DATA EXTRACTION AND QUALITY ASSESSMENT

Craig Campbell

Hierarchy of Models for setting Analytical Quality

- 1999 Stockholm Conference: *Strategies to set Global Quality Specifications in Laboratory Medicine*
- Over 100 participants from 27 countries
- Main outcome - agreement that the hierarchy be applied to set analytical quality specifications.

Hierarchy of models for setting analytical quality specifications

1. Evaluation of the effect of analytical performance on clinical outcomes in specific clinical settings
2. Evaluation of the effect of analytical performance on clinical decisions in general:
   a. Data based on components of biological variation
   b. Data based on analysis of clinicians' opinions
3. Published professional recommendations
   a. From national and international expert bodies
   b. From expert local groups or individuals
4. Performance goals set by
   a. Regulatory bodies
   b. Organizers of External Quality Assessment (EQA) schemes
5. Goals based on the current state of the art
   a. As demonstrated by data from EQA or Proficiency Testing scheme
   b. As found in current publications on methodology.

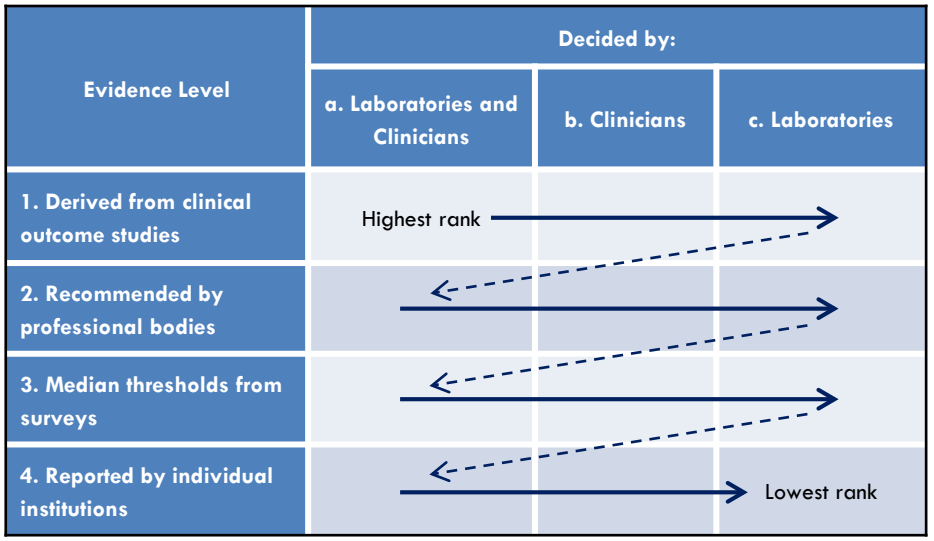
The Stockholm Hierarchy applied to reference intervals and clinical decision limits.

1. Clinical decision limit based on clinical outcome study
   e.g. HbA_{1c} cut-off based on the presence of diabetes outcome (retinopathy).
2. Other methods of determining reference interval or clinical decision limit
   a. Reference intervals derived from apparently healthy populations e.g. NORIP, CALIPER.
   b. Clinical decision limits based on clinicians' opinions of disease e.g. thyroid-stimulating hormone (TSH) upper reference limit (2.5 mIU/L) from NACB.
3. Published professional recommendations
   a. National or international expert bodies e.g. national urine protein cut-offs.
   b. Expert local groups or individuals e.g. ARQAG, SONIC.
4. Reference limits set by
   a. Regulatory bodies e.g. prostate-specific antigen (PSA) cut-offs.
   b. Formal Reference Interval Survey e.g. UK Harmony Survey.
5. Referance interval based on the current state of the art
   a. Reference interval used in postanalytical external quality assurance e.g. pathology interpretation exercises.
   b. Current publications on methodology e.g. textbooks or kit inserts.

- 15 years on: quality expectations were found to be too ambitious considering performance of most tests.
- In 2012, Ken Sakaris described how the Stockholm hierarchy could be applied to setting clinical decision limits.

Alert Threshold Evidence Ranking System

- 2016: Two-dimensional alert threshold evidence ranking system created, inspired by Ken's application of the hierarchy.
- 1st Dimension: Four evidence levels based on **where** the threshold evidence was sourced from.
- 2nd Dimension: Each evidence level has three subcategories, based on **who** the thresholds were derived by.



The Stockholm Hierarchy applied to reference intervals and clinical decision limits.

