

# Assessing Toxicity in Drug Trials in Liver Disease

Morris Sherman, MB, BCh, PhD, FRCP(C)<sup>1</sup>

<sup>1</sup> Research Administration, University Health Network, Toronto, Canada

**Address for correspondence** Morris Sherman, MB, BCh, PhD, FRCP(C), 560 Briar Hill Avenue, Toronto, M5N 1M9, Canada (e-mail: morris.sherman@uhn.on.ca).

Semin Liver Dis

## Abstract

Since the early trials in viral hepatitis, more and more new drugs are being tested for use in various liver diseases. Since drug hepatotoxicity is a major cause of drugs under investigation not making it to market, the assessment of drug-induced liver injury in clinical trials of new drugs is crucial. This review will focus on the systems that are used to assess drug-induced liver injury in clinical trials and will discuss how some of these criteria are inappropriate or inaccurate in this function together with suggestions for improvement.

## Keywords

- ▶ clinical trials
- ▶ drug-induced liver injury
- ▶ adverse event

Large-scale phase III trials in liver disease really started with the use of interferons and nucleoside(tide) analogues for hepatitis B and later hepatitis C. This development was accompanied by a renewed interest in drug toxicity during these trials because drug hepatotoxicity is the most common toxicity that results in drugs under development not making it to market.<sup>1</sup> Furthermore, although in liver disease drug trials hepatotoxicity was always an issue, it became a major concern after the fialuridine incident.<sup>2</sup> Fialuridine was the first nucleoside analogue to be tested against hepatitis B. Unfortunately, it caused mitochondrial injury that led to liver failure. Several participants died as a result.

In nonliver disease therapeutic studies, hepatotoxicity is easy to discern because the transaminases would go up if the drug being tested caused liver injury. However, in drug trials in liver disease when the toxicity of the new agent is unknown rising transaminases may be due to progressive liver disease or to toxicity. Furthermore, in some liver diseases, even with effective treatment the transaminases may not predictably improve and may be intermittently or persistently elevated above baseline, for example, in hepatocellular carcinoma (HCC) and nonalcoholic fatty liver disease (NAFLD). This means that physicians who enter patients into those trials may have difficulty in assessing whether, for example, the development of ascites in a patient in a trial of HCC treatment is due to drug or disease, or whether an alanine aminotransferase (ALT) elevation in a patient on a PD-1 inhibitor or similar agent is due to autoimmune hepatitis or to progression of either liver disease or cancer. All of this means that distinguishing between disease pro-

gression and toxicity of study treatment requires careful assessment.

This review will discuss these issues in light of the standard way in which liver toxicity is reported in drug trials and will offer some suggestions to more accurately assess severity of adverse events (AEs) and to attribute changes in liver diseases status to drug or to disease.

Most drug hepatotoxicity in clinical practice is idiosyncratic and rare in comparison to the frequency of use for each drug.<sup>3</sup> However, so many drugs cause idiosyncratic reactions that this is not an uncommon clinical presentation. Only a few drugs have predictable dose-related toxicity, acetaminophen and alcohol being prime examples. This means that in a drug trial of a few hundred participants rare idiosyncratic hepatotoxicity is unlikely to be identified. Not all idiosyncratic reactions are rare. In HCC therapy, the immune modulators such as nivolumab and pembrolizumab and others acting at cellular targets such as PD1, PDL-1, and CTLA-4 all cause an immune-mediated liver injury with incidence ranging from 0.7 to 16%.<sup>4</sup> This is nonetheless an idiosyncratic reaction and is not dose-related. In contrast, in the studies of new drugs dose-related toxicity is much more frequent. These are less common in clinical practice because agents with dose-related hepatotoxicity are weeded out before coming to market.

