Expert Judgment for Causality Assessment

Paul H. Hayashi, MD, MPH
Associate Professor
University of North Carolina, Chapel Hill, NC
June 3, 2017
Outline

• DILIN Expert Opinion
  • What is it exactly?
  • How reliable is it?
  • How does it compare to RUCAM?

• Using DILIN Expert Opinion Cases toward a better diagnostic tool
  • DILI CAT 1.0
  • RUCAM revision
Toward better causality assessment

• Maturing registries around the globe
• “Big data” from electronic medical records
• New inroads and technologies in biomarkers for diagnosis and risk
• We have some important, basic building blocks
  • RUCAM
  • Expert opinion process
  • LiverTox
  • Histology
Expert Opinion
US Drug-Induced Liver Injury Network (DILIN)

1. University of Southern California/University of California, Los Angeles, Los Angeles, CA
2. Indiana University, Indianapolis, IN
3. University of Michigan, Ann Arbor, MI
4. University of North Carolina, Chapel Hill, NC
5. Einstein Medical Center, Philadelphia, PA
6. Mount Sinai Hospital, New York, NY
7. Duke University, Data Coordinating Center, Durham, NC
8. Liver Disease Research Branch, NIDDK, Bethesda, MD

Hayashi PH. IJMS 2016
DILIN Expert Consensus Opinion

• 7 enrolling center across US & 1 Data Coordinating Center
• Expert opinion process done after 6 months of follow-up obtained.
• Case scored overall and for each agent relative to each other.

<table>
<thead>
<tr>
<th>Score – Category</th>
<th>Likelihood</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – Definite</td>
<td>&gt;95%</td>
<td>Beyond reasonable doubt</td>
</tr>
<tr>
<td>2 – Highly likely</td>
<td>75-95%</td>
<td>Clear and convincing, but less than definite</td>
</tr>
<tr>
<td>3 – Probable</td>
<td>50-74%</td>
<td>Preponderance of evidence supports</td>
</tr>
<tr>
<td>4 – Possible</td>
<td>25-49%</td>
<td>Not supported by preponderance of evidence but cannot exclude</td>
</tr>
<tr>
<td>5 – Unlikely</td>
<td>&lt;25%</td>
<td>Highly unlikely based on evidence</td>
</tr>
</tbody>
</table>
DILIN
Expert Consensus Opinion Process
Hayashi PH. IJMS 2016
DILIN Expert Consensus Opinion: Reliability
Hayashi, PH et al. Liver Int 2014

• 100 cases randomly chosen from DILIN registry
• Re-adjudicated by 3 DILIN investigators new to each case.
  • Inter-raters, test-retest.
• Median time between adjudications: 938 days (140-2352)
DILIN Expert Consensus Opinion: Reliability

Hayashi, PH et al. Liver Int 2014

• Weighted Kappas
  • 0.60 (95% CI 0.5 – 0.7) for overall score and individual agent scores

• 14% crossed probable (≥50% likelihood) vs. possible (<50%) line
  • Most due to unconvincing timing and/or competing diagnosis

Adjudicated with 6 month follow-up
Expert Consensus Opinion vs. RUCAM
Rockey DC, et al. Hepatology 2010

• 187 single agent cases randomly taken from the DILIN registry
• All cases given expert opinion score and RUCAM score by 3 investigators
• Modest correlation across 5 category scales.
  • Spearman $R = 0.42$ ($p < 0.01$)
# Expert Consensus Opinion vs. RUCAM

Rockey DC, et al. Hepatology 2010

<table>
<thead>
<tr>
<th>Expert Opinion</th>
<th>RUCAM Category</th>
<th>95% PPV</th>
<th>23% NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Highly Probable</td>
<td>Probable</td>
<td>Possible</td>
</tr>
<tr>
<td>Definite</td>
<td>80</td>
<td>72</td>
<td>41</td>
</tr>
<tr>
<td>Highly Likely</td>
<td>38</td>
<td>100</td>
<td>70</td>
</tr>
<tr>
<td>Probable</td>
<td>10</td>
<td>35</td>
<td>33</td>
</tr>
<tr>
<td>Possible</td>
<td>2</td>
<td>10</td>
<td>18</td>
</tr>
<tr>
<td>Unlikely</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>132</strong></td>
<td><strong>220</strong></td>
<td><strong>167</strong></td>
</tr>
</tbody>
</table>