1. Clinical decision limit based on clinical outcome study
   e.g. HbA_{1c} cut-off based on the presence of diabetes outcome (retinopathy).
2. Other methods of determining reference interval or clinical decision limit
   a. Reference intervals derived from apparently healthy populations e.g. NORIP, CALIPER.
   b. Clinical decision limits based on clinicians' opinions of disease e.g. thyroid-stimulating hormone (TSH) upper reference limit (2.5 mIU/L) from NACB.
3. Published professional recommendations
   a. National or international expert bodies e.g. national urine protein cut-offs.
   b. Expert local groups or individuals e.g. ARQAG, SONIC.
4. Reference limits set by
   a. Regulatory bodies e.g. prostate-specific antigen (PSA) cut-offs.
   b. Formal Reference Interval Survey e.g. UK Harmony Survey.
5. Referance interval based on the current state of the art
   a. Reference interval used in postanalytical external quality assurance e.g. pathology interpretation exercises.
   b. Current publications on methodology e.g. textbooks or kit inserts.

Why do we need test specific literature reviews?

- A systematic review of alert threshold evidence has already been done.
- However, the search terms were exclusively laboratory terminology.

Clinical Chemistry 62:11
1445–1457 (2016)

Review

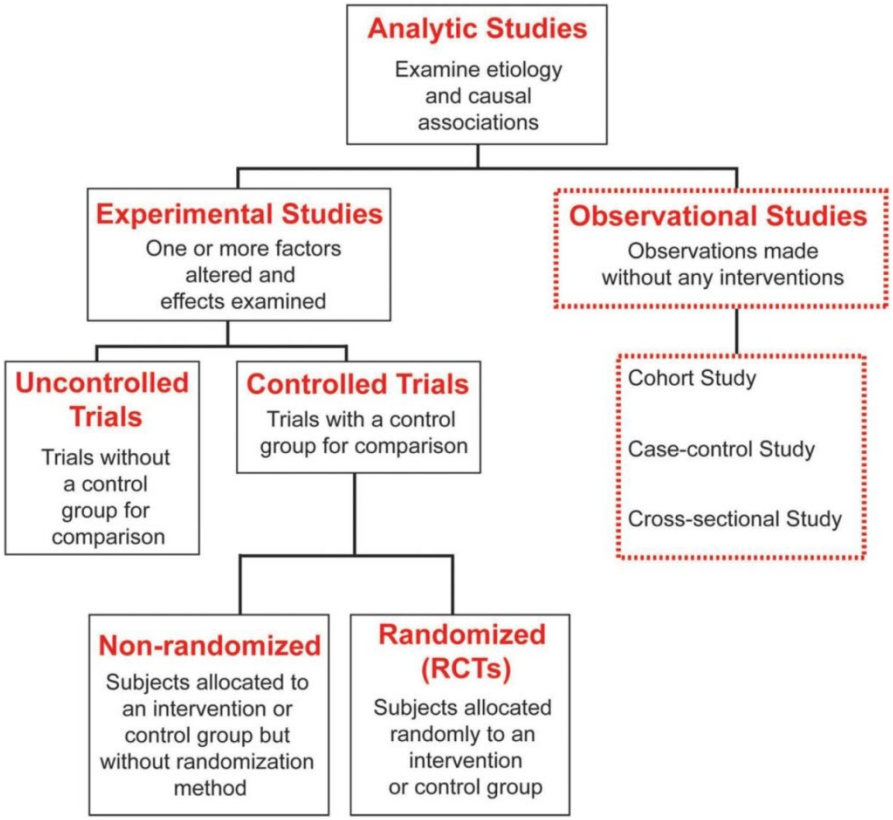
What Alert Thresholds Should Be Used to Identify Critical Risk Results: A Systematic Review of the Evidence

Craig A. Campbell,[1,2*] Andrew Georgiou,[1] Johanna I. Westbrook,[1] and Andrea R. Horvath[2]

- We (the RCPA-AACB working party) believe there may be relevant published outcome studies that do not use laboratory terminology.
- Test-specific literature searches can target specific outcome measures.

Critical Appraisal of Outcome Studies

Types of Outcome Studies



- Studies that measure the harm associated with high risk results can only be observational.
- Retrospective studies are quicker and cheaper than prospective studies.
- The quality of a study is dependent on the rigour of it's design.

Critical Appraisal of Outcome Studies

JAMA, May 25, 1994—Vol 271, No. 20

The Medical Literature

Users' Guides to the Medical Literature

IV. How to Use an Article About Harm

Mitchell Levine, MD, MSc; Stephen Walter, PhD; Hui Lee, MD, MSc; Ted Haines, MD, MSc;
Anne Holbrook, MD, PharmD, MSc; Virginia Moyer, MD, MPH; for the Evidence-Based Medicine Working Group

Are the results of the study valid?

