Monitoring for Hepatotoxicity: What Is the Predictive Value of Liver "Function" Tests?

JR Senior[1]

Drug-induced liver injury (DILI) is a major reason drugs fail during development or are withdrawn from the market.[1] The ability to predict, detect, and avoid DILI through appropriate patient selection and effective monitoring has proved to be an elusive goal. Many approved drugs have labeling recommendations for serum enzyme monitoring intended to detect and prevent hepatotoxicity, but such monitoring is often seen as inconvenient, uncomfortable, costly, and inefficient by both patients and doctors, and thus monitoring recommendations are poorly followed, if at all. This review considers whether monitoring works to prevent DILI, whether monitoring recommendations are derived from data or opinions, and whether any better alternatives exist.

WHAT HEPATOTOXICITY ARE WE LOOKING FOR?

Drug-induced liver injury (DILI) occurs at five levels of severity (**Figure 1**),[2] from mild, asymptomatic, and usually spontaneously reversible elevations of serum enzyme activities (level 1) to fatal liver failure or the need for a transplant (level 5). Less severe DILI occurs more frequently but may not be predictive of more severe levels of injury. Different populations and drugs show great variation in relative incidence of DILI. For example, tacrine,[3] which is used to treat Alzheimer's disease, caused about half of those exposed to show elevated serum alanine aminotransferase (ALT) activity but never caused level 3–5 DILI. In contrast, troglitazone,[4] the first peroxisome proliferator–activated receptor-γ agonist for treating adult-onset diabetes, less often led to elevated ALT levels but caused many cases of fatal liver failure and was withdrawn in March 2000.

The term "drug-induced liver injury" implies that we know that a drug or other agent (rather than a particular disease) caused the observed liver injury. It is not easy to be certain about this. There are no pathognomonic features, biomarkers, or findings (including liver biopsy) prove drug causality, so attribution of liver injury to a specific drug becomes a challenging differential diagnosis of exclusion of other causes and estimation of likelihood.[5] Absent any confirmed and accurate quantitative method, the likelihood estimation defaults to opinion, which is best formed by expert consensus. A categorization of such estimates is proposed in **Table 1**. This categorization does not specify the phenotype of the liver injury (hepatocellular, cholestatic, hypersensitivity, mitochondrial, or other) or address issues of histopathology, pathophysiology, or mechanisms of injury.

In evaluating clinical trials, US Food and Drug Administration reviewers look for the potential of a new drug to cause severe DILI (level 5) in at least some people,[1] recognizing that certain people react idiosyncratically and may show significant DILI at doses and exposure durations well tolerated by most. Even when mild liver injury follows initial exposure to a drug, most people can overcome injury to liver cells, recover, adapt[6] and become tolerant to continued exposure to the drug or resumed exposure after interruption. This is exemplified by isoniazid, which is valuable in preventing tuberculosis. Isoniazid caused significant ALT elevation in ~20% of recipients initially, yet only 1 in 1,000 patients failed to adapt and was taken off the drug to avoid serious or fatal DILI from continued exposure.[7] In looking at the potential of a drug to cause severe DILI, lower levels of DILI (especially level 1) occur far more frequently, but level 2 DILI raises much more significant concern and may restrict development of the drug.[1]

The most dangerous type of DILI is the rapid onset of hepatocellular injury extensive enough to impair the liver's capacity to function normally. The liver performs myriad functions in maintaining the body's internal chemical homeostasis, but regulation of the plasma activities of enzymes is not among them (hence the quotation marks around "function" in the title of this article). It is likely that the activities of serum enzymes are determined by a dynamic balance between the rate of their release from cells and the rate of their inactivation by reticuloendothelial system proteolytic degradation, although careful study is needed to support this assertion. Activity levels are volatile, changing over hours and days. When release rates from injured cells exceed clearance rates, the increases in activity levels may offer a rough measure of the rate at which cells are injured, but they are not measures of the functional capacity of the liver. When enough liver cells are injured to reduce the organ's ability to perform its normal functions, such as clearance of bilirubin