Discordant 31% of the time
## Going forward

### Goals for a new diagnostic method

- Reliable
  - Clear, simple and objective
  - Computer based
- Valid
  - Leverage registry data to determine criteria & weighting.
  - Leverage LiverTox data.
  - Test across registries, clinicians and investigators

### Challenges

- Diversity in DILI signatures
- Herbals & Dietary Supplements
- Lack of a true gold standard
Computerized Causality Instruments

Improving Causality Assessment in DILI

Tillmann HL, Barnhart HX, Serrano J, Rockey DC.
A Novel Computerized Drug Induced Liver Injury Causality Assessment Tool
(DILI-CAT)
Hepatology 64: 320A (2016)
DILI-Causality Assessment Tool (CAT)

- Clinical Signature: Point allocation

<table>
<thead>
<tr>
<th>Characteristic for particular agent</th>
<th>Maximum points obtainable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latency from start to onset</td>
<td>30</td>
</tr>
<tr>
<td>Latency from stop to onset</td>
<td>10</td>
</tr>
<tr>
<td>Dechallenge (washout)</td>
<td>20</td>
</tr>
<tr>
<td>R-value</td>
<td>20</td>
</tr>
<tr>
<td>History of allergy to medications</td>
<td>7</td>
</tr>
<tr>
<td>Rash</td>
<td>3</td>
</tr>
<tr>
<td>Fever</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>91</strong></td>
</tr>
</tbody>
</table>

Normalized to a percentage likelihood of 55%
Core interval concept for continuous variables

The Core Interval Concept. Shown are examples of how the core interval process is conceptualized. In (A) is shown an “ideal” normative distribution, with 50% of the core highlighted in blue. In (B), is shown the actual distribution of latencies for INH. Using a 50% core interval iteration and a eventual negative point scale for cases outside the range on definite/highly likely cases, the points changed to:

- 30 points for 54-104 days.
- 20 points for 48-54 or 104-162 days
- 10 points for 33-48 or 162-264 days

0 points for 4-33 or 264-272 days
-10 points for <4 or >272 days (outside the range for definitive/highly likely or probable cases).
DILI-CAT

- Other features make up the remaining 45 percentage points.
  - 20 percentage points: Prior literature (hepatotoxicity risk)
    - Based on LiverTox grading
  - 25 percentage points: Competing causes ruled out/

- 100% likelihood maximum score
  - 55% clinical signature score (normalized)
  - 20% hepatotoxicity risk
  - 25% competing causes evaluation
DILI-CAT for Isoniazid
Clinical presentation scoring already normalized to 55

<table>
<thead>
<tr>
<th>Definite</th>
<th>Highly likely</th>
<th>Probable</th>
<th>Possible</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

![Graph showing clinical presentation scoring for Isoniazid](image-url)
DILI-CAT for Isoniazid
Signature score + Hepatotoxicity score for INH
DILI-CAT for Isoniazid
Signature score + Hepatotoxicity + Deductions for competing causes
DILI-CAT for Isoniazid
Signature + Hepatotoxicity + Deductions + Competing causes excluded
DILI-CAT 1.0 compared to RUCAM - INH

**DILI-CAT**

- P=0.0076
- AUC=0.77 for Definite/Highly Likely/Probable
- AUC=0.62 for Definite/Highly Likely

**RUCAM**

- AUC=0.67 for Definite/Highly Likely/Probable
- AUC=0.62 for Definite/Highly Likely

