area under Receiver Operating Characteristic
BM	Biomarker
BSEP	Bile Salt Efflux Pump
ccK18	Caspase-cleaved keratin 18
CIOMS	Council for International Organizations of Medical Science
CLSI	Clinical and Laboratory Standards Institute
CoU	Context of use
DILI	Drug-induced liver injury
DILIN	Drug-induced Liver Injury Network
EMA	European Medicines Agency
FDA	Food and Drug Administration
GGT	Gamma glutamyl transferase
GLDH	Glutamate Dehydrogenase
GSTα1	Glutathione S-transferase 1
HMGB1	High mobility group box 1
HPD	4-hydroxyphenyl pyruvate dehydrogenase
HV	Healthy Volunteer
HWG	Hepatotoxicity Working Group
IMI-JU	Innovative Medicines Initiative-Joint Undertaking
IPRG	Interdisciplinary Pharmacogenomic Review Group
K18	Total Keratin 18
KCC	King's College Criteria
LECT2	Leukocyte cell-derived chemotaxin 2
LLoQ	Lower limit of quantification
LoS	Letter of Support
MCSFR1	Macrophage colony stimulating factor receptor 1

MELD	Model for End-stage Liver Disease
Micro RNA122	Micro-ribonucleic acid 122
mRNA	Messenger ribonucleic acid
MSMS	Tandem mass spectrometry
OLT	Orthotopic liver transplantation
OPN	Osteopontin
PELD	Pediatric End-stage Liver Disease
QC	Quality Control
PSTC	Predictive Safety Testing Consortium
ROC	Receiver Operating Characteristic
RUCAM	Roussel Uclaf Causality Assessment Method
SAFE-T	Safer and Faster Evidence-based Translation
SAWP	Scientific Advice Working Party
SDH	Sorbitol dehydrogenase
SDTM	Study Data Tabulation Model
ST6Gal1	Beta-galactoside alpha-2,6-sialyltransferase 1
UPLC	Ultra-Performance Liquid Chromatography
TBIL	Total bilirubin
WP	Work Package

1 Introduction

1.1 Overview of SAFE-T and PSTC

The Safer and Faster Evidence-based Translation (SAFE-T) Consortium, a non-profit, public private partnership, started its work in June 2009 under the European (EU) Innovative Medicines Initiative-Joint Undertaking (IMI-JU). The objective of the IMI-JU is to support projects for the development of tools and methodologies to address key "bottlenecks" in the pharmaceutical research and development process, similar to the Food and Drug Administration (FDA)'s Critical Path Initiative. The overall objective of the IMI SAFE-T consortium is the regulatory qualification of clinical safety biomarkers of drug-induced injury to three organs; kidney (DIKI), liver (DILI) and vasculature (DIVI) in humans using peripheral samples such as blood and urine (1).

The Predictive Safety Testing Consortium (PSTC) was formed in 2006, and brings together pharmaceutical companies to share and validate innovative safety testing methods under advisement of the FDA and the European Medicine Agency (EMA), and submit them for formal regulatory qualification when appropriate. The SAFE-T consortium has collaborated from the very start with the PSTC based on shared objectives, and in 2014 a legal agreement which formalized the collaborative efforts was signed. The collaboration between PSTC and SAFE-T addressed among others the selection of biomarkers and setting normal ranges for new biomarkers as defined in healthy volunteers.

The IMI SAFE-T project was finalized in June 2015. In light of the data gathered the initial objective of regulatory qualification of DIKI, DILI and DIVI biomarkers had to be reconsidered and a Letter of Support (LoS) was considered a more realistic goal.

1.2 Drug-Induced Liver Injury Work Package 3 Objectives

The Drug-Induced Liver Injury (DILI) Work Package 3 (WP3) of the SAFE-T consortium specifically aimed to address the current lack of sensitive and specific clinical tests to diagnose, predict and monitor drug-induced injury to the liver, which is a major hurdle in drug development.

The primary objectives of DILI WP3 were:

1) To gain scientific acceptance and ultimately regulatory endorsement for the use of new DILI biomarkers in defined clinical contexts
2) To characterize the biomarkers with respect to the:
   a) predictivity of DILI outcome, with particular emphasis on severe DILI/acute liver failure
   b) monitoring of prognosis, progression and regression of DILI
   c) differentiation between patients who incur true drug-induced liver injury from those who recover from the initial injury despite ongoing drug treatment (adaptors)

The objective 2c) of differentiation between patients that adapt from those who remain susceptible had to be abandoned due to lack of DILI cases in the study specifically designed to address this objective (protocol 4).

1.2.1 DILIN collaboration

Collaboration with the Drug-Induced Liver Injury Network (DILIN) was initiated in order to address the lack of patients developing severe DILI in the SAFE-T clinical studies. The DILIN network was established to advance understanding and research into DILI by initiating a prospective registry of patients with bona fide DILI (2). Overall, DILIN provided in total 166 samples from patients with acute DILI, at sample volumes of 350 µL per sample.

2 Proposed Context of Use

2.1 Context of Use Statements

The biomarkers studied in the SAFE-T clinical DILI studies were rated according to their performance for the following three Contexts of Use.

Context-of-use statement "A":

Based on preliminary data, the following biomarkers have potential as clinical DILI biomarkers that sponsors may choose to incorporate into their clinical trials to provide additional information beyond the diagnostic value of Alanine transaminase (ALT) and Total bilirubin (TBIL) according to the following pathophysiological mechanisms (including the detection of severe DILI as defined by Hy's law criteria):

a) *markers of hepatocyte necrosis (total Keratin 18(K18), miR-122, High Mobility Group Box 1(HMGB1), Glutamate Dehydrogenase (GLDH), Sorbitol Dehydrogenase (SDH))*
b) *apoptosis (caspase-cleaved keratin 18 (ccK18))*
c) *immune activation (hyperacetylated HMGB1, Macrophage Colony Stimulating Factor Receptor 1(MCSFR1))*

Context of Use A is essentially a validation of the biomarker's sensitivity and specificity in comparison to ALT, which was the benchmark for inclusion into the acute DILI studies. ALT already offers a high degree of sensitivity in detecting hepatocyte injury. However, ALT does not yield information as to the underlying mechanism of DILI and only identifies potentially severe cases of DILI in combination with bilirubin. Furthermore, mild elevations of ALT due to alternative causes are commonly seen in drug development, e.g. due to fatty liver disease or due to enzyme induction and in these cases further mechanistic information may be of value.

Context-of-use statement "B":

Based on preliminary data, the biomarkers hyperacetylated HMGB1, Osteopontin, Total Keratin 18 and MCSFR1 have potential as clinical DILI biomarkers that sponsors may choose to incorporate into their clinical trials to anticipate a risk for progression of hepatocellular injury to severe DILI in patients in whom an initial DILI diagnosis has been established based on elevations of the standard marker ALT alone or in combination with TBIL.

Context of Use B aims at separating patient groups according to risk profile. A major challenge that sponsors, regulators and physicians are facing is assessing DILI with respect to the severity of the injury and the risk of deterioration, despite providing best supportive care including cessation of the causative drug. The magnitude of an ALT elevation in the acute

DILI setting does not predict the patient's subsequent course. The Hy's law paradigm indicates a 10% risk of liver failure / fatality in patients in the setting of hepatocellular DILI, the increase in bilirubin being the result of progressive liver failure, with an impairment of the liver's overall capacity to excrete bilirubin into bile. Due to the large reserve capacity of the liver, however, bilirubin increases occur at a relatively late stage during the course of liver failure. Moreover, the degree of hepatocellular dysfunction may not be the only driver of clinical outcome; the level of immune activation may play a key role as well in determining prognosis in a given patient.

Current management of DILI would benefit greatly from novel biomarkers that could separate patients with a likelihood of recovering spontaneously from those who are at risk of worsening to a state of advanced liver injury with a requirement for liver transplantation. Examples of idiosyncratic DILI resulting in liver failure include lumiracoxib (3), troglitazone (withdrawn in May 2000 because of liver toxicity), and bromfenac (withdrawn in June 1998 because of liver toxicity). A biomarker that could predict the clinical course of a patient in whom a drug-induced elevation of Alanine transaminase (ALT) has been observed could help to scale the risk of subsequent deterioration, which would also impact on the overall clinical management.

Context-of-use statement "C":

Based on preliminary data, the following biomarkers: total HMGB1, total and caspase-cleaved keratin 18, miR-122 and GLDH have potential as clinical DILI biomarkers that sponsors may choose to incorporate into early stage clinical trials for the assessment of suspected intrinsic liver injury before elevation of the standard marker of ALT (within the first 24 hours following drug exposure).

Context of Use C targets patients in early clinical studies taking a potentially hepatotoxic drug in whom close monitoring of liver integrity is warranted, e.g. in the face of an unclear liver signal from pre-clinical studies. Due to the well-characterized temporal profile of liver injury induced by acetaminophen (APAP), this model was chosen based on published data (4) to assess the temporal performance of biomarkers in detecting hepatocyte damage before ALT begins to rise above normal.

2.2 Letters of Support issued by FDA and EMA

Regulatory support in the form of a Letter of Support (LoS) was obtained from the FDA on July 25, 2016 for CoU B for the following biomarkers: CK-18, HMGB1 (total and hyperacetylated), osteopontin and MCSFR1 to encourage the further development and exploratory use of the aforementioned biomarkers alone or in combination. Regulatory support from EMA was also in the form of a LoS on 30 September 2016, with recommendations for further development and exploration of biomarkers graduated according to the results achieved. Accordingly, a clear and unconditional support to encourage further research was given for the biomarker candidates included in the CoU B. For CoU C, the EMA acknowledged the results achieved for the biomarkers explored as promising, albeit not as relevant to drug development as CoU B. For CoU A, the EMA has recommended lowest priority since the added value of biomarkers studied is regarded as less apparent.

3 Background and History

3.1 Overview of specific organ injury

Drug-induced liver injury (DILI) in the general Western population occurs with an annual incidence of about 14 to 19 per 100,000 inhabitants according to surveys in France and Iceland (5, 6). The reported frequency of adverse drug liver reactions as a proportion of all adverse reactions to drugs ranges from 4 to 10 % (3, 7, 8) and hepatic adverse reactions accounted for about 8% of all fatal adverse drug reactions (9). DILI was responsible for 11 out of 77 drugs withdrawn from the market over a 32 year period from more than 6,000 compounds (10). It is thought that about 2,000 cases of acute liver failure (ALF) occur annually in the US. ALF leads to death in about 30% of patients receiving aggressive therapy including orthotopic liver transplantation (OLT) (11, 12) while at times without OLT option the mortality rate has been 60 – 80% (13, 14). According to these overviews (15-17), about half of ALF cases are due to DILI. The leading drug causing ALF is APAP overdose (~40%) with a fairly predictable dose-response to liver toxicity and therefore called intrinsic hepatotoxicity. All non-APAP cases represent about 10 to 15% of DILI cases (12). In the latter group, amoxicillin/clavulanate, isoniazid and NSAIDS in several surveys are the leading suspect drugs based on absolute numbers, based on number of prescriptions the incidence e.g. for amoxicillin/clavulante is about 1.7 per 10,000 prescriptions (18). Of note, with wide regional variability, herbals and dietary supplements are increasingly being recognized as potential causes of liver injury (19-21).

Given the ratio of increase in aminotransferase relative to ALP increase, the DILI pattern is classified into hepatocellular, cholestatic and mixed (22, 23). In most registries, the hepatocellular type dominates with a prevalence of about 40 to 50%, followed by the cholestatic type with about 30 to 40 % and the mixed type with a prevalence close to 20% (24). In general, the hepatocellular type has a higher mortality / OLT risk (10 to 15% compared to 6 to 8% reported for the cholestatic type and 2 to 3% for the mixed type (25, 26), the cholestatic type has a higher risk for prolonged resolution after terminating the suspect drug (27-29). Evidence for an immune-allergic component as evidenced by fever, rash, peripheral eosinophilia or tissue biopsy sample is seen in about 25% to 30% of DILI patients (30) and may be associated with a better prognosis (31). Of note, no histological feature has been found to be pathognomonic for DILI (31).

3.2 Use and limitations of current tools

Clinical symptoms of acute DILI may include fatigue, abdominal pain, jaundice, nausea, pale stool and dark urine. Acute DILI may progress to acute liver failure (ALF) with emergence of encephalopathy diagnosed into grades 0 (no symptoms) through 4 (coma) and coagulopathy. King's College Criteria (KCC) are used to estimate prognosis and need of OLT in the ALF setting. The MELD (model for end-stage liver disease) / PELD (pediatric end-stage liver disease) scores were originally developed to predict mortality in patients with chronic liver disease. The parameters included in these scoring systems applied in the clinical setting (KCC / MELD / PELD) contrast with the use of ALT, AST and bilirubin in clinical trials for drug development to detect early signals of relevant drug hepatotoxicity and to predict a new drug's potential to cause severe idiosyncratic DILI when applied to a larger population post-marketing.

Based on his observations collected mainly with anti-infectious and CNS drugs more than 50 years ago, Dr Hyman Zimmermann concluded that the combination of elevated ALT (usually > 8 x ULN) with jaundice constitutes – in the absence of alternative potential causes for hyper-bilirubinemia – a serious entity and was associated with mortality rates between 10 and more than 50% (8). Dr Robert Temple from the FDA translated this finding into a combination of ALT exceeding 3 x ULN, TBIL exceeding 2 x ULN with ALP staying below 2 x ULN and in the absence of potential alternative causes and named it "Hy's Law" (32). Analyses referring to Spanish and Swedish DILI registries (25, 26) with data mainly collected during the 1990s and early 2000s found that this combination was associated with a mortality of about 10%. In 2009, the FDA Guidance for industry outlined the relevance of observing Hy's law cases in clinical drug development trials assuming a severe DILI rate of 10% of Hy's Law cases.

The FDA guidance states: "*The specificity of this finding appears very high if two or more cases are seen (...). We are not aware of the occurrence of false positive Hy's Law findings for a drug that was subsequently found not to cause severe DILI in a larger treatment population*."

The FDA guidance however also states that "*Failure to find a case, however, does not imply that a drug with AT elevations is free of a risk of severe DILI.*" When Hy's law was applied to two surveys attempting to predict ALF as a result of DILI, the following statistical features were obtained (33):

Table 3-1 Prediction of ALF as result of DILI using Hy's law

Hy's Law statistics (95%CI)	specificity	negative predictive value	sensitivity	positive predictive value
Survey I with 22 ALF events in 15,345 patients with a DILI diagnosis	0.92 (0.91-0.93)	0.99 (0.99-1.00)	0.68 (0.45-0.86)	0.02 (0.01-0.03)
Survey II, 76 DILI patients	0.27 (0.17-0.39)	0.9 (0.68-0.99)	0.78 (0.4-0.97)	0.13 (0.05-0.24)

In another study from the Netherlands the positive predictive value for Hy's law to confirm acute liver injury as defined in ICD-9-CM code was calculated to be 22% (34).

In published results from the US-DILIN network including 660 patients with definite, highly likely or probable DILI and followed up for 6 months after DILI onset, Hy's Law was present at DILI onset in 63% of patients who received OLT, in 35.3% with liver related death and in 26.7% of those with non-hepatic death, a rate similar to those without OLT / death (26.2%) (29). As of today, Hy's law remains the best available predictor of severe DILI requiring OLT or leading to death later on. A significant caveat for the use of Hy's law relates to its poor positive predictive value and limited sensitivity. Moreover, from a pathophysiological standpoint, it is desirable to better understand the association between clinical chemistry parameters used to monitor liver integrity and clinical symptoms. Indeed, several registries and information sharing platforms have been created and initiatives started to bring more light into the pathophysiology of DILI, e.g. the US DILI Network started in 2004 (2), the

"LiverTox" Web site – [http://www.livertox.nih.gov] – and the Prospective European Drug-Induced Liver Injury Registry [http://www.spanishdili.uma.es/proeuro/].

The Hy's Law algorithm is being re-evaluated in the context of special populations such as oncology trials (35) by setting different thresholds for ALT and TBIL and by adding clinical chemistry and hematological parameters such as AST and platelet count (33, 36). To move the field forward, novel biomarkers such as microRNAs, CK-18, HMGB1 and / or various bile acids need to be evaluated in the DILI setting (37, 38). Promising results have been reported, as new biomarkers not only correlate with ALT levels, but may also precede and even predict ALT increases (39). However, the usefulness of biomarkers in predicting outcome of severe DILI in terms of mortality and the need for OLT remains to be evaluated. The ideal novel biomarker would provide information beyond what can be deduced pathophysiologically from available algorithms such as Hy's Law and the KCC/MELD/PELD criteria, thus allowing improved hepatotoxic risk assessment and monitoring in clinical drug development.

3.3 Methodology for the selection of biomarkers

Originally, the overall strategy for biomarker selection consisted of three steps:

Step 1: A "stage gate analysis" was performed by comparing performance in patients with confirmed DILI vs healthy volunteers (HV) to exclude biomarkers considered least promising from further exploration.

Step 2: The remaining biomarkers entered an exploratory phase to derive hypothesis generating data and to further narrow the biomarker list.

Step 3: A confirmatory phase to assess biomarker performance in more depth and to validate hypotheses generated in the exploratory phase was planned as the final qualification step.

However, given time constraints and the limited number of patients available by the end of 2014, the DILI-WP decided to investigate the new biomarkers selected from the first stage gate analysis in one subsequent analysis using all available datasets and to no longer separate an exploratory from a confirmatory phase. True confirmatory data are currently not available and all results are considered exploratory for the purpose of a LoS for further research and future confirmatory studies.

3.4 Biological rationale for each candidate biomarker selection

The biomarkers taken forward into the full exploratory sample set and which were analyzed in the final dataset are listed in Table 3-2. In addition, data generated across four cohorts ([i] healthy subjects treated with acetaminophen (N=58); [ii] patients with HIV and / or tuberculosis and ALT exceeding 3 x ULN (N=38); [iii] patients with DILI (N=10); [iv] healthy subjects taking heparin (N=48)), two additional biomarkers - cadherin 5 and liver fatty acid binding protein – that were not explored at the time of the stage gate analysis, were included into the biomarker panel and were available for the final analyses.

