


ORIGINAL ARTICLE

A revised electronic version of RUCAM for the diagnosis of DILI

Paul H. Hayashi¹  | M. Isabel Lucena^{2,3} | Robert J. Fontana⁴ |
 Einar S. Bjornsson^{5,6} | Guruprasad P. Aithal⁷ | Huiman Barnhart⁸ |
 Andres Gonzalez-Jimenez^{2,3} | Qinghong Yang⁸ | Jiezhun Gu⁸ |
 Raul J. Andrade^{2,3} | Jay H. Hoofnagle⁹

¹Food and Drug Administration, Silver Spring, Maryland, USA

²Servicio de Aparato Digestivo y de Farmacología Clínica, Hospital Universitario Virgen de la Victoria, Instituto de Investigación Biomédica de Málaga-IBIMA, Facultad de Medicina, Universidad de Málaga, Málaga, Spain

³Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas, UICEC-IBIMA, Plataforma Investigación Clínica, Madrid, Spain

⁴University of Michigan, Ann Arbor, Michigan, USA

⁵Landspítali University Hospital, Reykjavik, Iceland

⁶Faculty of Medicine, University of Iceland, Reykjavik, Iceland

⁷NIHR Nottingham BRC, Nottingham University Hospitals NHS Trust and the University of Nottingham, Nottingham, UK

⁸Duke University, Durham, North Carolina, USA

⁹National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, Maryland, USA

Correspondence

Paul H. Hayashi, MD, MPH, FAASLD, Physician Lead, DILI Team, Division of Hepatology and Nutrition, Food and Drug Administration, 10903 New Hampshire Avenue, WO-22/5387, Silver Spring, MD 20993, USA.

Email: Paul.Hayashi@fda.hhs.gov

Funding information

The Drug-Induced Liver Injury (DILN) Network is structured as a U01 cooperative agreement with funds provided by the National Institute of Diabetes and Digestive and Kidney Diseases (U24-DK065176, U01-DK065201, U01-DK065184, U01-DK065211, U01DK065193, U01-DK065238, U01-DK083023, U01-DK083027, U01-DK082992, U01-DK083020, and U01-DK100928). Additional support is provided by CTSA grants (UL1 RR025761, UL1TR000083, UL1 RR024134, UL1 RR024986, UL1 RR024982, UL1 RR024150), the Intramural Research Program of the National Institutes of Health, the National Cancer Institute (ClinicalTrials.gov number: NCT00345930), and the 2016

Abstract

Background and Aims: Roussel Uclaf Causality Assessment Method (RUCAM) for DILI has been hindered by subjectivity and poor reliability. We sought to improve the RUCAM using data from the Drug-Induced Liver Injury Network (DILIN) and the Spanish DILI Registry, published literature, and iterative computer modeling.

Approach and Results: RUCAM criteria were updated, clarified, and computerized. We removed criteria 3 (risk factors) for lack of added value and criteria 4 because we felt it more useful to assess each drug separately. Criteria 6 (drug-specific risk) was anchored to LiverTox likelihood scores. Iterative testing in subsets of 50–100 single-agent, nonherbal cases from both registries was done to optimize performance. We used classification tree analysis to establish diagnostic cutoffs for this revised electronic causality assessment method (RECAM) and compared RECAM with RUCAM for correlation with expert opinion diagnostic categories in 194 DILI cases (98 DILIN, 96 Spanish DILI). Area under receiver operator curves for identifying at least probable DILI were the same at 0.89 for RECAM and RUCAM. However, RECAM

SEE EDITORIAL ON PAGE 3

Abbreviations: AIH, autoimmune hepatitis; ALT, alanine aminotransferase; ALP, alkaline phosphatase; AST, aspartate aminotransferase; DILIN, DILI Network; HDS, herbal and dietary supplement; RECAM, Revised Electronic Causality Assessment Method; RUCAM, Roussel Uclaf Causality Assessment Method; ULN, upper limit of normal.

Paul H. Hayashi and M. Isabel Lucena contributed equally to this work.

© 2022 American Association for the Study of Liver Diseases.

American Association for the Study of Liver Disease (AASLD) Innovations Fund. The Spanish DILI Registry is funded by grants from Instituto de Salud Carlos III, cofunded by Fondo Europeo de Desarrollo Regional - FEDER (PI 18/01804 and PT 20/00127) and Agencia Española del Medicamento. Plataforma ISCIII de Investigación Clínica and CIBERehd are funded by ISCIII

diagnostic categories have better observed overall agreement with expert opinion (0.62 vs. 0.56 weighted kappa, $p = 0.14$), and had better sensitivity to detect extreme diagnostic categories (73 vs. 54 for highly likely or high probable, $p = 0.02$; 65 vs. 48 for unlikely/excluded, $p = 0.08$) than RUCAM diagnostic categories.

Conclusions: RECAM is an evidence-based update that is at least as capable as RUCAM in diagnosing DILI compared with expert opinion but is better than RUCAM at the diagnostic extremes. RECAM's increased objectivity and clarity will improve precision, reliability, and standardization of DILI diagnosis, but further refinement and validation in other cohorts are needed.

INTRODUCTION

The diagnosis of DILI is primarily based on clinical judgment and the elimination of alternate diagnoses. Lack of an evidence-based and a reliable diagnostic tool is a significant hindrance to clinical care and research. In 1993, Danan and Benichou published the Rousell Uclaf Causality Assessment Method (RUCAM, also credited to Council of International Organizations of Medical Sciences), which is a diagnostic scorecard based on seven clinical criteria.^[1] It is the most widely used and accepted DILI diagnostic tool.^[2,3] However, clinical and research usefulness is still debated.^[4,5]

Since 1993, there have been three major problems with RUCAM: (1) unclear operating instructions and subjectivity leading to poor reliability and usability, (2) unclear validity due to lack of an accepted gold standard, and (3) domain criteria that are not evidence-based.^[6] Even the updated RUCAM, which is quite similar to the original, retains a significant degree of subjectivity in terms of ruling out competing diagnoses.^[7] Nevertheless, RUCAM's criteria include most of the critical elements needed to make a diagnosis of DILI, thus providing a framework for evaluation. Despite its limitations, this framework has led to RUCAM's durability in publications. However, establishing a causal relationship between exposure to an agent and the appearance of liver injury remains the Achilles heel in DILI research, and improved standardization, automation, and reproducibility in causality assessment are needed.

Using an evidence-based approach, we sought to revise the RUCAM, with an aim of having an instrument that not only had criterion and construct validity against the current RUCAM but improved precision and reproducibility. We used data from two large prospective DILI registries of well-vetted cases, the US Drug-Induced Liver Injury Network (DILIN), and the Spanish DILI Registry, to refine and develop instrument domains and scoring. We then piloted the instrument in randomly selected cases to determine the

instruments performance properties in comparison to RUCAM.

METHODS

Process overview

Since 2015, the authors met regularly to modify the RUCAM criteria using data from the DILIN and Spanish Registry cases. In addition, a review of the published literature and expert opinion were used when robust data were lacking. The development was restricted to provide assessment of single medication cases because full separate assessment for each competing agent would have been needed to achieve reliable scoring. Herbal and dietary supplement (HDS) product cases were also excluded due to the uncertainty of product contents and less well-established causality assessment methods.

