Taylor & Francis
Taylor & Francis Group

Expert Opinion on Drug Metabolism & Toxicology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/iemt20

Biomarkers of idiosyncratic drug-induced liver injury (DILI) - a systematic review

Edmond Atallah, Cristiana Freixo, Ismael Alvarez-Alvarez, F.J Cubero, Alexander L. Gerbes, Gerd A. Kullak-Ublick & Guruprasad P. Aithal

To cite this article: Edmond Atallah, Cristiana Freixo, Ismael Alvarez-Alvarez, F.J Cubero, Alexander L. Gerbes, Gerd A. Kullak-Ublick & Guruprasad P. Aithal (2021) Biomarkers of idiosyncratic drug-induced liver injury (DILI) - a systematic review, Expert Opinion on Drug Metabolism & Toxicology, 17:11, 1327-1343, DOI: 10.1080/17425255.2021.1999410

To link to this article: https://doi.org/10.1080/17425255.2021.1999410

EXPERT OPINION ON DRUG METABOLISM & TOXICOLOGY
2021, VOL. 17, NO. 11, 1327–1343
https://doi.org/10.1080/17425255.2021.1999410

Taylor & Francis
Taylor & Francis Group

REVIEW

Biomarkers of idiosyncratic drug-induced liver injury (DILI) - a systematic review

Edmond Atallah [a,b], Cristiana Freixo[c], Ismael Alvarez-Alvarez [d,e], F.J Cubero [e,f], Alexander L. Gerbes[g], Gerd A. Kullak-Ublick [h,i] and Guruprasad P. Aithal [a,b]

[a]School of Medicine, Nottingham Digestive Diseases Centre, University of Nottingham, Nottingham, UK; [b]National Institute for Health Research (NIHR) Nottingham Biomedical Research Centre, Nottingham University Hospitals Nhs Trust and the University of Nottingham, Nottingham, UK; [c]Cintesis, Center for Health Technology and Services Research, Faculdade De Medicina Da Universidade Do Porto, Porto, Portugal; [d]Servicio De Farmacología Clínica and Ugc Aparato Digestivo, Hospital Universitario Virgen De La Victoria, Instituto De Investigación Biomédica De Málaga-IBIMA, Universidad De Málaga, Málaga, Spain; [e]Centro De Investigación Biomédica En Red De Enfermedades Hepáticas Y Digestivas (Ciberehd), Madrid, Spain; [f]Department of Immunology, Ophthalmology, and ENT, Complutense University School of Medicine, Madrid, Spain; [g]Department of Medicine, Liver Centre Munich, University Hospital Munich, Munich, Germany; [h]Department of Clinical Pharmacology and Toxicology, University Hospital Zurich, University of Zurich, Zurich, Switzerland; [i]Mechanistic Safety, CMO & Patient Safety, Global Drug Development, Novartis Pharma, Basel, Switzerland

ABSTRACT

Introduction: Idiosyncratic drug-induced liver injury (DILI) is an unpredictable event, and there are no specific biomarkers that can distinguish DILI from alternative explanations or predict its clinical outcomes.

Areas covered: This systematic review summarizes the available evidence for all biomarkers proposed to have a role in the diagnosis or prognosis of DILI. Following a comprehensive search, we included all types of studies in humans. We included DILI cases based on any threshold criteria but excluded intrinsic DILI, commonly caused by paracetamol overdose. We classified studies into diagnostic and prognostic categories and assessed their methodological quality. After reviewing the literature, 14 studies were eligible.

Expert Opinion: Diagnostic studies were heterogeneous with regard to the study population and outcomes measured. Prognostic models were developed by integrating novel biomarkers, risk scores, and traditional biomarkers, which increased their prognostic ability to predict death or transplantation by 6 months. This systematic review highlights the case of need for non-genetic biomarkers that distinguish DILI from acute liver injury related to alternative etiology. Biomarkers with the potential to identify serious adverse outcomes from acute DILI should be validated in independent prospective cohorts with a substantial number of cases.

ARTICLE HISTORY
Received 20 July 2021
Accepted 21 October 2021

KEYWORDS
Drug-induced liver injury; DILI; hepatotoxicity; biomarkers; systematic review

1. Introduction

Idiosyncratic drug-induced liver injury (DILI) is an unpredictable and serious adverse event with an annual incidence estimated to be 19.1 and 23.8 per 100,000 in Iceland and mainland China, respectively [1,2]. Nonetheless, the incidence of DILI secondary to commonly used medications is significantly higher, reaching 43 per 100,000 users of amoxicillin-clavulanate [1]. Liver biochemistry remains the mainstay of DILI detection followed by causality assessment to identify a temporal relationship between the suspected drug and the liver injury, taking multiple factors into account. They include the time and course of injury in relation to the medication, concomitant drugs, patient's risk factors, exclusion of other etiologies, drug's hepatotoxicity profile, and response to re-administration when applicable. Therefore, an acute elevation of serum liver enzymes (alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), and gamma-glutamyl transferase (GGT)) following drug administration without other clear clinical reasons is suggestive of DILI [3]. Elevated total bilirubin, specifically conjugated bilirubin, in addition to liver enzymes, can also indicate severe liver injury. However, these markers lack specificity and do not distinguish DILI from other liver etiologies, which can cause a similar elevation in liver enzymes. Moreover, an acute rise in ALT and AST, which may be interpreted as hepatocellular injury, can be secondary to muscle or cardiac injury. Additionally, asymptomatic elevation in transaminases associated with medications such as cholestyramine and heparin do not reflect clinically relevant liver injury [4,5], while variable proportion of ALT elevations temporally related to drug exposure resolve spontaneously despite continued medication [6]. Yet, there is no universal gold standard for the diagnosis of DILI [7]. When multiple medications are taken simultaneously, it is frequently difficult to single out the agent that has caused DILI [8]. Therefore, diagnosis of

CONTACT Guruprasad P. Aithal guru.aithal@nottingham.ac.uk Nottingham Digestive Diseases Centre, School of Medicine, University of Nottingham, Nottingham, UK

This article has been republished with minor changes. These changes do not impact the academic content of the article.

Supplemental data for this article can be accessed here.

Article Highlights

- Idiosyncratic drug-induced liver injury (DILI) is an unpredictable and challenging clinical event. There is a critical need for biomarkers to detect it, differentiate DILI from other liver injuries and predict its outcomes to guide clinical management. Multiple novel biomarkers have been studied but failed to reach full qualification and application in clinical practice.
- This is the first systematic review to summarize the available evidence on the diagnostic and prognostic applications of biomarkers in DILI.
- This review demonstrates significant heterogeneity in study designs and lack of validity for novel biomarkers. Predictive models that combine pre-existing risk scores, liver enzymes, and new biomarkers showed high accuracy for predicting death/transplantation at 6 months following DILI.
- This review highlights the need for a concerted effort from a multidisciplinary international collaboration focused on idiosyncratic DILI to design and coordinate the discovery, evaluation, and validation of novel biomarkers that are suitable for use in the clinical phase of drug development as well as clinical practice.

This box summarizes key points contained in the article.

idiosyncratic DILI remains challenging in new drug development and clinical practice, and there is an urgent need for sensitive and specific DILI biomarkers.

Definition of DILI based on liver enzymes elevation lacked standardization until 2011, when the international consensus and the European Association for the Study of the Liver proposed threshold criteria for the definition of DILI, which have been implemented since in clinical practice [3,9]. The pattern of DILI is based on the earliest identified liver chemistry elevations above the upper limit of normality (ULN) that fit DILI criteria and is defined using R ratio, where R = (ALT/ULN)÷(ALP/ULN). There are three patterns of DILI: hepatocellular (R ≥ 5), mixed (R > 2 and < 5), and cholestatic (R ≤ 2).

The degree of elevation of liver enzymes does not accurately reflect the severity of the liver injury or predict clinical outcome. However, elevated liver enzymes and total bilirubin indicate a worse prognosis, which was observed by Hyman Zimmerman several decades ago and became the basis of Hy's law, which is defined by drug-induced liver injury with ALT > 3 times ULN and total bilirubin >2 ULN after excluding other causes [10]. Furthermore, multiple registries demonstrated mortality/liver transplant rates exceed 10% in DILI patients with a hepatocellular injury with jaundice [11,12]. This usually leads to permanent discontinuation of the investigated drug in clinical trials. Since the risk is 10%, most patients who meet Hy's law criteria will spontaneously recover without liver transplantation. Therefore, although this makes Hy's law a useful tool for initial risk assessment and a predictor of a drug's potential to cause severe hepatotoxicity, it lacks the specificity required for a decision-making algorithm [10].

Due to the clinical need for sensitive and specific biomarkers for DILI, multiple new biomarkers have been studied in the last few decades. Most studies were mainly in the context of paracetamol-induced DILI, which may differ from idiosyncratic DILI due to different pathogenesis of liver injury and early clinical presentation. The development of biomarkers has received regulatory support from the Food and Drug Adminstration (FDA) [13,14]. Despite the efforts put by the Predictive Safety Testing Consortium in the USA and the former Safer and Faster Evidence-based Translation Biomarker Consortium in Europe in the last few decades, novel biomarkers have failed to reach full qualification and application in clinical practice [15,16]. Due to the increasing interest and importance in clinical practice, a systematic review is warranted to summarize the available evidence of biomarkers for idiosyncratic DILI in humans.

2. Methods

This systematic review was structured in accordance with the PRISMA checklist and Cochrane handbook for systematic reviews of diagnostic test accuracy [17]. It was registered in PROSPERO (registration number: CRD42020168708). We included all types of studies published in English regardless of publication status or whether data were collected prospectively or retrospectively. We included studies that provided information comparing one or more diagnostic or prognostic biomarkers against traditional index tests in patients with idiosyncratic DILI.