Table 3-2 Biomarkers & rationale for selection

Marker	Origin of Biomarker	Summary
Micro RNA 122	Liver-specific	Micro RNA 122 is an early marker of hepatocellular injury, possibly preceding ALT on a temporal scale And is a specific marker of hepatocellular injury. It has been reported as a sensitive DILI marker in multiple clinical studies. (4, 40-44)
High mobility group box 1 (HMGB1)	Detectable in almost all tissues	HMGB1 predicts patient prognosis following APAP overdose. Hyperacetylated HMGB1 is significantly elevated in patients that die/require a liver transplant, whereas in spontaneous survivors it is not significantly different from healthy volunteers. (39)
Cadherin 5	Endothelial cells	CDH5 is a calcium-dependent cell adhesion protein (also called VE-cadherin), that is specific to endothelial cells and a major component of endothelial adherens junctions and was identified as a potential biomarker for DILI susceptibility within SAFE-T (40).
Cytokeratin 18 full length	Epithelial cells;	The full-length protein is released from necrotic cells. It is significantly elevated in acetaminophen overdose patients that die/require a liver transplant compared to spontaneous survivors. (4, 39, 42, 43)
Cytokeratin 18 caspase cleaved fragment (cc Keratin 18)	Epithelial cells;	The caspase-cleaved fragment is released from apoptotic cells and helps define the type of cytotoxicity. cc Keratin 18 fragments in blood predict severity of disease in NASH and in hepatitis C. (4, 42, 43, 45)
Liver Fatty Acid Binding Protein (L-FABP)	Primarily liver; lower levels in the kidneys and small intestines	Primarily liver specific, in lower levels in the kidneys and small intestines. L-FABP is a sensitive marker for hepatocellular injury following liver transplant (46)
Glutamate dehydrogenase (GLDH)	Mitochondrial matrix; primarily in the centriolobular region of the liver; lower levels in the kidney and brain	A sensitive biomarker of liver toxicity with hepatocellular damage in preclinical species; shown to be elevated in humans with hepatic ischemia or hepatitis; shown to correlate with ALT in patients with a broad range of clinically demonstrated liver injuries including acetaminophen-induced liver injury and to detect mild hepatocyte necrosis in patients treated with heparin. Marker for mitochondrial injury or cellular injury in multiple clinical DILI and acute liver failure studies (42, 47, 48)
Glutathion S-Transferase (GST-alpha)	Centrilobular region of the liver; multiple tissues	Hepatotoxicity biomarker shown in rats to have enhanced specificity and sensitivity compared to ALT; humans with acetaminophen overdose show elevated GSTα levels earlier than ALT; GSTα may offer a better assessment of rapid changes in liver damage due to the shorter half-life of plasma GSTα compared to ALT or AST. (49, 50)
Alpha-fetoprotein (AFP)	Liver progenitor cells	Increase of AFP has been detected in many types of liver disease including APAP overdose. Data from literature suggest that AFP is expressed after the onset of liver injury and during regeneration with increased serum / plasma levels. AFP may have value as a prognostic marker in liver injury. (51, 52)

Marker	Origin of Biomarker	Summary
Arginase-1 (ARG1)	Primarily in the liver cytosol; lower levels in erythrocyte	ARG1 has been shown to be highly sensitive for acute liver damage (leakage marker). Circulating concentrations increase in patients with various hepatic disorders, such as hepatoma and viral or alcoholic hepatitis. Sensitive biomarker in the clinic for liver injury following liver transplant. A significant correlation with AST and ALT activities was described following partial resection and orthotopic liver transplantation. (50, 53)
Osteopontin (OPN)	Multiple tissue and cell types including liver	Elevated serum levels of OPN are detectable in patients with severe liver damage. Increased levels of serum OPN are associated with a poor prognosis. Plasma OPN levels in fulminant hepatic failure patients were higher than those of acute hepatitis patients and healthy adults. OPN is associated with inflammatory cell activation and with liver regeneration due to activation of hepatic stem cells. (54)
Macrophage colony stimulating factor receptor 1(MCSFR1 or CSF1R)	Cytokine receptor on macrophages/ monocytes	Data from the ximelagatran biomarker discovery study suggest that MCSF-R is shed from macrophages during DILI. CSFR1 serum/plasma levels may have value as a prognostic marker for liver disease associated with inflammation. (55)
Paraoxonase 1 (PON1)/Prothrombin	Primarily liver; lower levels in multiple tissues	PON1 is not a leakage enzyme, but is constitutively released into the circulation. Decreases in serum PON1 reflect liver injury or dysfunction and have been linked to chronic hepatic damage. The biomarker serves two purposes: 1) as a diagnostic marker for depressed liver function; 2) ratio together with prothrombin as a marker to differentiate between healthy controls and subjects with all types of NAFLD and NASH. (56)
Leukocyte cell-derived chemotaxin2 (LECT2)	Primarily liver; lower levels in testes	Prognostic indicator of liver regeneration and injury. Serum LECT2 levels are inversely proportional to ALT and decrease at the peak of liver regeneration after hepatectomy. (57)
Sorbitol dehydrogenase (SDH)	Multiple tissue and cell types including liver	Sensitive enzymatic serum marker of liver toxicity increasing with hepatocelluar damage in preclinical species. Shown to be elevated in humans with various liver diseases and to detect mild hepatocyte necrosis in patients treated with heparin. The biomarker serves two purposes: 1) as an early marker of hepatocellular injury, possibly preceding ALT on a temporal scale 2) as a specific marker of hepatocellular injury. (47)
Conjugated/unconjugated Bile acids	Synthesized by the liver	1) early markers of cholestasis, possibly preceding ALP and ALT on a temporal scale 2) sensitive marker of inhibition of the bile salt export pump (BSEP), known to be inhibited by several drugs (58, 59) 3) marker of liver synthetic function

The nine biomarkers selected from the above Table 3-2 to support issuance of a LoS for the specified CoUs are detailed in section 2.1 under the respective CoU statements A, B and C.

4 Preclinical Studies

Results of relevant published studies relating to biomarkers that form part of the CoUs proposed are summarized above in Table 3-2.

A joint work plan between SAFE-T work package 3 and the Predictive Safety Testing Consortium Hepatotoxicity Working Group (PSTC HWG) was created in 2014. Studies performed by the PSTC primarily focus on the validation and performance assessment of preclinical assays in the rat. The PSTC also conducts studies in rats, and in some cases dogs, to assess the performance of potential DILI biomarkers. These studies are intended to address three areas of need for DILI biomarkers. First, potential biomarkers are being evaluated as alternatives to ALT for detection of hepatocellular necrosis. These biomarkers would have improved specificity compared to ALT and are anchored on histopathology. The second set of studies is aimed at discovering a biomarker that can discriminate between whether an increase in ALT is due to potential liver toxicity or if it is due to factors not related to toxicity. These biomarkers would not track with increased ALT in the absence of microscopic hepatic lesions, and the synthesis and clearance of the new biomarker would not be regulated by the same mechanisms as ALT. The third area of interest for PSTC is to investigate specific bile acids, or combinations of bile acids, that can be biomarkers for BSEP-mediated liver injury. SAFE-T WP3 and PSTC have also collaborated on methods for analysis of some biomarkers in order to have one method available for all species when possible.

5 Clinical Studies

5.1 Methods

SAFE-T WP3 prioritized among a number of novel biomarkers possibly indicating drug-induced liver toxicity. Performance criteria of these new biomarkers included comparison against traditional liver markers such as ALT and TBIL and Hy's law, or performance in relation to clinical outcome of DILI. Biosamples from healthy volunteers, patients with acute DILI and patients taking potentially hepatotoxic drugs were used.

5.2 Study design

The various clinical studies in patients can be divided into (i) protocols that recruited patients diagnosed with DILI and (ii) protocols that recruited patients without a diagnosis of DILI but who were on treatment with potentially hepatotoxic drugs and were prospectively monitored for several months. For all studies, cases with suspected DILI were ascertained by clinical judgment of the investigators and, subsequently, by the evaluation of an adjudication committee. All cases meeting the trial enrollment criteria were adjudicated, the great majority of those fulfilled the consensus criteria for DILI as published by Aithal et al. (23) (ALT ≥ 5xULN, OR ALP ≥ 2xULN; OR ALT ≥ 3xULN with simultaneous elevation in total bilirubin > 2xULN). The "Adjudication Committee" assessed drug causality of liver dysfunction according to the following criteria: 1) an appropriate temporal relationship between the intake of the drug and the onset of the event, 2) the improvement of liver damage following the withdrawal of the drug, 3) exclusion of other causes of liver disease, 4) relapse following re-exposure when applicable, and 5) previous reports of the adverse reaction. Cases were further evaluated for causality assessment, by application of the Council for

International Organizations of Medical Science (CIOMS)/Roussel Uclaf Causality Assessment Method (RUCAM) scale.

Apart from one patient, only cases that met the consensus definition at the date of baseline or within the last 4 weeks before baseline and that were adjudicated as DILI were used in the analyses for CoU A. The patient that did not meet the consensus definition of DILI (ALT < 5xULN), was nonetheless adjudicated as DILI caused by chemotherapy for gastric cancer and was therefore included in the analyses.

Trials that enrolled patients diagnosed with DILI:

- **Protocol 3A**: A 12-wk follow-up study investigating the prognostic value of new biomarkers in patients with DILI. Samples from 98 patients adjudicated as DILI patients were available for final analyses. No patient progressed to severe DILI during the observation period.
- **Swiss DILI study**: A 8-wk follow-up study investigating the prognostic value of new biomarkers in patients with DILI. Samples from 28 patients adjudicated as DILI patients were available for analyses. No patient progressed to severe DILI during the observation period.
- **DILIN**: A US prospective registry study including patients within 6 months of DILI onset (2). Samples from 166 patients taken at a single timepoint within a mean of 2 weeks after diagnosis of DILI were made available to WP3 by the US DILIN network [http://www.dilin.org/]. All cases had been adjudicated previously by DILIN. Samples for selected biomarkers were analyzed and results compared with the standard markers of ALT and bilirubin. The major asset of the DILIN samples is that a subgroup of patients developed acute liver failure, thus providing a basis for assessing Context of Use B.
- **Liverpool study (4)**: A study conducted at two UK hospitals to assess the potential of novel biomarkers to identify patients with acetaminophen-induced acute liver injury at first presentation to the hospital. The majority of patients were participants in an ongoing randomized, controlled study, SNAP (EudraCT number 2009-017800-10). From a population of 129 subjects who were known to have taken an overdose of acetaminophen, 100 subjects had ALT values below 3 x ULN at the time of hospital admission. Samples were analyzed for HMGB1, CK18, caspase-cleaved CK18, microRNA122, and GLDH with respect to their prognostic value in predicting any subsequent changes in ALT and bilirubin. Raw data from this study were provided by Liverpool and were analyzed by WP3. The results from this study provide the basis of Context of Use C. The biomarker HMGB1 (total and hyperacetylated form) was measured at the MRC Centre for Drug safety Science, University of Liverpool, UK (Head: Prof. Kevin B. Park) (4). Values measured in the SAFE-T acute DILI samples were compared with a group of healthy volunteers from Liverpool, for whom HMGB1 values were already available. The SAFE-T healthy volunteer samples were not additionally measured.

Trials that enrolled patients with normal ALT values who were followed up while taking drugs known to be potentially hepatotoxic:

- **Protocol 4**: A 9 month follow-up study in tuberculosis patients starting anti-tuberculosis drug therapy. From this study, 81 patients were included in biomarker analyses. No patient

in this protocol developed DILI defined as ALT above 5 x ULN during the observation period.

- **Protocol 5**: A 3-yr follow-up study of rheumatoid arthritis patients with normal ALT at start. From this study, 92 patients were included in biomarker analyses. No patient developed DILI defined as ALT above 5 x ULN during the observation period.

Protocols 4 and 5 provided data from a non-DILI patient population for comparison with healthy volunteers and patients with DILI with respect to biomarker performance.

Healthy volunteer protocols:

- **Tel Aviv (TASMC) HV study**: 192 HVs were included in SAFE-T from this study.
- **Liverpool HV group**: for biomarker HMGB1 (total and hyperacetylated), measured in Liverpool, a separate HV control group comprising 154 HVs was already available.

5.3 Biomarker assays

The primary focus of the IMI SAFE-T consortium was on the clinical qualification of soluble blood and urine protein biomarkers. Whenever possible, commercial assays and materials were used. The assays used for measurement of the SAFE-T sample set along with the respective assay platforms are listed in Table 5-1 and rely on qPCR, LC-MS, enzyme activity or sandwich immunoassays. In certain cases it was necessary to generate new specific assay material. Assay development and validation was coordinated, overseen and approved by a dedicated group of experts within IMI SAFE-T (WP5).

For ensuring assay quality, a SAFE-T validation procedure (SVP) was developed based on the fit-for-purpose concept (60, 61) where technical performance is evaluated against the predefined purpose and consequently, the stringency of performance verification varies with the intended use. The validation procedures described in the document are based on guidelines issued by the regulatory authorities (e.g. EMA, 2009; FDA 2013) (62, 63), but also consider the guidelines available from the Clinical and Laboratory Standards Institute (CLSI) for the most extensive phase of assay validation (proof of performance testing). During assay validation the following parameters were tested: limit of detection, limit of quantification, intra-/inter-assay precision, parallelism and/or dilutional linearity, parallelism, analyte stability, assay dynamic range, and spike-in recovery, when possible. Validation criteria were set following common assay validation standard procedures. Appropriate Quality Control (QC) controls were applied during the sample screening procedure to ensure data reliability and data comparability over the different phases of SAFE-T. The full SAFE-T assay validation protocol and SAFE-T QC guidance's are provided as supplementary material. A summary of the validation results for each assay is shown in Table 5-1.

Table 5-1 Assays used for measurement of SAFE-T samples & associated assay platforms

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 (24 h at RT and 4°C)	F/T stability, 3 cycles
acetyl. HMGB1	LC-MS	Serum	ng/ml	0.2	0.06	30	2.7 - 13.7	2.1 - 13.6	1:2 - 1:16	90 - 102	ND	yes
AFP	Immunoassay	Serum	ng/mL	0.367	0.367	584	2 - 16	7 - 13	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
ccK18	ELISA	Serum	U/L	16.2	62.5	1000	2.2	5.7 - 7.9	up to 1:16	112 - 118	yes	yes
CDH5	ELISA	Serum	ng/mL	0.36	3.13	100	6.0	4.7 - 7.2	1:40 - 1:640	50 - 83	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
GST-alpha	Immunoassay	Serum	ng/mL	1.79	1.82	373	1 - 14	9 - 11	1:5 - 1:10	77 - 94	yes	yes
K18	ELISA	Serum	U/L	20	100	5000	3.7	6.1 - 9.4	up to 1:32	83 - 107	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
L-FABP	Immunoassay	Serum	pg/mL	3.1	15.6	16000	5.6	6.7 - 18.1	1:2 - 1:2048	110 - 115	yes	yes
MCSF-R	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
miR-122	qPCR	Serum	copies/µL	NA	384	5089837	1.3 - 12.1	0.5 - 25.4	ND	ND	2 hrs at RT, 5 hrs at 4°C	yes
OPN	Immunoassay	Serum	ng/mL	1.25	1.25	1149	1 - 5	6 - 11	1:5 - 1:10	81 - 85	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	24 h at 4°C, 4 h at RT	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
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

5.4 Clinical data management

Clinical Data Management was performed by Koehler eClinical for 7 studies with an eCRF defined in OpenClinica. In addition Data Management tasks for 4 trials outside Open Clinica were covered.

Data Management included the cleaning process for the variables to be analyzed.

In addition, data from several external sources were mapped to the clinical eCRF data. Thus, biomarker results were directly transferred from the screening laboratories to Data Management for mapping.

For analysis, relevant data were mapped to CDISC / Study Data Tabulation Model (SDTM) format for analysis.

5.5 Statistical analysis

The main statistical analysis method used is the calculation of Receiver Operating Characteristics (ROC) for distinguishing the two outcome groups being considered and more specifically the calculation of the area under the ROC curve. This has been done using logistic regression for single predictor variables and single predictor variables with key covariates added in. The area under the ROC curve has also been calculated when classification trees have been fitted using many predictor variables and key covariates. To visualize comparison between two or more outcome groups boxplots have been plotted.

The Random Forest approach was used to give an additional assessment of predictor importance to correlate with the results from using single predictor variables.

Summary tables of subject numbers, biomarker and clinical laboratory data have been drawn up to aid in the interpretation of the data.

6 Results

6.1 Assay validation

A summary of the validation results for each assay is shown in Table 5-1.

6.2 Clinical sensitivity and specificity

As shown in Table 6-1 98 patients from protocol 3A and 28 cases from the Swiss DILI study were adjudicated as being DILI. Of these 126 acute DILI cases, 90 (71%) were classified as hepatocellular, 11 (9%) as cholestatic and 25 (20%) as mixed-type injury (Table 6-2). This classification was based on the standard Council for International Organizations of Medical Sciences (CIOMS) criteria (64).

Table 6-1 Summary of Number of Patients by Study- Studies: Protocol 3A, Swiss DILI, HV (Tel Aviv), Protocol 4, Protocol 5 and HV (Liverpool)

Study	N
Protocol 3A	98
Swiss DILI	28
HV (Tel Aviv)	192
Protocol 4	81
Protocol 5	92
HV (Liverpool)	154

The type of DILI and frequency of causative drugs according to arbitrary categories are shown in Table 6-2.

Table 6-2 Acute DILI Information: Type of DILI and Causative Drug Class- Studies: Protocol 3A and Swiss DILI

Study	N	Percent
Protocol 3A	98	100.0
Swiss DILI	28	100.0

Study=Protocol 3A

DILI Type	N	Percent
Cholestatic	5	5.10
Hepatocellular	69	70.41
Mixed	24	24.49

Study=Swiss DILI

DILI Type	N	Percent
Cholestatic	6	21.43
Hepatocellular	21	75.00
Mixed	1	3.57

Study=Protocol 3A

DILI Drug Class	N	Percent
APAP	13	13.27
APAP + NSAID	1	1.02
NSAID	4	4.08
antibiotics	23	23.47
chemotherapy	7	7.14
flupirtine	14	14.29
others	35	35.71
others + APAP	1	1.02

Study=Swiss DILI

DILI Drug Class	N	Percent
APAP	6	21.43
anti-Tbc	3	10.71
antibiotics	9	32.14
others	10	35.71

As expected, the largest group was antibiotics, followed by acetaminophen and flupirtine, a non-opioid, non-NSAID, non-steroidal centrally acting analgesic. 14 cases of flupirtine-induced DILI were recorded, making this drug the single most frequent cause of DILI in the two acute DILI protocols. In 2013, EMA imposed a restriction on the use of flupirtine due to the risk of liver injury.