The instrument was developed through five sequential stages: (1) Each of the seven RUCAM criteria were separately analyzed and revised to optimize diagnostic scoring. Registry data for latency and dechallenge were robust and well-suited to optimize cutoff values and scores. Contrary to expectations, distinction between hepatocellular and cholestatic/mixed injuries was not necessary for latency and dechallenge scoring. Other criteria changes were based on a combination of registry data, expert opinion, and available literature. (2) The revised criteria, renamed domains, were tested for ability to detect at least probable DILI cases in the DILIN. During this stage, revisions were made including elements added or discarded based on performance contribution. (3) Computer programming was applied to extract data directly from the DILIN database and Spanish Registry with single-agent DILI cases of varying levels of prior causality scoring. We assessed concordance of computer scoring with human scoring to ensure proper computer programming. (4) The revised electronic causality assessment method (RECAM) scored groups of

50–100 single medication cases from the DILIN stratified equally on DILIN's five expert diagnostic categories (see next paragraph). Scoring outputs were used to revise the RECAM and programming to optimize performance. (5) RECAM was then applied to groups of 50–100 single-agent, non-HDS cases randomly selected from the Spanish DILI Registry to assess instrument performance including domain validity and comparison of scoring obtained with RUCAM. Through this final phase, the RECAM went through modifications by an iterative process of testing both DILIN and Spanish DILI cases. RECAM was applied across the range of DILI likelihood categories used by the DILIN and Spanish DILI Registry. Throughout the process, an emphasis was placed on clarity, performance, and precise language that would be adaptable to a clinically useful website application with minimal subjective opinion from the user.

Likelihood categories and causality assessment in the DILIN

DILIN uses a consensus expert opinion method of causality assessment previously described.^[8] Each case was evaluated by three DILIN hepatologists who independently assigned an ordinal causality score or category representing percent likelihood of attribution (1 = definite or >95% likelihood, 2 = highly likely or 75–95%, 3 = probable or 50%–74%, 4 = possible or 25%–49%, and 5 = unlikely or <25%). Consensus was reached by e-mail and monthly conference calls. The enrolling DILIN investigator also provides a RUCAM score for each case.

Likelihood categories and causality assessment in the Spanish DILI Registry

Each case referred to the Spanish Registry was independently assessed and adjudicated by at least three expert investigators. Expert opinion is used to assess whether DILI consideration was reasonable and whether further data should be requested from the referring providers as needed. Case likelihood categorization is based on traditional RUCAM categories, but expert opinion can override the RUCAM assigned category as necessary (e.g., drugs with long half-lives and known long latencies after drug stop, mandatory testing of hepatitis E).^[9,10]

RECAM and RUCAM performance in diagnosing DILI in DILIN and Spanish DILI Registry

A total of 100 and 96 single-agent, non-HDS cases from the DILIN and Spanish-DILI, respectively, were

randomly chosen for testing the 12th and final version of RECAM. We used the R -value ($[\text{alanine aminotransferase (ALT)/upper limit of normal (ULN)}] + [\text{alkaline phosphatase (ALP)/ULN}]$) to categorize cases as hepatocellular ($R \geq 5$), cholestatic ($R \leq 2$), or mixed ($2 < R < 5$). The DILIN cases were stratified equally across its five likelihood categories. One DILIN case was excluded due to data-entry error in DILIN adjudication, requiring re-assessment, and another DILIN case was excluded due to an indirect, atypical liver injury of drug-induced sphincter of Oddi dysfunction. Therefore, 98 DILIN cases were used for RECAM scoring.

RECAM scoring was undertaken via semi-automated computer data extraction and scoring from both registries. Computer programming used software version 9.4 and R language version 4.02 for the DILIN cases in which R language version 3.5.0 was used for Spanish Registry cases. However, both registry databases contain free text fields (e.g., imaging, histology findings) that required some human interpretation and input for the computer to score the RECAM correctly.

Area under the curves and diagnostic cutoffs for RECAM

For the purposes of comparing performance between registries and combining data, the DILIN definite and highly likely cases were combined and considered equivalent to the highly probable Spanish cases. Similarly, the unlikely and excluded cases in the Spanish Registry were combined and considered equivalent to the DILIN unlikely cases. The other category labels of probable and possible are the same in both registries. Area under the curve (AUC) values were generated for both RECAM and RUCAM scores. RECAM and RUCAM AUC values for identification of at least high probable (or at least highly likely), at least probable, and at least possible DILI were determined for both registries. Overall, correlation of RECAM and RUCAM to DILIN and Spanish Registry expert opinion diagnostic categories was assessed by using Spearman's Rho coefficient.

Using the combined DILIN and Spanish Registry data, we built a classification tree^[11] based on RECAM scores to obtain three cutoffs for classifying each case into four categories: highly likely/highly probable, probable, possible, and unlikely/excluded. Performance of RECAM classification based on these cutoffs was compared with the performance of RUCAM classification based on its published cutoff scores of highly probable (≥ 9), probable (8 to 6), possible (5 to 3), and unlikely/excluded (≤ 2). We tested the overall percent agreement and Cohen's weighted kappa coefficient between the RECAM and RUCAM scales with expert's opinion. Diagnostic performance, sensitivity, and specificity values were calculated for

the diagnostic categories. *P* values are reported for testing the equality of agreement metrics (overall agreement, sensitivity, and specificity) of RECAM and RUCAM diagnostic categories with expert's opinion via the generalized estimating equations and for testing equality of weighted Kappa statistics of RECAM and RUCAM diagnostic categories with expert's opinion via the bootstrap approach, to account for correlation of RECAM and RUCAM diagnostic categories within the same subject.

Role of funding source and institutional review board

An American Association for the Study of Liver Diseases Innovations Fund awarded in 2016 supported this project. The DILIN Network is structured as a U01 cooperative agreement with funds provided by the National Institute of Diabetes and Digestive and Kidney Diseases. Separate institutional review board approvals were maintained at each center throughout study participation. The Spanish DILI Registry is funded by competitive grants from the National Health Institute–FEDER and Spanish Medicines Agency. All patients enrolled in both registries provided written informed consent.

RESULTS

RUCAM modifications for RECAM development

The seven original RUCAM criteria (Table S1) were modified, reordered, and renamed as domains. The resulting five RECAM domains are given in Table 1. Here we describe how the RUCAM criteria were modified and resulted in the five RECAM domains.

Criteria 1 (Time to onset)

We retained latency from both drug start and stop to form Domains 1a and 1b, but time intervals for scoring were revised, and the need to stratify by type of liver injury (determined by *R*-value) was eliminated (Table 1). The original RUCAM was unclear as to whether both or only one latency is to be scored. Unlike the updated RUCAM, which scores one or the other,^[7] the RECAM requires both latencies from drug start (Domain 1A) and stop (Domain 1B) to be scored. However, latency after stopping drug can only hurt the case for DILI by subtracting up to 6 points. Some drugs (e.g., monoclonal antibodies) clearly have long half-lives and long latency to DILI. For these drugs, Domain 1b is passed over with no points taken. Cutoffs for point allocations

were based on the DILIN and Spanish Registry latency data across likelihood categories. Time intervals were expanded creating a wider range of scores compared with RUCAM.

Criteria 2 (Course)

In RUCAM, dechallenge time cutoffs and scoring are different for hepatocellular and cholestatic/mixed cases. Based on analysis of DILIN and Spanish Registry data, dechallenge timing is similar for hepatocellular and cholestatic/mixed cases in terms of causality. Therefore, dechallenge time cutoffs are the same regardless of *R*-value and were based on the observed distribution of dechallenge times across definite to unlikely DILIN cases. *R*-value still defines which liver biochemistry to use for dechallenge scoring. Hepatocellular injury cases follow the course of ALT, whereas cholestatic and mixed-injury cases follow ALP or total bilirubin, whichever yields a higher score (Table 1). This modified dechallenge criteria became Domain 2.