[1]Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, Maryland, USA. Correspondence: JR Senior (john.senior@fda.hhs.gov)

Received 30 October 2008; accepted 15 November 2008; advance online publication 7 January 2009. doi:10.1038/clpt.2008.262

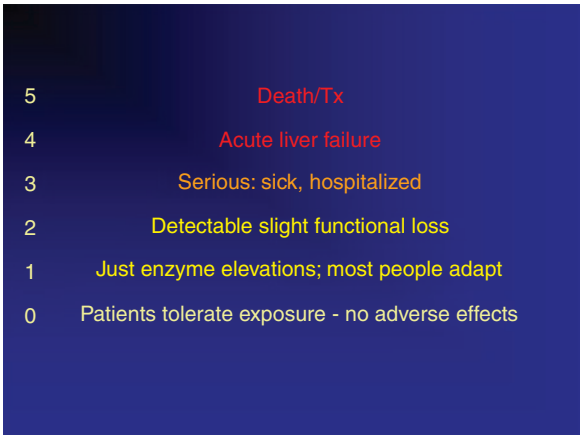

Figure 1 Levels of severity of drug-induced liver injury (DILI). Lower levels of severity generally occur much more frequently than higher levels. Levels 1 and 2 are determined primarily by serum chemistry measures of enzyme activity and total bilirubin concentration, but determinations of level 3–5 DILI are based primarily on additional clinical information. The relative rates of the five levels differ for any given drug and population sample exposed.

Table 1 Categories of likelihood of drug-induced liver injury

Category	Definition
0	*Almost certainly not*, some other explanation is probable; ER <5%
1	*Unlikely*, other possibilities greater; ER = 5–25%
2	*Possible*, no other cause probable; ER >25–50%
3	*Probable*, most likely of all possible causes; ER >50–75%
4	*Very likely*, other causes unlikely; ER >75–95%
5	*Definite*, almost certain, no other cause apparent; ER ≥95%

No precise, accurate, objective measure is available to estimate the likelihood that a given case of liver injury was caused by a drug. Therefore, the best current method is based on expert consensus and relies on additional clinical information to rule out alternative causes.

ER, estimated range; not intended to imply numerical precision but as an adjective to modify the category definition and thus increase the consistency of estimates among evaluators.

from plasma into bile, the DILI is more severe (levels 2–5); these are the so-called Hy's law cases.

ETIOLOGY AND DERIVATION OF HY'S LAW

In 1968, the late Hyman Zimmerman stated that "hepatocellular drug-induced jaundice is a grave illness with an estimated mortality rate of 10 to 50 per cent."[8] Zimmerman repeated this statement in his textbooks of 1978[9] and 1999,[10] although he did not specify how he determined that a case was drug-induced, what level of serum enzyme elevation indicated hepatocellular DILI, or what serum bilirubin concentration defined jaundice. His observation, which appeared accurate over time, was dubbed "Hy's law" by Robert Temple of the US Food and Drug Administration.[11] Temple defined Hy's law in terms of serum abnormalities (ALT > 3 × the upper limit of normal (ULN) and total bilirubin (TBL) > 2 × ULN, after appropriate exclusion of cholestatic problems, Gilbert's syndrome, and other details), rather than in general clinical terms, and on the basis of somewhat more severe cases than Zimmerman used. Epidemiologic confirmation of Zimmerman's concept was recently reported in Sweden[12] and Spain.[13] Elevation of serum ALT activity has proved to be a quite sensitive but not entirely specific test for liver injury; the older TBL test is much more specific but rather insensitive for assessing liver dysfunction. The combination of these two measures has proved very useful. More severe levels of liver injury (levels 3–5) are determined by clinical observation and not just by serum chemical measurements. For >30 years, Hy's law has proved to be valid in predicting that some people showing drug-induced hepatocellular injury with jaundice will progress to fatal hepatotoxicity or require a liver transplant.