In diseases that do not involve the liver, the development of drug hepatotoxicity is easily identified. Among the many blood tests that are performed to look for side effects, the liver enzymes and function tests are always included. Thus, any elevation in transaminases is likely due to drug-induced injury. Some caution in interpretation is warranted because

for some drugs the enzymes become elevated, only to improve with time (e.g., statins and isoniazid). This is not considered to be an important injury and is not a reason for a drug to be withdrawn (as an aside, most of this data was derived in an era before the natural history and frequency of NAFLD was known, and it is possible, perhaps even likely for statins, that these were not true drug injuries, but the effects of NAFLD causing intermittent ALT elevations). Short-term dose-related injuries are usually identified in the drug development phase. However, acute drug trials lasting a few months cannot identify longer term slow injury. Estrogen-induced hepatic adenoma is an example that was not identified in the initial oral contraceptive trials, but was only identified by epidemiological studies. Since hepatic adenoma was an uncommon to rare tumor, once the increased incidence in women was identified and correlated with the temporal introduction of the birth control pill, the association was made.<sup>5</sup>

## Assessment of Drug Toxicity

As hepatologists we are asked to assist in determining whether a patient is suffering from a liver injury due to a drug. We see this often in cancer patients and in intensive care unit patients, both categories of patients frequently on multiple drugs. Several techniques have been developed to help assess whether a particular drug is causing toxicity in a particular patient. An early algorithm was the Roussel UCLAF Causality Assessment method (RUCAM) (►Fig. 1, ►Tables 1 and 2).<sup>6</sup> This was developed in 1993 by expert opinion. Toxicity was defined as either hepatocellular or cholestatic/mixed depending on a calculation based on ALT and alkaline phosphatase (►Fig. 1). This was then combined with other factors in the score (►Table 1) that gave a probability that that an individual drug was the offending agent (►Table 2). The RUCAM method has several drawbacks that limit its accuracy, including poor interobserver variability and is not widely used.<sup>7</sup> The Drug-Induced Liver Injury Network (DILIN) uses a different process that also relies on expert opinion in individual cases as well as some of the criteria enumerated in RUCAM.<sup>4</sup> Both the RUCAM method and the DILIN method require that the pattern of liver injury caused by a particular drug be known from previous literature reports. This is of no help in the situation of a new drug being tested that does not have the experience of thousands of patients in the background to help assess whether liver injury is due to drug or not. In HCC trials, many of the drugs being tested, but not all, have been used in other cancers

beforehand, so the toxicity potential is known (e.g., immunomodulating agents such as nivolumab and pembrolizumab), although the extent to which toxicity might be modified by the underlying liver disease is not known.

In assessing whether an AE is drug related in a clinical trial attribution is important, but so is assessment of severity. In clinical practice, Hy's Law<sup>8</sup> can be used to predict whether a drug reaction has a high risk of a fatal outcome. This requires consideration of three factors: there should be a hepatocellular injury as manifested by ALT >3xULN; a bilirubin >2xULN without significant cholestasis; and no other explanation to cause elevated ALT such as viral hepatitis, alcohol, ischemia or congestion and no other drug known to cause such an injury. In drug trials for liver disease, Hy's Law cannot be applied because of the presence of pre-existing liver disease. Therefore, an alternative standardized method of assessing severity of liver injury was required.

## Standardized Reporting of Drug Toxicity in Clinical Trials

There are several conventions that have been developed to aid reporting of drug-induced AEs in clinical trials. All AEs, whether treatment related or not, have to be reported. The reporting tools used include the Common Terminology Criteria for AEs (CTCAE),<sup>9</sup> the Medical Dictionary for Regulatory Activities (MedDRA),<sup>10</sup> and the Council for International Organizations of Medical Sciences (CIOMS) reports.<sup>11</sup> The MedDRA defines the organ systems that are assessed and the groupings of symptoms and definitions within those organ systems (►Table 3) There are several classification levels going from nonspecific to more and more specific. The bottom and most specific term is LLT (lowest level term), which is how investigators report the AE. The preferred terms (PT) are the standardized terms used to classify AEs. The higher level terms, HLT (high level term), HLTG (high level group term), and SOC (system organ class) (►Table 3) are more used to help classification and retrieve AEs than to help precisely identify the AE. Study management maps the event described by the LLT to the PT (one level up). The PTs are the terms in which the AEs are reported to the Data Monitoring Committees (DMC) and the regulators. The CTCAE defines the severity of these symptoms or signs (►Table 4). The CTCAE terms correspond to the PT of the MedDRA system. The CIOMS reports provide a narrative that describes the clinical history leading up to the AE and its sequelae.

$$R = (\text{ALT}/\text{ULN ALT})/(\text{ALP}/\text{ULN ALP})$$

R> 5= hepatocellular injury (e.g., acetaminophen)

R<2= cholestatic injury (e.g., estrogens)

R>2<5= mixed injury (e.g., most idiosyncratic antibiotic induced liver injury)

Abbreviations: ALT alanine amino transferase; ALP alkaline phosphatase; ULN upper limit of normal

**Fig. 1** Calculation of the R value in the RUCAM system. The R value indicates whether the injury is hepatocellular or cholestatic or mixed. ALT, alanine aminotransferase; ALP, alkaline phosphatase; RUCAM, Roussel UCLAF Causality Assessment method; ULN, upper limit of normal.