2.1. Study design and search strategy

We searched MEDLINE via OvidSP (January 1946 to 10/03/2021) and Embase via OvidSP (January 1947 to 10/03/2021) and restricted our search results to English language and adult population. We designed structured search strategies using controlled search terms appropriate for each database as well as free-text search terms. The search strategy for MEDLINE is shown in Supplementary Material 1. We also searched Scopus, Cochrane Controlled Register of Trials, OpenGrey databases, and clinical trial registers for additional trials (EU Clinical trials register (www.clinicaltrialsregister.eu) and Clinicaltrials.gov (www.clinicaltrials.gov)) within the same temporal framework as above mentioned. We screened the reference lists of all relevant papers to retrieve additional studies and searched for similar articles related to the final included studies. We contacted relevant authors for further details about the studies when required. We did not perform hand-searching, as there is little published evidence of the benefits of hand-searching for reports of diagnostic test accuracy studies [18].

2.2. Inclusion and exclusion criteria

The inclusion criteria were adult population with suspected DILI or hepatotoxicity with raised liver enzymes (ALT, AST, and ALP) based on any threshold criteria. Exclusion criteria were cases with intrinsic (direct) DILI, commonly caused by paracetamol overdose. The index tests studied were all non-genetic biological markers in humans, and we excluded purely mechanistic studies that were done exclusively *in vitro*. There

is no reference standard available for DILI, and clinical diagnosis is usually based on biochemical alteration and causality assessments. European and American DILI registries (Pro-Euro DILI Network and Drug Induced Liver Injury Network (DILIN)) have an established adjudication process with a panel of experts in the field to make the final decision on the diagnosis [19].

2.3. Outcome definitions

The main outcomes in the diagnostic category were performance characteristics of non-genetic biomarkers, alone or in combination, to distinguish DILI from other etiologies that form competing diagnoses. In the prognostic category, the primary outcome was accuracy in predicting clinical endpoints 6 months after DILI onset (recovery, persistent DILI, acute liver failure, liver transplant, and death).

2.4. Study selection

Three review authors (EA, CF, and IAA) independently identified relevant studies. We retrieved studies from references that at least one of the review authors judged as relevant. Two review authors independently assessed the full-text articles. We resolved any differences in study selection by discussion. For data extraction, only data from studies that meet the inclusion criteria were used. We included all types of studies published in English regardless of publication status or whether data were collected prospectively or retrospectively. We included studies that provided information comparing one or more diagnostic or prognostic biomarkers against traditional index tests in patients with idiosyncratic DILI. We considered data from abstracts if they meet the inclusion criteria and contained sufficient relevant data.

2.5. Data extraction and quality assessment

Two review authors (EA and IAA) independently extracted the following data from each included study: first author, year of publication, study design (prospective or retrospective; cross-sectional studies or case-control studies that reported results of diagnostic accuracy of biomarkers in people with suspected DILI); inclusion and exclusion criteria for individual studies; the total number of participants; the number of males and females; mean age of the participants; severity of DILI; participants' risk factors of liver disease; tests carried out before testing biomarkers; biomarkers tested (index tests); reference standard; and true positive, false positive, true negative, and false-negative data with receiver operator characteristic curve. When necessary, we sought further information from the authors of the studies. Disagreements between the review authors were resolved by discussion. We used the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool to evaluate the quality of diagnostic studies and Quality in the Prognostic Studies (QUIPS) tool for prognostic studies [20,21]. We evaluated the two segments of QUADAS-2 evaluation separately, i.e. the risk of bias and applicability. We adjusted signaling questions according to our review. Each of the included studies was independently assessed by two review authors (EA and IAA).

2.6. Statistical analysis and data synthesis

We summarized the findings from the included studies in a narrative synthesis and classified them based on the type of DILI biomarkers (diagnostic and prognostic), following FDA definitions [22]. We aimed to perform a quantitative analysis to evaluate the diagnostic and prognostic accuracy of the biomarkers, but it was not adequate due to heterogeneity of the study population and outcomes measured. We summarized the performance characteristics of all the diagnostic and prognostic biomarkers and calculated unreported sensitivity, specificity, and the area under the receiver operating characteristic curve (AUROC) with 95% confidence intervals (CI) when possible.

3. Results

3.1. Study characteristics

Out of 5,809 records (following removal of duplicated), we excluded 5,596 irrelevant papers. We retrieved and reviewed a total of 213 full-text reports to assess their eligibility for inclusion in the review as illustrated in the flow diagram. The identified studies that investigated urine-based biomarkers were done in animal models and in the context of intrinsic DILI; therefore, they were excluded. We excluded 34 studies in humans that did not meet our eligibility criteria. The most common reasons for exclusion were intrinsic DILI secondary to paracetamol overdose and conference abstracts from included papers with insufficient data. We also identified seven registered trials on www.clinicaltrials.gov with unpublished results and two trials that are currently ongoing (NCT04269486 and NCT02353455). Figure 1 presents a schematic overview of the study selection process.

We finally included 14 studies in the review and classified them into two main categories: diagnostic and prognostic. All studies investigated blood-based biomarkers; we provided a summary of all included studies in Table 1.

4. Findings

We have sub-grouped the results into diagnostic and prognostic domains. Out of 14 included full texts, we classified eight studies as diagnostic, four studies as prognostic, and two studies as both diagnostic and prognostic.

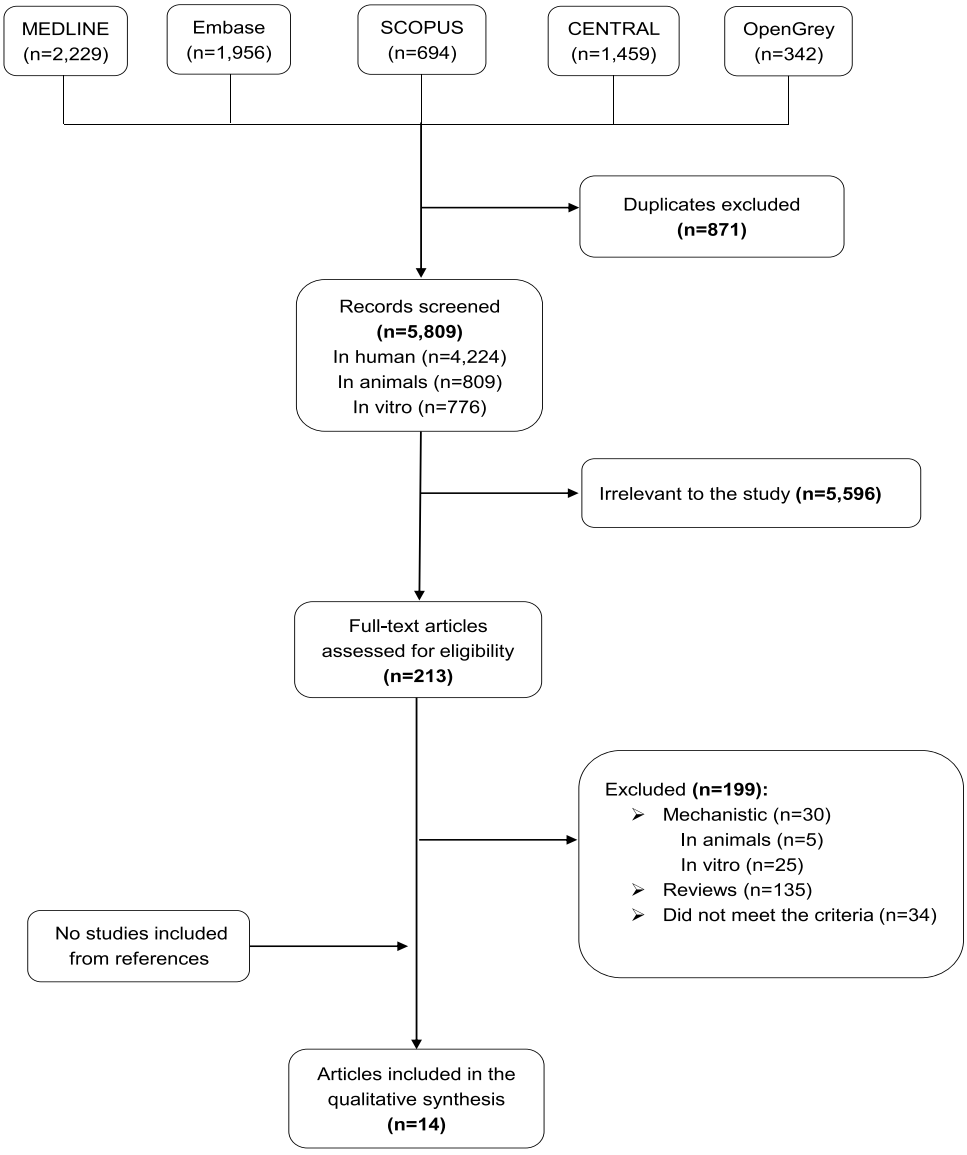

Figure 1. Study selection flow diagram.

4.1. Diagnostic studies

4.1.1. Study design

Ten studies investigated the diagnostic ability of biomarkers in DILI. All studies were case control in design, with healthy controls being the main comparator to DILI. Five studies compared DILI with patients receiving drugs or herbal medicine known to cause DILI without developing liver injury [23–27]. Three studies included patients with liver injury from other etiologies. Dragoi *et al*. had patients with acute liver injury unrelated to drug [23], whereas Soga *et al*. had patients with chronic liver diseases and Huang *et al*. included patients with autoimmune hepatitis and hepatitis-B virus infection during an acute flare [28,29].

The DILI population varied across the diagnostic studies. Five studies investigated biomarkers for liver injury secondary to specific drugs; one small study investigated dicolofenac-induced liver injury [23], two prospective studies focused on HIV/tuberculosis medications [24,26] and two Chinese studies investigated liver injury secondary to the herbal medicine, *Polygonum multiflorum* Thunb (PM) [25,29]. In contrast, three studies used samples from DILI patients recruited as part of the DILIN prospective multicenter study in the USA. Bell *et al*. and Steuerwald *et al*. included samples from the same DILIN cohort [30,31], whereas Church *et al*. combined DILIN samples with two other cohorts from SAFE-T network and had the largest number of DILI samples overall [27].

The threshold of liver enzymes used to define acute liver injury varied across studies as per Table 1. One study, Ma *et al*., defined DILI as ALT > 40 U/L and total bilirubin (TBL) >20 µmol/L [32], whereas Thulin *et al*. defined it as ALT > 3 times ULN similar to SAFE-T cohorts [24]. Rupprechter *et al*. described DILI as ALT > 3 times ULN with symptoms or > 5

Table 1. Summary of included diagnostic and prognostic studies.