Criteria 3 (Risk factors)

For the standard RUCAM and new RECAM, these three variables did not contribute significantly to logistic regression modeling to diagnose at least probable DILIN cases (age, OR 1.12 [95% CI 0.71–1.76], *p* = 0.62; alcohol and pregnancy, OR 0.90 [0.47–1.73], *p* = 0.75). This lack of Domain 3 contribution coincided with expert opinion and clinical experience of the group.^[12,13] Therefore, Criteria 3 were eliminated.

Criteria 4 (Concomitant drugs)

We reasoned that concomitant medications of clinical significance should be scored separately for simplicity and reliability of scoring. The assessment of competing drugs in RUCAM is prone to subjectivity (e.g., “suggestive” timing, “known as hepatotoxin”) and does not provide detailed assessment for these agents (Table S1).^[1] Therefore, we limited this revised RUCAM to assess drugs individually, and these concomitant drug criteria are not included in the RECAM.

Criteria 5 (Search for nondrug causes)

This RUCAM criteria became RECAM Domain 4 (Table 1). All competing diagnoses in the RUCAM were retained, but HEV, congestive hepatopathy, infiltrating cancer, and cholestasis of sepsis based on what is considered necessary evaluation testing in the literature were added (Table 2).^[14,15] We chose to only penalize

TABLE 1 RECAM algorithm (Domains 1–5)

Domain 1a & 1b:	Points
Score both sections 1a (onset after drug start) and 1b (onset after drug stop)	
1a: Onset after drug start (points given)	
Days after drug start where <i>day 1 is first day drug taken</i>	
≤1 day	-6
2–9 days (inclusive)	3
10–60 days (inclusive)	4
61–90 (inclusive)	2
>90 days	0
1b: Onset after drug stop (points taken) (for long 1/2 life agents^a, enter zero points for Domain 1b)	
Days after drug stop where <i>day 1 is the first day the drug is not taken</i>	
≤30 days	0
31–60 days (inclusive)	-1
61–90 days (inclusive)	-2
91–120 days (inclusive)	-4
>120 days	-6
Domain 2: Dechallenge or washout	
Initial R-value ≥ 5: Apply washout criteria below to serum ALT	
Initial R-value < 5: Apply washout criteria below to either ALP or bilirubin, whichever gives a higher score	
ALT, ALP, or bilirubin (whichever used by R-value criteria) declines to less than 50% of peak	
If drug still taken when greater than 50% of peak, decline occurs	-6
Days from peak value to less than 50% of peak (assumes drug was discontinued)	
1–30 days (inclusive)	4
31–90 days (inclusive)	3
91–182 days (inclusive)	2
183–365 days (inclusive)	1
>365 days	0
All other instances in which ALT, ALP, or bilirubin does not decline, or has not yet declined to less than 50% of peak	0
ALT, ALP, or bilirubin (whichever used by R-value criteria) is >90% of peak value at any time >182 days and before any transplant without other explanation recurrent or persistent elevation	-6
Domain 3: Literature supporting liver injury	
LiverTox Category (reference: https://www.ncbi.nlm.nih.gov/books/NBK548392/)	
A or B	3
C, D, or E*	1
E or X	0
Domain 4: Exclusion of competing diagnoses	
Hepatitis A	
Missing anti-HAV IgM data	-3
Anti-HAV IgM negative (if total anti-HAV is negative, consider IgM negative as well)	0
Anti-HAV IgM positive	-6*
Hepatitis B	
Missing anti-HBc IgM (note: [-] anti-HBc <i>total</i> means IgM is negative, but [+] anti-HBc <i>total</i> does <i>not inform IgM result</i>)	-3
HBsAg <i>and</i> anti-HBc IgM negative (if total anti-HBc is negative, consider IgM negative; anti-HBc IgG may be + or -)	0
HBsAg positive and anti-HBc IgM negative (if total anti-HBc is negative, consider IgM negative; anti-HBc IgG may be + or -)	-1
Anti-HBc IgM positive regardless of HBsAg result or HBsAg result missing	-6*

TABLE 1 (Continued)

Domain 4: Exclusion of competing diagnoses	
Hepatitis C	
Missing anti-HCV or HCV RNA	-3
Anti-HCV and HCV RNA both negative	0
Anti-HCV and/or HCV RNA (+) then score according to initial <i>R</i> -value and history:	
<i>R</i> < 5 HCV RNA (-) and anti-HCV (+)	0
<i>R</i> < 5 HCV RNA (+) and anti-HCV (+) or HCV RNA (+) and anti-HCV (-)	-1
<i>R</i> ≥ 5 with known chronic infection	-1
<i>R</i> ≥ 5, no known chronic infection and no exposure risk in ≤100 days before onset	-1
<i>R</i> ≥ 5, no known chronic infection and history of exposure risk in ≤100 days before onset	-6*
HEV (IgM serologies)	
Missing anti-HEV IgM data	-3
Anti-HEV IgM negative	0
Anti-HEV IgM positive	-6*
Alcohol (AST and ALT values at onset)	
AST:ALT ≥ 2 with AST ≤ 500 and missing alcohol history	-3
AST:ALT < 2 and/or AST > 500	0
AST:ALT ≥ 2 with AST ≤ 500 then score according to following alcohol history:	
Average of ≤2 standard drinks/day for women, ≤3 standard drinks/day for men within 6 weeks of injury onset	0
Average of >2 and ≤4 standard drinks/day for women, >3 and ≤6 standard drinks/day for men within 6 weeks of injury onset	-3
Average of >4 standard drinks/day for women, >6 standard drinks/day for men within 6 weeks of injury onset	-6*
Biliary or parenchymal disease assessed by imaging (US, CT, MRI, MRCP, or cholangiogram)	
Missing imaging data	-3
Imaging shows no biliary stenosis(es) or obstruction, no or <50% malignant infiltration of the liver	0
Imaging shows biliary stenosis(es) or obstruction or infiltrating malignancy occupying ≥50% of the liver	-6*
AIH: Use either (a) or (b) as follows:	
(a) AIH assessment for non-minocycline and non-nitrofurantion cases	
Missing ANA and ASMA and IgG	-3
ANA < 1:80, ASMA < 1:80, IgG < 1.1 ULN. Can be missing 1–2 of these, but those obtained must be below these levels	0
ANA ≥ 1:80 or ASMA ≥ 1:80 or IgG ≥ 1.1 ULN	-1
(ANA ≥ 1:80 or ASMA ≥ 1:80) and IgG ≥ 1.1 ULN, and liver biopsy with typical features of AIH	-6*
(b) AIH assessment for minocycline and nitrofurantion cases	
Missing ANA and ASMA and IgG	-3
ANA < 1:80, ASMA < 1:80, IgG < 1.1 ULN. Can be missing 1–2 of these, but those obtained must be below these levels	0
ANA ≥ 1:80 or ASMA ≥ 1:80 or IgG ≥ 1.1 ULN	1
Liver injury due to ischemic liver injury (shock liver) and/or acute congestive hepatopathy^b	
No information on possible hypoxia, hypotension, shock, or acute congestive hepatopathy (history incomplete or inadequate)	-1
No known or suspected prolonged hypoxia, hypotension, shock, or acute congestive hepatopathy within 1 week prior	0
Known or suspected episodes of prolonged hypoxia, hypotension, shock, or acute congestive hepatopathy within 1 week prior	-2
Sepsis causing cholestasis	
No information on sepsis or SIRS, and <i>R</i> -value < 5	-1
<i>R</i> -value < 5 but no sepsis or SIRS, or <i>R</i> -value ≥ 5	0
Sepsis or SIRS present and <i>R</i> -value < 5	-2
Domain 5: Additional data	
The following information may be available in the evaluation, but is not required:	
Retrospective rechallenge: History of DILI with jaundice due to same drug	
No history of prior exposure or no DILI with jaundice after exposure to this drug in the past	0
Positive history of DILI with jaundice after exposure to drug or agent; no documentation by lab results necessary	1