**Table 1** RUCAM scoring system (simplified). Points are awarded for each criterion. These may be different if the injury is cholestatic, mixed or hepatocellular

Criteria	Hepatocellular			Cholestatic or mixed		
	Initial exposure	Subsequent exposure	Points	Initial exposure	Subsequent exposure	Points
Time to onset	5–90 d	1–15 d	2	5–90 d	1–90 d	2
	<5–>90 d	>15 d	1	<5–>90 d	>90 d	1
Time to remission	≤15 d	≤15 d	1	≤30 d	≤30 d	1
Decrease in ALT/ALP	≥50% in 8 d		3	≥50% in 180 d		2
	≤50% in 30 d		2	<50% in 180 d		1
	<50% in > 30 d		–2			
Other risk factors	Ethanol		1	Ethanol or pregnancy		1
	Age > 55		1	Age > 55		1
Other drugs	With suggestive timing		–1	With suggestive timing		–1
	Known hepatotoxin with suggestive timing		–2	Known hepatotoxin with suggestive timing		–2
	Drug with other evidence as cause, e.g., re-challenge		–3	Drug with other evidence as cause, e.g., rechallenge		–3
Other liver disease	All other causes ruled out		2	All other causes ruled out		2
	All acute causes ruled out		1	All acute causes ruled out		1
	Some liver disease cannot be ruled out		–2	Some liver disease cannot be ruled out		–2
	Nondrug cause probable		–3	Nondrug cause probably		–3
Known hepatotoxin	In product label		2	In product label		2
	Published reports only		1	Published reports only		1
Re-challenge	Positive		3	Positive		3
	Compatible		1	Compatible		1
	Negative		–2	Negative		–2

Abbreviations: ALP, alkaline phosphatase; ALT, alanine aminotransferase; RUCAM, Roussel UCLAF Causality Assessment Method.

**Table 2** Use of RUCAM scoring system

Score	Probability of drug causing liver injury
≥8	Highly likely
6–7	Probable
3–5	Possible
≤2	Unlikely

Abbreviation: RUCAM, Roussel UCLAF Causality Assessment Method.

### Attributing Cause of and Assessing Liver Injury in Clinical Trials in Liver Disease

The issue to be addressed is how to determine whether a change in liver status is a treatment-related AE and second, to accurately describe its severity. In assessing the cause of the AE initially, the investigator makes an attribution as whether this change was due to disease or to drug. The AE is reported to the sponsors. This is the LLT that may be a MedDRA term or may be the investigator's own terminology. A CIOMS form may be completed if the event is severe enough to warrant a narrative report. The medical monitor at the sponsor company reviews

**Table 3** Frequency of MedDRA terms at the different MedDRA levels (March 2019)

Abbreviation	Number of terms	Uses	Examples
SOC	27	Terms used in classification and in tables of AEs to group organ systems	Gastrointestinal disorders
HLGT	337	Terms used for data retrieval (populated automatically from lower level terms)	Gastrointestinal signs and symptoms
HLT	1,737	Terms used for data retrieval (populated automatically from lower level terms)	Nausea and vomiting symptoms
PT	23,708	Terms translated from LLT by study management	Nausea
LLT	80,262	Terms used by investigators	Feeling queasy

Abbreviations: AE, adverse event; HLGT, high level group term; HLT, high-level term; LLT, lowest level term; MedDRA, Medical Dictionary for Regulatory Activities; PT, preferred terms; SOC, system organ class.