In addition to the acute DILI cases, two additional protocols prospectively recruited patients receiving potentially hepatotoxic medications (Table 6-1). Protocol 4 monitored patients with a diagnosis of tuberculosis who received anti-tuberculosis combination treatment including isoniazid and rifampicin. 81 patients from protocol 4 were included in the final analysis: however, no case of acute DILI as defined by the consensus criteria according to Aithal et al. (23) was observed. A subgroup of patients developed mild transient elevations of ALT, but not to the extent that the Aithal criteria for DILI were fulfilled. Protocol 5 recruited patients with a diagnosis of rheumatoid arthritis who received continuous treatment with disease-modifying antirheumatic drugs including methotrexate. 92 patients from protocol 5 were included in the final analysis, however – as in protocol 4 – no case of acute DILI was observed. Thus the patients recruited within protocols 4 and 5 were not considered DILI, but in fact served as an additional control group of patients with a chronic inflammatory condition of non-hepatic origin and receiving potentially hepatotoxic drugs.

The actual control group consisted of 192 HV recruited in the Tel Aviv study. Solely for the evaluation of the two biomarkers HMGB1 and acetylated HMGB1, a different control group was used, comprising 154 healthy volunteers from Liverpool - the academic partner where these biomarkers were measured.

See Table 6-3, Table 6-4, and Table 6-5 for some demographic results per study.

Table 6-3 Demography by Study– Gender

Study=Protocol 3A

Gender	N	Percent
Female	57	58.16
Male	41	41.84

Study=Swiss DILI

Gender	N	Percent
Female	13	46.43
Male	15	53.57

Study=HV (Tel Aviv)

Gender	N	Percent
Missing	1	.
Female	88	46.07
Male	103	53.93

Study=Protocol 4

Gender	N	Percent
Female	32	39.51
Male	49	60.49

Study=Protocol 5

Gender	N	Percent
Female	61	66.30
Male	31	33.70

Study=HV (Liverpool)

Gender	N	Percent
Female	89	57.79
Male	65	42.21

Table 6-4 Demography by Study - Studies: Protocol 3A, Swiss DILI, HV (Tel Aviv), Protocol 4, Protocol 5 and HV (Liverpool) – Ethnicity

Study=Protocol 3A

Ethnicity	N	Percent
Asian	5	5.10
Black or African American	1	1.02
Other	2	2.04
White	90	91.84

Study=Swiss DILI

Ethnicity	N	Percent
Black or African American	2	7.14
Other	1	3.57
White	25	89.29

Study=HV (Tel Aviv)

Ethnicity	N	Percent
Missing	1	0.52
White	191	99.48

Study=Protocol 4

Ethnicity	N	Percent
Asian	1	1.23
Black or African American	21	25.93
Other	2	2.47
White	57	70.37

Study=Protocol 5

Ethnicity	N	Percent
Asian	1	1.09
Black or African American	11	11.96
Other	10	10.87
Unknown	2	2.17
White	68	73.91

Table 6-5 Demography by Study - Studies: Protocol 3A, Swiss DILI, HV (Tel Aviv), Protocol 4, Protocol 5 and HV (Liverpool) - Age (years)

Study	N	Mean	SD	Min	Q1	Median	Q3	Max
Protocol 3A	98	51.9	16.0	19	38	53.0	66	83
Swiss DILI	28	54.6	14.8	24	42	56.0	67	84
HV (Tel Aviv)	191	52.7	14.1	24	42	52.0	62	90
Protocol 4	81	36.5	13.7	18	26	32.0	44	75
Protocol 5	92	51.6	12.8	23	44	52.5	61	88
HV (Liverpool)	154	34.6	9.9	18	25	35.0	42	66

6.3 Results supporting the CoU statements

6.3.1 Results pertinent to Context of Use A:

"Based on preliminary data, the following biomarkers have potential as clinical DILI biomarkers that sponsors may choose to incorporate into their clinical trials to provide additional information beyond the diagnostic value of ALT and bilirubin according to the following pathomechanisms (including detection of severe DILI as defined by Hy's Law):

a) markers of hepatocyte necrosis (total keratin 18, miR-122, total HMGB1, GLDH, SDH)

b) apoptosis (caspase-cleaved keratin 18)

c) immune activation (hyperacetylated HMGB1, MCSFR1)"

The SAFE-T acute DILI studies, protocol 3A and the Swiss DILI study, support CoU A. Inclusion into protocol 3A and Swiss DILI study was based on ALT activity exceeding 3 x ULN ***or*** ALP > 2 x ULN, within the last 4 weeks before the baseline visit. As detailed in section 3.1, apart from one subject, Context of Use A evaluated biomarker performance only in those patients fulfilling the consensus definition for DILI (23).

To address these claims pertaining to CoU A, the subsequent box plots (Figure 6-1 to 6-12) show a selection of biomarkers as measured in the initial blood sample obtained at baseline, i.e. at a point in time when acute liver injury in protocol 3A and Swiss DILI study was either still evident and ALT was elevated or at most 4 weeks following the acute DILI episode. Each box plot contains four different panels that are defined as follows:

DILI: acute DILI cases recruited in protocol 3A and Swiss DILI study

HV: healthy volunteers recruited in Tel Aviv study 22BIZY

Protocol 4: patients from protocol 4 (non-DILI)

Protocol 5: patients from protocol 5 (non-DILI)

HV Liverpool: healthy volunteers recruited in Liverpool (for HMGB1 only) (39)

Figure 6-1 Alanine aminotransferase (ALT) – Boxplot of baseline biomarker data by study

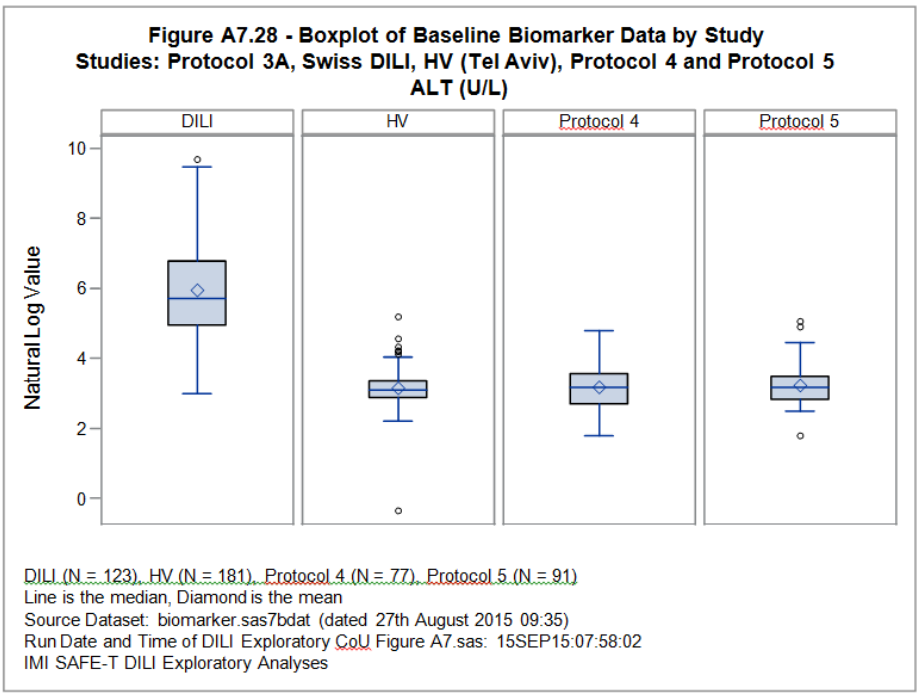


Figure 6-2 Aspartate aminotransferase (AST) Boxplot of baseline biomarker data by study

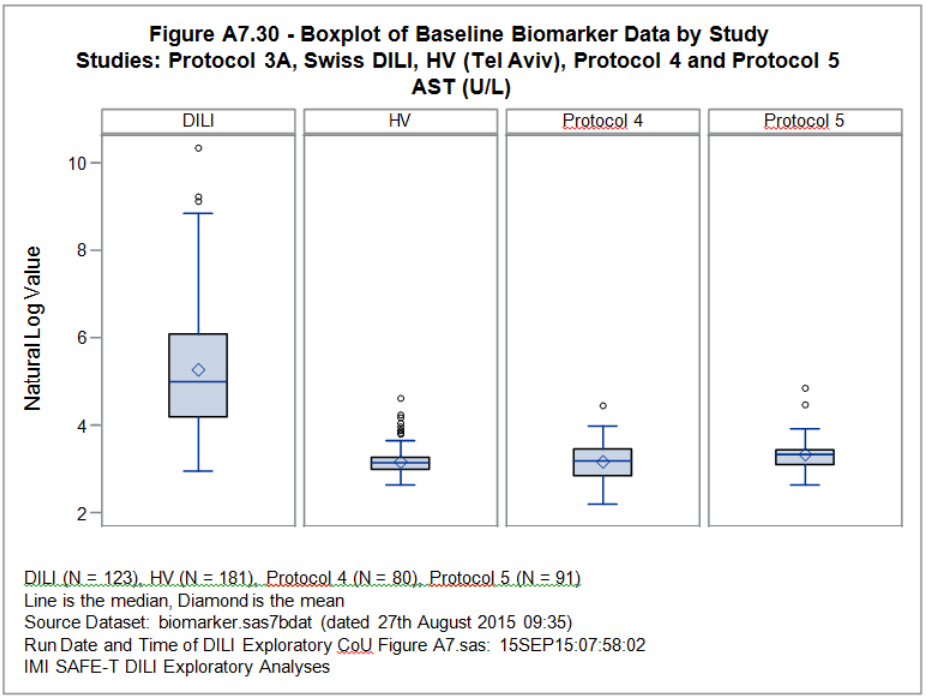


Figure 6-3 Alkaline phosphatase - Boxplot of baseline biomarker data by study

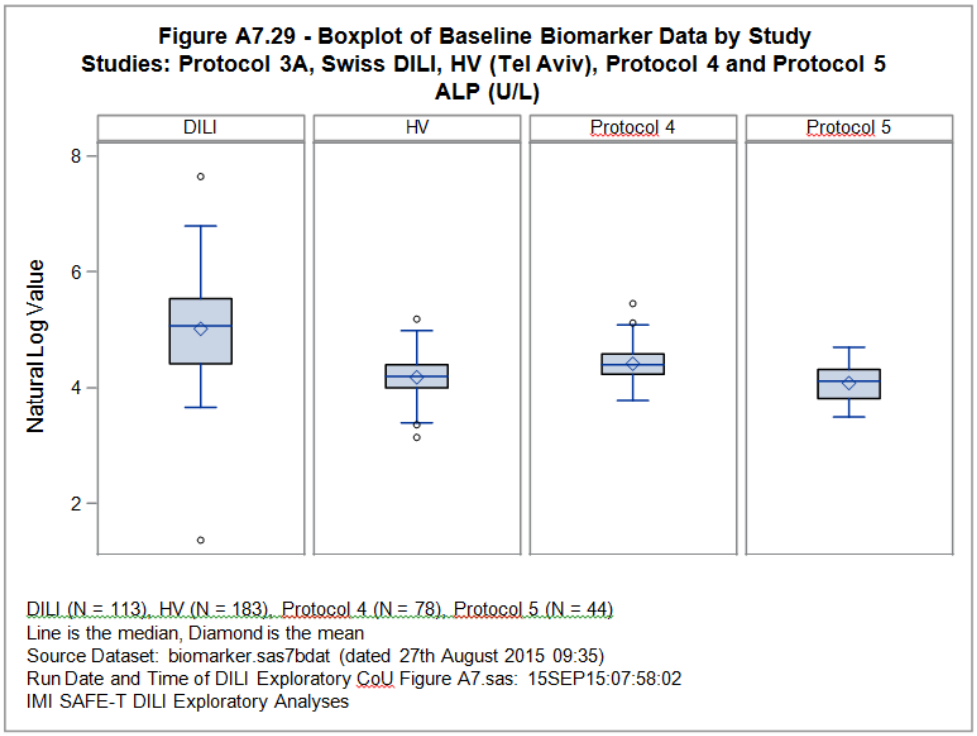


Figure 6-4 Total bilirubin- Boxplot of baseline biomarker data by study

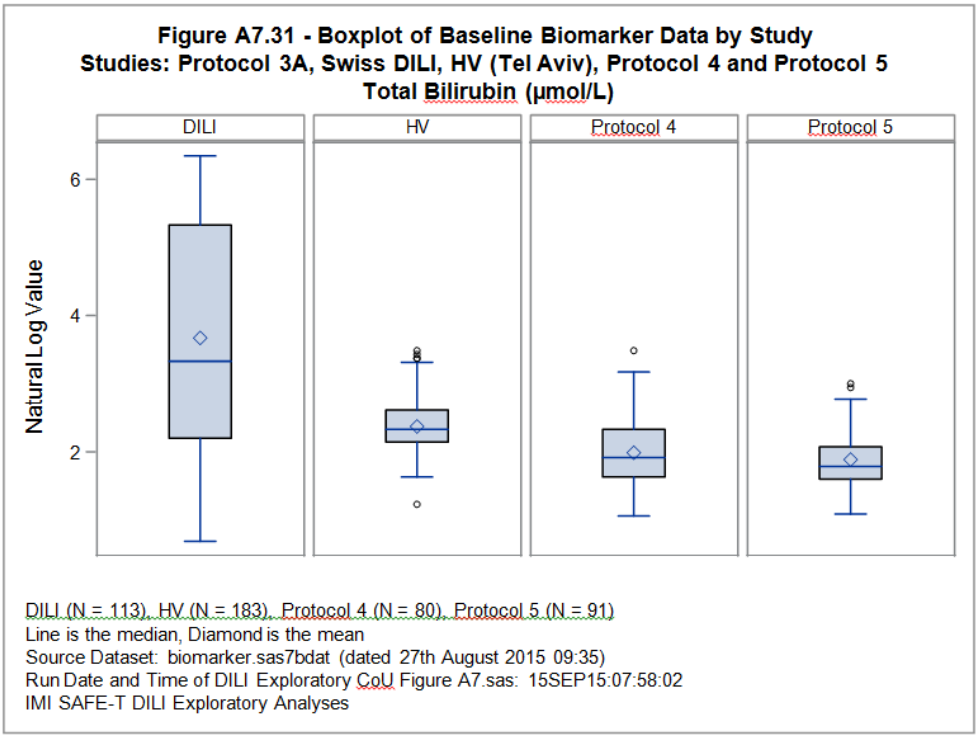


Figure 6-5 Total keratin 18 - Boxplot of baseline biomarker data by study

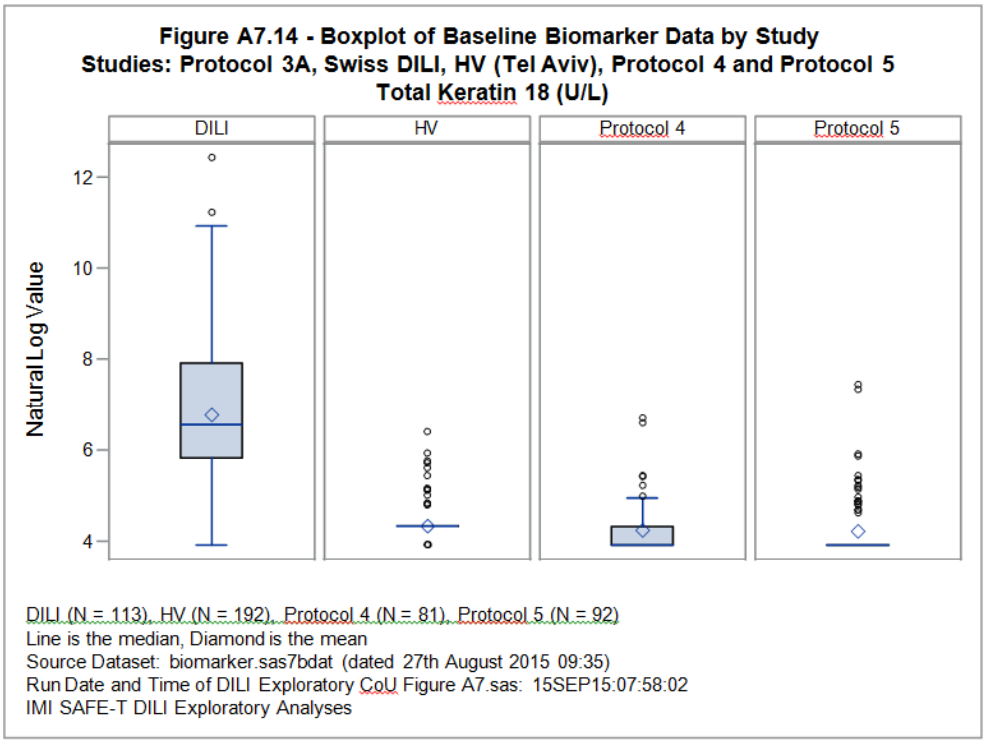


Figure 6-6 Caspase-cleaved keratin 18 (ccK18) - Boxplot of baseline biomarker data by study

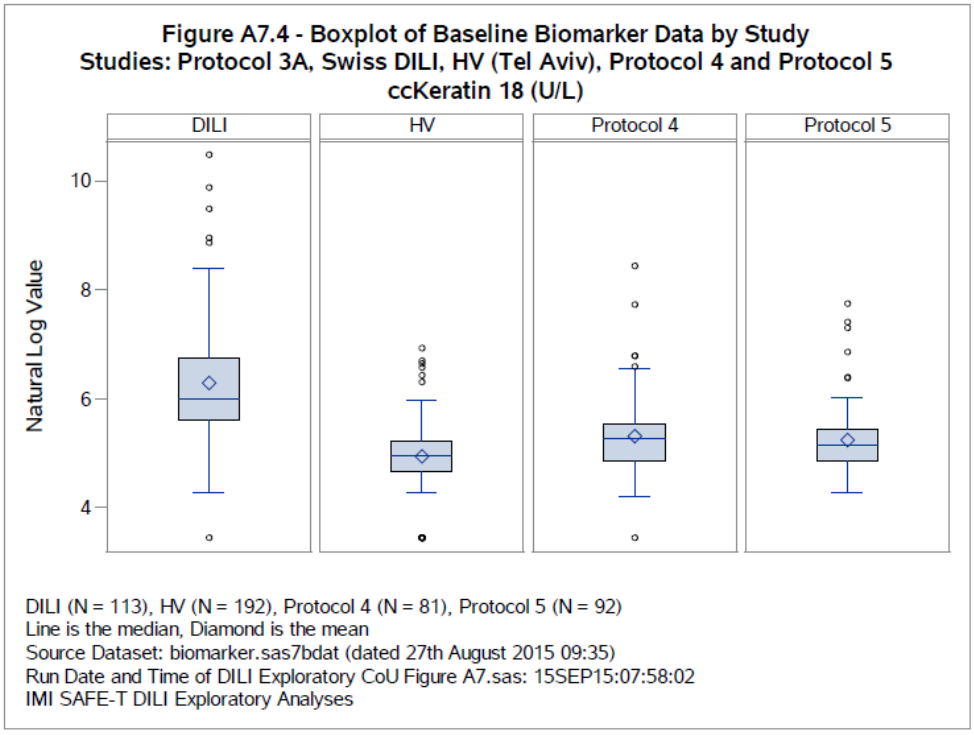


Figure 6-7 Glutamate dehydrogenase (GLDH) - Boxplot of baseline biomarker data by study