PRECLINICAL ASSESSMENTS FOR PREDICTING HUMAN DRUG-INDUCED HEPATOTOXICITY

Preclinical animal and *in vitro* testing has been used quite successfully to screen out dangerously hepatotoxic drugs but has not always predicted the rare but sometimes serious idiosyncratic hepatotoxicity that occurs in some individuals when hundreds of thousands or millions of people are treated. This failure has been attributed to the great genetic and experiential diversity of humans, which cannot be modeled in traditional toxicological studies using purebred or standard strains of animals, controlled conditions of exposure, and an absence of confounding conditions or treatments in the experimental animals. However, it may be possible, using innovative methods, to show that idiosyncratic toxicity is not always rare, and that it is dose-related, is fairly predictable, and can be demonstrated in modest numbers of standard laboratory animals. Data from such studies may allow identification of differentially susceptible individual animals and lead to truly predictive tests or procedures that can be used to identify which patients should not be exposed to a particular drug so it can be given safely to others. These issues are complex and beyond the scope of this brief discussion.

HOW IS HEPATOTOXICITY DETECTED IN CLINICAL TRIALS?

Because increases in serum enzyme concentrations are not uncommon and are a major concern in drug development when more severe levels of DILI occur, most protocols require that sera be tested for biomarkers before and during exposure of participants to investigative and control agents. Among the biomarkers usually tested for are ALT, aspartate aminotransferase (AST), alkaline phosphatase, and TBL. The first three are enzymes and are measured by activity level; TBL is measured by serum concentration. Only TBL is a measure of liver function; increased serum enzyme activity levels reflect injury to hepatocytes or cholangiocytes. Transamination is a basic biochemical process that links carbohydrate and protein metabolism in many tissues.[14] Although ALT is highly concentrated in hepatocytic cytoplasm, it is also present in the mitochondria[15] and many other tissues (heart and skeletal muscle, gut epithelial cells, and renal tubular cells). Isoforms have recently been described.[16] Clinical use of serum enzyme measurements was jump-started in 1955 by the brilliant work of Arthur Karmen,[17] who developed a rapid spectrophotometric assay to replace the tedious methods used previously. Karmen's test used AST (then known as serum glutamic-oxalacetic transaminase) to diagnose

myocardial infarction. The use of AST and ALT measurements to detect acute hepatocyte injury soon became more widespread, and ALT is now the most frequently used indicator, with AST testing usually rising in parallel redundantly.

There is no agreement on a standard for the level of ALT activity that should trigger action in a clinical trial, although the expert consensus of the 1978 Fogarty Conference was that 3 × ULN was "markedly abnormal."[18] This level was arbitrary, not data-based, and indicates that injury has already taken place rather than predicting what will or might happen. Single measurements do not determine "peaks" for this volatile assay; only fairly close, serial measurements can do so. Most people with abnormal levels of ALT activity will resolve them spontaneously by robust liver repair, regeneration, and recovery mechanisms, whether or not the drug is stopped, reduced, or interrupted. Even initially susceptible people will tend to adapt and become tolerant of the drug. It is not clear whether higher ALT activity, such as 5 × ULN or 8 × ULN, would be a better level for consideration. Moreover, ALT elevation is not a very specific indicator of serious liver injury, resulting in many false-positives that prompt costly and useless additional studies and do not distinguish between drug-induced and disease-induced injury.

If we use ALT > 3 × ULN as a positive test for potentially serious liver injury in a sample population exposed to a drug for which the true incidence of DILI is 1%, with test sensitivity of 95% and specificity of 85% (reasonable estimates), the hard reality is that most of the positive test results will be false (**Table 2**). The timing of elevated ALT varies, and thus many periodic tests may be required to detect the elevation. If we add in the requirement for baseline measures for comparison, then perhaps an average of 10 tests may be necessary. If the true incidence of DILI is 1%, then almost 300 people must undergo 3,000 tests for a 95% chance of finding at least one positive result, and there is a 94% chance that this result will be false. Not all serious cases will be drug-induced; some will be disease-induced. Thus, to find one bonafide case of DILI, even at a relatively high incidence of 1%, many thousands of tests must be performed, depending on the incidence and prevalence of liver diseases among the study population. This is not only inefficient but costly, discouraging, and needlessly alarming. Elevated ALT alone in an individual, whether 3 × ULN or 10 × ULN, is not reliably predictive of serious liver injury.