**Table 4** Selected CTCAE criteria regarding possible liver injury that are pertinent in liver disease

CTCAE term	Grade I	Grade II	Grade III	Grade IV	Grade V
Hepatic failure	–	–	Asterixis, mild encephalopathy; drug-induced liver disease, limited self-care ability	Life-threatening; moderate-to-severe encephalopathy, coma	Death
Hepatic hemorrhage	Mild symptoms, intervention not indicated	Moderate symptoms, intervention indicated	Transfusion indicated; invasive intervention indicated, hospitalization	Life-threatening; urgent intervention indicated	Death
Hepatic necrosis	–	–	–	Life-threatening consequences, urgent invasive intervention indicated	Death
Hepatic pain	Mild pain	Moderate pain limiting ADL	Severe pain, limiting ADL	–	–
Portal Hypertension	–	Decreased portal vein flow	Retrograde portal vein flow; associated with varices or ascites	Life-threatening consequences; urgent intervention indicated	Death
Portal vein thrombosis	–	Intervention not indicated	Medical intervention indicated	Life-threatening; intervention indicated	Death
Sinusoidal obstruction syndrome	–	Bilirubin 2–5mg/dL, minor intervention required (e.g., blood products diuretics, oxygen)	Bilirubin > 5 mg/dL. Pharmacological intervention required; reversal of flow on ultrasound	Life-threatening (e.g., ventilatory support, dialysis, plasmapheresis, paracentesis)	Death
Ascites	Asymptomatic, intervention not required	Symptomatic, medical intervention required	Severe symptoms, invasive intervention required	Life-threatening; urgent operative intervention required	Death
Chylous ascites	Asymptomatic, intervention not required	Symptomatic, medical intervention required (e.g., fat restricted diet or paracentesis)	Severe symptoms, operative intervention required	Life-threatening; urgent operative intervention required	Death
Variceal bleeding	–	Self-limited; no intervention required	Transfusion indicated invasive intervention indicated; hospitalization	Life-threatening, urgent intervention required	Death
CTCAE term	Grade I	Grade II	Grade III	Grade IV	Grade V
Alanine aminotransferase	>ULN–3.0 x ULN if baseline was normal. 1.5–3.0 x ULN if baseline abnormal	>3.0 x ULN–5.0 x ULN if baseline was normal. >3.0 x ULN–5.0 x ULN if baseline abnormal	>3.0 x ULN–20.0 x ULN if baseline was normal. >3.0 x ULN–20.0 x ULN if baseline abnormal	> 20 x ULN	
Aspartate amino transferase	>ULN–3.0 x ULN if baseline was normal. 1.5–3.0 x ULN if baseline abnormal	>3.0 x ULN–5.0 x ULN	>3.0 x ULN–20.0 x ULN	> 20 x ULN	
Alkaline phosphatase	>ULN–2.5 x ULN if baseline was normal. 2.0–2.5 x ULN if baseline abnormal	>2.5 x ULN–5.0 x ULN	>5.0 x ULN–20 x ULN	>20 x ULN	
Bilirubin	>ULN–1.5 x ULN	> 1.5 ULN–3.0 ULN	>3.0 x ULN–10 x ULN	> 10x ULN	
INR	1.2–1.5	1.5–2.5	> 2.5	–	

Abbreviations: ALD, activities of daily living; CTCAE, Common Terminology Criteria for AEs; INR, international normalized ratio; ULN, upper limit of normal.

the information. The review may be limited to the medical monitor or may involve other members of the sponsor company, who then either agree with the investigator's assessment or disagree. This conclusion is reported on the CIOMS form that ultimately goes to the regulators. There is a process in place in most pharmaceutical company-sponsored trials (less rigorously applied in investigator-initiated trials) to bring changes in liver disease blood tests or AEs to attention (e.g., called Hepatic Events of Special Interest). The criteria that might call a result to attention in this way in any particular clinical trial may or may not mirror the CTCAE. There may be some adjustment for the presence of liver disease at the start of the trial.