Study ID (type)	Participants (N)	Controls (N)	Index tests	Reference test	Outcomes	Derivation of DILI threshold & validation
Bell et al. [30] (Diagnostic)	DILIN cohort (n = 74)	Healthy (n = 40)	Mass-spectrometry-based proteomics for serum protein expression	RUCAM and adjudication	• Comparison of DILI and control • Comparison of DILI patients based on pattern, severity, and causality assessment	• The DILIN threshold criteria for DILI • No validation cohort
Dragoi et al. [23] (Diagnostic)	Suspected DILI (n = 16)	• Healthy (n = 6) • Patients with DILI due to another drug (n = 6) • Patients with acute liver injury not due to DILI (n = 6)	Proteomic analysis of monocyte-derived hepatocyte-like (MH) cells	RUCAM and adjudication	Identifying specific biomarker for Diclofenac-DILI	• Undefined threshold • No validation cohort
Ma et al. [32] (Diagnostic)	DILI (n = 38)	Healthy (n = 30)	• Serum metabolomic profiling • Bile acids	RUCAM	Comparison of DILI and control	• ALT > 40 U/L and total bilirubin TBIL > 20 mmol/L • No validation cohort

Soga et al. [28] (Diagnostic)	DILI (n = 10)	• Healthy (n = 53)
		• Patients with other liver diseases (n = 99)
			Serum metabolomic profiling	Unclear	• Comparison of patients with different liver diseases and control
					• Developing and validating MLR model from metabolites to discriminate different liver diseases
						• Undefined threshold
						• Validation cohort:
						  • DILI (n = 17)
						  • Healthy controls (n = 4)
						  • Patients with other liver diseases (n = 54)

Zhang et al. [25] (Diagnostic)	Female patients received Polygonum multiflorum Thunb. (PM) with ALT elevation > 2ULN (n = 6)	Matched female patients received PM without ALT elevation (n = 30)	Serum metabolomic profiling prior exposure to PM	ALT (serum)	Identifying baseline metabolomic profile to predict PM-DILI	• Susceptible PM-DILI patients defined as ALT > ULN after PM exposure • No validation cohort
Huang et al. [29] (Diagnostic)	PM-induced liver injury (PM-DILI) (n = 13)	• Healthy (n = 9) • Acute auto-immune hepatitis flare (n = 12) • Acute HBV flare (n = 24)	Plasma metabolomic profiling	Unclear	• Identifying metabolomic profile associated with PM-DILI • Developing decision tree for PM-DILI diagnosis • Comparing the metabolomic profile of PM-DILI with other liver injuries (AIH and HBV)	• Undefined threshold • No validation cohort
Thulin et al. [24] (Diagnostic)	TB/HIV patients received treatment with ALT > 3ULN (n = 38)	TB/HIV patients received treatment without ALT > 3ULN (n = 38)	• K18, ccK18 (plasma) • miR-122 (plasma) • GLDH (plasma) • AFP (plasma)	ALT (plasma)	• Comparison of sensitivity/specificity of novel biomarkers against ALT	• ALT values exceeding three times baseline levels (elevation >30 U/L) at any time point during the study period. • No validation cohort

(*Continued*)

Table 1. (Continued).

Study ID (type)	Participants (N)	Controls (N)	Index tests	Reference test	Outcomes	Derivation of DILI threshold & validation
Rupprechter et al. [26] (Diagnostic)	TB/HIV patients received treatment with ALT > ULN (50 U/L) (total n = 15): • Active tuberculosis (n = 2) • Latent TB (n = 2) • Nontuberculous mycobacteria infection (n = 1) • HIV-TB coinfection (n = 10)	• Healthy volunteers (n = 28) • Active tuberculosis (n = 42) • Latent TB (n = 140) • Nontuberculous mycobacteria infection (n = 38) • HIV-TB coinfection (n = 231)	• miR-122 (serum) • K18 (serum)	• ALT (serum) • RUCAM	• Difference of miR-122 and K18 levels in healthy individuals compared to patients with infection • Diagnostic accuracy of miR-122 and K18 in detecting ALT elevation in TB treated patients	• ALT > 3ULN in the presence of symptoms or > 5ULN in the absence of symptoms. • No validation cohort
Church et al. [27] (Diagnostic & Prognostic)	• 2 SAFE-T cohorts: • Swiss DILI (n = 28) • Protocol 3A DILI (n = 98) • DILIN cohort (n = 141)	• Healthy • PSTC (n = 81) • SAFE-T (n = 192) • Treated with known DILI drugs without developing injury • Protocol 4 cohort (n = 55) • Protocol 5 cohort (n = 92)	• AFP (serum) • ARG1 (serum) • CDH5 (serum) • FABP1 (serum) • GSTa (serum) • K18, ccK18 (serum) • MCSFR (plasma) • OPN (serum) • GLDH (serum) • LECT2 (plasma) • PON 1 (plasma) • SDH (serum) • miR-122 (serum)	RUCAM and adjudication	• Diagnostic accuracy for DILI detection • Prognostic accuracy for unresolved DILI in 6 months • Prognostic accuracy for Death/transplantation • Developing a prognostic model with MELD score for death/transplantation	• SAFE-T cohorts: ALT > 3ULN or ALP > 2ULN, within the last four weeks before the baseline visit • DILIN cohort: the DILIN threshold criteria • No validation cohort
Steuerwald et al. [31] (Diagnostic & Prognostic)	DILIN cohort (n = 78)	Healthy (n = 40)	• Serum cytokines, chemokine and growth factors	RUCAM and adjudication	• Comparison of DILI and control • Developing a prognostic model to predict death within 6 months	• The DILIN threshold criteria • No validation cohort
Bonkovsky et al. [35] (Prognostic)	• DILIN cohort (n = 127) • Acute Liver Failure Study Group (n = 118)	Healthy (n = 63)	• Serum cytokines, chemokines, and growth factors	RUCAM and adjudication	• Comparison of DILI, acute liver failure due to different etiologies (DILI, auto-immune hepatitis, viral hepatitis, acetaminophen overdose) and control • Developing a prognostic model to predict death within 6 months	• The DILIN threshold criteria • No validation cohort
Russo et al. [36] (Prognostic)	DILIN cohort (n = 78)	Healthy (n = 40)	Serum miRNA	Unclear	• Comparison of DILI and control • miRNA levels and death within 6 months of DILI onset • Developing a prognostic model with albumin to predict death within 6 months	• The DILIN threshold criteria • No validation cohort

(*Continued*)

Table 1. (Continued).

Study ID (type)	Participants (N)	Controls (N)	Index tests	Reference test	Outcomes	Derivation of DILI threshold & validation
Xie et al. [37] (Prognostic)	DILI (n = 56)	Healthy (n = 34)	Serum metabolites, cytokines, and K18/ccK18	RUCAM	• Comparison of severe/non-severe DILI and control • Developing model from cytokines/metabolites to predict DILI severity	• The DILIN threshold criteria • No validation cohort
Peta et al. [34] (Prognostic)	SAFE-T cohort (n = 154 including 29 APAP)	Acute liver injury (n = 22)	Serum apolipoprotein-A1 (Apo-A1), haptoglobin (HAPTO), alpha-2 macroglobulin (A2M) and GGT	RUCAM and adjudication	• Prediction of recovery at 12 weeks (ALT < 2ULN, and TBIL < 2ULN. • Specific drug signature • Risk of fibrosis in DILI	• SAFE-T threshold: • ALT > 3ULN or ALP > 2ULN, within 4 weeks before the inclusion visit (D0) • An increase of at least 2-fold the pre-treatment level to D0 was required when pre-treatment ALT or ALP activity was available and > ULN • No validation cohort

Abbreviations: Multiple logistic regression models (MLR); *Polygonum multiflorum* Thunb (PM); total cytokeratin 18 (K18); caspase cleaved cytokeratin 18 (ccK18); microRNA-122 (miR-122); glutamate dehydrogenase (GLDH); alpha fetoprotein (AFP); arginase 1 (ARG1); cadherin 5 (CDH5); liver fatty acid binding protein (L-FABP); glutathione-S-transferase α (GSTα); macrophage colony stimulating factor receptor (M-CSF-R); osteopontin (OPN); leukocyte cell-derived chemotaxin 2 (LECT2); paraoxonase 1 (PON1, normalized to prothrombin protein); sorbitol dehydrogenase (SDH); apolipoprotein-A1 (ApoA1); haptoglobin (HAPTO); alpha-2 macroglobulin (A2M); gamma-glutamyl transpeptidase (GGT).

times ULN without symptoms; however, due to a low number of DILI cases (2 cases), the comparison was performed with a lower threshold of ALT elevation (50 U/L, ULN) [26]. Zhang *et al*. defined patients as susceptible to DILI following an elevation of ALT > 2 ULN. DILIN cohort had the highest threshold of ALT (>5 times ULN) [25].

4.1.2. Performance characteristics of diagnostic biomarkers

Studies were heterogeneous, assessed different biomarkers in different populations, thus precluded the combination of the results into a meta-analysis. Alternatively, we provided a narrative summary of their main findings and grouped them by the type of diagnostic biomarkers explored. The area under the receiver operating curve (AUROC) was commonly used to assess the performance characteristics of biomarkers. We summarized the diagnostic accuracies of biomarkers studied in Table 2. We have divided the studies according to the type of index tests studied into four subgroups: proteomics, metabolomics, immune-analytes, and candidate biomarkers.

4.1.2.1. Proteomics. Dragoi *et al*. concentrated on diclofenac-induced DILI and studied the proteome of monocyte-derived hepatocyte-like (MH) cells, then validated the findings in the whole blood to identify potential individual susceptibility to diclofenac-induced liver injury [23]. Briefly, monocytes were isolated from patients' blood samples and cultivated under serum-free conditions for 10 days, generating cells with some hepatocyte features, such as cytochrome P450 activities. These cells, MH cells, were incubated for 48hours in 96-well plates using $1xC_{max}$ and $10xC_{max}$ of the implicated drugs the respective patient had consumed. Then, toxicity was measured with a standardized algorithm based on the release of lactate dehydrogenase in the supernatant and cell lysate [33]. The study revealed that the cell adhesion molecule ITGB3 was four-fold up-regulated in the MH cells from diclofenac-induced liver injury patients and reduced in the whole blood compared to healthy subjects, DILI due to other drugs and patients with other acute liver injuries. Also, ITGB3 correlated inversely with liver biochemistry and clinical outcomes, raising the possibility of its role as a diagnostic and potentially prognostic biomarker for diclofenac-induced liver injury.