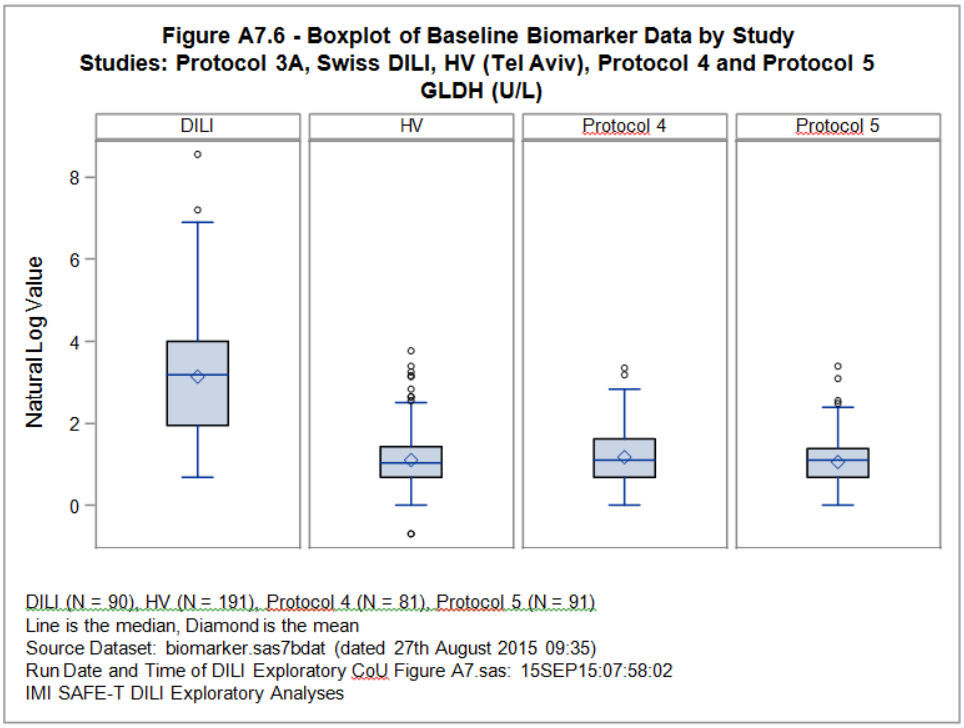


Figure 6-8 Sorbitol dehydrogenase (SDH) - Boxplot of baseline biomarker data by study

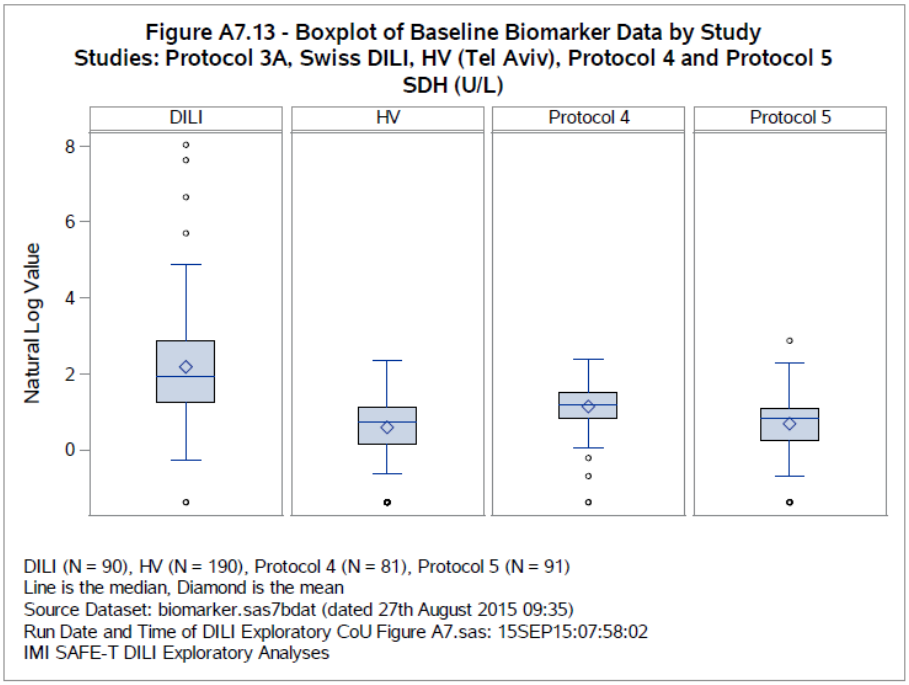


Figure 6-9 Macrophage colony-stimulating factor receptor 1 (MCSFR1) - Boxplot of baseline biomarker data by study

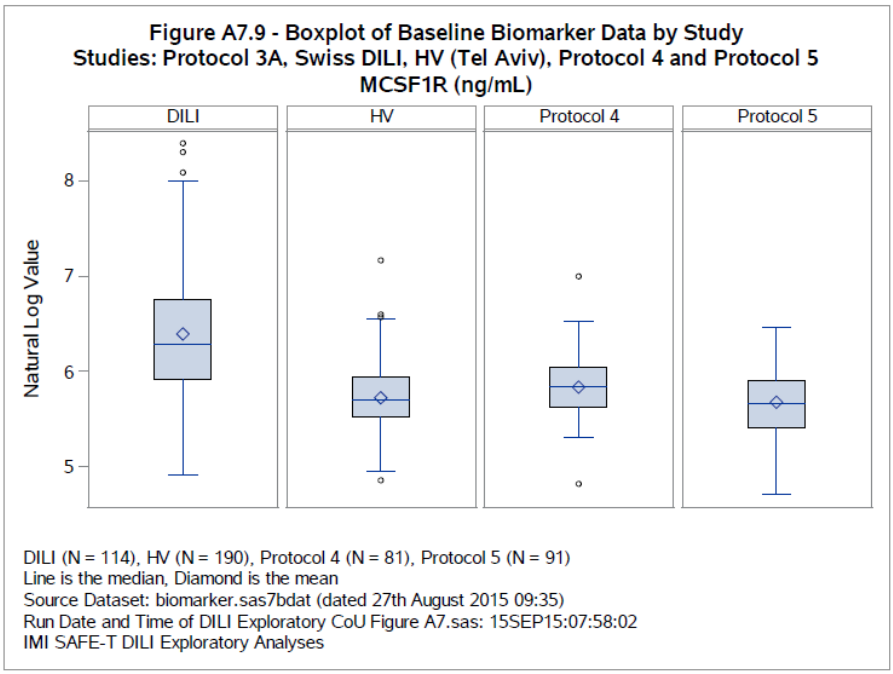


Figure 6-10 microRNA-122 - Boxplot of baseline biomarker data by study

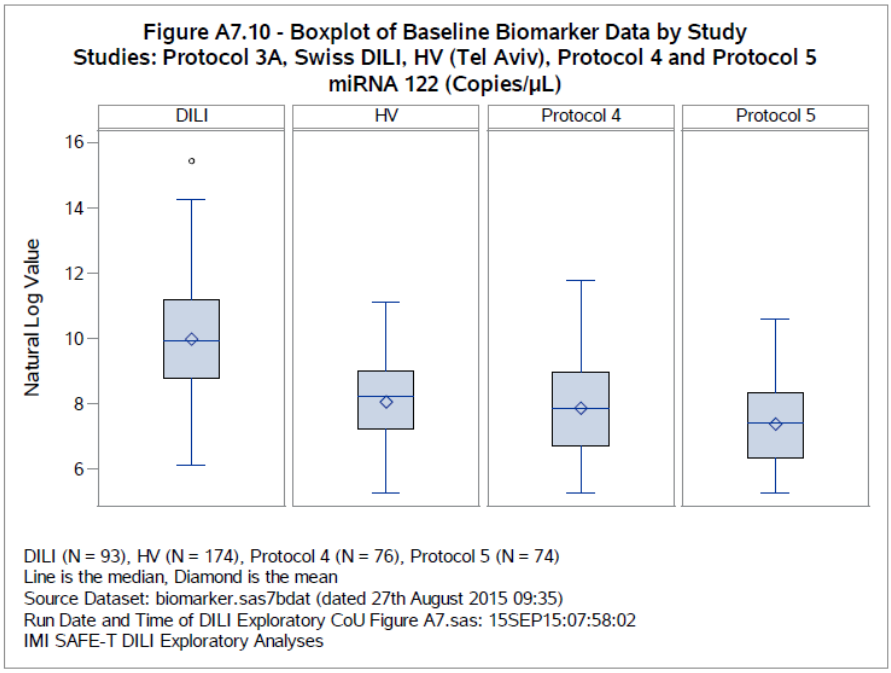


Figure 6-11 Hyperacetylated HMGB1 -- Boxplot of baseline biomarker data by study

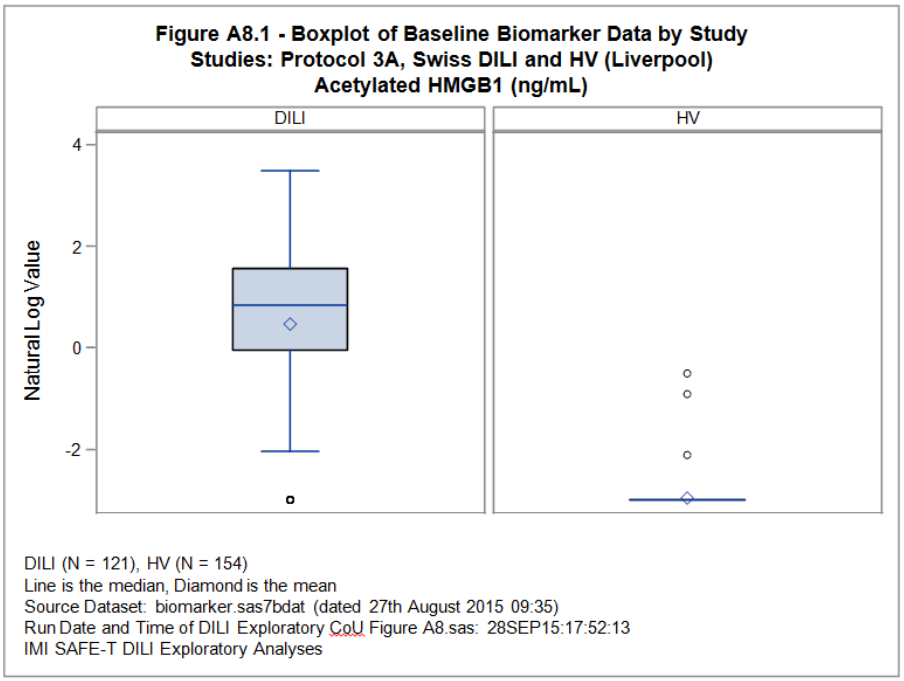
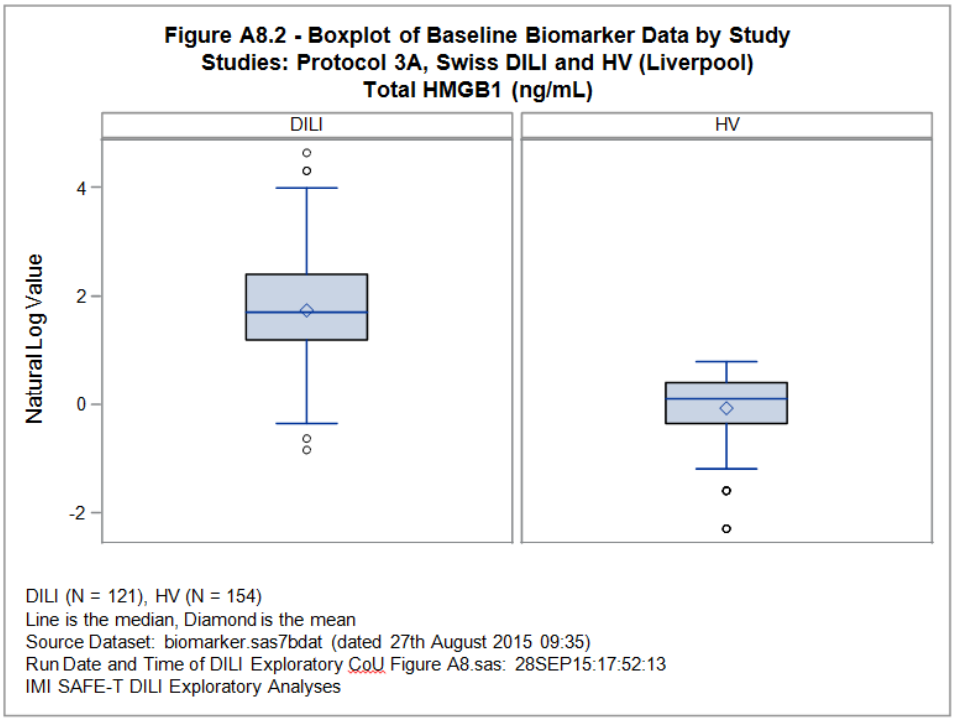


Figure 6-12 Total HMGB1 - Boxplot of baseline biomarker data by study



The performance of each biomarker for detecting acute DILI was calculated as the AUC under the receiver operator characteristic (AUROC) and the following values were obtained: Table 6-6, comparison of acute DILI cases vs. HV:

Table 6-6 Biomarkers ranking and associated AUROC value

Biomarker Ranking	Biomarker Name	AUROC	95% CI
Benchmark	ALT	0.99	(0.98, 1.00)
Benchmark	AST	0.97	(0.96, 0.99)
BM 1	BA64	0.93	(0.89, 0.97)
BM 2	GLDH	0.91	(0.87, 0.95)
BM 3	BA75	0.91	(0.86, 0.95)
BM 4	Total Keratin 18	0.90	(0.85, 0.94)
BM 5	BA19/67	0.89	(0.84, 0.95)
BM 6	ccKeratin 18	0.89	(0.85, 0.93)
BM 7	BA11	0.88	(0.84, 0.93)
BM 8	BA57/32	0.88	(0.83, 0.93)
BM 9	BA56	0.88	(0.83, 0.93)
BM 10	BA61	0.88	(0.82, 0.94)
BM 11	BA30	0.87	(0.82, 0.92)
BM 12	SDH	0.85	(0.80, 0.91)
BM 13	FABP1	0.85	(0.80, 0.90)
BM 14	BA25	0.84	(0.77, 0.91)
BM 15	MCSFR1	0.81	(0.76, 0.87)
BM 16	GST alpha 1*	0.81	(0.75, 0.86)

Biomarker Ranking	Biomarker Name	AUROC	95% CI
BM 17	BA22	0.81	(0.74, 0.88)
BM 18	miRNA 122 (Copies/µL)	0.81	(0.75, 0.86)
BM 19	Alpha-Fetoprotein	0.76	(0.71, 0.82)
BM 20	BA37	0.76	(0.68, 0.83)
BM 21	BA59	0.73	(0.66, 0.81)
BM 22	BA78	0.72	(0.64, 0.80)
BM 23	Cadherin 5	0.71	(0.64, 0.77)
BM 24	Osteopontin	0.67	(0.60, 0.74)
BM 25	LECT2	0.51	(0.43, 0.58)
BM 26	Arginase 1~	0.44	(0.37, 0.52)
BM 27	Paraoxonase 1~	0.37	(0.30, 0.44)

* Non-Convergence due to Quasi-Complete Separation
~ Both AUROC and CI need subtracting from 1 due to low values predicting DILI

Because HMGB1 and hyperacetylated HMGB1 were compared to a different group of healthy volunteers (HV Liverpool), ROC values were calculated separately:

Table 6-7, Comparison of HMGB1 levels in acute DILI cases (n=121 for HMGB1) vs. HV (n=154)

Table 6-7 AUROC values for HMGB1 levels in acute DILI vs. HV

Biomarker Ranking	Biomarker Name	AUROC	95% CI
Benchmark	ALT	0.99	(0.98, 1.00)
Benchmark	AST*	0.98	(0.97, 1.00)
BM 1	Acetylated HMGB1	0.94	(0.92, 0.97)
BM 2	Total HMGB1	0.93	(0.89, 0.96)

* Non-Convergence due to Quasi-Complete Separation

Statistical significance can be assumed when AUROC values exceed 0.50 and the lower value for the 95% CI is > 0.50. Biomarkers (where decreased levels reflect liver injury i.e. AUROC values are less than 0.50) also achieve statistical significance but the values need to be subtracted from 1 (and the confidence interval reversed) to be comparable with the other results.

As part of the exploratory analyses performed, the following biomarkers were chosen to support CoU A based on AUROC values; **GLDH, K18, ccK18, SDH** and **MCSFR1** (i.e. top five biomarkers from Table 6-6), together with **total HMGB1** and **acetylated HMGB1.** In addition**, microRNA-122** was selected despite not being among the top five performers, considering the amount of additional published evidence indicating liver specificity and the additional support from analyses supporting CoU C.

Despite good performance, some of the BMs were ***not*** selected for the CoU A in view of the lack of supportive data from the literature.

Biomarker levels in the non-DILI patient population receiving potential hepatotoxic drugs (protocol 4 and 5) did not differ from levels measured in healthy volunteers.

Additional analyses supporting some of the selected BMs for CoU statement A

Several biomarkers have previously been characterized in the context of acetaminophen (APAP) induced DILI. These include miR-122, K18, ccK18 and hyperacetylated_HMGB1 (39, 41). APAP induced DILI is an intrinsic, dose-dependent form of toxicity that is triggered by formation of the toxic metabolite NAPQI. To assess whether idiosyncratic forms of DILI also show alterations in the levels of these biomarkers, we compared DILI caused by APAP with non-APAP drug classes and HVs. Non-APAP drug classes were the following: flupirtine (N=14), amoxicillin (N=11), antibiotics excluding amoxicillin (N=21), chemotherapy (N=6), NSAID (N=4) and Other (N=44), see Table 6-2.