The problem of misattribution is particularly acute in drug trials in HCC, where liver disease progression, drug toxicity, and cancer progression may all manifest in the same

way. Part of the problem may be that the investigators may not be expert hepatologists nor have particular experience in dealing with progressive liver disease. Many centers recruiting for HCC trials are oncology centers. There are many oncology centers with a great deal of experience in dealing with HCC, but this is not always the case. The medical monitor may also not be expert in liver disease. Members of data monitoring committees in HCC trials report numerous instances where the investigator attributed AEs to drug that should have been attributed to progressive HCC. Some of these patients turned out to be in the placebo arm. Part of the confusion arises because of the way that progressive cancer is classified in oncology clinical research. Although the tumor may enlarge, it is not called progression until the tumor mass has increased radiologically by 20% (using RECIST criteria)<sup>12</sup>

or 25% (using WHO criteria).<sup>13</sup> Any lesser enlargement of the tumor is not called progression. However, lack of radiologic progression does not mean that the cancer is not progressing clinically. Progression within the liver may be miliary and not detected radiologically, liver function may deteriorate due to the underlying liver disease and a critical destruction of functioning liver disease by tumor, or possibly by a nonmetastatic effect on liver function. Disease progression of this nature would result in the development of ascites or hepatic encephalopathy without radiological evidence of tumor progression by standard criteria. Thus, it would seem that in the absence of visible tumor progression the AE must be labeled as due to drug, whereas in fact the cause was cancer progression. Similarly, if the tumor enlarges but not enough to be called progression the AE might still be labeled as drug-induced. Improving attribution will require expert knowledge of the underlying liver disease and its progression.

The CTCAE describes the severity of AEs and does not deal with their attribution to drug or disease. Therefore, it should encompass all possible liver-related AEs, whether due to drug or not. Severity of the AE is classified by Grades I–V, with V being death. The CTCAE version 5 has defined values of liver blood tests and liver-related symptoms and signs that qualify as AEs. The liver-related CTCAE terms and the description of severity applied to each grade are given in ▶Table 4. However, some of the so-called AEs may simply reflect disease variability and are not necessarily an indication of either progression of disease or of toxicity. For example, CTCAE calls an increase in ALT of 1.5 times over an elevated baseline as an AE (▶Table 3). Thus, if the baseline ALT is 50 U/L and ALT of 75 U/L is considered a Grade I AE. This degree of variation is common in NAFLD or HCC. A Grade II AE would be if the ALT rose to 150 U/L, also not an unheard-of fluctuation of ALT in NAFLD. Similarly, a gamma-glutamyl transpeptidase (GGT) of 2.5–5× normal is a Grade II AE. Yet GGT can vary widely with alcohol or, particularly in HCC or biliary disease, with no apparent trigger. These are still classified as AEs regardless of cause.

Another problem is with the CTCAE classification itself. In ▶Table 4, the description of the severity of the AE that corresponds to the severity grade in some cases is not consistent with current understanding of liver disease and its progression and how toxicity might affect this. Furthermore, criteria that would differentiate one grade from another are not clearly specified and it is left to the investigator to determine severity, whereas better defined criteria would improve interobserver variability. Inspection of the criteria listed in ▶Table 4 highlights several problem areas. Hepatic encephalopathy is not listed as a separate CTCAE term that is distinct from other forms of encephalopathy. It is listed as a feature of more severe grades of Budd–Chiari syndrome and of hepatic failure, but not as a standalone term. Therefore, there is no way of grading severity according to current understanding of hepatic encephalopathy. According to CTCAE liver failure is apparently never Grade I or II severity. Only a grading of III or IV is possible. There is no distinction made between acute liver failure (as might be seen with massive necrosis) and chronic liver failure (progressive

HCC or cirrhosis). Changes in INR and bilirubin are listed separately from liver failure. How such changes might correlate with the listed severity grades of liver failure is not clear. Since ascites and encephalopathy are mentioned elsewhere presumably “hepatic failure” refers to acute liver failure, but this is not clear. The development of hepatic encephalopathy, bleeding varices, and ascites are also manifestations of liver failure, but again there appears to be no way to link these events to the severity of “liver failure.” Bacterial peritonitis is not listed as a possible AE. Drug-induced liver injury is listed as a criterion for Grade III liver failure. Since drug-induced liver injury is not necessarily associated with liver failure, this can easily lead to misclassification. It also implies that any Grade III liver failure is due to drug toxicity, which is not the case. Hepatic hemorrhage may be mild, moderate, or severe, but no criteria are provided to help make this assessment. This ensures interobserver variability that may be major.