(Continues)

TABLE 1 (Continued)

Domain 5: Additional data	
Prospective rechallenge (documented with labs)	
No rechallenge or no data regarding rechallenge	0
Re-exposure results in rise in ALT, AST 2–3× ULN (or baseline) and/or ALP 1–2× ULN (or baseline)	0
Re-exposure: same R-value category, latency <60 days, ALT, AST > 3× ULN (or baseline) or ALP > 2× ULN (or baseline)	6
Re-exposure to same drug results in no injury with ALT, AST <2× ULN (or baseline), and ALP normal	–3
Liver biopsy	
No biopsy done	0
Nondiagnostic (can be suggestive of DILI, but not diagnostic)	0
Biopsy carries features consistent with a specific DILI	1
Diagnostic of non-DILI diagnosis (e.g., infiltrating cancer, ischemic injury, alcohol-associated hepatitis)	–6*
CMV (IgM = anti-CMV IgM)	
Missing both IgM and PCR	0
Negative (both IgM and PCR negative or at least one negative and other not done)	0
Positive IgM or PCR	–2
Positive IgM and PCR	–6*
EBV (Acute serologies can be any anti-EBV IgM antibody, heterophile test, monospot or EBV early antigen)	
Missing acute serologies and PCR	0
Negative (both acute serologies and PCR negative or at least one negative and other not done)	0
Positive acute serologies or PCR	–2
Positive acute serologies and PCR	–6*
HSV (IgM = anti-HSV IgM)	
Missing IgM and PCR	0
Negative (both IgM and PCR negative or at least one negative and other not done)	0
Positive IgM or PCR	–2
Positive IgM and PCR	–6*
DRESS or Stevens-Johnson syndrome	
Absent or no information	0
Present	1

Notes: Before using RECAM, the user should rule out non-liver-related sources for enzyme elevations (e.g., muscle, hemolysis, bone) and acetaminophen liver injury, for which this tool is not designed.

–6 points: Data entered suggests a DILI is not explanatory of liver injury. User should consider this case as excluded or unlikely DILI with a total score of –6. If user chooses to proceed, 6 points will be deducted from the running score, and user should recognize that DILI as the cause of liver injury is questionable due to inconsistent latency or dechallenge, regardless of total score obtained.

Domain 4, –3 points: When critical data are missing in Domain 4, –3 points are assessed, but user should consider obtaining these data before proceeding.

–6* points: Data entered suggests a non-DILI explanation for liver injury. User should consider the case as excluded DILI with a total score of –6. If user chooses to continue, 6 points will be deducted from the running score, and user should recognize that DILI as sole cause of liver injury is questionable due to a competing explanation, regardless of total sum score obtained.

LiverTox categories of DILI risk: A: well-known, well-described, and characteristic signature (more than 50 well-reported cases in the literature); B: known or highly likely to cause DILI with characteristic signature (12–49 cases in the literature); C: probable cause of DILI (no characteristic signature; less than 12 cases in the literature); D: possible cause of DILI (fewer than three cases in the literature); E: unlikely to cause DILI due to extensive use (cases in the literature may exist but are unconvincing); E*: unproven but suspected to cause DILI (suggestion of liver injury exists outside of published literature, e.g., trial data reported to regulatory agencies); X: unknown (agents recently approved or rarely used). For complete information, go to LiverTox online.^[17]

Abbreviations: AIH, autoimmune hepatitis; ALP, alkaline phosphatase; ALT, alanine aminotransferase; ANA, anti-nuclear antibody; ASMA, anti-smooth muscle antibody; AST, aspartate aminotransferase; CMV, cytomegalovirus; DRESS, drug reaction with eosinophilia and systemic symptoms; EBV, Epstein-Barr virus; HSV, herpes simplex virus; LDH, lactate dehydrogenase; SIRS, systemic inflammatory response syndrome; ULN, upper limit of normal; US, ultrasound.

^aAgents with estimated half-life or pharmacodynamic effect ≥15 days.

^bConsider ischemia or shock when transaminases are extremely high (e.g., >7500 U/l) with elevated LDH and AST > ALT.

for competing diagnoses, because DILI is a diagnosis of exclusion where competing causes should only hurt the case. All diagnoses in this domain should be addressed. At this point, the RECAM will suggest obtaining these data before proceeding. Otherwise, points are taken away for missing such information. Specific

tests and scoring instructions are provided to minimize subjectivity. Viral tests are specified, including HEV antibodies. Evaluation for acute hepatitis C include HCV RNA, history of prior hepatitis C, and risk factors. The RECAM provides scores based on prespecified test results. Consideration for alcoholic hepatitis

TABLE 2 Critical clinical elements for the diagnosis of DILI

Element	Comments
Minimum liver test elevations ^[14]	
ALT ≥ 5× ULN	
ALP ≥ 2× ULN	
ALT > 3× ULN + total bilirubin > 2× ULN	ULN may be replaced by the mean baseline values obtained before exposure to drug if baseline values are abnormal.
Temporal sequence for latency and dechallenge (RECAM Domains 1 and 2)	Consider temporal relationship among drug exposure, injury onset, and improvement.
Competing medications	Obtain thorough pharmacologic history of other drugs that have appropriate temporal relationship among drug exposure, injury onset, and improvement. Consider obtaining a separate RECAM score for these drugs.
Alternative diagnoses (RECAM Domain 4)	
Viral hepatitis A, B, C, and E	For chronic hepatitis B or C, try to establish a baseline and course for liver enzymes, bilirubin, and viral load to help exclude disease exacerbation.
Alcohol-associated hepatitis	Obtain detailed alcohol intake history
Biliary obstruction	Imaging studies needed
AIH	Testing for ANA, ASMA, total IgG
Hypotension due to shock and/or heart failure	Clinical diagnosis
Cholestasis of sepsis	Clinical diagnosis
Malignant infiltration of the liver	Imaging studies needed. Biopsy may be needed.

Abbreviations: ANA, anti-nuclear antibody; ASMA anti-smooth muscle actin; ULN, upper limit of normal.

TABLE 3 Clinical characteristics of 98 DILIN and 96 Spanish DILI Registry cases

Patient Characteristics from DILIN and Spanish DILI Registries		
Characteristic	DILIN (n = 98)	Spanish Registry (n = 96)
Age in years, mean (SD)	48 (18.4)	58 (17.3)
Women	56 (57%)	48 (50%)
Race		
Caucasian	80 (82%)	95 (99%)
Black	9 (9%)	0 (0%)
Asian	4 (4%)	0 (0%)
Other	5 (5%)	1 (1%)
Injury pattern ^a		
Cholestatic	22 (23%)	17 (18%)
Mixed	22 (23%)	21 (22%)
Hepatocellular	51 (54%)	58 (60%)
Likelihood category:		
Definite/highly likely or highly probable	38 (39%)	10 (10%)
Probable	20 (20%)	49 (51%)
Possible	20 (20%)	17 (18%)
Unlikely or excluded	20 (20%)	20 (21%)

^aBased on R-value (ALT/ULN + ALP/ULN). R-value ≥ 5 hepatocellular, 2 < R-value < 5 mixed, R-value ≤ 2 cholestatic.^[1]

diagnosis is prompted by the AST-to-ALT ratio and AST less than 500 U/L. Only if prompted will the user need to enter information about the amount of alcohol

use. Imaging data are clarified with three binary questions based on evidence of pancreaticobiliary disease and cancer infiltration. Autoimmune marker interpretation was aligned more closely with the simplified autoimmune hepatitis (AIH) score^[16] but also scored differently for certain medications known to cause DILI with autoimmune marker positivity.