In a 5-year study of 3,248 patients on a placebo, ALT elevations >3 × ULN (with redundant AST rises) were found in 44 patients.[19] Most were not due to serious or diagnosable liver disease, but six patients who also had TBL elevations to >2 × ULN showed liver disease on hospitalization: two had acute viral hepatitis, two had common duct gallstones, and two had fatal cases of liver infiltration with amyloid or metastatic colon cancer. Hence, the specificity of the combined test for serious liver injury was 100%, but the specificity of the ALT elevation test alone was only 14%. This work needs to be replicated.

Table 2 Sensitivity and specificity of ALT elevations

Test result	Serious liver injury: Present	Serious liver injury: Absent	Total	Predictive value
ALT > 3 × ULN	95	1,485	1,580	0.060
Not	5	8,415	8,420	0.999
Incidence 1%	100	9,900	10,000	
	Sensitivity 95%	Specificity 85%	Accuracy 85.1%	

Clinical interpretation of alanine aminotransferase (ALT) elevations. Using a cutoff for a positive test result of ALT > 3 × ULN to detect serious liver injury in a population sample of 10,000 patients exposed to a given drug with a true incidence of 1% showing treatment-emergent hepatotoxicity during exposure, test sensitivity of 95%, and test specificity of 85%. Note the poor predictive value of positive test results: only 6% (94% wrong).

ULN, upper limit of normal.

DOES PERIODIC MONITORING OF SERUM ALT PREVENT SERIOUS DILI?

Although scores of drugs include recommendations for monitoring on their labels,[20] the question arises as to whether monitoring works, even when done properly. Moreover, as monitoring is often not done well in practice, it has little chance of working. The basis for the continued recommendation of monitoring has been challenged.[21] In contrast to efficacy indications, which require data on which to base statements on approved labeling, there are almost no such data on which to base recommendations for monitoring, resulting in default to consultant opinions.[22] Only very sparse evidence is available to show that periodic ALT monitoring prevents serious DILI,[23,24] and there are no systematic studies to prove that it does. This point was advanced by DeLeve at a 2003 workshop on screening in liver disease.[25] Delay between initiating drug treatment and the onset of hepatotoxicity is extremely variable, both for individual drugs and for individual subjects within a treatment group. Monitoring could work for drugs for which the pace of injury is relatively slow, by allowing detection of early injury in time to prevent worsening, but even that has not been proved. For drugs in which the onset of hepatotoxicity is rapid (as reported for telithromycin[26] and, in some studies, troglitazone[27]), hepatic injury may progress to irreversible liver failure within less than a reasonable monitoring interval.

Monitoring during controlled clinical trials is another matter. When a drug's efficacy has not been established and safety problems remain to be discovered, monitoring makes sense. It is not clear what frequency or duration of testing or intervals between tests may be optimal, and these parameters may vary with both the drug and the study population. Obviously, peak ALT values may be missed if they are brief, asymptomatic, and spontaneously reversible, but there are practical limitations on how often and for how long tests can be performed. Clinical data on which to base recommendations are sorely needed.

IS THERE ANY ALTERNATIVE TO PERIODIC SERUM TESTING OF ALT?

Two large studies of isoniazid monotherapy to prevent active tuberculosis were conducted in similar populations. The earlier study relied on monthly monitoring for symptoms or signs of hepatic injury among 13,838 people who took isoniazid for a year. There were eight deaths from liver failure.[28] In the more recent study, patients were carefully and repeatedly instructed

to check themselves daily and to interrupt treatment and immediately have their serum ALT and TBL measured if they noticed even mild symptoms of possible liver injury.[7] Prompt testing after symptoms were perceived was effective: only 11 of 11,141 patients had to have their isoniazid treatment stopped permanently, and there were no deaths. This approach was both clinically effective and cost-effective in that somewhat younger group, but it cannot necessarily be generalized.