Correlation between Rucam and DILI-CAT score: r=0.65 (P=0.74)
## DILI-CAT 1.0

### AUC for predicting causality for 8 drugs

#### AUC definite/highly likely

<table>
<thead>
<tr>
<th></th>
<th>Augmentin</th>
<th>INH</th>
<th>Bactrim</th>
<th>Nitrofurantoin</th>
<th>Minocycline</th>
<th>Cefazolin</th>
<th>Azithromycin</th>
<th>Ciprofloxacin</th>
</tr>
</thead>
<tbody>
<tr>
<td>70% core</td>
<td>0.68</td>
<td>0.88</td>
<td>0.67</td>
<td>0.55</td>
<td>0.67</td>
<td>0.7</td>
<td>0.5</td>
<td>0.65</td>
</tr>
<tr>
<td>50% core</td>
<td>0.75</td>
<td><strong>0.92</strong></td>
<td>0.7</td>
<td>0.53</td>
<td><strong>0.67</strong></td>
<td>0.7</td>
<td>0.45</td>
<td><strong>0.69</strong></td>
</tr>
<tr>
<td>RUCAM</td>
<td><strong>0.77</strong></td>
<td>0.62</td>
<td><strong>0.71</strong></td>
<td><strong>0.64</strong></td>
<td>0.56</td>
<td><strong>0.69</strong></td>
<td><strong>0.87</strong></td>
<td>0.65</td>
</tr>
</tbody>
</table>

#### AUC definite/highly likely/probably

<table>
<thead>
<tr>
<th></th>
<th>Augmentin</th>
<th>INH</th>
<th>Bactrim</th>
<th>Nitrofurantoin</th>
<th>Minocycline</th>
<th>Cefazolin</th>
<th>Azithromycin</th>
<th>Ciprofloxacin</th>
</tr>
</thead>
<tbody>
<tr>
<td>70% core</td>
<td>0.78</td>
<td>0.74</td>
<td>0.63</td>
<td><strong>0.77</strong></td>
<td><strong>0.84</strong></td>
<td>0.75</td>
<td>0.87</td>
<td>0.78</td>
</tr>
<tr>
<td>50% core</td>
<td><strong>0.84</strong></td>
<td><strong>0.77</strong></td>
<td><strong>0.77</strong></td>
<td>0.68</td>
<td><strong>0.95</strong></td>
<td><strong>0.77</strong></td>
<td><strong>0.9</strong></td>
<td>0.78</td>
</tr>
<tr>
<td>RUCAM</td>
<td>0.8</td>
<td>0.67</td>
<td>0.75</td>
<td>0.43</td>
<td>0.69</td>
<td>0.7</td>
<td>0.8</td>
<td><strong>0.79</strong></td>
</tr>
</tbody>
</table>
Revising the RUCAM
International Working Group

**USA**
- Robert J. Fontana, MD
- Paul H. Hayashi, MD, MPH
- Jay H. Hoofnagle, MD

**Europe**
- Guru Aithal, MD
- Raul Andrade, MD
- Einar Bjornsson, MD
- Isabel Lucena, MD

**Statistician**
- Huiman Barnhart, PhD
Goals and Methods

Goals
• To revise the current RUCAM into an easier to use clinical and research tool with clear operation instructions and in a computerized application platform
• Keep the tool generalizable to all agents.

Methods
• Expert opinion to revise parameters and scores.
• Wordsmith instructions for maximal clarity
• Use of DILIN cases to beta test.
### Domain 1: Latency (Maximum/minimum points: 4/3)

**Score both Onset after drug start and Onset after drug stop**

<table>
<thead>
<tr>
<th>Days after drug start where day 1 is first day drug taken</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 to 7</td>
<td>2</td>
</tr>
<tr>
<td>8 to 90</td>
<td>4</td>
</tr>
<tr>
<td>91 to 182</td>
<td>3</td>
</tr>
<tr>
<td>183 to 365</td>
<td>2</td>
</tr>
<tr>
<td>&gt; 365</td>
<td>1</td>
</tr>
</tbody>
</table>

**Onset after drug stop (points taken)**

<table>
<thead>
<tr>
<th>Days after drug stop where day 1 is the first day the drug is not taken</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 to 7</td>
<td>0</td>
</tr>
<tr>
<td>8 to 14</td>
<td>-1</td>
</tr>
<tr>
<td>15 to 30</td>
<td>-2</td>
</tr>
<tr>
<td>31 to 60</td>
<td>-3</td>
</tr>
<tr>
<td>&gt; 60</td>
<td>-4</td>
</tr>
</tbody>
</table>