Criteria 6 (Previous information on hepatotoxicity of the drug)

We moved these criteria to Domain 3, reasoning that most clinicians seek this information early in their consideration of DILI. To increase objectivity and reliability, scoring was anchored to LiverTox Likelihood Categories,^[17] which are loaded into the RECAM. Based on iterative performance testing, the likelihood scores were grouped into three categories of LiverTox Likelihoods (Table 1A), and the RECAM will automatically input the corresponding score after entering the implicated medication. If an agent is not listed in LiverTox (e.g., flucloxacillin), then the user will be given the opportunity to assign a score of 0, 1, or 2 (Table 1).

Criteria 7 (Response to re-administration)

Because rechallenge was so infrequent in both registries and clinical practice, these criteria were incorporated as part of Domain 5 of additional (optional) data (Table 1). We distinguish between a rechallenge

prospectively documented with laboratory testing and a retrospective rechallenge that is elicited in a patient history only and laboratory data may be lacking. We provide specifics on scoring each. Rechallenge is infrequent, but a positive prospective rechallenge is highly indicative of DILI and awarded more points than any other component in the RECAM (+6).

RECAM Domain 5 (additional data)

Besides rechallenge, liver histology, atypical viral testing, and presence of severe skin reactions were included in Domain 5. Liver histology is uncommonly diagnostic of DILI, so points awarded were limited. However, the case is penalized heavily if the biopsy findings yielded an obvious competing diagnosis (Table 1). The presence of severe cutaneous drug reactions adds a point. The presence of non-hepatotropic viral infection, for which testing should be done according to clinical context (e.g., fever, lymphadenopathy, immunocompromise), leads to loss of points.

RECAM warnings and stops

When a firm alternate diagnosis or inconsistent timing for DILI is evident, the user is warned to stop with a -6 final score rendered. The user may override this warning, but -6 points will be deducted from the overall score, and the user should recognize that DILI as sole cause of liver injury is questionable due to a competing explanation or inconsistent timing, regardless of total score obtained (Table 1).

RECAM and RUCAM performance

RECAM went through 12 versions based on iterative testing of cases and meetings. The RUCAM and final version of RECAM scoring was done on 98 DILIN and 96 Spanish DILI cases. Characteristics of each cohort are given in Table 3. Spanish cases were older and had a greater proportion of probable cases. The DILIN had more definite and highly likely cases compared with the Spanish Registry's highly probable cases. Table S2 lists the most common medications implicated.

Both RECAM and RUCAM had similarly high statistical correlation between the resulting scores and the

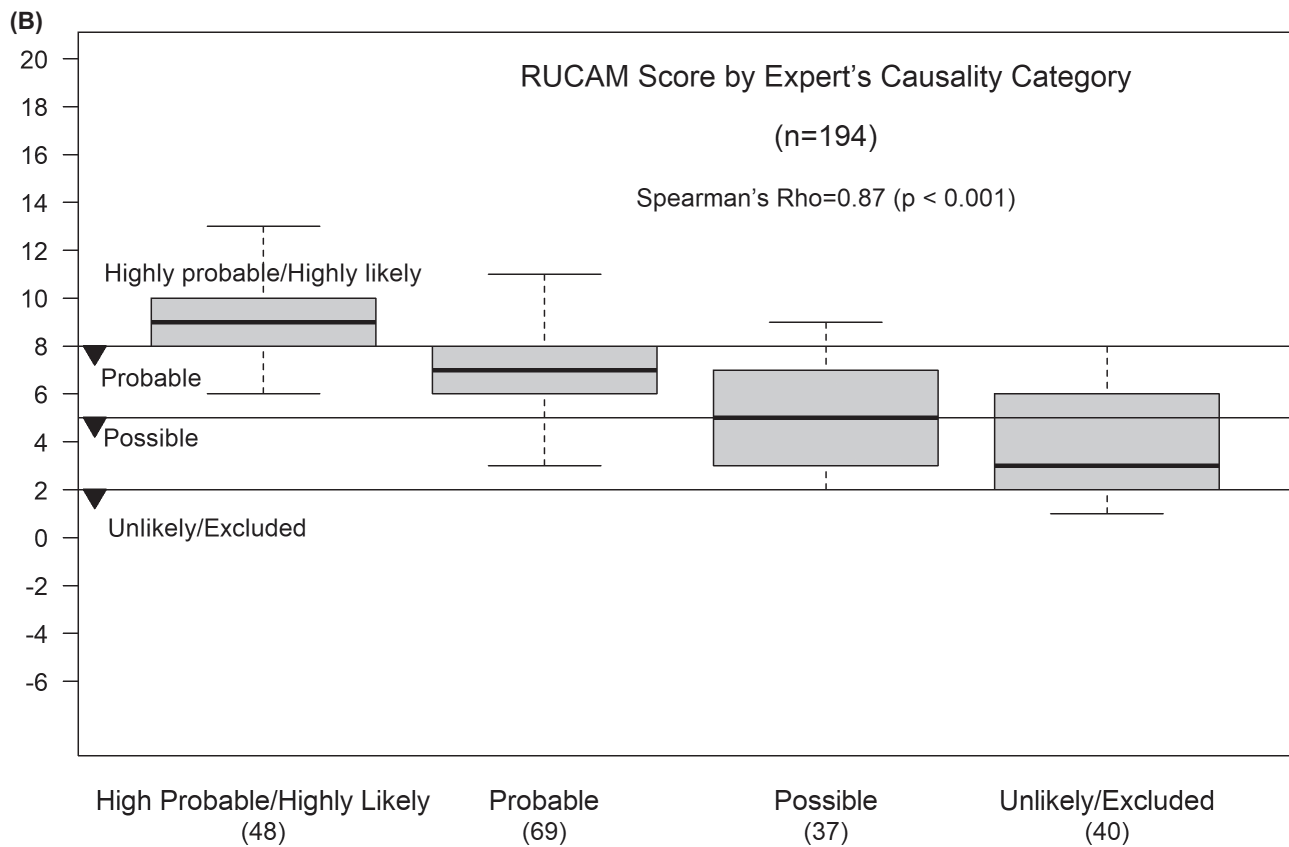
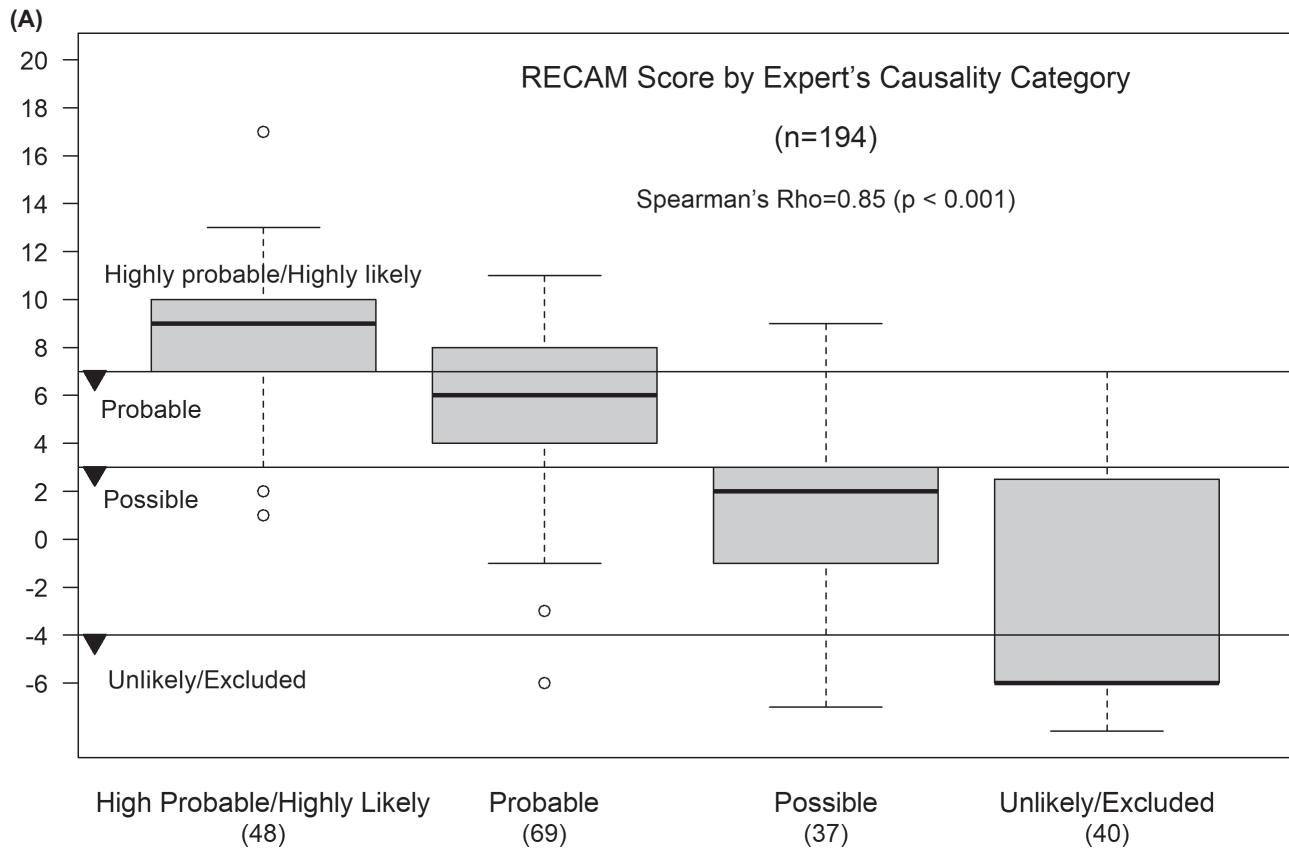
four ordinal diagnostic categories provided by experts (Spearman Rho 0.85; $p < 0.001$ and 0.87, $p < 0.001$, respectively). By using the classification tree approach, we estimated RECAM diagnostic cutoffs of ≥ 8 for highly likely/high probable, 7 to 4 for probable, 3 to -3 for possible, and ≤ -4 for unlikely/excluded DILI, respectively. Classification of combined DILIN and Spanish Registry cases along diagnostic categories using the RECAM and traditional RUCAM cutoffs are shown via boxplots in Figure 1. In a stratified analysis by separate cohorts, the 96 Spanish DILI cases were better classified when using the RECAM compared with the 98 DILIN cases (Figure S1). The AUCs for cumulative cutoffs in the likelihood category for both cohorts combined are given in Table 4. RECAM and RUCAM performed similarly well across all three cutoffs (AUC > 0.8 in all likelihood categories). In a stratified analysis by cohort, the RECAM and the RUCAM scale AUCs showed better performance in Spanish DILI cases compared with DILIN cases. For the Spanish cases, the RECAM AUCs ranged from 0.95 (at least probable cases) to 0.99 (at least possible cases), whereas in DILIN cases the AUCs ranged from 0.80 (at least possible cases) to 0.86 (at least probable cases) (Table S3).