Bell *et al*. investigated serum protein expression patterns in patients with idiosyncratic DILI using a mass spectrometry-based quantitative proteomic approach [30]. Priority proteins were classified according to the quality of peptide identification with priority 1 proteins having the greatest likelihood of correct identification. The association between phenotype of DILI, DILI severity, and its role in causality assessment was determined. The diagnostic accuracy of priority 1 proteins and clinical characteristics was explored by linear discriminant analysis and assessment of AUROC. Apolipoprotein E had the greatest power to differentiate DILI patients from healthy controls (AUROC = 0.97; 89% correctly classified as DILI). Furthermore, consideration of expression of several additional

Table 2. Summary of performance characteristics of diagnostic biomarkers.

Study ID	Biomarker/model	Reported measure	Traditional biomarker	Reported measure
Bell et al [30].	Priority 1 proteins	AUROC, Percentage correctly classified as DILI (accuracy)		AUROC, Percentage correctly classified as DILI (accuracy)
	Apolipoprotein E	0.97, 89%	ALT	0.99, 73%
	+ inter-alpha-trypsin inhibitor (heavy chain H3, isoform 1)	0.98, 91%	AST	0.99, 67%
	+ gelsolin	0.99, 92%	ALP	0.96, 68%
	+ complement C7	0.99, 93%	TBL	0.94, 77%
	+serum amyloid P	0.99, 95%	ALT+AST+ ALP+ TBL	0.99, 81%
	+ age	0.99, 96%		
Ma et al. [32]	Bile acids	AUROC	Traditional biomarkers	AUROC
	GCA	0.978	ALT	0.97
	TCA	0.985	AST	0.97
	TUDCA	0.909	GGT	0.97
	GCDCA	0.954	ALP	0.85
	GCDCS	0.946	TBL	0.91
	TDCA	0.976		
	DCA	0.77		
	LCA	0.66		
	CDCA	0.67		
Soga et al [28].	C-glutamyl dipeptides	AUROC (95% CI)		
	MLR using (ALT, γ-Glu-Citrulline) in the training cohort	0.817 (0.639–0.995)		
	MLR using (ALT, γ-Glu-Citrulline) in the validation cohort	0.849 (0.763–0.934)		
Zhang et al. [25]	Serum metabolites	AUROC (95% CI)		
	PE 22:6	0.939 (0.822 −1.0)		
	Crotonyl-CoA	0.933 (0.764 −1.0)		
	Indole-5,6-quinone	0.917 (0.789– 0.989)		
	2E-tetradecenoyl-CoA	0.911 (0.789 −1.0)		
	Phenyllactic acid	0.906 (0.767 −0.983)		
	Phosphoribosyl-ATP	0.900 (0.767 −0.978)		
Huang et al. [29]	Decision tree classification model	AUROC, accuracy		
	P-cresol sulfate/phenylalanine ratio followed by inosine/bilirubin ratio	0.931, 89.8%		
Rupprechter et al. [26]	Candidate biomarkers	AUROC (95% CI)		
	miR-122	0.93 (0.88–0.98)		
	K18	0.80 (0.72–0.87)		
Church et al [27].	Candidate biomarkers	AUROC (95% CI)	Traditional biomarkers	AUROC (95% CI)
	K18	0.947 (0.928–0.966)	ALT	0.99 (0.984–0.996)
	FABP1	0.916 (0.890–0.941)	AST	0.975 (0.963–0.987)
	ccK18	0.911 (0.887–0.935)	ALP	0.902 (0.873–0.930)
	GLDH	0.907 (0.870–0.945)	TBL	0.857 (0.821–0.892)
	MCSFR	0.854 (0.822–0.887)		
	miR-122	0.831 (0.779–0.883)		
	AFP	0.826 (0.793–0.859)		
	GSTα	0.827 (0.792–0.862)		
	SDH	0.819 (0.763–0.876)		
	OPN	0.758 (0.718–0.799)		
	CDH5	0.658 (0.614–0.701)		
	PON1	0.612 (0.542–0.682)		
	ARG1	0.564 (0.519–0.609)		
	LECT2	0.519 (0.450–0.588)		

Abbreviation: GCA: glycocholic acid, TCA: taurocholic acid, TUDCA: tauroursodeoxycholic acid, GCDCA: glycochenodeoxycholic acid, TDCA: taurodeoxycholic acid, TBL: total bilirubin, DCA: deoxycholic acid, LCA: lithocholic acid, CDCA: chenodeoxycholic acid, total cytokeratin 18 (K18); liver fatty acid-binding protein (L-FABP); caspase cleaved cytokeratin 18 (ccK18); glutamate dehydrogenase (GLDH); macrophage colony-stimulating factor receptor (M-CSF-R); microRNA-122 (miR-122); alpha-fetoprotein (AFP); glutathione-S-transferase α (GSTα); sorbitol dehydrogenase (SDH); osteopontin (OPN); cadherin 5 (CDH5); (PON1, normalized to prothrombin protein); arginase 1 (ARG1); leukocyte cell-derived chemotaxin 2 (LECT2).

proteins (inter-alpha-trypsin inhibitor (heavy chain H3, isoform 1), gelsolin, complement C7, and serum amyloid P) and age increased the AUROC to 0.99 with 96% of DILI cases correctly detected, performing better than ALT (AUROC = 0.99; 73% correctly classified DILI). When severity analysis was performed, the expression of 9 priority 1 proteins involved in acute-phase response, activation of the complement cascade, and peroxisome proliferator-activated receptor (PPAR)-α was significantly different between groups of different DILI severity.

4.1.2.2. Metabolomics. Four included studies explored metabolomics as diagnostic tests in DILI. Ma *et al.* used ultra-high performance liquid chromatography-tandem mass spectrometry (UHPLC-MS/MS) for metabolomic profiling and selected reaction monitoring (SRM) to measure 15 targeted

bile acid metabolites [32]. They identified six bile acid metabolites that significantly differentiated DILI from controls and correlated with DILI severity for glycocholic acid (GCA) (AUROC= 0.978) and for taurocholic acid (TCA) (AUROC= 0.985) as per Table 2. Soga and colleagues took a different approach by analyzing metabolites from patients with nine different types of liver diseases using capillary electrophoresis and liquid chromatography-mass spectrometry [28]. They discovered c-glutamyl dipeptides as potential diagnostic biomarkers in liver injury. Consequently, they developed a multiple logistic regression model that differentiated DILI from other liver pathologies by using ALT and γ-Glu-Citrulline (AUROC= 0.817; 95% CI 0.639–0.995 in a training cohort, and AUROC = 0.849; 95% CI 0.763–0.934 in a validation cohort).

Following the rising number of DILI cases from *Polygonum multiflorum* Thunb. (PM), Zhang *et al*. and Huang *et al*. studied the metabolomic profile of PM-induced liver injury [25,29]. Zhang *et al*. aimed to predict PM-induced liver injury by studying baseline metabolites of female patients who developed liver injury following PM exposure [25]. They identified 25 major differential serum metabolites in patients susceptible to PM-induced liver injury, involving sphingolipid, glycerophospholipid, fatty acid, histidine, and aromatic amino acid metabolism. The diagnostic accuracy of six metabolites, PE 22:6, indole-5,6-quinone, 2E-tetradecenoyl-CoA, crotonoyl-CoA, phenyllactic acid, phosphoribosyl-, ATP, to differentiate between groups was significant with AUROC ≥ 0.9 as shown in Table 2. On the other hand, Huang *et al*. examined the metabolomic characteristics of patients with PM-induced liver injury compared to healthy volunteers and patients with auto-immune hepatitis (AIH) and hepatitis-B virus infection (HBV) [29]. They reported changes in multiple metabolic pathways in PM-induced liver injury group, including metabolisms of essential amino acids (tryptophan, valine, phenylalanine), glycerophospholipid metabolism, primary bile acid biosynthesis, and sphingolipid metabolism. The authors used the ratios of P-cresol sulfate/phenylalanine and inosine/bilirubin in a decision tree analysis to differentiate PM-induced liver injury from AIH and HBV with sensitivity of 92.3% and specificity of 88.9%.

4.1.2.3. Immune analytes. Steuerwald *et al.* explored immune profiles by measuring cytokines, chemokines, and growth factors in serum at DILI onset and 6-month follow-up [31]. They found a significant difference in 19 out of 27 analytes studied with a strong association with jaundice alone regardless of ALT or AST levels. Interestingly, there was no significant association with DILI severity or drug class. In addition, when DILI patients were grouped according to their immune profiles, most patients were fitted with adaptive or innate immune profiles.

4.1.2.4. Candidate biomarkers. Three studies explored the diagnostic abilities of candidate biomarkers. Thulin *et al.* and Rupprechter et al. focussed on DILI in prospectively treated cohorts with TB/HIV [24,26]. Both studies measured miRNA-122 (miR-122) and total keratin18 and correlated changes with ALT activity. However, Rupprechter *et al*. assessed serum biomarkers with only two cases of pre-defined DILI [26], whereas Thulin *et al*. measured plasma biomarkers including caspase-cleaved keratin 18 (ccK18), GLDH, and AFP [24]. All biomarkers in both studies were correlated with ALT except AFP, with miR-122 being the most sensitive biomarker with an 8-fold increase in samples with an elevated ALT > ULN [26]. It was apparent that K18 showed a less significant correlation with ALT with a transient elevation in the first week of injury compared to other biomarkers that showed a persistent elevation for a few weeks following liver injury [24]. Moreover, miR-122 showed a superior accuracy of detecting ALT elevation > ULN compared to K18 (AUC= 0.93 and 0.80, respectively). The specificity of K18, miR-122, and GLDH to the liver compared to muscle was demonstrated by their stable levels in a muscle injury cohort compared to ALT.