The results showed that selected biomarkers performed exceptionally well in flupirtine-induced DILI, a prototype for severe idiosyncratic DILI. All flupirtine cases fulfilled Hy's law criteria. In the case of acetaminophen, Hy's law was fulfilled in some but not all cases, despite ALT levels reaching very high values even in some of the non-Hy's law cases. Analysis of biomarkers in flupirtine-induced DILI compared to healthy volunteers yielded the following ROC values (Table 6-8):

Table 6-8 Biomarker performance in flupirtine-induced DILI compared to healthy volunteers from Tel Aviv study

Biomarker Ranking	Biomarker Name	AUROC	95% CI
Benchmark	Total Bilirubin#	1.00	(1.00, 1.00)
Benchmark	ALT	1.00	(1.00, 1.00)
Benchmark	AST	1.00	(1.00, 1.00)
Benchmark	ALP	0.97	(0.94, 1.00)
*BM 1**	*BA19/67#*	*1.00*	*(1.00, 1.00)*
*BM 2**	*BA30#*	*1.00*	*(1.00, 1.00)*
*BM 3**	*BA56#*	*1.00*	*(1.00, 1.00)*
*BM 4**	*BA57/32#*	*1.00*	*(1.00, 1.00)*
*BM 5**	*BA61*	*1.00*	*(1.00, 1.00)*
*BM 6**	*BA64#*	*1.00*	*(1.00, 1.00)*
*BM 7**	*BA75#*	*1.00*	*(1.00, 1.00)*
BM 8	Total Keratin 18	1.00	(1.00, 1.00)
BM 9	Alpha-Fetoprotein	1.00	(0.99, 1.00)
*BM 10**	*BA22*	*1.00*	*(0.99, 1.00)*
BM 11	MCSFR1	1.00	(0.99, 1.00)
*BM 12**	*BA59*	*0.99*	*(0.97, 1.00)*
BM 13	*BA37*	*0.98*	*(0.97, 1.00)*
*BM 14**	*BA11*	*0.98*	*(0.94, 1.00)*
*BM 15**	*BA25*	*0.98*	*(0.95, 1.00)*
BM 16	Osteopontin	0.96	(0.92, 1.00)
BM 17	BA78	0.94	(0.88, 1.00)
BM 18	FABP1	0.93	(0.84, 1.00)
BM 19	Cadherin 5	0.93	(0.87, 1.00)
BM 20	GLDH	0.90	(0.76, 1.00)
BM 21	ccKeratin 18	0.89	(0.70, 1.00)

Biomarker Ranking	Biomarker Name	AUROC	95% CI
BM 22	LECT2	0.81	(0.62, 1.00)
BM 23	SDH	0.77	(0.59, 0.96)
BM 24	Paraoxonase 1	0.76	(0.57, 0.96)
BM 25	miRNA 122 (Copies/µL)	0.73	(0.51, 0.94)
BM 26	GST alpha 1	0.72	(0.53, 0.92)
BM 27	Arginase 1	0.58	(0.37, 0.79)

Non-Convergence due to Complete Separation

Acetylated and total HMGB1 are derived from a separate analysis since a different group of healthy volunteers (Liverpool HV) was used for the comparison (Table 6-9)

Table 6-9 Biomarker performance in flupirtine-induced DILI compared to healthy volunteers from Liverpool study

Biomarker Ranking	Biomarker Name	AUROC	95% CI
Benchmark	ALT #	1.00	(1.00, 1.00)
Benchmark	AST#	1.00	(1.00, 1.00)
BM 1	hyperacetylated HMGB1	1.00	(1.00, 1.00)
BM 2	Total HMGB1	0.95	(0.85, 1.00)

Non-Convergence due to Complete Separation

MCSFR1, a marker reported to indicate immune activation, performed best in flupirtine DILI with an ROC value of 1.00, suggesting that MCSFR1 could be a good marker of severe idiosyncratic DILI. **Hyperacetylated HMGB1** performed equally well in APAP and flupirtine DILI indicating that it may not discriminate between severe intrinsic and severe idiosyncratic DILI.

AFP – reported in the literature as a marker of regeneration – was higher in APAP and flupirtine DILI than in other drug classes, suggesting activation of regenerative processes in APAP and flupirtine DILI. Accordingly, **LECT2** - also a marker of regeneration that shows ***decreased*** serum levels at the peak of liver regeneration (57) - was lower in these two drug classes. Since APAP is considered to reflect intrinsic rather than idiosyncratic DILI, LECT2 and AFP show no specificity in identifying idiosyncratic DILI.

Biomarker performance in patients fulfilling Hy's Law

In the final dataset of SAFE-T acute DILI studies (protocols 3A and Swiss DILI study), 39 patients fulfilled the Hy's law criteria at baseline (31%), 73 did not, and for 14 patients the data were missing (Table 6-10). According to the Hy's law paradigm, there is a 10% risk of fatality due to hepatocellular DILI with jaundice, i.e. fulfilling Hy's law criteria. However, all but 2 of the 39 SAFE-T acute DILI patients that met the Hy's law criteria had full or partial recovery by week 12. No patient progressed to severe DILI.

Table 6-10 Number of Patients with Hy's Law at Baseline - Studies: Protocol 3A and Swiss DILI

Outcome	N	Percent
Hy's Law at Baseline	39	30.95
No Hy's Law at Baseline	73	57.94

Outcome	N	Percent
Missing	14	11.11

Hy's Law defined as ALT > 3 x ULN and Total Bilirubin > 2 x ULN

To illustrate comparability of the data between the SAFE-T acute DILI cases and the DILIN cohort, we performed the same analysis in the DILIN cohort, in which 125 out of 166 patients developed Hy's law criteria between the date of baseline sampling and month 6. In total 96 cases fulfilled Hy's Law criteria at baseline, 29 cases developed Hy's law criteria during the 6-month follow-up. Of the 125 Hy's law cases, 14 cases developed liver failure and required liver transplantation (n=6) or died from liver-related complications (n=8), 19 had developed chronic DILI by month 6 and 91 recovered spontaneously.

Performance of biomarkers in patients with Hy's Law baseline (SAFE-T dataset) and up to 6 months (DILIN dataset) were compared to No Hy's Law baseline (SAFE-T dataset) and up to 6 months (DILIN dataset), see Table 6-11 and Table 6-12.

Table 6-11 SAFE-T dataset: ROC analysis of Hy's law at baseline versus No Hy's law at baseline

Biomarker Ranking	Biomarker Name	AUROC	95% CI
Benchmark	Total Bilirubin	0.93	(0.89, 0.98)
Benchmark	ALT	0.72	(0.63, 0.82)
BM 1	BA19/67	0.94	(0.89, 0.99)
BM 2	BA57/32	0.91	(0.84, 0.98)
BM 3	BA30	0.90	(0.83, 0.98)
BM 4	BA75	0.90	(0.82, 0.98)
BM 5	BA56	0.89	(0.81, 0.97)
BM 6	BA64	0.87	(0.79, 0.96)
BM 7	BA37	0.85	(0.75, 0.95)
BM 8	MCSFR1	0.84	(0.75, 0.93)
BM 9	BA11	0.82	(0.71, 0.92)
BM 10	BA59	0.81	(0.70, 0.91)
BM 11	BA61	0.80	(0.69, 0.90)
BM 12	Alpha-Fetoprotein	0.78	(0.66, 0.89)
BM 13	BA25	0.74	(0.61, 0.87)
BM 14	Total HMGB1	0.73	(0.62, 0.84)
BM 15	Cadherin 5	0.72	(0.61, 0.83)
BM 16	BA22	0.71	(0.59, 0.83)
BM 17	Osteopontin	0.71	(0.60, 0.83)
BM 18	Total Keratin 18	0.71	(0.60, 0.81)
BM 19	Acetylated HMGB1	0.70	(0.59, 0.81)
BM 20	BA78	0.68	(0.54, 0.82)
BM 21	LECT2	0.66	(0.53, 0.78)
BM 22	ccKeratin 18	0.64	(0.53, 0.76)
BM 23	Paraoxonase 1	0.64	(0.52, 0.76)
BM 24	FABP1	0.60	(0.49, 0.72)
BM 25	GST alpha 1	0.58	(0.46, 0.70)

Biomarker Ranking	Biomarker Name	AUROC	95% CI
BM 26	miRNA 122 (Copies/µL)	0.56	(0.42, 0.69)
BM 27	SDH~	0.48	(0.33, 0.62)
BM 28	Arginase 1~	0.44	(0.33, 0.56)
BM 29	GLDH~	0.44	(0.30, 0.57)

~ Both AUROC and CI need subtracting from 1 due to low values predicting DILI

Table 6-12 DILIN dataset: ROC analysis of Hy's law up to month 6 versus No Hy's law up to month 6

Biomarker Ranking	Biomarker Name	AUROC	95% CI
Benchmark	Total Bilirubin	0.79	(0.69, 0.90)
Benchmark	ALT	0.76	(0.67, 0.85)
Benchmark	ALP	0.68	(0.59, 0.77)
BM 1	Total Keratin 18	0.80	(0.72, 0.88)
BM 2	ccKeratin 18	0.78	(0.70, 0.86)
BM 3	MCSFR1	0.75	(0.67, 0.84)
BM 4	Total HMGB1	0.74	(0.65, 0.83)
BM 5	FABP1	0.73	(0.64, 0.82)
BM 6	Cadherin 5	0.71	(0.62, 0.80)
BM 7	Alpha-Fetoprotein	0.68	(0.60, 0.76)
BM 8	Arginase 1	0.65	(0.56, 0.75)
BM 9	GST alpha 1	0.63	(0.53, 0.73)
BM 10	Acetylated HMGB1	0.60	(0.53, 0.66)
BM 11	Osteopontin	0.59	(0.49, 0.69)

Hy's law is defined as a >2-fold elevation of bilirubin in patients with an ALT >3x ULN, thus a biomarker can by definition not outperform bilirubin in assessing Hy's law. However, several biomarkers outperformed ALT in discriminating DILI severity in Hy's law cases from non-Hy's law cases, as shown in Table 6-11 and Table 6-12 above (**MCSFR1** in the case of SAFE-T, **total K18** and **ccK18** in the case of DILIN). This is of importance, since bilirubin is not a specific marker of severe DILI but could also result from interference of a drug with bilirubin metabolism/excretion. In these situations, biomarkers could help "de-risk" a biochemical Hy's Law situation.

The prognostic value of these biomarkers in comparison to standard biomarkers of ALT and bilirubin (Hy's Law) is further addressed in the Context of Use B analysis that included a sufficient number of cases with progressive liver failure recruited with the DILIN cohort outside of SAFE-T (see section 6.3.2 below).

Combinations of biomarkers together with potentially key covariates, sex, age, BMI and alcohol consumption, were included in a classification tree analysis for DILI subjects versus healthy volunteers, DILI subjects versus protocol 4 subjects and DILI subjects versus protocol 5 subjects to see if the AUC could be improved compared with that obtained for the biomarkers with the best results. No real evidence of improved performance was obtained.

Key covariates were also added to each individual biomarker in a logistic regression for the comparison of DILI subjects with healthy volunteers. Again there was no real evidence that

discrimination performance was improved for individual biomarkers with the best AUC results.

Random Forest analyses were carried out for all biomarkers. These plots tended to show that the same biomarkers identified in univariate analyses as important for discriminating between DILI subjects and healthy volunteers were identified by this multivariate analysis method.

6.3.2 Results pertinent to CoU B

"Based on preliminary data, the biomarkers MCSFR1 and hyperacetylated HMGB1 have potential as clinical DILI biomarkers that sponsors may choose to incorporate into their clinical trials to anticipate a risk for progression of hepatocellular injury to severe DILI in patients in whom an initial DILI diagnosis has been established based on elevations of the standard marker ALT alone or in combination with TBIL."

No patient in the SAFE-T acute DILI studies progressed to severe DILI. Thus, to assess the value of biomarkers in predicting progression from acute DILI to severe DILI (i.e. liver failure resulting in transplantation or liver-related death) in patients in whom an initial DILI diagnosis had been established based on elevations of standard markers ALT and TBIL, samples from 166 patients recruited in the US DILIN cohort were analyzed . The mean time intervals between the date of DILI diagnosis and date of sampling were as follows (note there is one patient with 51 days and one with 38, all the others were 25 days and less):

Table 6-13 Time intervals (in days) between the date of DILI diagnosis and date of sampling in the DILIN cohort

Outcome	N	Mean	SD	Median	Min	Max
Liver-related death / transplantation	16	6.13	3.12	6.00	1.00	11.0
Chronic DILI	22	7.77	4.58	8.00	0.00	14.0
Recovery	128	8.52	6.60	7.00	0.00	51.0

All cases were considered idiosyncratic DILI; APAP-induced DILI was excluded. Out of 166 cases, 16 patients progressed to severe DILI with liver failure leading to liver-related death or transplantation.

To correlate biomarker levels in the initial baseline sample with actual clinical outcome, the following three groups were compared for the standard biomarkers of DILI (components of Hy's Law) and novel biomarkers.

Liver-related death/transplant: patients with acute DILI who progressed to severe DILI (n=16)

Chronic DILI: patients with acute DILI who still showed elevated ALT levels at month 6 (n=22)

Recovery: patients with acute DILI who showed full recovery at month 6 (n= 128)

Figure 6-13 to 19 shows boxplots for the standard biomarkers of ALT and bilirubin and the final set of biomarkers selected for inclusion in CoU B (DILIN cohort), i.e. total and acetyl HMBGB1, MCSFR1, osteopontin and cytokeratin-18.

Figure 6-13 Alanine aminotransferase (ALT) – Boxplot of baseline biomarker data by outcome up to 6 months

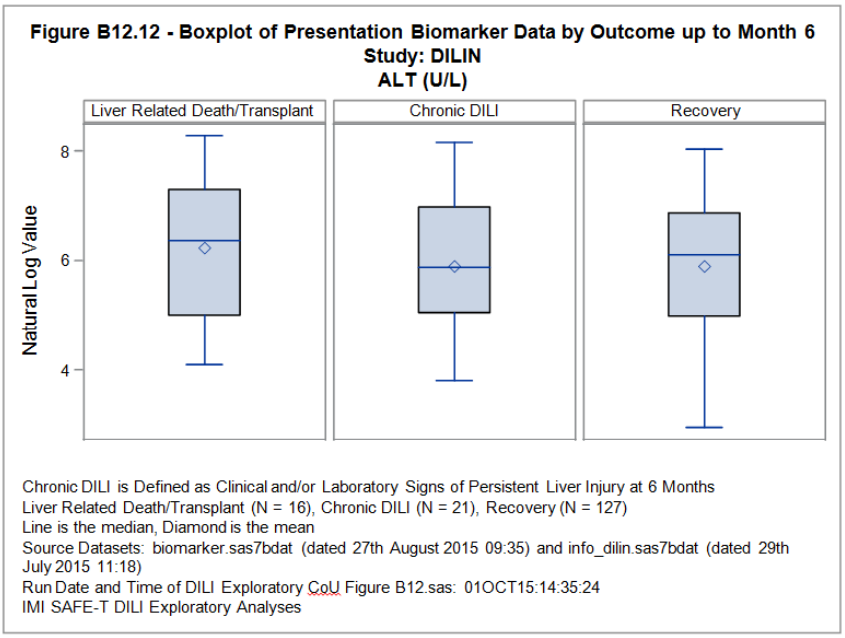


Figure 6-14 Total bilirubin – Boxplot of baseline biomarker data by outcome up to 6 months

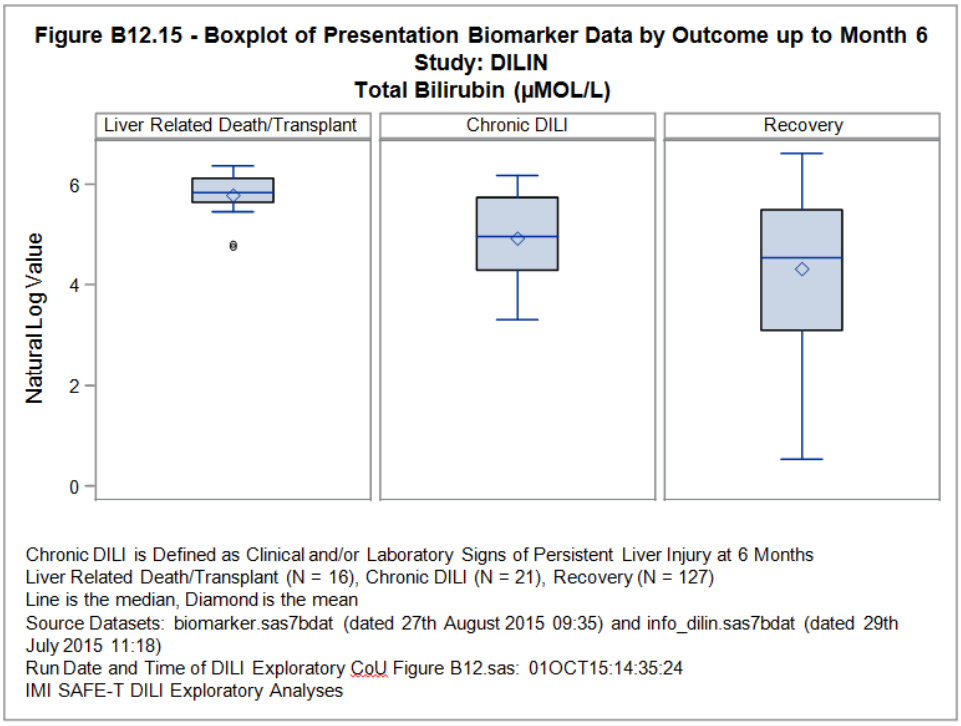


Figure 6-15 Total HMGB1– Boxplot of baseline biomarker data by outcome up to 6 months