Hepatic necrosis is one of the CTCAE terms. There is only one grade here, Grade IV. Presumably, this refers to events such as might be seen with acetaminophen injury or acute ischemia. This would also be reflected in the ALT and AST criteria, so it's not clear that this term is necessary. The term “portal hypertension” is also poorly described. A Grade II AE is described as decreased portal vein flow, but it is not specified how this should be determined or how much decreased. If the ultrasound shows reversed flow, perhaps this should be a Grade III AE rather than Grade II. The category of life-threatening portal hypertension is hard to envision unless they are referring to variceal bleeding. However, variceal bleeding is a separate CTCAE term. Furthermore, no criteria are given for determining whether an event is life-threatening or not. As for variceal hemorrhage, a Grade II AE is described as being self-limited with no intervention. Yet is it hard to imagine a scenario when variceal bleeding, whether self-limited or not, does not require at least a gastroscopy and banding or injection. In fact, in the absence of a gastroscopy the diagnosis of variceal bleeding cannot be made with certainty. Chylous ascites is included, presumably because of its occurrence in nonliver malignancies, but it certainly occurs in cirrhosis as well. The diagnosis of chylous ascites requires paracentesis, but the criteria for severity for Grade I specify that no intervention is required. Grade II severity requires intervention in the form of drainage (may be indicated) or fat free diet (not indicated). Grade III AE includes operative management, which is clearly not applicable to liver disease.

These inaccuracies and lack of precision with regard specifically to liver disease make assessment of the severity of the AE a hit and miss process, makes attribution more difficult and less likely to be accurate, and finally may be misleading (e.g., no intervention for bleeding varices) to the extent that impacts patient safety.

There is very little literature on the role of liver biopsy in assessing drug hepatotoxicity in the setting of new drug development. Yet liver biopsy could be helpful in determining the cause of a change in liver status. Yet this cannot be widely applied because the biopsy appearances of drug injury can be nonspecific. In HCC drug development however, this can be



particularly useful, given that additional immune active agents with mechanisms of action similar to nivolumab pembrolizumab are likely to be trialed in the future and since the biopsy appearances of the immune mediated injury are typical (fibrin ring granuloma and endothelial damage),<sup>14</sup> this might be very helpful in distinguishing progressive HCC from drug injury. This does not mean that, for example, Grade 1 and 2 toxicity require biopsy, but more severe changes might.

### Consequences of Misattribution

Incorrect attribution may have significant effects on trial outcomes. In most trials, one of the management options when the AE is an increased ALT or AST is to withdraw the drug, with the possibility of restarting later. If the ALT elevation is due to disease variability, the subsequent ALT might be improved, even if it has nothing to do with drug-induced injury. Yet, it looks as if withdrawal of the drug led to improvement. This would strengthen the assessment that this was due to drug, whereas it might simply be due to disease variability. There is also the concern that reintroduction of a toxic agent might accelerate and worsen the toxicity.

One might argue that as long as misattribution is similar in both treated and control this does not matter. However, misattribution may lead to premature withdrawal of the study drug in individual patients. This would have no effect on the outcome in the placebo group but might affect the outcome in the experimental group. Furthermore, even if the misattribution rate was the same in treated and control groups, a high treatment withdrawal in both groups may result in the sample size not being adequate to demonstrate a significant difference between the two arms. Misattribution may also lead to an unwarranted “black box” warning when the product is eventually approved.

The only way that toxicity can definitively be attributed to the drug being tested is to look at the frequency of the liver-related AE in treated and control groups to determine if there is a difference. For this assessment to be effective, the classification has to be accurate. If the same event was to be classified as, for example, liver failure by one investigator and ascites or jaundice by another, this may decrease the ability to correctly assess the frequency and severity of the AE. During the conduct of the study, the DMC, which has access to the unblinded data, has the task of assessing the AE frequencies. The DMC meets at specified intervals to consider safety in patients treated to that point. The DMC may be able to identify toxicity before the end of the study, but if the frequency of the AE in the treated group and the placebo group is not very different and the underlying disease itself contributes to the AE rate, the difference in specific liver-related AEs might only become evident when the final results from the full dataset are analyzed.

### Measures to Improve Reporting of Liver Toxicity in Studies in Liver Disease

Several measures to better assess hepatotoxicity of experimental drugs in liver disease have been instituted in some

studies. Changes in liver blood tests or in liver disease severity may be highlighted as events of special interest and provided to the DMC in separate tables, attribution can be improved by the sponsor setting up a small committee of expert hepatologists to review attributions and advise the sponsor as to whether the attribution was reasonable. This committee would be independent of the sponsor in the same way that the DMC is independent. This might be more effective in studies of HCC performed primarily by oncologists or nonhepatologists, but may not be necessary in studies in PBC or PSC where most of the investigators are likely to be hepatologists or gastroenterologists. This is not to cast aspersions on oncologists, but they are likely to have little exposure to chronic liver disease and liver failure.

