On the need for evidence-based clinical biochemistry

It can be argued that there are 3 distinct phases in clinical biochemistry. First is a preanalytical stage, which involves research into new or better technology of potential importance in helping particular investigations. Relatively few practicing clinical biochemists have the opportunity to contribute to this area, increasingly the domain of academic or industrial research groups. Most clinical biochemists spend most of their time in the second stage of high-volume clinical laboratory analysis, which is driven by issues of management, productivity, quality, and cost. There is no reason, however, why all biochemists could not contribute to the third, postanalytical stage in which laboratory tests can be harnessed to facilitate both increased clinical effectiveness and cost-effectiveness. This stage (which logically should precede rather than follow the second stage) is the focus of many of the articles about diagnosis that appear in this journal.

Read, Lachs, and Feinstein set down 7 methodological standards for evaluating diagnostic tests (Table) and applied them to reports of the studies published in the Lancet, BMJ, New England Journal of Medicine, and Journal of the American Medical Association from 1978 through 1993 (1). They found 112 articles, predominantly about radiological tests and immunoassays. Few of these articles met the 7 standards: Only 46% avoided work-up bias, just 8% reported accuracy in subgroups, and only half mentioned the sex or age of the patients being investigated. The more recent period studied (1990 to 1993) showed some improvement, with 24% of later reports meeting up to 4 standards, and 6% meeting up to 6. It is not surprising then that so few articles about diagnostic tests appear in this journal.

By contrast, several studies have now shown clear links between the intensity of diagnostic testing and subsequent therapy, both geographically (2) and temporally (3): Test more, treat more. But we simply do not know whether, by testing more and treating more, we improve the quality and efficiency of the care we provide.

I asked David Sackett, as a physician practicing evidence-based medicine, what he wanted from evidence-based clinical biochemistry. Sackett expressed 3 desires: to be able to discuss a patient’s illness with a colleague; to be able to abandon reporting of normal ranges; and to have evidence available to support the validity, importance, and clinical usefulness of biochemical tests. The first and third requests can be filled by the thoughtful generation and use of likelihood ratios (LRs), the odds that a given finding would occur in a patient with, as opposed to without, the target disorder or condition.

LRs are easily calculated from the sensitivity and specificity of a diagnostic test. They can be combined with clinical judgement and pretest probabilities through the use of a nomogram, which generates the post-test probability of a patient having the target disease or disorder. A good illustration of the use of LRs can be shown in testing for hypothyroidism in general practice, where the prevalence is low but simple clinical scoring can increase pretest probabilities. Suppose that a serum thyroid-stimulating hormone (TSH) test result of > 15 mU/L had both a sensitivity and specificity of 98% for detecting hypothyroidism, so that the LR for a high result is 56. When a general practitioner sends a sample from a patient with no overt clinical indication of hypothyroidism to the laboratory for a TSH test, the pretest probability is less than 1% (the prevalence of hypothyroidism in an ambulatory population). Even when such a patient has a positive test result, their post-test probability of having hypothyroidism is still less than 30%. On the other hand, if a patient had several signs and symptoms of hypothyroidism, with a 30% pretest probability, the same TSH test result would have a post-test probability of over 90%. Such post-test probabilities simplify decisions on whether to treat, refer, or wait. The prevalence, sensitivity, and specificity data needed for these calculations can be obtained from a clinical audit of general practitioners’ requests for hypothyroid tests. In such a collaboration between clinicians and clinical biochemists, general practitioners could generate clinical scoring systems that document the rising pretest probability of hypothyroidism at increasing

<table>
<thead>
<tr>
<th>Standard</th>
<th>Conditions to be met</th>
<th>Percentage meeting this standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectrum</td>
<td>Provide information about age, sex, presenting symptoms or stage, and eligibility criteria</td>
<td>27%</td>
</tr>
<tr>
<td>Pertinent subgroups</td>
<td>Separate results by pertinent subgroups</td>
<td>8%</td>
</tr>
<tr>
<td>Avoidance of work-up bias</td>
<td>All patients have to undergo both the diagnostic test and the gold standard verification</td>
<td>46%</td>
</tr>
<tr>
<td>Avoidance of review bias</td>
<td>Confirmation that the test and standard were evaluated independently</td>
<td>38%</td>
</tr>
<tr>
<td>Precision</td>
<td>Reporting confidence intervals or standard errors</td>
<td>11%</td>
</tr>
<tr>
<td>Indeterminate results</td>
<td>Reporting them and how they were handled in the analyses</td>
<td>23%</td>
</tr>
<tr>
<td>Reproducibility</td>
<td>Reporting observer variation for any tests requiring interpretation</td>
<td>23%</td>
</tr>
</tbody>
</table>
levels of clinical suspicion. The use of information technology could easily generate additional useful ways of integrating the information, and it might even simplify the way in which thyroid function tests are now used (4).