In an international collaborative study, Church *et al*. found that, among 14 biomarkers studied in the largest cohort to date, only four biomarkers (K18, ccK18, FABP1, and GLDH) showed high accuracy to detect DILI with AUROC > 0.9 [27]. GLDH demonstrated the strongest correlation with ALT in Church's cohort (GLDH, r = 0.88; miR-122, r = 0.66).

4.2. Prognostic studies

4.2.1. Study design

Six case-control studies evaluated the prognostic ability of novel biomarkers in DILI. One study had patients with acute liver injury from other etiologies as a comparator (n = 22) [34], another study compared acute DILI patients, acute liver failure patients due to several etiologies (idiosyncratic DILI (n = 39), auto-immune hepatitis (n = 38), viral hepatitis (n = 28), and acetaminophen overdose (n = 13)), and healthy controls (n = 63) [35]. In contrast, healthy volunteers were the main control arm in the other studies.

Four studies were conducted in the DILIN cohort [27,31,35,36]. One study, Peta *et al*. included patients from the SAFE-T-DILI project [34], and the remaining study population was recruited from a single study center in China [37]. Five out of six studies defined DILI using the thresholds proposed in DILIN prospective study (ALT or AST > 5 times ULN or ALP > 2 times ULN in the absence of jaundice or coagulopathy, or total bilirubin ≥2.5 mg/dL or INR > 1.5 and elevations of ALT, AST, or ALP) [27,31,35–37]. Peta *et al*. defined DILI using the less stringent SAFE-T criteria (ALT > 3 times ULN or ALP > 2 times ULN) [34].

Prognostic biomarkers varied across studies as well as endpoints. Three studies developed prognostic models to predict death at 6 months of DILI onset [31,36] compared to Church *et al*. who combined death and liver transplantation at 6 months as their endpoint [27]. Xie *et al*. evaluated the DILI severity [37], whereas Peta *et al*. focused on the DILI recovery at 12 weeks [34].

Table 3. Summary of performance characteristics of prognostic biomarkers.

Study ID	Biomarker/model	Outcome endpoint	Sensitivity	Specificity	PPV	NPV	AUROC
Church et al [27].	International Normalized Ratio (INR, traditional biomarker)	Death/transplantation at 6 months	NA	NA	NA	NA	0.920 (0.864–0.977)
	Total bilirubin (traditional biomarker)	Death/transplantation at 6 months	NA	NA	NA	NA	0.821 (0.733–0.909)
	Aspartate aminotransferase (AST, traditional biomarker)	Death/transplantation at 6 months	NA	NA	NA	NA	0.700 (0.587–0.814)
	Alanine aminotransferase (ALT, traditional biomarker)	Death/transplantation at 6 months	NA	NA	NA	NA	0.606 (0.433–0.780)
	Alkaline phosphatase (ALP, traditional biomarker)	Death/transplantation at 6 months	NA	NA	NA	NA	0.597 (0.433–0.760)
	Osteopontin (OPN, candidate biomarker)	Death/transplantation at 6 months	NA	NA	NA	NA	0.858 (0.759–0.957)
	Cytokeratin 18 (K18, candidate biomarker)	Death/transplantation at 6 months	NA	NA	NA	NA	0.832 (0.737–0.927)
	Macrophage colony-stimulating factor receptor (MCSFR, candidate biomarker)	Death/transplantation at 6 months	NA	NA	NA	NA	0.775 (0.654–0.896)
	Caspase cleaved cytokeratin 18 (ccK18, candidate biomarker)	Death/transplantation at 6 months	NA	NA	NA	NA	0.778 (0.676–0.881)
	Fatty acid binding protein-1 (FABP1, candidate biomarker)	Death/transplantation at 6 months	NA	NA	NA	NA	0.721 (0.608–0.833)
	Alpha-fetoprotein (AFP, candidate biomarker)	Death/transplantation at 6 months	NA	NA	NA	NA	0.687 (0.566–0.809)
	Cadherin-5 (CDH-5, candidate biomarker)	Death/transplantation at 6 months	NA	NA	NA	NA	0.623 (0.498–0.748)
	Arginase-1 (ARG1, candidate biomarker)	Death/transplantation at 6 months	NA	NA	NA	NA	0.588 (0.436–0.741)
	Glutathione S-transferase alpha (GST-α, candidate biomarker)	Death/transplantation at 6 months	NA	NA	NA	NA	0.536 (0.359–0.713)
	Apoptotic index (AI, candidate biomarker)	Death/transplantation at 6 months	NA	NA	NA	NA	0.761 (0.627–0.895)
	MELD score value 20–29 + K18 ≥7.98 and MCSFR ≥6.94 MELD score value 20–29 + K18 <7.98 or MCSFR <6.94	Death/transplantation at 6 months	0.933	0.899	0.438 †	0.991 †	0.895 †

(*Continued*)

Table 3. (Continued).

Study ID	Biomarker/model	Outcome endpoint	Sensitivity	Specificity	PPV	NPV	AUROC
Steuerwald et al. [31]	Serum albumin ≤2.8 g/dL/>2.8 g/dL	Death at 6 months	1.0	0.60 (0.48–0.72)	0.26 (0.11–0.40)	1.0	NA
	IL-9, IL-17, PDGF-bb, RANTES All below the median/at least one above the median	Death at 6 months	0.80 (0.55–1.0)	0.94 (0.89–1.0)	0.67 (0.40–0.67)	0.97 (0.93–1.0)	
	All four immune analytes below the median and albumin ≤2.8 g/dL/at least one immune analyte above the median or albumin >2.8 g/dL	Death at 6 months	0.78 (0.51–1.0)	0.99 (0.96–1.0)	0.88 (0.65–1.0)	0.97 (0.93–1.0)	
Bonkovsky et al. [35]	RANTES below the median (11,349 pg/mL) and serum albumin ≤2.8 g/dL‡	Death at 6 months	0.39 (0.30–0.47)	0.91 (0.86–0.96)	0.41 (0.33–0.50)	0.90 (0.85–0.95)	0.65 (0.53–0.77) *
Russo et al. [36]	miRNA-122 <7.89 RFU and serum albumin ≤2.8 g/dL	Death at 6 months	1.0	0.81	0.38	1.0	NA
Xie et al. [37]	Model including 31 metabolites and five cytokines (PDGF-bb, IP-10, IL-1Rα, MIP-1b, and TNF-α)	Severe DILI	NA	NA	NA	NA	0.983
Peta et al [34].	Apolipoprotein A1	Recovery at 12 weeks	0.94 †	0.08 †	0.69 †	0.38 †	0.663 (0.536–0.760)
	Haptoglobin	Recovery at 12 weeks	1.0†	0†	0.69†	NA †	0.619 (0.496–0.718)
	ActiTest components (excluding ALT and total bilirubin, used as reference), i.e. apolipoprotein A1, haptoglobin, GGT, A2M, plus age and gender	Recovery at 12 weeks	0.85†	0.31†	0.73†	0.48†	0.723 (0.610–0.806)

Abbreviations: PPV: positive predictive value; NPV: negative predictive value; AUROC: area under the ROC curve; NA: data not available; MELD: Model for End-Stage Liver Disease; IL-9: interleukin 9; IL-17: interleukin 17; PDGF: Platelet-derived growth factor; RANTES: Regulated upon Activation, Normal T Cell Expressed and Presumably Secreted, also known as chemokine ligand 5 (CCL5); RFU: relative fluorescence units; IP-10: interferon gamma-induced protein 10; IL-1Rα: interleukin 1 receptor alpha; MIP-1b: macrophage inflammatory protein 1 beta; TNF-α: tumor necrosis factor-alpha; GGT: gamma-glutamyl transferase; A2M: alpha-2 macroglobulin. Severe DILI was defined as per DILIN study (Elevations in serum ALT and/or ALP levels, and the total serum bilirubin level is ≥2.5 mg/dL + at least one of the following: hepatic failure (INR ≥1.5, ascites, or encephalopathy) or other organ failure believed to be due to a DILI event (i.e. renal, or pulmonary). Recovery was defined as an ALT <2 ULN and BILI <2 ULN attained between 8 and 12 weeks.

NA: not reported.

†These values were not reported in the original studies and were calculated based on other reported measures.

‡Only including 127 acute DILI cases from the DILIN cohort.

4.2.2. Performance characteristics of prognostic biomarkers

Peta and colleagues [34] assessed the prognostic value of some of the ActiTest components (apolipoprotein-A1 [ApoA1], haptoglobin [HAPTO], alpha-2 macroglobulin [A2M] and GGT) as predictors of recovery outcome at 8 to 12 weeks (defined as <2 times ULN for both ALT and TBL). High levels of ApoA1 and HAPTO were found as predictive biomarkers of recovery (AUROC = 0.663; 95% CI 0.563–0.760, and AUROC = 0.619; 95% CI 0.496–0.718, respectively). Indeed, a model including all four evaluated ActiTest components plus age and sex showed a significant predictive value for recovery, with an AUROC of 0.723 (95% CI 0.610–0.806). Moreover, they tried to evaluate the risk of liver fibrosis following DILI using FibroTest [38] and transient elastography (TE), but the small sample size and short follow-up limited their assessment.

One study, Xie *et al*., aimed to evaluate DILI-related changes in metabolic and immune pathways, using gas chromatography-mass spectrometry and UHPLC-MS/MS techniques to identify biomarkers of DILI severity [37]. A total of 31 metabolites with a different expression between severe and non-severe DILI patients were identified, jointly with five cytokines (PDGF-bb, IP-10, IL-1Rα, MIP-1b, and TNF-α) whose serum levels were significantly lower in severe compared to non-severe patients. Indeed, a model developed to differentiate severe from non-severe DILI cases, including both metabolites and immune mediators, yielded an AUROC of 0.983. In addition, differences in the serum levels of K18 were also studied. Caspase cleaved K18 concentrations (ccK18) were higher in severe DILI patients, though the authors did not find differences in total K18 concentration or ccK18/total K18 ratio between severe and non-severe patients.