Primary guides:

- Were there clearly identified comparison groups that were similar with respect to important determinants of outcome, other than the one of interest?
- Were the outcomes and exposures measured in the same way in the groups being compared?
- Was follow-up sufficiently long and complete?

Secondary guides:

- Is the temporal relationship correct?
- Is there a dose-response gradient?

What are the results?

- How strong is the association between exposure and outcome?
- How precise is the estimate of the risk?

Will the results help me in caring for my patients?

- Are the results applicable to my practice?
- What is the magnitude of the risk?
- Should I attempt to stop the exposure?

- Comparison groups:
  - **Exposure group** – sample of patients with a result exceeding the alert threshold
  - **Non-exposure (control) group** – sample of patients whose result does not exceed the alert threshold
  - Follow both groups forward in time, monitor for an occurrence of the outcome.

- Are both groups similar with respect to other important determinants of outcome?
  - Investigators should document the characteristics of all subjects, and either:
    - demonstrate the comparability of the groups, or
    - use statistical techniques to adjust for differences between the groups.

Critical Appraisal of Outcome Studies

JAMA, May 25, 1994—Vol 271, No. 20

The Medical Literature

Users' Guides to the Medical Literature

IV. How to Use an Article About Harm

Mitchell Levine, MD, MSc; Stephen Walter, PhD; Hui Lee, MD, MSc; Ted Haines, MD, MSc;
Anne Holbrook, MD, PharmD, MSc; Virginia Moyer, MD, MPH; for the Evidence-Based Medicine Working Group

Are the results of the study valid?
- Primary guides:
  - Were there clearly identified comparison groups that were similar with respect to important determinants of outcome, other than the one of interest?
  - Were the outcomes and exposures measured in the same way in the groups being compared?
  - Was follow-up sufficiently long and complete?
- Secondary guides:
  - Is the temporal relationship correct?
  - Is there a dose-response gradient?

What are the results?
- How strong is the association between exposure and outcome?
- How precise is the estimate of the risk?

Will the results help me in caring for my patients?
- Are the results applicable to my practice?
- What is the magnitude of the risk?
- Should I attempt to stop the exposure?

- Were exposures/outcomes measured the same way for both groups?
  - Surveillance bias: clinicians may search more diligently for disease in exposed patients.
  - Mortality is not subjective.
- Was the time period for follow-up appropriate?
  - Critical risk results indicate an imminent risk of serious harm.
  - Monitoring for outcomes for a month or year post exposure does not determine "imminent" risk.
  - Can the outcome be reasonably attributed to a laboratory result that was measured months ago?

Critical Appraisal of Outcome Studies

JAMA, May 25, 1994—Vol 271, No. 20

The Medical Literature

Users' Guides to the Medical Literature

IV. How to Use an Article About Harm

Mitchell Levine, MD, MSc; Stephen Walter, PhD; Hui Lee, MD, MSc; Ted Haines, MD, MSc;
Anne Holbrook, MD, PharmD, MSc; Virginia Moyer, MD, MPH; for the Evidence-Based Medicine Working Group

Are the results of the study valid?
- Primary guides:
  - Were there clearly identified comparison groups that were similar with respect to important determinants of outcome, other than the one of interest?
  - Were the outcomes and exposures measured in the same way in the groups being compared?
  - Was follow-up sufficiently long and complete?
- Secondary guides:
  - Is the temporal relationship correct?
  - Is there a dose-response gradient?

What are the results?
- How strong is the association between exposure and outcome?
- How precise is the estimate of the risk?

Will the results help me in caring for my patients?
- Are the results applicable to my practice?
- What is the magnitude of the risk?
- Should I attempt to stop the exposure?

- Is the temporal relationship correct?
  - We should be sure that the study subjects did not have the outcome of interest until after the exposure.
- Is there a dose-response gradient?
  - The outcome is more likely attributable to the exposure if – a rise in the quantity or duration of the exposure corresponds with a rise in the risk of the outcome.

Critical Appraisal of Outcome Studies

JAMA, May 25, 1994—Vol 271, No. 20

The Medical Literature

Users' Guides to the Medical Literature

IV. How to Use an Article About Harm

Mitchell Levine, MD, MSc; Stephen Walter, PhD; Hui Lee, MD, MSc; Ted Haines, MD, MSc;
Anne Holbrook, MD, PharmD, MSc; Virginia Moyer, MD, MPH; for the Evidence-Based Medicine Working Group

- How strong is the association between exposure and outcome?
  - Association is typically measured as a risk ratio (relative risk) or an odds ratio.
  - RR (or OR) > 1 = an increase in risk of harm associated with the exposure
  - RR (or OR) < 1 = a reduction in risk
  - Observational studies are prone to bias and confounding factors, thus higher RR (or OR) values are needed (compared to RCTs) to confirm a true association.