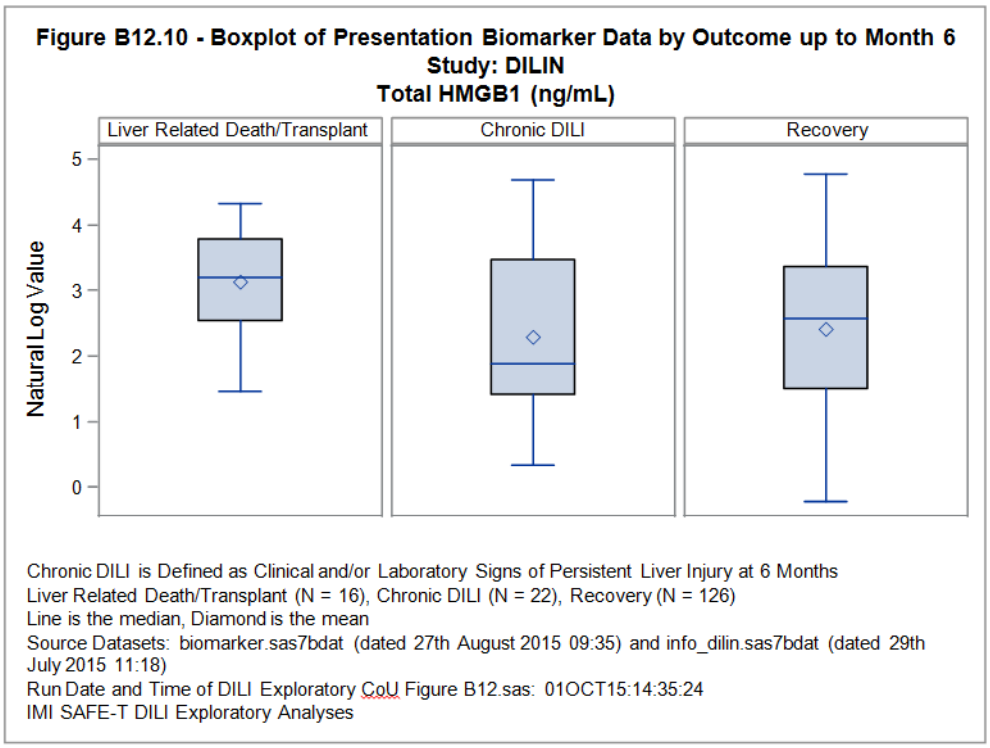


Figure 6-16 Acetylated HMGB1– Boxplot of baseline biomarker data by outcome up to 6 months

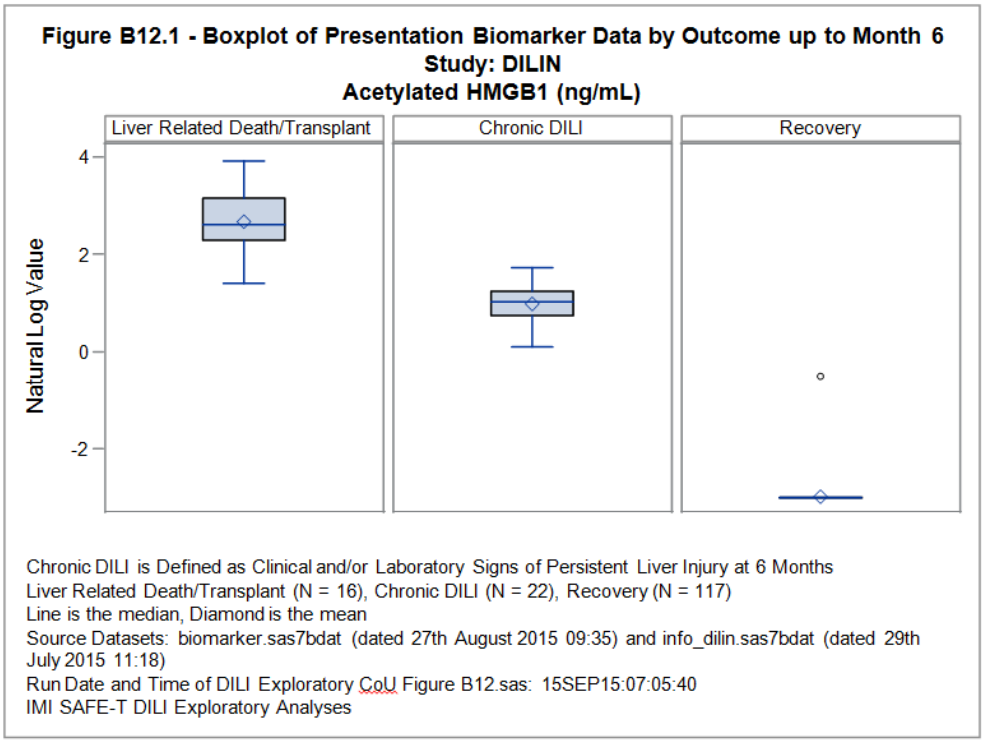


Figure 6-17 MCSFR1 - Boxplot of baseline biomarker data by outcome up to 6 months

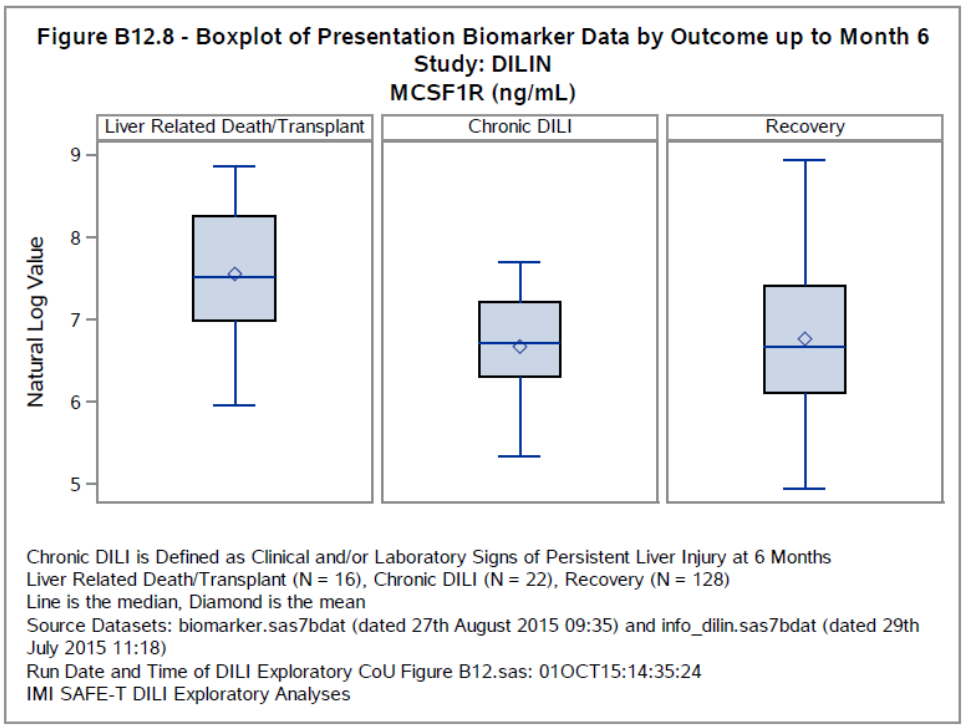


Figure 6-18 Osteopontin - Boxplot of baseline biomarker data by outcome up to 6 months

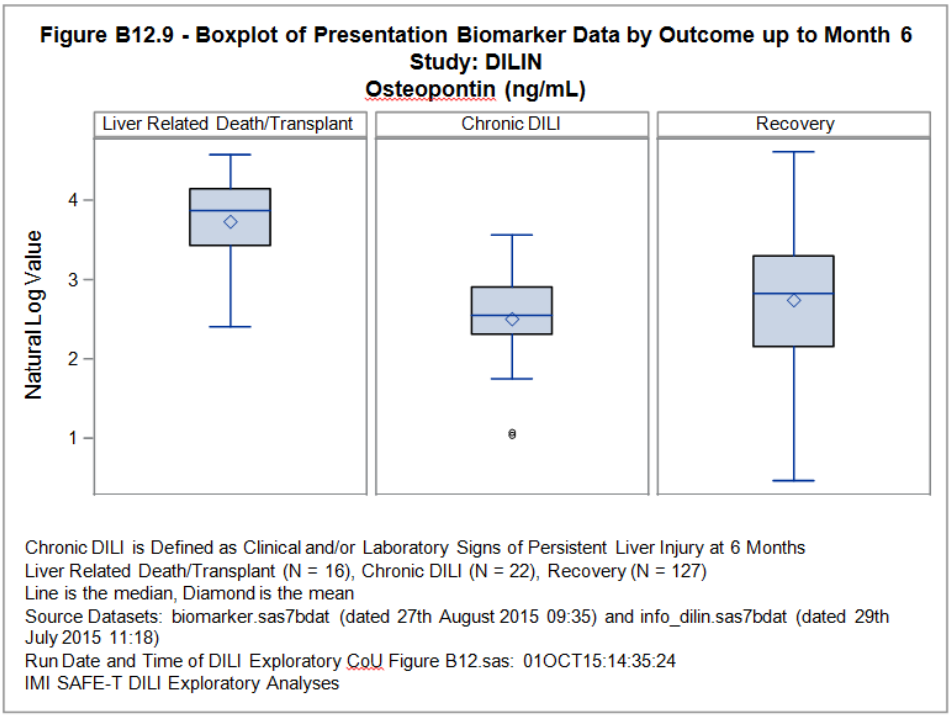
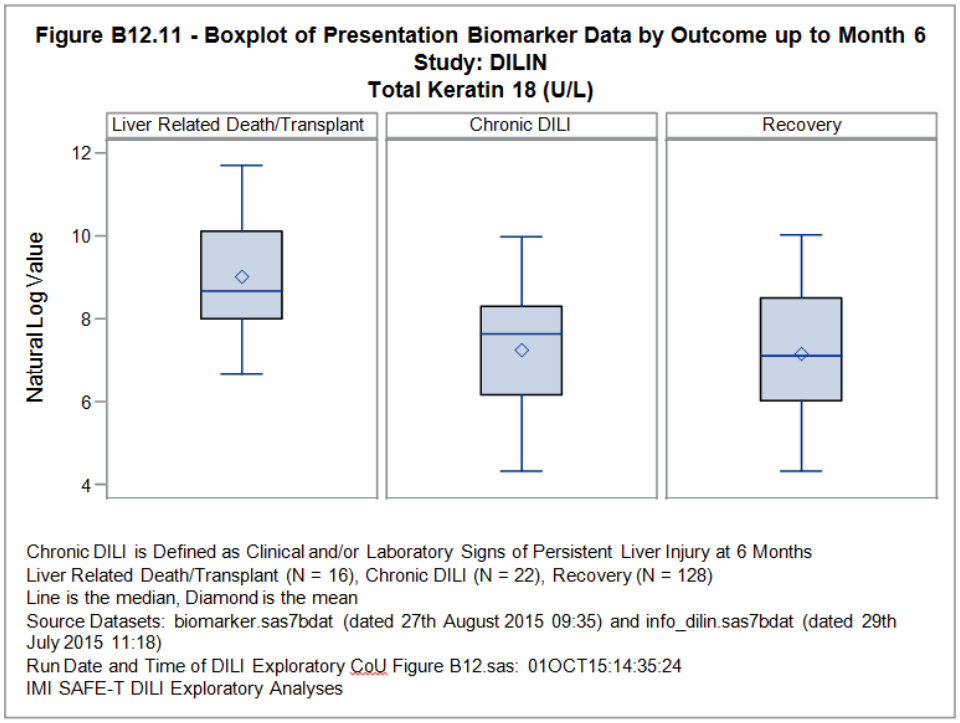


Figure 6-19 Total CK-18 - Boxplot of baseline biomarker data by outcome up to 6 months



The AUROC values comparing patients with acute DILI who recovered spontaneously with acute DILI patients who progressed to liver failure are shown in Table 6-14).

Table 6-14 Comparison of biomarkers in DILIN patients who progressed to liver failure (n=16) with DILIN patients who recovered (n=128)

Biomarker Ranking	Biomarker Name	AUROC	95% CI
Benchmark	Total Bilirubin	0.85	(0.77, 0.93)
Benchmark	ALP	0.60	(0.45, 0.74)
Benchmark	ALT	0.57	(0.40, 0.74)
BM 1	Acetylated HMGB1#	1.00	(1.00, 1.00)
BM 2	Osteopontin	0.83	(0.73, 0.93)
BM 3	Total Keratin 18	0.80	(0.69, 0.90)
BM 4	MCSFR1	0.75	(0.64, 0.87)
BM 5	ccKeratin 18	0.74	(0.63, 0.86)
BM 6	FABP1	0.70	(0.58, 0.81)
BM 7	Total HMGB1	0.67	(0.54, 0.80)
BM 8	Cadherin 5	0.65	(0.53, 0.77)
BM 9	Arginase 1	0.58	(0.43, 0.73)
BM 10	GST alpha 1~	0.45	(0.28, 0.63)
BM 11	Alpha-Fetoprotein~	0.33	(0.21, 0.45)

Non-Convergence due to Complete Separation

~ Both AUROC and CI need subtracting from 1 due to low values predicting Liver Related Death/Transplant

The values shown in Table 6-14 above were almost identical if patients who developed liver failure are compared to [recovery + chronic DILI] patients as a single comparator group, i.e. adding the 22 patients with chronic DILI to the recovery group.

The most striking finding in this analysis was the **perfect separation** between liver failure and recovery that was obtained with **acetylated HMGB1** outperforming both standard biomarkers of ALT and bilirubin (see Figure 6-17 above).

Biomarkers vs Hy's Law in predicting liver failure

DILIN comprised 96 subjects with Hy's Law at baseline and 70 subjects without Hy's Law at baseline. Sixteen subjects died or underwent transplant, the remaining 150 subjects developed chronic DILI (n=22) or recovered (n=128):

- of the 16 who died or underwent transplantion, Hy's Law was satisfied at baseline for 13, thus the sensitivity of Hy's Law as a predictor of liver failure is 13/16 = 0.81.
- of the 150 who developed chronic DILI or recovered, 67 did not fulfill Hy's Law criteria at baseline, thus the specificity for non-Hy's law as a predictor of recovery is 67/150 = 0.45.
- the AUROC for acetylated HMGB1 and total HMGB1 are given in Table 6-13 as 1.00 and 0.68 respectively. It can be concluded that -
  a) the performance of **acetylated HMGB1** in distinguishing death/transplant from chronic DILI/recovery was **better than that of Hy's Law** because acetylated HMGB1 demonstrates perfect separation between the two groups.
  b) the performance of **total HMGB1** was also **better than that of Hy's Law** because the combination of sensitivity and specificity for Hy's Law as a predictor lies below the ROC curve for total HMGB1.

Table 6-15 ROC Analysis of Presentation Biomarker Data by Outcome: Liver Related Death/Transplant vs Recovery/Chronic DILI - Study: DILIN

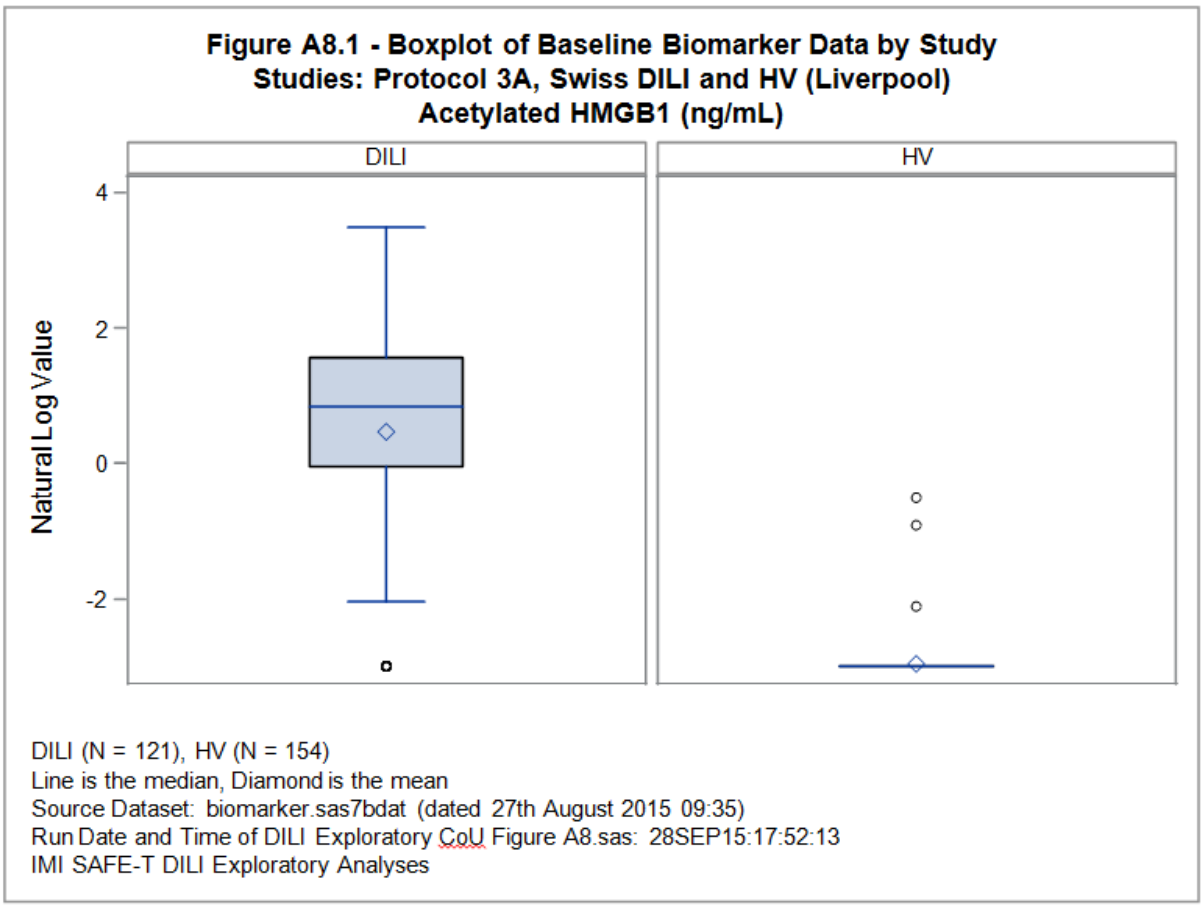
Biomarker Ranking	Biomarker Name	AUROC	95% CI
Benchmark	Total Bilirubin	0.84	(0.75, 0.92)
Benchmark	ALP	0.63	(0.48, 0.77)
Benchmark	ALT	0.57	(0.40, 0.74)
BM 1	Acetylated HMGB1	1.00	(0.99, 1.00)
BM 2	Osteopontin	0.85	(0.75, 0.94)
BM 3	Total Keratin 18	0.80	(0.70, 0.90)
BM 4	MCSF1R	0.76	(0.64, 0.88)
BM 5	ccKeratin 18	0.73	(0.62, 0.85)
BM 6	FABP1	0.70	(0.59, 0.82)
BM 7	Total HMGB1	0.68	(0.55, 0.80)
BM 8	Cadherin 5	0.64	(0.52, 0.77)
BM 9	Arginase 1	0.59	(0.44, 0.75)
BM 10	GST alpha 1~	0.46	(0.29, 0.63)
BM 11	Alpha-Fetoprotein~	0.32	(0.21, 0.44)

~ Both AUROC and CI need subtracting from 1 due to low values predicting Liver Related Death/Transplant

Thus, both acetylated and HMGB1 and total HMGB1 are better than Hy's Law in predicting serious injury resulting in death/transplantation.