The overall percent agreements between the RECAM and RUCAM scales with expert's opinion were 62% and 59%, respectively ($p = 0.44$). By Cohen's weighted Kappa coefficient, RECAM had better observed overall agreement compared with RUCAM (0.62 vs. 0.56), although statistical significance was not reached ($p = 0.16$) (Table 4). The RECAM had a markedly greater sensitivity for classifying extreme likelihood categories of highly likely/highly probable and unlikely/excluded. Both scales showed similarly good specificity along likelihood categories, except for probable cases, in which the RECAM scale showed better performance (Table 4).

DISCUSSION

This revised electronic causality assessment method, RECAM, provides an evidence-based update of RUCAM. Both RECAM and RUCAM had good diagnostic performance in classifying cases across varying cutoffs in likelihood of DILI based on expert opinion in two large DILI registries. However, RECAM tended to have better observed overall agreement with expert opinion and to better discriminate diagnostic categories especially at the extremes (i.e., highly likely/probable and

FIGURE 1 Box and whisker plots showing median, interquartile, outliers, and Spearman's rho values for revised electronic causality assessment method (RECAM) (A) and Roussel Uclaf Causality Assessment Method (RUCAM) (B) scores by expert opinion diagnostic categories (98 Drug-Induced Liver Injury Network [DILIN] and 96 Spanish Registry cases combined [$n = 194$]). Horizontal lines represent diagnostic score cutoffs for RECAM and RUCAM. Downward pointing arrowheads indicate that the cutoff integer value is included in the category below the line. DILIN categories of definite and highly likely were combined and considered equivalent to the Spanish Registry highly probable category (labeled highly probable/highly likely). Spanish Registry unlikely and excluded categories were combined and considered equivalent to the DILIN unlikely category (labeled -unlikely/excluded).



unlikely). It also had greater specificity to correctly classify probable cases. These differences were likely due to a wider scoring range for latency and dechallenge that

was developed from case data and the heavier penalization for lack of data or data indicating a non-DILI diagnosis. RECAM also offers an automated scoring with less

TABLE 4 Diagnostic performance of RECAM and RUCAM compared with expert opinion for DILIN and Spanish Registry cases combined ($n = 194$)

Performance category	RECAM	RUCAM	<i>p</i> value
AUC (95% CI)			
At least highly likely or highly probable	0.87 (0.81, 0.92)	0.85 (0.80, 0.91)	0.73
At least probable	0.89 (0.84, 0.93)	0.89 (0.84, 0.93)	0.92
At least possible	0.88 (0.81, 0.94)	0.87 (0.81, 0.93)	0.90
Overall agreement (95% CI)			
Percent agreement	62.4 (55.6–69.2)	58.8 (51.8–65.7)	0.44
Weighted kappa	0.62 (0.53, 0.70)	0.56 (0.48, 0.65)	0.16
Sensitivity (95% CI)			
Highly probable, definite, or highly likely	72.9 (60.4–85.5)	54.2 (40.1–68.3)	0.02
Probable	49.3 (37.5–61.1)	68.1 (57.1–79.1)	0.03
Possible	70.3 (55.5–85.0)	59.5 (43.6–75.3)	0.20
Unlikely or excluded	65.0 (50.2–79.8)	47.5 (32.0–63.0)	0.08
Specificity (95% CI)			
Definite, highly likely, or highly probable	86.3 (80.7, 91.9)	89.0 (84.0, 94.1)	0.41
Probable	82.4 (75.7, 89.1)	63.2 (54.8, 71.7)	<0.01
Possible	82.8 (76.9, 88.7)	89.2 (84.3, 94.0)	0.08
Unlikely or excluded	97.4 (94.9, 99.9)	99.4 (98.1, 1.00)	0.18

subjective input, which should lead to better interrater reliability.

Computerization of RECAM (<http://gihep.com/dili-recam/>) is important because RUCAM's poor interrater reliability has limited its adaptation in clinical practice and research. The RECAM categorically scores test results, latency, dechallenge, medication-specific DILI risk, and most competing diagnoses without the need for subjective user opinion or knowledge. The user merely enters the objective data of dates, lab values, and test results. The only subjective information needed for Domains 1 through 4 are the presence of biliary obstruction, >50% malignant liver infiltration on imaging, sepsis, shock, or congestive hepatopathy, as these defined consistent objective parameters for computer entry. Similarly, subjective opinion in Domain 5 is limited to histology and presence of drug reaction with eosinophilia and systemic symptoms or Stevens-Johnson syndrome.