Are the results of the study valid?
- Primary guides:
  - Were there clearly identified comparison groups that were similar with respect to important determinants of outcome, other than the one of interest?
  - Were the outcomes and exposures measured in the same way in the groups being compared?
  - Was follow-up sufficiently long and complete?
- Secondary guides:
  - Is the temporal relationship correct?
  - Is there a dose-response gradient?

What are the results?
- How strong is the association between exposure and outcome?
- How precise is the estimate of the risk?

Will the results help me in caring for my patients?
- Are the results applicable to my practice?
- What is the magnitude of the risk?
- Should I attempt to stop the exposure?

Estimate of Relative Risks and Odds Ratios for Exposed and Unexposed Patients

Patient	Adverse Event (Case)	No Adverse Event (Control)
Exposed	a	b
Not exposed	c	d

Relative risk = [a/(a+b)]/[c/(c+d)].
Odds ratio = (a/c)/(b/d).

Critical Appraisal of Outcome Studies

JAMA, May 25, 1994—Vol 271, No. 20

The Medical Literature

Users' Guides to the Medical Literature

IV. How to Use an Article About Harm

Mitchell Levine, MD, MSc; Stephen Walter, PhD; Hui Lee, MD, MSc; Ted Haines, MD, MSc;
Anne Holbrook, MD, PharmD, MSc; Virginia Moyer, MD, MPH; for the Evidence-Based Medicine Working Group

Are the results of the study valid?
- Primary guides:
  - Were there clearly identified comparison groups that were similar with respect to important determinants of outcome, other than the one of interest?
  - Were the outcomes and exposures measured in the same way in the groups being compared?
  - Was follow-up sufficiently long and complete?
- Secondary guides:
  - Is the temporal relationship correct?
  - Is there a dose-response gradient?

What are the results?
- How strong is the association between exposure and outcome?
- How precise is the estimate of the risk?

Will the results help me in caring for my patients?
- Are the results applicable to my practice?
- What is the magnitude of the risk?
- Should I attempt to stop the exposure?

- How precise is the estimate of risk?
  - Confidence intervals should be calculated for the risk estimate.
  - The lower limit of the interval provides a minimal estimate of the strength of the association.
- Which clinical setting are these results applicable to?
  - Need to consider whether the study population is representative of the general population.

  e.g., in relation to morbidity, age, race, treatments, exposures.
  - May need to specify the clinical setting in which the thresholds can be used.

Critical Appraisal of Outcome Studies

JAMA, May 25, 1994—Vol 271, No. 20

The Medical Literature

Users' Guides to the Medical Literature

IV. How to Use an Article About Harm

Mitchell Levine, MD, MSc; Stephen Walter, PhD; Hui Lee, MD, MSc; Ted Haines, MD, MSc;
Anne Holbrook, MD, PharmD, MSc; Virginia Moyer, MD, MPH; for the Evidence-Based Medicine Working Group

Are the results of the study valid?
- Primary guides:
  - Were there clearly identified comparison groups that were similar with respect to important determinants of outcome, other than the one of interest?
  - Were the outcomes and exposures measured in the same way in the groups being compared?
  - Was follow-up sufficiently long and complete?
- Secondary guides:
  - Is the temporal relationship correct?
  - Is there a dose-response gradient?

What are the results?
- How strong is the association between exposure and outcome?
- How precise is the estimate of the risk?

Will the results help me in caring for my patients?
- Are the results applicable to my practice?
- What is the magnitude of the risk?
- Should I attempt to stop the exposure?

1. Review the literature to identify appropriate alert thresholds
2. Rate the quality of the evidence on which these thresholds are based
3. Perform risk analysis to assess threshold suitability
4. Assess transferability and consider the pre- and postanalytical aspects of the alert threshold
5. Assess the impact of the selected thresholds on the frequency of critical alerts
6. Seek endorsement for selected thresholds from laboratories and clinical groups

DE GRUYTER Clin Chem Lab Med 2019; 57(1): 89–94

Opinion Paper

Craig A. Campbell*, Que Lam and Andrea R. Horvath

An evidence- and risk-based approach to a harmonized laboratory alert list in Australia and New Zealand

- What is the clinical significance of the risk?
- Is this a high risk result?

These are questions for the working party

Take Home Messages

- Outcome studies are regarded the best source of evidence for deciding alert thresholds.
- Critical appraisal of outcome studies is required to confirm the quality of the study design and findings.
- Clinical guidelines for the management of disease associated with high risk results may provide key information for deciding alert thresholds, especially when high quality outcome study evidence is lacking.