Some rules can be developed to aid in attribution. For HCC trials, the development of ascites in the presence of visible tumor progression should only be ascribed to drug with good reason, whether the progression meets RECIST criteria or not. In the absence of visible tumor progression in trials of BCLC stage C disease ascites is almost always due to cancer progression or liver disease progression and should be interpreted as such unless there is a compelling reason not to do so. Similarly, in these patients the development of hepatic encephalopathy, spontaneous bacterial peritonitis, or variceal bleeding should be considered disease related unless evidence suggests otherwise. If, in fact, some of these events turn out to be drug-related and if they occur with any frequency, the DMC should pick this up.

For more accurate and more standardized assessment of severity of AEs, the CTCAE need to be revised. ► **Table 5** is a personal view of how this should look. These suggestions are based on experience in several DMCs in trials of HCC therapy and in other trials in liver disease. These suggested changes can hopefully be a starting point for discussion to improve the CTCAE.

### Other Concerns in Clinical Trials of HCC

The presence of advanced liver disease in the majority of these patients contributes to the death rate and has to potential to affect the outcome. One of the other possible assessments that the DMC can make is to assess whether the liver-related AE has been properly managed. Inappropriate management can be reported to the sponsor so the sponsor can provide education or set specific actions to be taken. This is to both preserve trial integrity and to protect participants. For example, if a participant develops ascites it should be part of study procedures that an abdominal paracentesis be performed to look for tumor cells or bacterial peritonitis. All trials that include participants with cirrhosis should require a mandatory gastroscopy to assess varices and all big varices should be treated prior to starting study treatment. This is particularly important in HCC trials when participants enter through oncology when gastroscopies are often not done, or at any rate, not recorded in the case report form. This is likely to be less of a problem when the participants enter the study via hepatology or gastroenterology.

In summary, attributing liver-related AEs in a therapeutic trial in patients with liver disease is difficult. This can be

**Table 5** Suggestions for a revised list of CTCAE criteria specifically for drug trials in liver disease

CTCAE term	Grade I	Grade II	Grade III	Grade IV
Ascites	Detected radiologically only	Clinically detected. Paracentesis required for diagnosis but not for therapy. Diuretics may be needed	Diuretics required. Periodic paracentesis required (less than every 3 wk)	Requires regular paracentesis (weekly to every second week)
Chylous ascites		Incidental finding on paracentesis. Diuretics may be needed	Diuretics required. Periodic paracentesis required (less than every 3 wk)	Requires regular paracentesis (weekly to every second week)
Hepatic encephalopathy	Minimal hepatic encephalopathy (reports of odd behavior, memory lapses, driving errors)	Intermittent confusion. Asterix	Confused, drowsy, but rousable	Hepatic coma
Varices (gastric or esophageal)	–	Minor bleed, not requiring intravenous fluids. Diagnosed on gastroscopy. Treatment required	Severe bleed, but easily controlled. Treatment required	Life-threatening bleed
Acute liver injury (Increase in ALT to >10x ULN if previous ALT <2xULN)	No associated liver failure (no significant change in bilirubin, INR)	Mild liver function test abnormalities (bilirubin > ULN–1.5xULN and/or INR 1.0–1.5)	Moderately severe liver function abnormality (bilirubin 1.6xULN–3xULN and/or INR > 1.5–3.0)	Bilirubin > 5xULN and/or INR >3.0. Life-threatening
Acute mitochondrial injury (increased CK, lactic acidosis starting 2–3 mo after exposure to drug)	–	Biochemical abnormalities plus abdominal pain and peripheral paresthesia	Any one of: bilirubin > 1.5xULN Albumin < LLN INR > 1.0 Ammonia > ULN Lactate > ULN	Any two of: bilirubin > 1.5xULN Albumin < LLN INR > 1.0 Ammonia > ULN Lactate > ULN Life-threatening
Portal vein thrombosis	Branch portal vein partially occluded	Branch portal vein fully occluded	Main portal vein partially occluded	Main portal vein fully occluded

Abbreviations: ALT, alanine aminotransferase; CTCAE, Common Terminology Criteria for AEs; INR, international normalized ratio; LLN, lower limit of normal; ULN, upper limit of normal.