Monitoring in practice is meant to prevent severe hepatotoxicity, not just to detect ALT elevations in mild, reversible liver injury, and physicians should be fully aware of the limitations of monitoring. In clinical trials, the purpose of periodic ALT monitoring should be to detect cases for special attention in order to learn as much as possible about the effects of the drug under investigation[1] and to prompt appropriate closer observation, additional studies, and follow-up in the few cases of liver injury that occur.

Until better tests are developed, the detection of hepatocyte injury is suggested for serum transaminase activities >3 × ULN and liver function impairment using laboratory measures of excretory (TBL > 2 × ULN) and synthetic (prothrombin time, international normalized ratio >1.5) functions. Unexplained transaminase elevations combined with loss of liver function in a clinical trial are considered a worrisome finding by the US Food and Drug Administration and factor into the agency's assessment of the benefit–risk balance for a drug.

It remains to be seen whether better biomarkers can be found,[29] especially truly predictive indicators of which people should not be exposed to a particular drug because of a high likelihood that they will not be able to tolerate or adapt to it. Clinical trials, used to their fullest potential,[30] may help us discover a basis for individual idiosyncratic susceptibility that will lead to the safer use of new drugs in the right patients.

ACKNOWLEDGMENTS

Opinions expressed in this article are those of the author and do not represent policies or positions of the US Food and Drug Administration.

CONFLICT OF INTEREST

The author declared no conflict of interest.