### Domain 2: Dechallenge or Washout (Maximum/minimum points: 4/0)

**Score either Hepatocellular or Mix/Cholestatic based on R-ratio at onset**

**Hepatocellular (initial R ratio > 5)**

<table>
<thead>
<tr>
<th>Days from peak ALT to &lt; 50% of peak; max 4/min 0</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 to 30</td>
<td>4</td>
</tr>
<tr>
<td>31 to 90</td>
<td>3</td>
</tr>
<tr>
<td>91 to 182</td>
<td>2</td>
</tr>
<tr>
<td>183 to 365</td>
<td>1</td>
</tr>
<tr>
<td>&gt; 365</td>
<td>0</td>
</tr>
</tbody>
</table>

**Mixed/cholestatic (initial R ratio ≤ 5)**

<table>
<thead>
<tr>
<th>Days from peak Alkaline Phosphatase to &lt; 50% of peak</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 to 30</td>
<td>4</td>
</tr>
<tr>
<td>31 to 182</td>
<td>3</td>
</tr>
<tr>
<td>183 to 365</td>
<td>1</td>
</tr>
<tr>
<td>&gt;365</td>
<td>0</td>
</tr>
</tbody>
</table>
### Domain 3: Literature supporting liver injury (Maximum/minimum points: 4/0)

<table>
<thead>
<tr>
<th>LiverTox Category</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>4</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
</tr>
<tr>
<td>E</td>
<td>0</td>
</tr>
</tbody>
</table>

### Domain 4: Exclusion of competing diagnoses* (Maximum/minimum points: 8/0)

| Hepatitis A: IgM anti-HAV negative (if total anti-HAV is negative, consider IgM negative) | 1 |
| Hepatitis B: HBsAg and IgM anti-HBc negative (if total anti-HBc is negative, consider IgM negative) | 1 |
| Hepatitis C: H-anti-HCV and HCV RNA negative | 1 |
| Alcohol: <2 drinks/d for women, <3/d for men (within 6 weeks of injury onset); AST:ALT < ?—Look at transaminase ratio within 3-7 days of injury | 1 |
| Biliary disease: Imaging shows no evidence of biliary dilation or obstruction (if no imaging then 0) | 1 |
| Autoimmune hepatitis: ANA <1:80 and ASMA <1:80 and, if available, IgG ≤ 1.1 ULN | 1 |
| Ischemic hepatitis: No shock, ischemia or prolonged hypotension within 1 week prior to injury | 1 |
| CMV and EBV: Negative serologies (or PCR testing) for acute infection for both viruses | 1 |

*One may consider stopping the RUCAM and pursuing any of the above competing diagnoses should exclusion criteria not be met.

### Domain 5: Additional data (Maximum/minimum points: 0/6)

**Points**

The following information may be obtained, but none are necessary to complete the RUCAM.

**Rechallenge**

- Positive without obvious competing diagnoses: 3
- Negative rechallenge: -2

**Liver biopsy**

- Consistent with DILI: 1
- Suggestive or compatible with non-DILI diagnosis (e.g., ischemic injury, alcoholic hepatitis): -1 **

**Hepatitis E virus IgM (Should this be given more priority?)**

- Negative: 1

**Consider that DILI may be ruled out regardless of RUCAM scoring (i.e., consider stopping RUCAM)**

### Summary

**Domains 1 through 4: Maximum to minimum score:** 26 to -4

**Domains 1 through 5: Maximum to minimum score:** 25 to -7
Issues and next steps.

• HSV, CMV, EBV.
• Alcohol criteria
• Single drug tool---no competing medications/agents
• Scoring DILIN cases for beta testing and revision.
Push and Pull

Revised RUCAM → DILIN Registry → DILI-CAT(s)

Reliability ← Validity
End game for any new diagnostic tool

Reliability
• Computerized
• Beta testing
  • DILI hepatologists
  • Hepatologists
  • Generalists

Validation
• External validation
  • Tested on non-DILIN registry cases.
• True validation
  • Await biomarkers
  • “Unlikely” cases
  • SAEs not attributable to agent
Toward better causality assessment

- Maturing registries around the globe
- “Big data” from electronic medical records
- New inroads and technologies in biomarkers for diagnosis and risk
- We have some important, basic building blocks
  - RUCAM
  - Expert opinion process
  - LiverTox
  - Histology