It was noted that the levels of acetylated HMGB1 in the "Recovery" DILIN patients were no higher than in SAFE-T healthy volunteers (see Figure 6-16 and Figure 6-20).

Figure 6-20 Acetylated HMGB1 Boxplot of Baseline Biomarker Data by Study



This may be related to the timing of the samples, but further studies will be needed to elucidate the temporal dynamics of this biomarker in relation to the onset of DILI.

To test this hypothesis, acetylated-HMGB1 levels in SAFE-T samples obtained at week 1 were measured. The results showed that, in these week 1 samples, acetylated-HMGB1 levels (that were elevated in the baseline blood sample) were indeed significantly lower (i.e. closer to the level in healthy volunteers).

Combinations of biomarkers together with potentially key covariates, sex, age, BMI and alcohol consumption, were included in a classification tree analysis for acute liver injury subjects versus subjects that recover to see if the AUC could be improved compared with that obtained for the individual biomarkers with the best results. No evidence of improved performance was obtained.

6.3.3 Results pertinent to CoU C

"Based on preliminary data, the following biomarkers: total HMGB1, total and caspase-cleaved keratin 18, miR-122 and GLDH have potential as clinical safety biomarkers that sponsors may choose to incorporate into their clinical trials for the prediction of liver injury in patients who have taken an overdose of acetaminophen and who present at an early stage (within the first 24 hours) before ALT increases."

Published data have monitored the levels of biomarkers in patients who have taken an overdose of acetaminophen (APAP, paracetamol) but were immediately hospitalized at an early stage of DILI (4). The primary data were re-analyzed to assess the claim that biomarkers could detect DILI at an early stage before ALT increases. A total of approx. 100 APAP overdose patients were selected on the basis of the following two criteria:

- overdose of APAP
- normal ALT at the time of admission and initial clinical workup

Of these 100 patients, 15 subsequently developed liver injury as defined by an increase in ALT >3x ULN, whereas 85 patients did not develop liver injury, thus indicating a more benign clinical course.

Figure 6-21 to 25 compares patients with outcome liver injury to those that did not develop liver injury for a limited set of novel biomarkers and the standard marker ALT:

Figure 6-21 ALT - Boxplot by outcome study

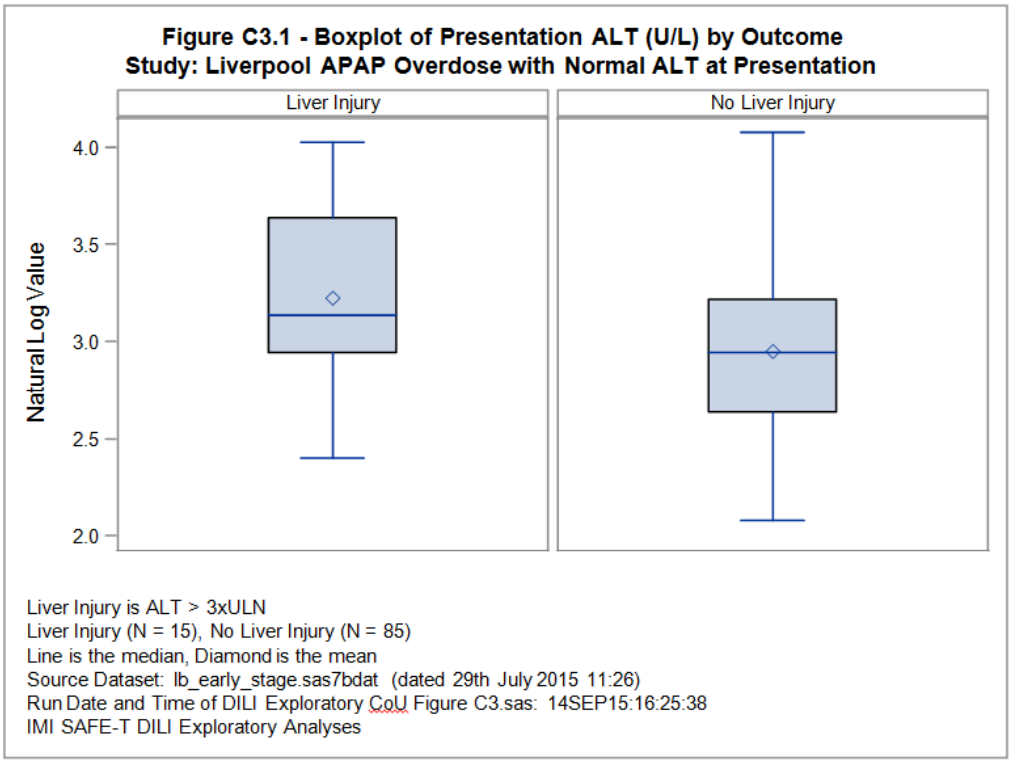


Figure 6-22 Total HMGB1 – Boxplot by outcome study

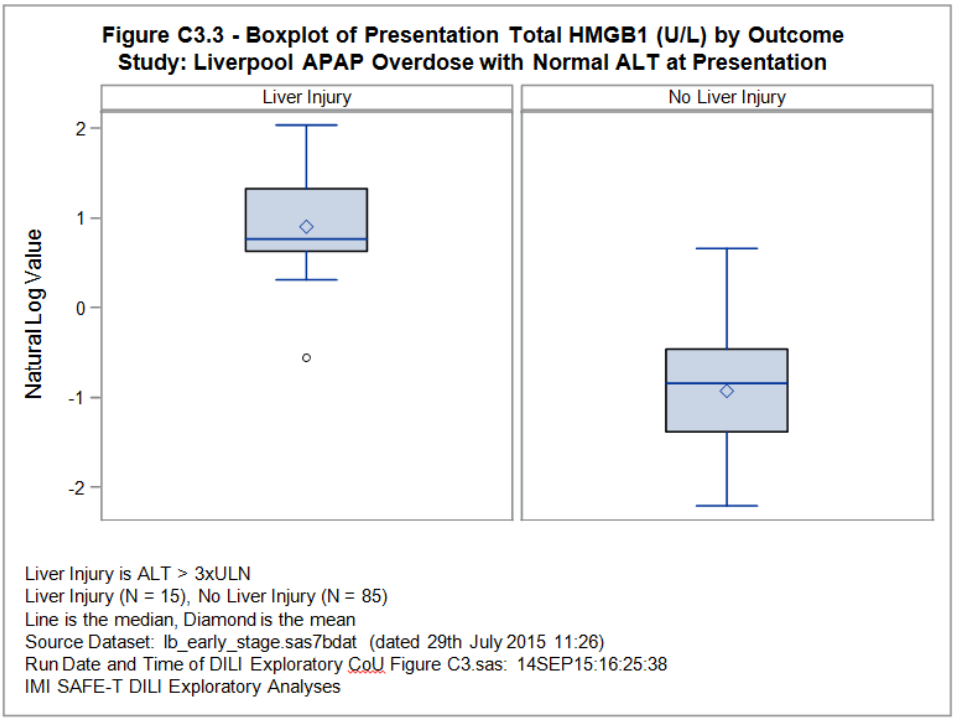


Figure 6-23 miR-122 – Boxplot by outcome study

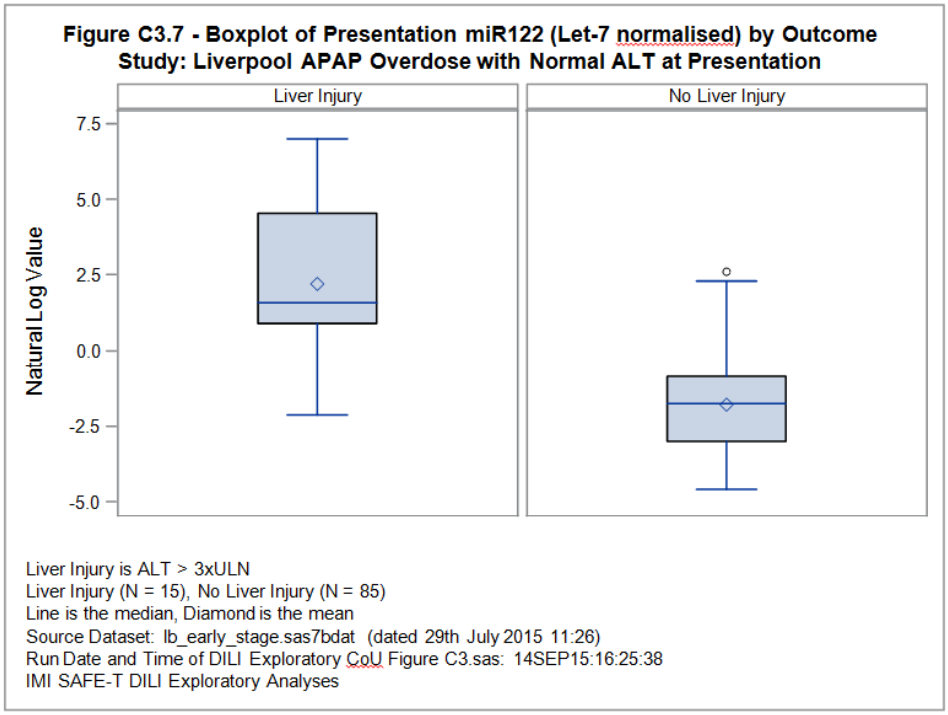


Figure 6-24 Total keratin 18 (K18) – Boxplot by outcome study

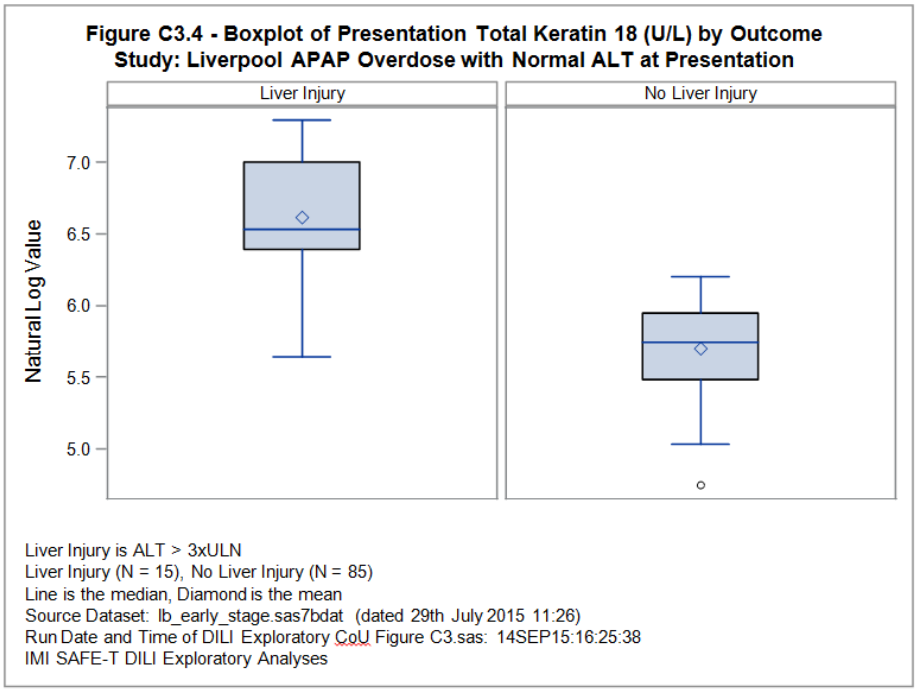
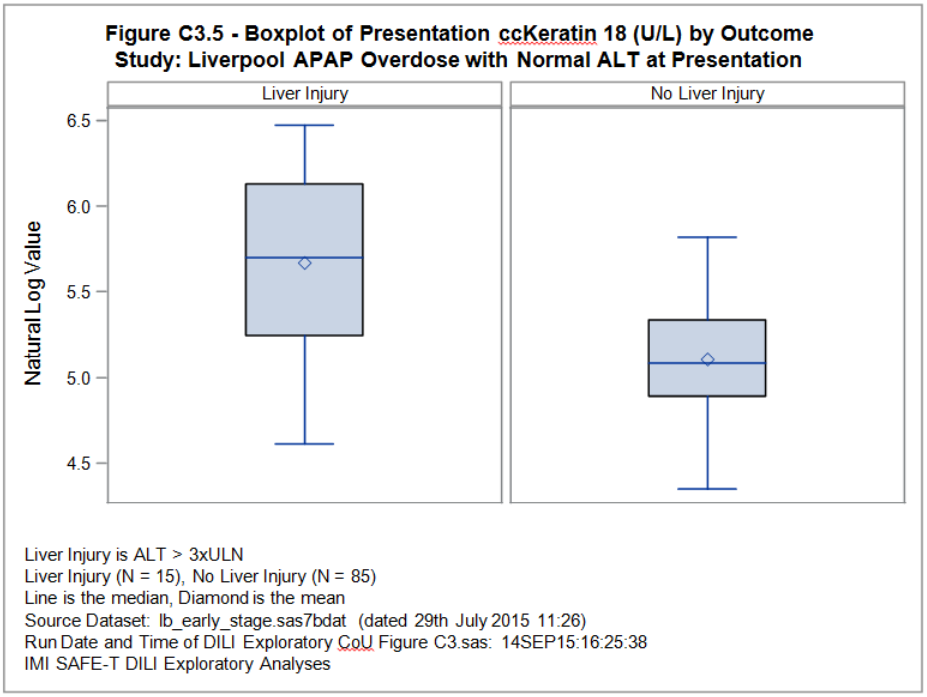


Figure 6-25 Caspase-cleaved keratin 18 - Boxplot by outcome study



The corresponding AUROC values are shown in Table 6-16.

Table 6-16 Biomarker ROC values in APAP overdose patients with outcome liver injury compared to those that did not develop liver injury

Biomarker Ranking	Biomarker Name	AUROC	95% CI
Benchmark	ALT	0.65	(0.48, 0.83)
Benchmark	INR	0.59	(0.44, 0.74)
BM 1	Total HMGB1	0.98	(0.94, 1.00)
BM 2	total K18	0.95	(0.87, 1.00)
BM 3	miR122 (Let-7 normalised)	0.94	(0.87, 1.00)
BM 4	GLDH	0.82	(0.69, 0.94)
BM 5	caspase-cleaved K18	0.79	(0.63, 0.96)

Liver Injury is ALT > 3xULN

The levels of biomarkers in the patients who subsequently developed liver injury were higher than in those who did not develop liver injury, but were lower than in the SAFE-T acute DILI patients with APAP overdose. It thus appears as if the difference between the two groups could reflect a difference in the degree of liver injury that anticipates the subsequent clinical course. In this respect the biomarkers addressed the Context of Use C with respect to early detection of the degree of liver injury as a determinant of the subsequent clinical course. However, defining thresholds would not yet be possible based on the current data. An additional limitation is the DILI causing drug, since APAP leads to intrinsic, dose-dependent liver injury and in this sense may not be representative of idiosyncratic DILI which typically is not predictable and becomes evident only once liver chemistry tests are in the pathological range. However, idiosyncratic DILI is not expected in early clinical trials due to its low incidence. Although often predicted by animal studies, intrinsic DILI may be seen in early drug development and a sensitive marker of liver injury would add value when intensified monitoring is required, e.g. based on uncertain findings in pre-clinical studies.

Combinations of biomarkers together with potentially key covariates, sex, age, BMI and alcohol consumption, were included in a classification tree analysis for liver injury subjects versus no liver injury subjects to see if the AUC could be improved compared with that obtained for the individual biomarkers with the best results. No evidence of improved performance was obtained.

6.3.4 Justification for not defining diagnostic cut-offs

No specific cut-offs were chosen in the above analyses with reliance instead being placed on AUROC as a general measure of discriminatory ability. Explicit cut-off values or a specific rule for determining a cut-off, such as that corresponding to 95% specificity, would need to be specified before attempting to obtain confirmatory evidence of the utility of a biomarker or combination of biomarkers. If a classification tree approach were ultimately used then the rules given by the tree combined with the probability level required for a classification of DILI would give a point on the AUROC as needed.

Ranges of ALT and bilirubin have been established previously in blood donors and the general population by one of the SAFE-T partners (Paris - APHP) (66, 67). The impact of ULN ALT variability on the definition of DILI, as well as the use of daily screening to increase the detection of DILI when centralized in a biochemistry department, has been assessed and published by the SAFE-T consortium (68).

Further information on the normal ranges for the PSTC (Table 6-17) and Tel Aviv (Table 6-18) data are included below:

For the PSTC HV data, a mixed effect model for log transformed data was to be used to obtain the variance components for between subject variation and within subject variation assuming a log-normal distribution. The 95th percentile was to be obtained using the estimated mean and standard deviation for the log-normal distribution. The 95th percentile was to be taken as the estimate of the ULN. For two of the markers (ccK18 and GSTA) 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 to be used to estimate the ULN. For one of the markers (K18) a nonparametric approach was to be used to obtain the 95th percentile because approximately 93% of the data were below LLoQ. Similar analyses were to be carried out to derive normal ranges for the Tel Aviv healthy volunteer data.

Table 6-17 PSTC HV study normal ranges

Biomarker	Unit	Estimated Geometric Mean	Intra-Subject CV (%)	Inter-Subject CV (%)	Estimated ULN (95% Percentile)
Alpha-Fetoprotein	ng/mL	0.68	31.93	61.53	1.98
Arginase 1	ng/mL	7.63	37.46	46.03	19.46
ccKeratin 18	U/L	90.65	.	.	260.16
Cadherin 5	ng/mL	2798.89	17.69	18.00	4225.87
FABP1	ng/mL	6.91	32.86	32.75	14.55
GLDH	U/L	2.71	34.53	52.74	7.24
GST alpha 1	ng/mL	6.31	.	.	60.00
Total Keratin 18	U/L	.	.	.	121.35
LECT2	ng/mL	252.27	28.64	20.97	447.96
MCSFR1	ng/mL	334.81	13.89	30.08	571.64
miRNA 122	Copies /µL	2152.98	93.56	90.89	13356.52
Osteopontin	ng/mL	4.13	26.61	52.15	10.31
Paraoxonase 1	ng/mL	294.97	31.67	43.33	690.82
Prothrombin	µg/L	60.07	13.48	17.58	86.29
SDH	U/L	3.02	41.01	43.43	7.75

Table 6-18 Tel Aviv study normal ranges

Biomarker	Unit	Estimated Geometric Mean	CV (%)	Estimated ULN (95% Percentile)
ALT	U/L	23.14	54.74	53.71
Alpha-Fetoprotein	ng/mL	0.99	90.21	3.54
Arginase 1	ng/mL	35.97	42.57	70.38
Cadherin 5	µg/mL	2287.79	39.52	4281.32
FABP1	ng/mL	9.21	44.55	18.54
GLDH	U/L	3.00	79.68	9.51
GST alpha 1	ng/mL	6.61	.	64.11

Biomarker	Unit	Estimated Geometric Mean	CV (%)	Estimated ULN (95% Percentile)
LECT2	ng/mL	177.96	47.50	373.74
MCSF1R	ng/mL	306.67	34.75	534.39
Osteopontin	ng/mL	6.54	58.56	15.99
Paraoxonase 1	ng/mL	639.81	48.05	1354.22
Prothrombin	µg/L	67.75	20.29	94.28
SDH	U/L	1.79	101.57	7.17
Total Bilirubin	µMOL/L	10.68	46.04	21.97
Total Keratin 18	U/L	.	.	151.14
ccKeratin 18	U/L	139.99	65.39	373.55
miRNA 122	Copies/µL	3173.64	213.51	27367.62

6.3.5 Potential impact of various intrinsic and extrinsic factors on expected test performance

A number of intrinsic factors were examined when all biomarkers were included in the classification tree analysis, as described in the previous sections. These did not have a significant effect in the sense that the AUROC results were not appreciably better than the results for the best biomarkers considered alone.