Steuerwald *et al*. analyzed 27 serum immune analytes in acute DILI cases from the DILIN cohort, including 14 cytokines, seven chemokines, and six growth factors, to elucidate the profiles associated with worsened prognosis [31]. Lower levels, below the median, of the four immune analytes (IL-9, IL-17, PDGF-bb, and RANTES) were predictive of 6-month mortality with a 92% accuracy (95% CI 86–98). Furthermore, when combining these lower levels of immune analytes and lower levels of albumin (below 2.8 g/dL), the model showed an improved accuracy to 96% (95% CI 92–100).

In a more recent study, Bonkovsky and colleagues [35] aimed to replicate the findings of Steuerwald *et al*. [31] in a different DILIN cohort, and acute liver failure patients enrolled in the Acute Liver Failure Study Group. Multicomparison analyses between acute DILI patients, acute liver failure patients with different etiologies and healthy controls did not reveal unique patterns of expression of immune analytes for a specific etiology. Nonetheless, when the authors analyzed levels of immune analytes on sera samples from the 127 acute DILI patients, the only independent and significant predictor of death at 6 months was the combination of low levels of serum albumin (below 2.8 g/dL) and low levels of RANTES (below the median value of 11,349 pg/mL). This model showed a high specificity (91%; 95% CI 86–96), but low sensitivity (39%; 95% CI 30–47). Also, in subjects enrolled in the DILIN cohort, Russo *et al*. studied the miRNA profile predictive power in death within 6 months of DILI onset [36]. They found in acute DILI cases, compared to control subjects, higher levels of eight miRNAs (miR-122, −1246, −4270, −4433, −4463, −4484, −4532, and pre-miR-4767) and decreased levels of three miRNAs (miR-455-3p, 1281, pre-miR-4274). Among these 11 miRNAs, lower values of three of them (miR-122, −4463, pre- miR-4270) were associated with 6-month mortality. Remarkably, no subjects with higher values (above the median) of miR-122 died within 6 months. Thus, the authors developed a model combining lower levels of albumin (below 2.8 g/dL) and miR-122 (below the median), which showed the highest sensitivity (100%) and reasonable specificity (81%) for 6-month mortality.

In an international collaborative effort, Church and colleagues evaluated the prognostic performance for death/liver transplantation in DILI patients of 5 traditional and 10 candidate prognostic biomarkers [27]. Among traditional biomarkers, they found that increased INR, total bilirubin, and AST levels were strongly associated with death/liver transplantation (AUROCs > 0.7). Among the candidate biomarkers, the higher levels of OPN, K18, MCSFR, ccK18, FABP1, and AFP showed their value as predictors of fatal outcome (Table 3). To improve the available prognostic models, a decision tree model was built combining the MELD score and two identified candidate biomarkers (K18 and MCSFR). This latter model yielded the same sensitivity (0.933) as the MELD score (using a threshold of 20 to 29 points) but showed an improved specificity (0.899).

4.3. Methodological quality of included studies

4.3.1. Diagnostic studies

We appraised the quality of diagnostic studies using QUADAS-2 with adjusted signaling questions to tailor the review. We assessed the quality of the studies in all four main domains when applicable. We included ten diagnostic studies, all of which were case control in methodological design. A summary of the quality assessment across the included studies is shown in Table 4.

The selection of DILI patients or samples was unclear in four diagnostic studies [27–30] compared to other diagnostic studies in which patients were recruited prospectively [23–26,32]. Steuerwald *et al*. specified that DILIN patients used in the study were recruited prospectively over a specific time frame, so we considered it as a low risk of bias [31]. Applicability was judged to be of low concern in all studies except Thulin *et al*. due to apparent baseline imbalances and different treatment regimes [24].

Due to the rarity of DILI and the nature of the diagnosis, which is based on biochemical alteration and causality assessment, there is potentially a risk of bias being aware of the reference standard prior to conducting new index

Table 4. Risk of bias summary for diagnostic studies: review authors' judgments about each domain for each included study.

Study ID	Patient selection	Index test	Reference standard	Flow and timing
Bell 2012 [30]	Unclear	High	Low	High
Dragoi 2018 [23]	Low	High	Low	Unclear
Ma 2019 [32]	Low	High	Low	Unclear
Soga 2011 [28]	Unclear	High	Unclear	Unclear
Zhang 2020 [25]	Low	High	High	Low
Huang 2020 [29]	Unclear	High	Unclear	Unclear
Thulin 2014 [24]	Low	High	High	Low
Rupprechter 2020 [26]	Low	Low	High	Low
Church 2019 [27]	Unclear	Low	Low	Low
Steuerwald 2013 [31]	Low	Low	Low	Unclear

4.3.2. Prognostic studies

We used the quality of prognostic study tool (QUIPS) to appraise the quality of six included studies that investigated prognostic biomarkers or models, as shown in Table 5. The population in Peta *et al*. included only mild cases of DILI with ALT ranging from 244 to 414 U/L and was therefore judged as a moderate risk of bias [34]. The remaining studies included patients with all grades of DILI severities and were at low risk of bias. Russo *et al*. and Steuerwald *et al*. were deemed as moderate risk of attrition bias as they included the same DILIN population with over half of the patients lost from follow-up at 6 months [31,36]. However, Bonkovsky *et al*. was considered as low risk of attrition bias as no subjects from the chosen DILIN cohort lost follow-up [35]. We judged all studies as moderate risk of bias in the prognostic factor measurement domain as prognostic models reported have not been validated and the cut-off values used were chosen following exploratory analysis of biomarkers. We judged all studies as low risk of bias in the outcome measurement domain except Peta *et al*. [34]. The authors used DILI recovery at 12 weeks as an endpoint (defined as ALT < 2 ULN and TBL < 2 ULN), which is not a well-validated clinical endpoint in DILI.

Table 5. Risk of bias summary for prognostic studies: review authors' judgments about each domain for each included study.

Study ID	Study participation	Study attrition	Prognostic factor measurement	Outcome measurement	Study confounding	Statistical analysis and reporting
Church 2019 [27]	Low	Low	Moderate	Low	Moderate	Low
Russo 2017 [36]	Low	Moderate	Moderate	Low	Moderate	Low
Peta 2017 [34]	Moderate	Low	Moderate	Moderate	Low	Low
Xie 2019 [37]	Low	Low	Moderate	Low	Low	Low
Steuerwald 2013 [31]	Low	Moderate	Moderate	Low	Moderate	Low
Bonkovsky 2019 [35]	Moderate	Low	Moderate	Low	Moderate	Moderate

diagnostic tests. Therefore, having a pre-specified threshold for the index test studied before analysis is important to avoid potential overfitting diagnostic accuracy that can limit the external validity of the results. The threshold values were pre-specified in Rupprechter *et al*. [26], Church *et al*. [27], and Steuerwald *et al*. [31], so we classified them as low risk of bias compared to other studies.

In the reference standard domain of risk of bias assessment, we considered undertaking causality assessment in addition to elevation of liver enzymes as a quality measure for identifying DILI cases. Five included studies followed the RUCAM causality assessment score as the reference standard to establish DILI cases and were classified as low risk of bias [23,27,30–32]. In contrast, three prospective studies defined cases based on ALT elevation alone following drug exposure and were therefore judged as high risk of bias [24–26]. Two studies, Soga *et al*. and Huang *et al*., did not specify a reference standard for defining DILI cases [28,29].

We have considered specifying the interval between DILI occurrence and blood sampling as a quality measure, which was not clearly defined in four studies [23,28,29,32]. It was unclear if all patients in Steuerwald *et al*. had causality assessment and adjudication [31]. In Bell *et al*. study, a proportion of patients did not receive the reference standard and was therefore classified as high risk of bias [30].

Four studies included samples from the DILIN cohort and reported mortality at 6 months [27,31,35,36]. However, confounders for death were not clearly specified with only half of deaths being due to liver disease [36]. Therefore, we classified the risk of bias as moderate in the above studies for the study confounding domain.

All included studies presented data sufficiently and built prognostic models based on a conceptual framework with low concern regarding selective reporting of results, except for Bonkovsky *et al*., who did not report the AUROC values though it was planned in the statistical analysis [35].

5. Discussion

In this first systematic review focused on the application of biomarkers for the diagnosis of idiosyncratic DILI and prognostic evaluation of the acute DILI event, we found only 14 studies overall that met pre-defined criteria for inclusion.

Study designs included both system biology approach and candidate biomarker studies to identify putative biomarkers. However, assessments of the role of biomarkers in the diagnosis of DILI were limited by the fact that only three studies included individuals with other liver injuries unrelated to drug exposure as a comparator [23,28,29]; hence, the potential of a biomarker to distinguish DILI from other alternative etiology of acute liver injury was not sufficiently evaluated. Moreover, studies investigating DILI biomarkers in patients with underlying liver diseases such as nonalcoholic fatty liver disease and chronic viral hepatitis are lacking, which highlights another important gap in the field.

One prospective DILI cohort was used as a discovery set for four different studies, each identifying a different combination of biomarkers associated with a particular phenotype or prognosis of DILI [27,31,35,36]. However, data from individual cases were not available for further modeling or meta-analysis. In the largest longitudinal cohort involving 141 DILI patients, of which 15 died or received transplantation, a combination of MELD score, serum K18 and MCSFR was able to identify 14 out of 15 with adverse outcomes [27]. The latter algorithm, as well as significant findings from other cohort studies, requires validation in a further independent cohort.

The few biomarkers that have been assessed in more than one study include K18, GLDH, AFP, and miR-122 [24,26,27,36,37]. Both K18 and ccK18 have been elevated during drug therapy of combination therapy for tuberculosis and HIV [24], and the relationship between DILI and the ratio between the two was inconsistent in different studies [27,37].