The heterogeneity of DILI phenotypes makes it difficult to develop a single, easy-to-use diagnostic tool for all medications. Thus, the RECAM did not completely mirror expert opinion for a variety of reasons. First, experts rely on knowledge of recent DILI research and emerging phenotypes that can be difficult to translate into algorithmic scoring. Second, some patients had symptoms but delayed seeking medical care, artificially lengthening the latency, or had DILI due to agents with prolonged latencies of months to years (e.g., nitrofurantoin, minocycline). Experts correctly adjusted their opinion of what RECAM considered a latency too long for DILI. Finally, death, transplant, and chronic DILI also prevented receipt of dechallenge points, while experts

accounted for typical cases of fatal or chronic DILI. Inability to capture such clinical factors into the RECAM led to the overall, complete agreement rate of just 62% (59% for RUCAM), but the AUCs of 0.87 to 0.89 across diagnostic category cutoffs are quite good and competitive with other clinical diagnostic tools. For the clinician, the cutoff of at least probable may be most useful when weighing the risks of rechallenge with a highly needed medication or need for further diagnostic evaluation. RECAM's AUC of 0.89 and better ability to separate diagnostic categories (Figure 1) provide a useful framework for such decision making. The improved stratification may better classify cases for genetic (e.g., HLA) and other DILI biomarker development, and increased consistency will make it a better teaching tool.

RECAM's remarkably high AUCs in the Spanish DILI Registry (Table S3) provide some criteria validity, as the Spanish experts rely more on RUCAM for their diagnostic categories. The high performance suggests enough retained similarity to support RECAM's application to that Registry and others currently based on RUCAM. The comparable AUCs for RUCAM and RECAM also confirm that the risk factors of age ≥ 55 , alcohol intake, and pregnancy do not add value to the diagnosis of DILI (Supporting Case 1) and suggests that the 5-domain RECAM without differentiation between hepatocellular and cholestatic/mixed injury is adequate. RECAM's separation of diagnostic categories, especially unlikely and excluded cases, was also better in Spanish cases (Figure S1C), possibly due to the fact that the DILIN often excludes cases that have definitive competing diagnoses before enrollment, whereas the Spanish group retains such cases in their data analyses.

The RECAM has several other notable improvements. The elimination of alternate diagnoses only prevents a loss of points, because ruling out competing etiologies does not directly support a DILI diagnosis in the same way as latency and dechallenge do. The RECAM has automatic warnings for data that are inconsistent with DILI, which is not a part of RUCAM. In the RUCAM, an alternate diagnosis or other data could rule out DILI, but the case would still gain points in other criteria (Supporting Cases 2 and 3). Even when data clearly diagnose acute viral hepatitis or AIH by simplified AIH score,^[16] points are still given for latency, dechallenge, or underlying hepatotoxicity risk of the drug. In these situations of highly implausible DILI, RECAM gives warnings to stop with an imputed total score of -6. One can override these warnings, if one believes DILI may be concurrent with the non-DILI diagnosis. However, -6 points are still assessed. Similarly, warnings to consider stopping or proceeding with a -3 penalty occur when critical data are missing. Such prompts firmly remind the user of tests needed during DILI evaluation. These stops and penalizations led to downward distribution of scores in both registries, particularly unlikely or excluded cases.

The RUCAM assigns a single point for any latency from 5 to 90 days after drug start, while the RECAM has three different scores within the span of 2 to 90 days, regardless the type of liver injury. Gradation of cutoffs was increased for latency times based on latencies in DILIN cases, expert opinion, and iterative testing of cases. This may have led to better identification of highly likely or high probable cases (Supporting Case 3). A pre-assessment DILI risk score (Domain 3) for specific medications is automatically assigned based on LiverTox likelihood score, thus clarifying one of the more ambiguous domains in RUCAM.^[18] These changes also may have helped RECAM better identify more of the highly probable cases.

Incorporating liver histology into a categorical scoring system was challenging. Certain findings may be quite consistent with a specific DILI episode (e.g., ring granulomas with allopurinol liver injury), but we felt even these readings are open to interpretation and need clinical context. Thus, only 1 point is awarded for histologic findings, but histology can hurt the case for DILI when a clear alternate diagnosis is found, like infiltrating cancer or ischemic injury. In these cases, a heavy penalty of -6 and warning are given. In both registries, liver biopsy was often not obtained, and pathognomonic signs of DILI or alternate diagnosis were even less common. Therefore, the impact of histology on RECAM performance was minimal. Nevertheless, the computer program used to develop the RECAM will allow us to adjust this variable as more data on how histology influences the diagnosis of DILI become available.^[19]

The RECAM also has several important limitations. It was developed in US and Spanish cohorts, so we do not know how it may perform in other regions, particularly

Asia. Also, both registries have minimum enrollment criteria for liver enzyme and bilirubin elevation, so it is unclear how the RECAM may perform in less severe cases.^[8,10] The RECAM also needs testing by a broader group of clinicians including non-hepatologists. It is currently limited to single-agent medication cases, leaving the user to score each medication individually in multi-drug cases. However, any competing medication causing loss of points in the RUCAM probably deserves its own RECAM score. The RECAM is also not designed nor tested for HDS liver injury, which is increasingly reported.^[20–22] While simplified with fewer Domains and clearer operating instructions, the web application increases the amount of data entry compared with the RUCAM. Yet, we believe the increased data entry will be offset by automated latency and dechallenge calculations by the computer. Moreover, users no longer need to render a subjective opinion on competing diagnoses. They simply choose test results regarding competing diagnoses from short dropdown menus. The RECAM retains a few parameters that need clinical judgment. Whether a biliary stricture is clinically insignificant is still left up to the user. Drugs not included in LiverTox must still be scored by opinion of labeling and available literature. Finally, the RECAM will need updating as DILI epidemiology and research evolve. For example, the cutoff of >90-day latency garnering 0 points was based on a broad range of cases with most having shorter latencies, but as longer-acting medications (e.g., monoclonal check point inhibitors) grow in use and latencies increase, this cutoff may need adjustment. Pharmacogenomic data and new biomarkers may also need to be incorporated with the computerization of RECAM, lending itself well to such modifications.

RUCAM has been a valuable clinical framework for DILI diagnosis since 1993. However, user subjectivity made it unreliable, and it was overdue for an evidence-based update. RECAM has better sensitivities at the extreme diagnostic categories and tends to have better overall agreement with expert opinion. It will likely have better interrater and intrarater reliability due to computerized categorical, data entry, and minimized subjective opinion. The RECAM also eliminates unnecessary variables that were not diagnostically helpful. Domains are based on data from well-vetted cases that were often followed for a minimum of 6 months. Accuracy of 80%–90% for identifying at least probable DILI compared with expert opinion is high, but not high enough to make the RECAM a standalone diagnostic tool. For now, nothing can replace good history taking, chart review, and thorough evaluation for competing causes. There will always be cases that defy proper scoring by any single algorithm that seeks to account for the extensive heterogeneity in DILI phenotypes and presentation (e.g., very long latency DILI, chronic DILI). Therefore, further refinement and validation are anticipated.

Indeed, the RECAM provides an opportunity to conduct causality assessment using standardized, quantitative, and categorical data fields, which should lead to improved case identification, earlier diagnosis, and medical management. The electronic, automated platform of the RECAM that is available for all to use on the Internet should also help with efforts at harmonization and standardization in DILI research.