1. US Food and Drug Administration. Draft guidance for industry. Drug-Induced Liver Injury: Premarketing Clinical Evaluation <http://www.fda.gov/cder/guidance/7507dft.htm> (2007).
2. Senior, J.R. Lessons from Isoniazid: Would It Be Approved Today? (see slide 12) <http://www.fda.gov/cder/livertox/presentations2008/C-Senior2.pdf> (2008).
3. Watkins, P.B., Zimmerman, H.J., Knapp, M.J., Gracon, S.I. & Lewis, K.W. Hepatotoxic effects of tacrine administration in patients with Alzheimer's disease. *JAMA* **271**, 992–998 (1994).
4. Watkins, P.B. & Whitcomb, R.W. Hepatic dysfunction associated with troglitazone. *N. Engl. J. Med*. **338**, 916–917 (1998).
5. Rockey, D.C. Was It the Drug, or a Disease? How to Determine It <http://www.fda.gov/cder/livertox/presentations2008/F-Rockey2.pdf> (2008).
6. Senior, J.R. Liver Adaptation to Xenobiotic Injury <http://www.fda.gov/cder/livertox/presentations2006/Senior.pdf> (2006).
7. Nolan, C.M., Goldberg, S.V. & Ruskin, S.E. Hepatotoxicity associated with isoniazid preventative therapy: a 7-year survey from a public health tuberculosis clinic. *JAMA* **281**, 1014–1018 (1999).
8. Zimmerman, H.J. The spectrum of hepatotoxicity. *Perspect. Biol. Med*. **12**, 135–161 (1968).
9. Zimmerman, H.J. *Hepatotoxicity: The Adverse Effects of Drugs and Other Chemicals on the Liver* 351–357 (Appleton-Century-Crofts, New York, 1978).
10. Zimmerman, H.J. *Hepatotoxicity: The Adverse Effects of Drugs and Other Chemicals on the Liver* 2nd edn. 428–430 (Lippincott, Williams & Wilkins, Philadelphia, PA, 1999).
11. Temple, R. Hepatotoxicity Through the Years: Impact on the FDA <http://www.fda.gov/cder/livertox/meeting2001.htm#Drug-Induced> (2001).
12. Björnsson, E. & Olsson, R. Outcome and prognostic markers in severe drug-induced liver disease. *Hepatology* **42**, 481–489 (2005).
13. Andrade, R.J. *et al*. Drug-induced liver injury: an analysis of 461 incidences submitted to the Spanish registry over a 10-year period. *Gastroenterology* **129**, 512–521 (2005).
14. Awapara, J. & Seale, B. Distribution of transaminases in rat organs. *J. Biol. Chem*. **194**, 497–502 (1952).
15. Swick, R.W., Barnstein, P.A. & Stange, J.L. The metabolism of mitochondrial proteins: I. Distribution and characterization of the enzymes of alanine aminotransferase in rat liver. *J. Biol. Chem*. **240**, 3334–3340 (1965).
16. Lindblom, P. *et al*. Isoforms of alanine aminotransferases in human tissues and serum—differential tissue expression using novel antibodies. *Arch. Biochem. Biophys*. **466**, 66–77 (2007).
17. Karmen, A. A note on the spectrometric assay of glutamic oxalacetic transaminase in human blood serum. *J. Clin. Invest*. **34**, 131–133 (1955).
18. Davidson, C.S., Leevy, C.M. & Chamberlayne, E.C. Guidelines for detection of hepatotoxicity due to drugs and chemicals. Fogarty Conference. US Department of Health, Education, and Welfare. National Institutes of Health Publication no. 79–313, 114 (Public Health Service, Washington, DC, 1978).
19. Tipping, R.W. Liver Test Elevations in Patients on Placebo <http://www.fda.gov/cder/livertox/presentations2008/B-Tipping2.pdf> (2008).
20. Willy, M.E. & Li, Z. What is prescription labeling communicating to doctors about hepatotoxic drugs? A study of FDA approved product labeling. *Pharmacoepidemiol. Drug Saf*. **13**, 201–206 (2004).
21. Sniderman, A.D. Is there value in liver function test and creatine phosphokinase monitoring with statin use? *Am. J. Cardiol*. **94**, 30F–34F (2004).
22. Lewis, J.H. "Hy's Law," the "Rezulin Rule," and other predictors of severe drug-induced hepatotoxicity: putting risk-benefit into perspective. *Pharmacoepidemiol. Drug Saf*. **15**, 221–229 (2006).
23. McNeill, L., Allen, M., Estrada, C. & Cook, P. Pyrazinamide and rifampin vs isoniazid for the treatment of latent tuberculosis: improved completion rates but more hepatotoxicity. *Chest* **123**, 102–106 (2003).
24. Beghetti, M. *et al*. Safety experience with bosetan in 146 children 2-11 years old with pulmonary arterial hypertension: results from the European postmarketing surveillance program. *Pediatr. Res*. **64**, 200–204 (2008).
25. Adams, P.C. *et al*. Screening in liver disease: report of an AASLD clinical workshop. *Hepatology* **39**, 1204–1212 (2004).
26. Clay, K.D., Hanson, J.S., Pope, S.D., Rissmiller, R.W., Purdum, P.P. 3rd & Banks, P.M. Brief communication: severe hepatotoxicity of telithromycin: three case reports and literature review. *Ann. Intern. Med*. **145**, 415–420 (2006).
27. Graham, D.J., Green, L., Senior, J.R. & Nourjah, P. Troglitazone-induced liver failure: a case study. *Am. J. Med*. **114**, 299–306 (2003).
28. Kopanoff, D.E., Snider, D.E. Jr. & Caras, G.J. Isoniazid-related hepatitis: a U.S. Public Health Service cooperative surveillance study. *Am. Rev. Respir. Dis*. **117**, 991–1001 (1978).
29. Ozer, J., Ratner, M., Shaw, M., Bailey, W. & Schomaker, S. The current state of serum biomarkers of hepatotoxicity. *Toxicology* **245**, 194–205 (2008).
30. Watkins, P.B., Seligman, P.J., Pears, J.S., Avigan, M.I. & Senior, J.R. Using controlled clinical trials to learn more about acute drug-induced liver injury. *Hepatology* **48**, 1680–1689 (2008).