The diagnostic value of the AFP was evaluated in two studies [24,27]. In the TB/HIV study, AFP was the only biomarker that did not correlate with ALT, which might be explained by its role as a cell regeneration marker rather than a liver injury. Its prognostic role in intrinsic DILI has been well described, and an increase in AFP level was shown to predict favorable outcome in paracetamol-induced liver injury [39]. In contrast, in an idiosyncratic DILI cohort, raised AFP levels significantly predicted death or liver transplantation at 6 months [27].

miR-122 was the most sensitive biomarker to elevate from baseline following hepatotoxic drugs and reached a sensitivity of 100% for predicting death in 6 months when combined with albumin [24,26,36]. Furthermore, a recent study explored its potential practical use as a point-of-care biomarker by measuring its level in capillary blood [40]. Despite its value, the significant interindividual variability of miR-122 shown by Church *et al*. and Rupprechter *et al*. limits its use as a liver-specific biomarker [26,27]. Therefore, PSTC has recently prioritized GLDH over mir-122 to pursue biomarker qualification [27]. The time between DILI detection and blood sampling might contribute to the variability of biomarkers and lacked standardization across studies. Blood sampling in the DILIN cohort was performed within 2 weeks of the liver injury compared to 4 weeks in the SAFE-T cohorts. Nonetheless, Church *et al.* did not find a significant correlation between the levels of biomarkers and the time between symptom onset and blood sampling [27]. It also generated a controversy whether degradation with time in the samples used by Church *et al*. played a part in this variability [27]. However, there was no correlation between sample age and miR-122 variation in the SAFE-T healthy volunteers [41,42].

Despite choosing a low threshold for case finding of DILI, targeted bile acid metabolites were strongly predictive of early diagnosis of DILI and significantly increased in proportion to the severity of the liver injury. However, these findings have not been validated in a second cohort, a limitation with most of the positive findings described. In contrast, the proteomics approach identified apolipoprotein E as a potential biomarker in DILI which was not differentially expressed in a nonalcoholic fatty liver disease (NAFLD) cohort [30]. Moreover, when Soga and colleagues studied metabolomic profiles in different liver diseases, c-glutamyl dipeptides were increased in all liver injuries compared to healthy controls [28]. This elevation may represent reduced hepatocellular glutathione (GSH) production; however, different types of c-glutamyl dipeptide showed variable elevation across liver pathologies for an unclear reason. In DILI patients, γ-Glu-Citrulline was significantly elevated and showed high diagnostic accuracy when integrated with ALT in a statistical model.

Following the rising number of DILI cases globally from the traditional Chinese medicine *Polygonum multiflorum* Thunb. [43,44], two Chinese studies explored metabolic profiles of PM-DILI at different time points [25,29]. One study investigated potential metabolic risk factors of PM-induced liver injury and found differences in baseline metabolites linked to multiple metabolic pathways suggesting low-grade inflammation and immune dysfunction in individuals susceptible to PM-induced liver injury [25]. This was consistent with previous data, which highlighted that patients with auto-immune diseases were more likely to develop PM-induced liver injury [45]. On the other hand, the other study identified metabolic signals in PM-induced liver injury cases, mainly in amino acids and sphingolipid metabolisms, compared to healthy and liver injuries secondary to auto-immune hepatitis or HBV infection [29].

Primary bile acid biosynthesis and alpha-linolenic acid metabolism pathways have been linked to the severity of DILI. The level of metabolites negatively correlated with pro-inflammatory (PDGF-bb, TNF-α, IP-10, and MIP-1b) and anti-inflammatory cytokines (IL-1Rα) [37]. However, the significant reduction in most of the cytokines' levels in severe DILI patients may represent a state of immune dysfunction or immune paresis that has been observed in patients with acute liver failure [46]. In the DILIN cohort, five cytokines (IL-12, IL-17, PDGF-bb, RANTES, and TNF-α) were lower in patients with acute liver failure; and in a Chinese cohort, levels of PDGF-bb and TNF-α were lower in the severe DILI group [31,37]. Therefore, cytokines may have a future role in predicting patients with a high risk of dying following DILI.

Systematic identification of DILI remains difficult and labor intensive even in hospitalized patients [47]. Despite the development of several diagnostic DILI biomarkers, clinicians still

face a challenge trying to identify the culprit drug in patients who take multiple medications. In vitro methods, using monocyte-derived hepatocyte-like (MH) cells from peripheral blood of DILI patients, have been developed to identify drugs that caused DILI with high specificity [33].

Besides its novelty and relevance in the field, the strengths of this review include the adoption of high methodological standards by performing a comprehensive literature search, detailed scrutiny of included studies, and rigorous independent risk of bias assessment. The review included a small number of studies, the majority of which in turn included a small number of participants. In addition, data from heterogeneous populations and outcomes, were unsuitable for pooling in quantitative analysis. This systematic review demonstrated the potential for a system biology approach in the derivation cohort with serum metabolome and targeted bile acid profiling, revealing the role of bile acid metabolism in DILI pathogenesis [32].

6. Conclusion

Our systematic review emphasizes that there is a clear case of need for research in this area. The low prevalence of DILI may explain the challenges of conducting such studies; therefore, larger prospective studies with collaborative efforts are required to qualify candidate biomarkers and suggest that a coordinated iterative process is needed.

7. Expert opinion

There is an important case of need both during drug-development and clinical practice for new tests that distinguish DILI from alternative etiology for acute liver injury and chronic liver diseases. Biomarkers that distinguish self-resolving elevation of liver enzymes (referred to as adaptation), and therefore recovery, from progression and therefore serious liver injury in DILI will transform monitoring in clinical trials and strengthen regulatory approval of novel molecular entities. These safety biomarkers are crucial to reduce the late attrition of drugs during their pre-clinical development and post-marketing withdrawals as well as for effective monitoring of drug therapy in clinical practice.

Interestingly, innovative research methodologies such as genome-wide association studies (GWAS) have been adopted in DILI research soon after their introduction. This delivered a major breakthrough [48] and triggered the formation of a large global research collaboration leading to identification of genetic markers associated with DILI secondary to over 20 of currently used medications [49,50]. Surprisingly, other system biology approaches and technologies are yet to yield similar success. One of the key reasons for lack of progress is that the study should be designed to identify patients with acute liver injury early in the course of the event, even before diagnosis is confirmed (for which the study should be seamlessly integrated into clinical pathways and enroll sufficient number of cases of uncommon, yet a serious adverse event to match the context of use of the candidate biomarkers. Only multicenter collaboration with harmonized protocols can deliver such a program effectively.

It was striking to note that the most relevant full texts identified in the search were review articles rather than primary studies. Hence, time is ripe for a step change in the field and to focus future academic efforts on primary research. First of all, candidate biomarkers identified through the most robust studies highlighted in our systematic review should be prioritized for validation. Currently, the combination of K18 and MCSFR, when used in conjunction with MELD score, is the only panel that appears to add value in prognostication of DILI event. As technologies mature investigations exploring the utility of new biomarkers such as microRNAs and extracellular vesicles will follow. It is important to recognize that several lines of evidence highlight the role of adaptive immune response as the common distal event in the development of DILI. Moreover, future drug pipeline is enriched by small molecules and biologics targeting the immune system. Therefore, characterization of immune mechanisms underpinning DILI may reveal biomarkers that typify the event. Unique features of circulating and infiltrating cell types may be carrying the hallmarks of DILI.

Author Contributions

E Atallah drafted the protocol, designed the search strategy, screened studies for eligibility, extracted data from included studies, assessed the methodological quality of studies and drafted the final review. C Freixo screened studies for eligibility, extracted data from included studies, and contributed to the text of the final review. I Alvarez-Alvarez screened studies for eligibility, extracted data from included studies, assessed the methodological quality of included studies, and contributed to the text of the final review. FJ Cubero, GA Kullak-Ublick, and AL Gerbes provided expert opinion and contributed the text of the final review. GP Aithal provided expert opinion on the whole review, involved in decision making, reviewed and revised the manuscript.

All authors approved the final version of the article, including the authorship list.

Declaration of Interest

GP Aithal has received consulting fees from Pfizer and GlaxoSmithKline paid to the University of Nottingham. C Frexio works as a Pharmaceutical Medicine Physician for Novo Nordisk A/S. I Alvarez-Alvarez holds a Sara Borrell contract (CD20/00083) funded by ISCIII. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Acknowledgement

We acknowledge the support of the Translational Safety Biomarker Pipeline (TransBioLine) consortium (GPA, GKU and ALG members) for their internal review process of the review. This article/publication is supported by COST Action CA17112 Prospective European Drug-Induced Liver Injury Network, supported by COST (European Cooperation in Science and Technology). COST is a funding agency for research and innovation networks. Our Actions help connect

research initiatives across Europe and enable scientists to grow their ideas by sharing them with their peers. This boosts their research, career and innovation (www.cost.eu).

Funding

This paper was funded by the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 821283 (www.imi.europa.eu). This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA. Translational Safety Biomarker Pipeline (TransBioLine): Enabling development and implementation of novel safety biomarkers in clinical trials and diagnosis of disease' — 'TransBioLine' ('action'). Grant Number: 821283.

Disclaimer

The communication reflects the authors' view and that neither IMI nor the European Union, EFPIA, or any Associated Partners are responsible for any use that maybe made of the information contained therein.