ACKNOWLEDGMENTS

This publication is partially based on work from COSTAction “CA17112—Prospective European Drug-Induced Liver Injury Network” supported by COST (European Cooperation in Science and Technology) www.cost.eu. MIL, ESB, GPA, AGJ, and RJA are members of COST Action CA17112. The authors thank Dr. Lauryna Aukštikalnė for her help in refining the online implementation of the RECAM. They also acknowledge the American Association for the Study of Liver Disease (AASLD) for supporting this research through the AASLD Innovations Fund (2016).

CONFLICT OF INTEREST

Dr. Fontana received grants from AbbVie and Gilead. Dr. Aithal consults for Pfizer and GlaxoSmithKline. Dr. Hayashi is employed by the US Food and Drug Administration (FDA), but conclusions in this paper do not reflect any opinions of the FDA. The FDA has not evaluated the findings, assessments or conclusions in this paper.

AUTHOR CONTRIBUTIONS

Paul H. Hayashi, M. Isabel Lucena, Robert J. Fontana, Guruprasad P. Aithal, Einar S. Bjornsson, Raul J. Andrade, and Jay H. Hoofnagle: substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; Paul H. Hayashi, M. Isabel Lucena, Robert J. Fontana, Guruprasad P. Aithal, Einar S. Bjornsson, Raul J. Andrade, Jay H. Hoofnagle, Huiman Barnhart, Andres Gonzalez-Jimenez, Qinghong Yang, and Jiezhun Gu: drafting the article or revising it critically for important intellectual content; Paul H. Hayashi, M. Isabel Lucena, Robert J. Fontana, Guruprasad P. Aithal, Einar S. Bjornsson, Raul J. Andrade, Jay H. Hoofnagle, Huiman Barnhart, Andres Gonzalez-Jimenez, Qinghong Yang, and Jiezhun Gu: final approval of the version to be published.

ORCID

Paul H. Hayashi  <https://orcid.org/0000-0002-5371-1020>

REFERENCES

- Danan G, Benichou C. Causality assessment of adverse reactions to drugs. 1: A novel method based on the conclusions of international consensus meetings: application to drug-induced liver injuries. *J Clin Epidemiol*. 1993;46:1323–30.
- Chalasanani NP, Maddur H, Russo MW, Wong RJ, Reddy KR, et al. ACG clinical guideline: diagnosis and management of idiosyncratic drug-induced liver injury. *Am J Gastroenterol*. 2021;116:878–98.
- Andrade RJ, Aithal GP, Björnsson ES, Kaplowitz N, Kullak-Ublick GA, Larrey D, et al. EASL clinical practice guidelines: drug-induced liver injury. *J Hepatol*. 2019;70:1222–61.
- Kaplowitz N. Causality assessment versus guilt-by-association in drug hepatotoxicity. *Hepatology*. 2001;33:308–10.
- Hayashi PH. Causality assessment in drug-induced liver injury. *Semin Liver Dis*. 2009;29:348–56.
- García-Cortés M, Stephens C, Lucena MI, Fernández-Castañer A, Andrade RJ, et al. Causality assessment methods in drug induced liver injury: strengths and weaknesses. *J Hepatol*. 2011;55:683–91.
- Danan G, Teschke R. RUCAM in drug and herb induced liver injury: the update. *Int J Mol Sci*. 2015;17:14.
- Fontana RJ, Watkins PB, Bonkovsky HL, Chalasanani N, Davern T, Serrano J, et al. Drug-Induced Liver Injury Network (DILIN) prospective study: rationale, design and conduct. *Drug Saf*. 2009;32:55–68.
- Andrade R, Lucena M, Fernandez M, Pelaez G, Pachkoria K, Garciaaruiz E, et al. Drug-induced liver injury: an analysis of 461 incidences submitted to the Spanish registry over a 10-year period. *Gastroenterology*. 2005;129:512–21. <https://doi.org/10.1016/j.gastro.2005.05.006>
- Stephens C, Robles-Diaz M, Medina-Caliz I, Garcia-Cortes M, Ortega-Alonso A, Sanabria-Cabrera J, et al. Comprehensive analysis and insights gained from long-term experience of the Spanish DILI registry. *J Hepatol*. 2021;75:86–97.
- Breiman LF, Olshen JH, Stone RA, Friedman CJ. *Classification and Regression Trees*. Belmont, California: Wadsworth; 1984.
- Lucena MI, Andrade RJ, Kaplowitz N, García-Cortés M, Fernández MC, Romero-Gomez M, et al. Phenotypic characterization of idiosyncratic drug-induced liver injury: the influence of age and sex. *Hepatology*. 2009;49:2001–9. Dr. Hayashi is employed by the US Food and Drug Administration (FDA), but conclusions in this paper do not reflect any opinions of the FDA. The FDA has not evaluated the findings, assessments or conclusions in this paper.
- Dakhoul L, Ghabril M, Gu J, Navarro V, Chalasanani N, Serrano J, et al. Heavy consumption of alcohol is not associated with worse outcomes in patients with idiosyncratic drug-induced liver injury compared to non-drinkers. *Clin Gastroenterol Hepatol*. 2018;16:722–9.e2.
- Aithal GP, Watkins PB, Andrade RJ, Larrey D, Molokhia M, Takikawa H, et al. Case definition and phenotype standardization in drug-induced liver injury. *Clin Pharmacol Ther*. 2011;89:806–15.
- Agarwal VK, McHutchison JG, Hoofnagle JH. Important elements for the diagnosis of drug-induced liver injury. *Clin Gastroenterol Hepatol*. 2010;8:463–70.
- Hennes EM, Zeniya M, Czaja AJ, Parés A, Dalekos GN, Krawitt EL, et al. Simplified criteria for the diagnosis of autoimmune hepatitis. *Hepatology*. 2008;48:169–76.
- LiverTox. Clinical and Research Information on Drug-Induced Liver Injury [Internet]. Categorization Of The Likelihood Of Drug Induced Liver Injury. Bethesda, Maryland: National Institute of Diabetes and Digestive and Kidney Diseases. 2012. <https://www.ncbi.nlm.nih.gov/books/NBK547852>. [Accessed October 30, 2020]
- Bjornsson ES, Hoofnagle JH. Categorization of drugs implicated in causing liver injury: critical assessment based on published case reports. *Hepatology*. 2016;63:590–603.
- Ahmad J, Barnhart HX, Bonacini M, Ghabril MS, Hayashi PH, Odin J, et al. Drug-Induced Liver Injury Network Investigators. Value of liver biopsy in the diagnosis of drug-induced liver

- injury. *J Hepatol*. 2022 Jan 21. <https://doi.org/10.1016/j.jhep.2021.12.043>. [Epub ahead of print].
20. Navarro VJ, Barnhart H, Bonkovsky HL, Davern T, Fontana RJ, Grant L, et al. Liver injury from herbals and dietary supplements in the U.S. Drug-Induced Liver Injury Network. *Hepatology*. 2014;60:1399–408.
 21. Medina-Caliz I, Garcia-Cortes M, Gonzalez-Jimenez A, Cabello MR, Robles-Diaz M, Sanabria-Cabrera J, et al. Herbal and dietary supplement-induced liver injuries in the Spanish DILI Registry. *Clin Gastroenterol Hepatol*. 2018;16:1495–502.
 22. Shen T, Liu Y, Shang J, Xie Q, Li J, Yan M, et al. Incidence and etiology of drug-induced liver injury in mainland China. *Gastroenterology*. 2019;156:2230–41.e11.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Hayashi PH, Lucena MI, Fontana RJ, Bjornsson ES, Aithal GP, Barnhart H, et al. A revised electronic version of RUCAM for the diagnosis of DILI. *Hepatology*. 2022;76:18–31. <https://doi.org/10.1002/hep.32327>