7 Follow-up actions toward qualification of DILI Biomarkers

The results of the SAFE-T DILI work package clearly show that selected biomarkers offer potential as diagnostic tools for the management of DILI. Given that patients with acute DILI of various causes were included on the basis of ALT, ALP and bilirubin levels, the biomarkers under study were by virtue of study design unable to outperform ALT and bilirubin in terms of diagnosing DILI. Several analyses yielded an advantage of biomarkers over existing diagnostic tools, for instance in the setting of Context of Use B where both forms of HMGB1 performed better than Hy's law in predicting progression to acute liver failure. The following two key follow-up initiatives should be pursued further:

1) Following the endorsement of regulators for exploratory measurement of biomarkers in development programs, the true value of the biomarkers can only be assessed in prospective trials with serial blood sampling before and at various timepoints following onset of drug treatment. This will allow conclusions to be made as to the optimal timing of biomarker measurements in clinical studies. In addition, performance of the biomarkers in various sub-populations, such as those with malignant diseases or underlying liver disease, as well as pediatric patients needs to be explored and respective reference ranges need to be established.
2) A second future activity should assess performance of biomarkers in other forms of liver injury such as viral hepatitis, alcoholic hepatitis, non-alcoholic fatty liver disease and steatohepatitis (NAFLD and NASH), and fibrosis. Drug-induced liver injury that leads to ongoing chronic liver damage with onset of fibrosis is an additional category that should be investigated.

To address these open issues, respective objectives have been incorporated into the scope of a planned follow-up project, the Translational Safety Biomarker Pipeline – TranSBioline. TranSBioline's key objectives will be to

- Establish a public-private framework to support continuous safety biomarker development, qualification, and exploitation
- Allow for continuous feed-in of promising markers and addition of new organ areas into a learn/confirm qualification pipeline
- Complete full qualification of safety biomarkers investigated in IMI SAFE-T
- Qualify new safety biomarkers for pancreatic injury
- Extend biomarker qualification to application in clinical practice
- For qualified biomarkers,
  - Develop point-of-care diagnostics for a subset of markers
  - Bridge preclinical and clinical biomarker assessment to in vitro and in silico models
- Establish a comprehensive reference safety database with biomarker profiles across relevant target patient populations and healthy volunteers, including data on new and established safety biomarkers along with demographic, adverse event, comed, and medical history information

The TranSBioline proposal is planned to be submitted to Call 11 under IMI2, with a tentative start date in fourth quarter of 2017 in case of project selection by the IMI.

8 References

1. Matheis K, Laurie D, Andriamandroso C, Arber N, Badimon L, Benain X, Bendjama K, et al. A generic operational strategy to qualify translational safety biomarkers. Drug Discov Today 2011;16:600-608.

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

3. Pillans PI, Ghiculescu RA, Lampe G, Wilson R, Wong R, Macdonald GA. Severe acute liver injury associated with lumiracoxib. J Gastroenterol Hepatol 2012;27:1102-1105.

4. Antoine DJ, Dear JW, Lewis PS, Platt V, Coyle J, Masson M, Thanacoody RH, et al. Mechanistic biomarkers provide early and sensitive detection of acetaminophen-induced acute liver injury at first presentation to hospital. Hepatology 2013;58:777-787.

5. Sgro C, Clinard F, Ouazir K, Chanay H, Allard C, Guilleminet C, Lenoir C, et al. Incidence of drug-induced hepatic injuries: a French population-based study. Hepatology 2002;36:451-455.

6. Bjornsson ES, Bergmann OM, Bjornsson HK, Kvaran RB, Olafsson S. Incidence, presentation, and outcomes in patients with drug-induced liver injury in the general population of Iceland. Gastroenterology 2013;144:1419-1425, 1425 e1411-1413; quiz e1419-1420.

7. Friis H, Andreasen PB. Drug-induced hepatic injury: an analysis of 1100 cases reported to the Danish Committee on Adverse Drug Reactions between 1978 and 1987. J Intern Med 1992;232:133-138.

8. Zimmerman HJ. Drug-induced liver disease: an overview. Semin Liver Dis 1981;1:93-103.

9. Bottiger LE, Furhoff AK, Holmberg L. Fatal reactions to drugs. A 10-year material from the Swedish Adverse Drug Reaction Committee. Acta Med Scand 1979;205:451-456.

10. Wysowski DK, Swartz L. Adverse drug event surveillance and drug withdrawals in the United States, 1969-2002: the importance of reporting suspected reactions. Arch Intern Med 2005;165:1363-1369.

11. Bower WA, Johns M, Margolis HS, Williams IT, Bell BP. Population-based surveillance for acute liver failure. Am J Gastroenterol 2007;102:2459-2463.

12. Ostapowicz G, Fontana RJ, Schiodt FV, Larson A, Davern TJ, Han SH, McCashland TM, et al. Results of a prospective study of acute liver failure at 17 tertiary care centers in the United States. Ann Intern Med 2002;137:947-954.

13. O'Grady JG, Alexander GJ, Hayllar KM, Williams R. Early indicators of prognosis in fulminant hepatic failure. Gastroenterology 1989;97:439-445.

14. Hoofnagle JH, Carithers RL, Jr., Shapiro C, Ascher N. Fulminant hepatic failure: summary of a workshop. Hepatology 1995;21:240-252.

15. Fisher K, Vuppalanchi R, Saxena R. Drug-Induced Liver Injury. Arch Pathol Lab Med 2015;139:876-887.

16. Leise MD, Poterucha JJ, Talwalkar JA. Drug-induced liver injury. Mayo Clin Proc 2014;89:95-106.

17. Lewis JH. The Art and Science of Diagnosing and Managing Drug-induced Liver Injury in 2015 and Beyond. Clin Gastroenterol Hepatol 2015;13:2173-2189 e2178.

18. Garcia Rodriguez LA, Stricker BH, Zimmerman HJ. Risk of acute liver injury associated with the combination of amoxicillin and clavulanic acid. Arch Intern Med 1996;156:1327-1332.

19. Wai CT. Presentation of drug-induced liver injury in Singapore. Singapore Med J 2006;47:116-120.

20. Navarro VJ, Barnhart H, Bonkovsky HL, Davern T, Fontana RJ, Grant L, Reddy KR, et al. Liver injury from herbals and dietary supplements in the U.S. Drug-Induced Liver Injury Network. Hepatology 2014;60:1399-1408.

21. Teschke R, Schwarzenboeck A, Eickhoff A, Frenzel C, Wolff A, Schulze J. Clinical and causality assessment in herbal hepatotoxicity. Expert Opin Drug Saf 2013;12:339-366.

22. Danan G, Benichou C. Causality assessment of adverse reactions to drugs--I. A novel method based on the conclusions of international consensus meetings: application to drug-induced liver injuries. J Clin Epidemiol 1993;46:1323-1330.

23. Aithal GP, Watkins PB, Andrade RJ, Larrey D, Molokhia M, Takikawa H, Hunt CM, et al. Case definition and phenotype standardization in drug-induced liver injury. Clin Pharmacol Ther 2011;89:806-815.

24. Fontana RJ, Seeff LB, Andrade RJ, Bjornsson E, Day CP, Serrano J, Hoofnagle JH. Standardization of nomenclature and causality assessment in drug-induced liver injury: summary of a clinical research workshop. Hepatology 2010;52:730-742.

25. Bjornsson E, Olsson R. Outcome and prognostic markers in severe drug-induced liver disease. Hepatology 2005;42:481-489.

26. Andrade RJ, Lucena MI, Fernandez MC, Pelaez G, Pachkoria K, Garcia-Ruiz E, Garcia-Munoz B, et al. Drug-induced liver injury: an analysis of 461 incidences submitted to the Spanish registry over a 10-year period. Gastroenterology 2005;129:512-521.

27. Andrade RJ, Lucena MI, Kaplowitz N, Garcia-Munoz B, Borraz Y, Pachkoria K, Garcia-Cortes M, et al. Outcome of acute idiosyncratic drug-induced liver injury: Long-term follow-up in a hepatotoxicity registry. Hepatology 2006;44:1581-1588.

28. Bjornsson E, Davidsdottir L. The long-term follow-up after idiosyncratic drug-induced liver injury with jaundice. J Hepatol 2009;50:511-517.

29. Fontana RJ, Hayashi PH, Gu J, Reddy KR, Barnhart H, Watkins PB, Serrano J, et al. Idiosyncratic drug-induced liver injury is associated with substantial morbidity and mortality within 6 months from onset. Gastroenterology 2014;147:96-108 e104.

30. Bjornsson E. Review article: drug-induced liver injury in clinical practice. Aliment Pharmacol Ther 2010;32:3-13.

31. Kleiner DE, Chalasani NP, Lee WM, Fontana RJ, Bonkovsky HL, Watkins PB, Hayashi PH, et al. Hepatic histological findings in suspected drug-induced liver injury: systematic evaluation and clinical associations. Hepatology 2014;59:661-670.

32. Senior JR. How can 'Hy's law' help the clinician? Pharmacoepidemiol Drug Saf 2006;15:235-239.

33. Lo Re V, 3rd, Haynes K, Forde KA, Goldberg DS, Lewis JD, Carbonari DM, Leidl KB, et al. Risk of Acute Liver Failure in Patients With Drug-Induced Liver Injury: Evaluation of Hy's Law and a New Prognostic Model. Clin Gastroenterol Hepatol 2015.

34. Udo R, Maitland-van der Zee AH, Egberts TC, den Breeijen JH, Leufkens HG, van Solinge WW, De Bruin ML. Validity of diagnostic codes and laboratory measurements to identify patients with idiopathic acute liver injury in a hospital database. Pharmacoepidemiol Drug Saf 2015.

35. Kullak-Ublick GA, Merz M, Griffel L, Kaplowitz N, Watkins PB. Liver safety assessment in special populations (hepatitis B, C, and oncology trials). Drug Saf 2014;37 Suppl 1:S57-62.

36. Robles-Diaz M, Lucena MI, Kaplowitz N, Stephens C, Medina-Caliz I, Gonzalez-Jimenez A, Ulzurrun E, 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 e105.

37. Beger RD, Bhattacharyya S, Yang X, Gill PS, Schnackenberg LK, Sun J, James LP. Translational biomarkers of acetaminophen-induced acute liver injury. Arch Toxicol 2015;89:1497-1522.

38.Vliegenthart AD, Antoine DJ, Dear JW. Target biomarker profile for the clinical management of paracetamol overdose. Br J Clin Pharmacol 2015;80:351-362.

39. Antoine DJ, Jenkins RE, Dear JW, Williams DP, McGill MR, Sharpe MR, Craig DG, 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;56:1070-1079.

40. Wang K, Zhang S, Marzolf B, Troisch P, Brightman A, Hu Z, Hood LE, et al. Circulating microRNAs, potential biomarkers for drug-induced liver injury. Proc Natl Acad Sci U S A 2009;106:4402-4407.

41. Starkey Lewis PJ, Dear J, Platt V, Simpson KJ, Craig DG, Antoine DJ, French NS, et al. Circulating microRNAs as potential markers of human drug-induced liver injury. Hepatology 2011;54:1767-1776.

42. Thulin P, Nordahl G, Gry M, Yimer G, Aklillu E, Makonnen E, Aderaye G, 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;34:367-378.

43. Singhal R, Harrill AH, Menguy-Vacheron F, Jayyosi Z, Benzerdjeb H, Watkins PB. Benign elevations in serum aminotransferases and biomarkers of hepatotoxicity in healthy volunteers treated with cholestyramine. BMC Pharmacol Toxicol 2014;15:42.

44. Starkey Lewis PJ, Merz M, Couttet P, Grenet O, Dear J, Antoine DJ, Goldring C, et al. Serum microRNA biomarkers for drug-induced liver injury. Clin Pharmacol Ther 2012;92:291-293.

45. Joka D, Wahl K, Moeller S, Schlue J, Vaske B, Bahr MJ, Manns MP, et al. Prospective biopsy-controlled evaluation of cell death biomarkers for prediction of liver fibrosis and nonalcoholic steatohepatitis. Hepatology 2012;55:455-464.

46. van den Broek MA, Bloemen JG, Dello SA, van de Poll MC, Olde Damink SW, Dejong CH. Randomized controlled trial analyzing the effect of 15 or 30 min intermittent Pringle maneuver on hepatocellular damage during liver surgery. J Hepatol 2011;55:337-345.

47. Harrill AH, Roach J, Fier I, Eaddy JS, Kurtz CL, Antoine DJ, Spencer DM, 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;92:214-220.

48. Schomaker S, Warner R, Bock J, Johnson K, Potter D, Van Winkle J, Aubrecht J. Assessment of emerging biomarkers of liver injury in human subjects. Toxicol Sci 2013;132:276-283.

49. Beckett GJ, Foster GR, Hussey AJ, Oliveira DB, Donovan JW, Prescott LF, Proudfoot AT. Plasma glutathione S-transferase and F protein are more sensitive than alanine aminotransferase as markers of paracetamol (acetaminophen)-induced liver damage. Clin Chem 1989;35:2186-2189.

50. Bailey WJ, Holder D, Patel H, Devlin P, Gonzalez RJ, Hamilton V, Muniappa N, et al. A performance evaluation of three drug-induced liver injury biomarkers in the rat: alpha-glutathione S-transferase, arginase 1, and 4-hydroxyphenyl-pyruvate dioxygenase. Toxicol Sci 2012;130:229-244.

51. Schmidt LE, Dalhoff K. Alpha-fetoprotein is a predictor of outcome in acetaminophen-induced liver injury. Hepatology 2005;41:26-31.

52. Kakisaka K, Kataoka K, Onodera M, Suzuki A, Endo K, Tatemichi Y, Kuroda H, et al. Alpha-fetoprotein: A biomarker for the recruitment of progenitor cells in the liver in patients with acute liver injury or failure. Hepatol Res 2014.

53. Ashamiss F, Wierzbicki Z, Chrzanowska A, Scibior D, Pacholczyk M, Kosieradzki M, Lagiewska B, et al. Clinical significance of arginase after liver transplantation. Ann Transplant 2004;9:58-60.

54. Ramaiah SK, Rittling S. Pathophysiological role of osteopontin in hepatic inflammation, toxicity, and cancer. Toxicol Sci 2008;103:4-13.

55. Andersson U, Lindberg J, Wang S, Balasubramanian R, Marcusson-Stahl M, Hannula M, Zeng C, et al. A systems biology approach to understanding elevated serum alanine transaminase levels in a clinical trial with ximelagatran. Biomarkers 2009;14:572-586.

56. Keskin M, Dolar E, Dirican M, Kiyici M, Yilmaz Y, Gurel S, Nak SG, et al. Baseline and salt-stimulated paraoxonase and arylesterase activities in patients with chronic liver disease: relation to disease severity. Intern Med J 2009;39:243-248.

57. Sato Y, Watanabe H, Kameyama H, Kobayashi T, Yamamoto S, Takeishi T, Hirano K, et al. Serum LECT2 level as a prognostic indicator in acute liver failure. Transplant Proc 2004;36:2359-2361.

58. Morgan RE, Trauner M, van Staden CJ, Lee PH, Ramachandran B, Eschenberg M, Afshari CA, et al. Interference with bile salt export pump function is a susceptibility factor for human liver injury in drug development. Toxicol Sci 2010;118:485-500.

59. Dawson S, Stahl S, Paul N, Barber J, Kenna JG. In vitro inhibition of the bile salt export pump correlates with risk of cholestatic drug-induced liver injury in humans. Drug Metab Dispos 2012;40:130-138.

60. Lee JW, Devanarayan V, Barrett YC, Weiner R, Allinson J, Fountain S, Keller S, et al. Fit-for-purpose method development and validation for successful biomarker measurement. Pharm Res 2006;23:312-328.

61. Lee JW, Hall M. Method validation of protein biomarkers in support of drug development or clinical diagnosis/prognosis. J Chromatogr B Analyt Technol Biomed Life Sci 2009;877:1259-1271.

62. European Medicines Agency, Committee for Medicinal Products for Human Use. Draft guideline on validation of bioanalytical methods. London, UK. 2009.

63. US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research. Guidance for Industry: Bioanalytical Method Validation. Revision 1. Rockville, MD, 2013.

64. Benichou C. Criteria of drug-induced liver disorders. Report of an international consensus meeting. J Hepatol 1990;11:272-276.

65. Fattinger K, Funk C, Pantze M, Weber C, Reichen J, Stieger B, Meier PJ. The endothelin antagonist bosentan inhibits the canalicular bile salt export pump: a potential mechanism for hepatic adverse reactions. Clin Pharmacol Ther 2001;69:223-231.

66. Piton A, Poynard T, Imbert-Bismut F, Khalil L, Delattre J, Pelissier E, Sansonetti N, et al. Factors associated with serum alanine transaminase activity in healthy subjects: consequences for the definition of normal values, for selection of blood donors, and for patients with chronic hepatitis C. MULTIVIRC Group. Hepatology 1998;27:1213-1219.

67. Poynard T, Lebray P, Ingiliz P, Varaut A, Varsat B, Ngo Y, Norha P, et al. Prevalence of liver fibrosis and risk factors in a general population using non-invasive biomarkers (FibroTest). BMC Gastroenterol 2010;10:40.

68. M'Kada H, Perazzo H, Munteanu M, Ngo Y, Ramanujam N, Fautrel B, Imbert-Bismut F, et al. Real time identification of drug-induced liver injury (DILI) through daily screening of ALT results: a prospective pilot cohort study. PLoS One 2012;7:e42418.