Reviewer Disclosures

Peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

ORCID

Edmond Atallah http://orcid.org/0000-0002-7341-9174
Ismael Alvarez-Alvarez http://orcid.org/0000-0001-6271-0604
F.J Cubero http://orcid.org/0000-0003-1499-650X
Gerd A. Kullak-Ublick http://orcid.org/0000-0002-0757-4408
Guruprasad P. Aithal http://orcid.org/0000-0003-3924-4830

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Björnsson ES, Bergmann OM, Björnsson HK, et al. Incidence, presentation, and outcomes in patients with drug-induced liver injury in the general population of Iceland. Gastroenterology. 2013;144(7):1419–1425.
2. Shen T, Liu Y, Shang J, et al. Incidence and etiology of drug-induced liver injury in Mainland China. Gastroenterology. 2019 Jun;156(8):2230–2241 e11.
3. Aithal GP, Watkins PB, Andrade RJ, et al. Case definition and phenotype standardization in drug-induced liver injury. Clin Pharmacol Ther. 2011 Jun;89(6):806–815.
•• **The recommendations from the DILI expert group to achieve consistency and homogeneity in the definition, characterization, and classification of DILI that are now accepted as standards. They include DILI case definition, threshold criteria, pattern of liver injury, causality assessment, severity assessment, chronicity, and an algorithm for evaluation of potential subjects.**
4. Singhal R, Harrill AH, Menguy-Vacheron F, et al. Benign elevations in serum aminotransferases and biomarkers of hepatotoxicity in healthy volunteers treated with cholestyramine. BMC Pharmacol Toxicol. 2014 Aug 3;15(1):42.
5. Harrill AH, Roach J, Fier I, 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 Aug;92(2):214–220.
6. Jiang F, Yan H, Liang L, et al. Incidence and risk factors of anti-tuberculosis drug induced liver injury (Dili): large cohort study involving 4652 Chinese adult tuberculosis patients. Liver Int. 2021 Jul;41(7):1565–1575.
7. Kullak-Ublick GA, Andrade RJ, Merz M, et al. Drug-induced liver injury: recent advances in diagnosis and risk assessment. Gut. 2017 Jun;66(6):1154–1164.
8. Tillmann HL, Suzuki A, Barnhart HX, et al. Tools for causality assessment in drug-induced liver disease. Curr Opin Gastroenterol. 2019 May;35(3):183–190.
9. European Association for the Study of the Liver. Electronic address eee, Clinical Practice Guideline Panel C, Panel m. EASL clinical practice guidelines: drug-induced liver injury. J Hepatol. 2019 Jun;70(6):1222–1261.
10. Temple R. Hy's law: predicting serious hepatotoxicity. Pharmacoepidemiol Drug Saf. 2006 Apr;15(4):241–243.
11. Andrade RJ, Lucena MI, Fernández MC, et al. Drug-induced liver injury: an analysis of 461 incidences submitted to the Spanish registry over a 10-year period. Gastroenterology. 2005;129(2):512–521.
12. Chalasani N, Bonkovsky HL, Fontana R, et al. Features and outcomes of 899 patients with drug-induced liver injury: the DILIN prospective study. Gastroenterology. 2015;148(7):1340–1352.e7.
13. FDA letter of support for drug-induced liver injury (Dili) biomarker(s). 2016 [cited 2021 09 03]. Available from: https://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/UCM517355.pdf
14. Roth SE, Avigan MI, Bourdet D, et al. Next-generation dili biomarkers: prioritization of biomarkers for qualification and best practices for biospecimen collection in drug development. Clin Pharmacol Ther. 2020 Feb;107(2):333–346.
15. Safer and faster evidence-based translation (SAFE-T). [cited 2021 09 03]. Available from: https://www.imi.europa.eu/projects-results/project-factsheets/safe-t
16. C-Path's predictive safety testing consortium (PSTC). [cited 2021 09 03]. Available from: https://c-path.org/programs/pstc/
17. Page MJ, McKenzie JE, and Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2020 ;372:n71 doi:10.1136/bmj.n71 https://www.bmj.com/content/bmj/372/bmj.n71.full.pdf.
18. Glanville J, Cikalo M, Crawford F, et al. Handsearching did not yield additional unique FDG-PET diagnostic test accuracy studies compared with electronic searches: a preliminary investigation. Res Synth Methods. 2012 Sep;3(3):202–213.
19. Fontana RJ, Watkins PB, Bonkovsky HL, et al. Drug-induced liver injury network (DILIN) prospective study: rationale, design and conduct. Drug Saf. 2009;32(1):55–68.
20. Hayden JA, Cote P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. Ann Intern Med. 2006 Mar 21;144(6):427–437.
21. Whiting PF, Rutjes AW, and Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011 Oct 18;155(8):529–536 doi:10.7326/0003-4819-155-8-201110180-00009.
22. FDA-NIH Biomarker Working Group BEST (Biomarkers, EndpointS, and other Tools) Resource. : Silver Spring (MD): Food and Drug Administration (US); 2016 https://www.ncbi.nlm.nih.gov/books/NBK326791/.
23. Dragoi D, Benesic A, Pichler G, et al. Proteomics analysis of monocyte-derived hepatocyte-like cells identifies integrin beta 3 as a specific biomarker for drug-induced liver injury by diclofenac. Front Pharmacol. 2018;9:699.
24. Thulin P, Nordahl G, Gry M, 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 Mar;34(3):367–378.
25. Zhang L, Niu M, Wei AW, et al. Risk profiling using metabolomic characteristics for susceptible individuals of drug-induced liver injury caused by Polygonum multiflorum. Arch Toxicol. 2020 Jan;94(1):245–256.

26. Rupprechter SAE, Sloan DJ, Oosthuyzen W, et al. MicroRNA-122 and cytokeratin-18 have potential as a biomarkers of drug-induced liver injury in European and African patients on treatment for mycobacterial infection. Br J Clin Pharmacol. 2021 Jan 11;87(8):3206–3217.
27. Church RJ, Kullak-Ublick GA, Aubrecht J, et al. Candidate biomarkers for the diagnosis and prognosis of drug-induced liver injury: an international collaborative effort. Hepatology. 2019;69(2): 760–773.
- **The first international study that characterized the performance characteristics of biomarkers in the diagnosis and prognosis of acute idiosyncratic DILI.**
28. Soga T, Sugimoto M, Honma M, et al. Serum metabolomics reveals gamma-glutamyl dipeptides as biomarkers for discrimination among different forms of liver disease. J Hepatol. 2011 Oct;55 (4):896–905.
29. Huang Y, Zhao X, Zhang ZT, et al. Metabolomics profiling and diagnosis biomarkers searching for drug-induced liver injury implicated to Polygonum multiflorum: a cross-sectional cohort study. Front Med (Lausanne). 2020;7:592434.
30. Bell LN, Vuppalanchi R, Watkins PB, et al. Serum proteomic profiling in patients with drug-induced liver injury. Aliment Pharmacol Ther. 2012:600–612. DOI:10.1111/j.1365-2036.2011.04982.x.
31. Steuerwald NM, Foureau DM, Norton HJ, et al. Profiles of serum cytokines in acute drug-induced liver injury and their prognostic significance. PLoS One. 2013;8(12):e81974.
32. Ma Z, Wang X, Yin P, et al. Serum metabolome and targeted bile acid profiling reveals potential novel biomarkers for drug-induced liver injury. Medicine (Baltimore). 2019 Aug;98(31):e16717.
33. Benesic A, Rotter I, Dragoi D, et al. Development and validation of a test to identify drugs that cause idiosyncratic drug-induced liver injury. Clin Gastroenterol Hepatol. 2018 Sep;16(9):1488–1494 e5.
34. Peta V, Tse C, Perazzo H, et al. Serum apolipoprotein A1 and haptoglobin, in patients with suspected drug-induced liver injury (Dili) as biomarkers of recovery. PLoS One. 2017;12(12):e0189436.
35. Bonkovsky HL, Barnhart HX, Foureau DM, et al. Cytokine profiles in acute liver injury-results from the US drug-induced liver injury network (DILIN) and the acute liver failure study group. PLoS One. 2018;13(10):e0206389.
36. Russo MW, Steuerwald N, Norton HJ, et al. Profiles of miRNAs in serum in severe acute drug induced liver injury and their prognostic significance. Liver Int. 2017 May;37(5):757–764.
37. Xie Z, Chen E, Ouyang X, et al. Metabolomics and cytokine analysis for identification of severe drug-induced liver injury. J Proteome Res. 2019 Jun 7;18(6):2514–2524.
38. Imbert-Bismut F, Messous D, Thibault V, et al. Intra-laboratory analytical variability of biochemical markers of fibrosis (fibrotest) and activity (actitest) and reference ranges in healthy blood donors. Clin Chem Lab Med. 2004 Mar;42(6):323–333.
39. Schmidt LE, Dalhoff K. Alpha-fetoprotein is a predictor of outcome in acetaminophen-induced liver injury. Hepatology. 2005 Jan;41 (1):26–31.
40. Vliegenthart ADB, Berends C, Potter CMJ, et al. MicroRNA-122 can be measured in capillary blood which facilitates point-of-care testing for drug-induced liver injury. Br J Clin Pharmacol. 2017 Sep;83 (9):2027–2033.
41. Dear JW. New biomarkers for drug-induced liver injury. Hepatology. 2018;67(6):2480–2481.
42. Church RJ, Watkins PB. Reply. Hepatology. 2018;67(6):2481–2482.
43. Lei X, Chen J, Ren J, et al. Liver damage associated with *Polygonum multiflorum* Thunb.: a systematic review of case reports and case series. Evid Based Complement Alternat Med. 2015. 2015 01 12;2015:459749.
44. Rao T, Liu YT, Zeng XC, et al. The hepatotoxicity of Polygonum multiflorum: the emerging role of the immune-mediated liver injury. Acta Pharmacol Sin. 2021 Jan;42(1):27–35.
45. Zhu Y, Liu SH, Wang JB, et al. [Clinical analysis of drug-induced liver injury caused by Polygonum multiflorum and its preparations]. Zhongguo Zhong Xi Yi Jie He Za Zhi. 2015 Dec;35(12):1442–1447.
46. Antoniades CG, Berry PA, Wendon JA, et al. The importance of immune dysfunction in determining outcome in acute liver failure. J Hepatol. 2008 Nov;49(5):845–861.
47. M'Kada H, Perazzo H, Munteanu M, 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(8): e42418.
48. Daly AK, Donaldson PT, and Bhatnagar P, et al. HLA-B*5701 genotype is a major determinant of drug-induced liver injury due to flucloxacillin. Nat Genet. 2009 Jul;41(7):816–819.
- **The first genome-wide association (GWA) study that demonstrated that possession of HLA-B* 5701 allele was strongly associated with an 81-fold increased risk of flucloxacillin DILI.**
49. Aithal GP. Pharmacogenetic testing in idiosyncratic drug-induced liver injury: current role in clinical practice. Liver Int. 2015 Jul;35 (7):1801–1808.
50. Kaliyaperumal K, Grove JI, Delahay RM, et al. Pharmacogenomics of drug-induced liver injury (Dili): molecular biology to clinical applications. J Hepatol. 2018 Oct;69(4):948–957.
- **This review explores DILI mechanisms, risk factors including the genetic susceptibility loci identified in GWAS and candidate gene studies and its application in clinical practice.**