There are 6 advantages to using LRs: 1) They are comprehensible and easier to handle than sensitivity and specificity data; 2) they can be used sequentially so that the post-test probability from one diagnostic test becomes the pretest probability for the next; 3) they combine clinical judgement with laboratory science; 4) they can be calculated at different levels of test results; 5) they show unequivocally the extreme standards needed for screening tests but suggest ways in which screening could be made more effective; and 6) they highlight the way in which clinical audits can be used to generate local data relevant for local use.

As can be seen from these examples, and as I have suggested elsewhere (5), the profession of clinical biochemistry clearly needs to grasp the nettle of evidence-based clinical biochemistry and apply itself in 6 areas:

1. Undertaking systematic reviews of the evidence in important areas of clinical biochemistry. This will not be easy because the field is large in number (about 1000 articles a year on prostate-specific antigen alone) and in breadth (from genetics and molecular biology to clinical effectiveness). Nevertheless, important information on the usefulness of tests can easily be overlooked unless the work is systematic.

2. Deriving and understanding the rules and methods to be used. Cooperation with the Cochrane Collaboration groups working in this area will facilitate this objective.

3. Ensuring that modules on evidence-based medicine, systematic reviews, and critical appraisal of the literature are included in educational programmes for those preparing for professional qualifications in clinical biochemistry.

4. Organising individuals and groups to pioneer examples of the application of evidence-based clinical biochemistry methods to specific diagnostic investigations. Concrete examples will serve to influence others.

5. Examining ways in which the clinical and cost-effectiveness information on laboratory tests can best be disseminated and implemented. Cost-effectiveness calculations should also benefit from the discipline of an evidence-based clinical biochemistry approach if it is to be really worthwhile.

6. Exploring the use of auditing in the development of local evidence-based clinical biochemistry guidelines. Sensitivity, specificity, and LRs can be applied to populations for which the tests are being used in practice, not just based on those used in the original research. Combining a definition of a clear clinical question with locally collected information and results should constitute an important way of influencing the use of diagnostic tests.

R.A. Moore, MA, DPhil, DN
University of Oxford
Oxford, England, UK

References

Journals Reviewed for This Issue* Core Journals
An J Med
Arch Pediatr Adolesc Med
Arch Surg
Arthritis Rheum
BMJ
Br J Gen Pract
Br J Obstet Gynaecol
Br J Surg
Circulation
Clin Pediatr

Arch Pediatr Adolesc Med
Arch Surg
Arthritis Rheum
BMJ
Br J Gen Pract
Br J Obstet Gynaecol
Br J Surg
Circulation
Clin Pediatr

Acta Obstet Gynecol Scand
Age Ageing
Am J Cardiol
Ann J Public Health
Am J Respir Crit Care Med
Ann Emerg Med
Ann Med
Arch Fam Med

Acta Obstet Gynecol Scand
Age Ageing
Am J Cardiol
Ann J Public Health
Am J Respir Crit Care Med
Ann Emerg Med
Ann Med
Arch Fam Med

Arch Pediatr Adolesc Med
Arch Surg
Arthritis Rheum
BMJ
Br J Gen Pract
Br J Obstet Gynaecol
Br J Surg
Circulation
Clin Pediatr

Journals for Continuing Review
Arch Neurol
Br J Psychiatry
Br J Rheumatol
Can Med Assoc J
Chest
Clin Invest Med
Crit Care Med
Fertil Steril
Gastroenterology
Gut
Heart (formerly Br Heart J)
J Am Coll Cardiol
J Am Geriatr Soc
J Clin Epidemiol
J Fam Pract
J Infect Dis
J Neurol Neurosurg Psychiatry
J Pediatr
J Vasc Surg
Lancet
N Engl J Med
Obstet Gynecol
Pediatrics
Surgery

Med Care
Med J Aust
Neurology
Pain
Spine
Stroke
Thorax

*Approximately 60 additional journals are reviewed. This list is available on request.