The criterion of death for listing as Grade IV remains, but is not included in the table. This table should serve as a starting point for discussion, and should not be considered a definitive suggested revision. Terms from ►Table 4 deleted here include variceal bleeding (changed to varices), portal hypertension (represented in the terms varices and ascites), hepatic failure (now called acute hepatic injury), and hepatic necrosis (no replacement). Chronic liver failure is captured in the terms ascites, hepatic encephalopathy, and the bilirubin and INR. There are other liver related terms that have not been extracted for ►Tables 4 and 5 (e.g., Budd–Chiari syndrome and sinusoidal obstruction syndrome). There are no suggested changes for these ►Table 5.

aided by highlighting any liver-related changes meeting specified criteria (not necessarily CTCAE) to the DMC and by a better understanding of how liver disease is viewed by experts in 2021. Along these lines, the CTCAE should be revised to more accurately describe the severity of liver AEs so that “inter-observer” variability is improved.

## Main Concepts and Learning Points

- Assessing drug toxicity in trials of new therapy for liver disease is difficult.
- Standard tools to help assess drug toxicity in new trials of liver disease do not reflect modern thinking about the progression of liver disease.
- Clinical trials of new drugs in liver disease should include an expert panel to evaluate interrogator provided attributions of adverse events.
- The CTCAE criteria concerning liver-related adverse events should be revised to improve attribution and assessment of severity of liver-related adverse events.

### Conflict of Interest

None.

## References

- 1 Regev A. Drug-induced liver injury and drug development: industry perspective. *Semin Liver Dis* 2014;34(02): 227–239
- 2 McKenzie R, Fried MW, Sallie R, et al. Hepatic failure and lactic acidosis due to fialuridine (FIAU), an investigational nucleoside analogue for chronic hepatitis B. *N Engl J Med* 1995;333(17):1099–1105
- 3 Chang CY, Schiano TD. Review article: drug hepatotoxicity. *Aliment Pharmacol Ther* 2007;25(10):1135–1151
- 4 Peeraphatdit TB, Wang J, Odenwald MA, Hu S, Hart J, Charlton MR. Hepatotoxicity from immune checkpoint inhibitors: a systematic review and management recommendation. *Hepatology* 2020;72(01):315–329
- 5 Rooks JB, Ory HW, Ishak KG, et al. Epidemiology of hepatocellular adenoma. The role of oral contraceptive use. *JAMA* 1979;242(07): 644–648
- 6 Hayashi PH. Overview of causality assessment in drug-induced liver injury. *Clin Liver Dis (Hoboken)* 2017;9(02):29–33
- 7 Rockey DC, Seeff LB, Rochon J, et al; US Drug-Induced Liver Injury Network. Causality assessment in drug-induced liver injury using a structured expert opinion process: comparison to the Roussel-UCLAF causality assessment method. *Hepatology* 2010;51(06):2117–2126
- 8 United States Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER) Center for Biologics Evaluation and Research (CBER): Guidance for Industry Drug-Induced Liver Injury: Pre-marketing Clinical Evaluation, Final, July 2009.

- 9 National Cancer Institute Division of Cancer Treatment and Diagnosis. Cancer Therapy Evaluation Program. [https://ctep.cancer.gov/protocolDevelopment/electronic\\_applications/ctc.htm](https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm). Accessed December 16, 2020
- 10 International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH). [www.MedDRA.org](http://www.MedDRA.org). Accessed December 16, 2020
- 11 Council for International Organizations of Medical Sciences. [https://cioms.ch/wp-content/uploads/2017/01/Mgment\\_Safety\\_Info.pdf](https://cioms.ch/wp-content/uploads/2017/01/Mgment_Safety_Info.pdf). Accessed December 16, 2020
- 12 Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009;45(02):228–247
- 13 Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 1981;47(01):207–214
- 14 De Martin E, Michot JM, Papouin B, et al. Characterization of liver injury induced by cancer immunotherapy using immune checkpoint inhibitors. *J Hepatol* 2018;68(06):1181–1190