Clinica Chimica Acta 432 (2014) 27–32

Contents lists available at ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/clinchim

How good is the evidence base for test selection in clinical guidelines?

Shivani Misra [a,*], Julian H. Barth [b]

[a] Clinical Biochemistry and Metabolic Medicine, Imperial Healthcare NHS Trust, London, UK
[b] Clinical Biochemistry, Leeds Teaching Hospitals NHS Trust, Leeds, UK

ARTICLE INFO

Article history:
Received 16 June 2013
Received in revised form 24 January 2014
Accepted 30 January 2014
Available online 7 February 2014

Keywords:
Guidelines
Test selection
Evidence-based
Quality
Audit

ABSTRACT

Clinical guidelines are ubiquitous, manifold and form an integral component of evidence-based clinical practice. Guidelines on test selection are often considered a useful adjunct to aid clinical decision-making, as test selection is a complex process that is influenced by many patient, clinician and laboratory factors. However, it is important to carefully evaluate several aspects of these guidelines, which include the context of the test in the guideline, the quality of the studies underpinning recommendations, the extent of the evaluation of effectiveness (or performance) of the specific test and in the clinical pathway, its applicability and ease of implementation. A robust evaluation of a diagnostic test should incorporate several stages including evaluation in healthy, symptomatic but unaffected and affected populations, and importantly a measurement of impact on patient outcomes. Few diagnostic studies meet these criteria, and therefore crucial aspects of test evaluation are overlooked prior to incorporation into clinical guidelines. Whilst efforts are made to standardise reporting of studies, strength of evidence and quality of guidelines, further work is required to improve the quality of the diagnostic studies that formulate these guidelines. It is important that clinicians using guidelines for test selection appreciate the limitations of the diagnostic test, and the guidelines themselves.

1. Introduction

The term evidence-based is defined as "the conscientious, explicit and judicious use of current best evidence, in making decisions about the care of patients" [1]. In laboratory medicine, this concept can be applied at several levels of decision-making, including for test selection.

Using evidence to guide test selection is an extremely important process that helps ensure laboratory tests are appropriately selected. Appropriate selection of tests in turn will be guided not only by the evidence base supporting use of the test, but also by the individual patient context, the clinical scenario, availability of the test, user preference and local economical considerations. Guidelines on test selection are an important facet of practicing evidence-based medicine. They should help to harmonise practice, reduce inappropriate test selection and reduce treatment variations secondary to inappropriate interpretation of test results due to analytical variations.

However, it is important to appreciate the type of evidence underpinning recommendations on test use in guidelines. In general terms diagnostic performance may be evaluated against a comparator that is deemed to be the reference standard and will thus evaluate diagnostic accuracy. Increasingly however, users of laboratory tests require some measure of the clinical utility, that is how well the test affects or changes health outcomes [2]. Therefore an evaluation of clinical utility in the form of randomised controlled trials is desirable, albeit not always achievable.

Evidence-based guidelines for test selection should ideally encompass many aspects of the process, but in practice may be limited in their scope. Whilst this does not negate the utility as they help standardise practice, users should understand the limitations of such an approach at an individual patient level.

One further aspect is also important to consider: the process by which guidelines are produced. Guidelines are produced by disparate groups. In the past this has led to inconsistency in quality and robustness of guidelines. Current initiatives such as AGREE, STARD and GRADE have provided a rigorous framework through which clinical guidelines can be constructed in a standardised way. Nevertheless the process is not infallible and the same body of evidence as interpreted by guideline developers may result in different recommendations.

This article reviews some of the salient issues and challenges to reflect on, when considering guidelines on test selection.

2. How are tests selected?

Test selection is a process undertaken by health-care professionals that can be influenced by laboratories, clinical guidelines and patients themselves. From a disease-perspective, tests may be selected for diagnosis, screening, monitoring, treatment selection, risk stratification and prognosis. But in practice, test selection is far more complex, with

* Corresponding author at: Diabetes, Endocrinology & Metabolism, Imperial College London, St Mary's Hospital, Ground Floor, Medical School Building, Norfolk Place, London W2 1PG, UK. Tel.: +44 20 7594 2460; fax: +44 20 7594 2432.
E-mail address: s.misra@imperial.ac.uk (S. Misra).

0009-8981/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cca.2014.01.040

non-disease related factors such as, patient choice, physician choice, test availability, cost and demand management strategies all impinging on the process.

Selecting a test is seldom an isolated process and laboratory tests are a component of a clinical pathway, relevant to a particular clinical scenario in a given patient [3]. The test results will usually be integrated with other clinical and laboratory findings as an investigative strategy related to the clinical scenario [4]. This is important to consider as studies evaluating tests often do so in the context of a clinical pathway, which may be affected by downstream events. These subsequent actions may affect the clinical course and thus the perceived usefulness of the test (see later).

The use of tests and their diagnostic utility alter at different time points during patient management. At the beginning of the pathway a diagnostic test requires sensitivity whereas confirmation of the test will require specificity, although it is expected that a single test will have both. Finally when a test is used for disease monitoring, analytical precision becomes the key property. However, it must be appreciated that the vast majority of laboratory tests are requested for routine assessment of unwell patients, during which the process of test selection becomes somewhat arbitrary and under these circumstances evidence is largely lacking.

3. Consequences of inappropriate test selection

The aim of appropriate test selection is to improve patient care, however, as discussed measuring clinical utility of testing can be challenging. Conceptually it is simpler to think about the consequences of inappropriate test selection, and to therefore consider guidelines on test selection as a means to preventing these.

Bossuyt and McCaffery published a framework to help define possible patient outcomes of a testing intervention [5]. These can be loosely categorised into clinical pathway effects, direct health effects and secondary non-clinical measures. It is therefore useful to think of these domains as those that can be affected by inappropriate test selection.

- Clinical pathway effects: defined as the clinical response to the test result, which may be to select, start, alter, stop or modify treatment. To order further tests, or to monitor with no further intervention.
- Direct health effects: defined as the physical effect to the patient in having the test. Venous blood tests are generally well-tolerated but other tests such as arterial puncture or cerebrospinal fluid sampling carry risks and complications to the patient.
- Secondary non-clinical effects: refers to the way in which the patient reacts to the result of the test, which can be further categorised as emotional, social, cognitive and behavioural.

Consequences at a patient level are therefore manifold and often not fully appreciated. Guidelines on appropriate test selection often focus on the clinical pathway, but it is worth considering these other potential outcomes when selecting tests.

Investigative tests are not perfect and carry the risks of false positive and negative diagnoses. Increasing numbers of investigations leads to increased interventions with their attendant risks [6]. This is particularly likely when a test is used outside its normal context. We have highlighted this in a patient with ascites due to hypothyroidism in whom multiple investigations including an explorative laparotomy were undertaken due to the misinterpretation that an elevated CA125 was caused by ovarian cancer rather than by ascitic fluid [7].

Part of the problem lies in the definition of what constitutes an appropriate test, which has led to wide variations in estimates of inappropriate testing [8,9]. Therefore guidelines on test selection are a positive step in helping to prevent any of these potential consequences through standardisation of practice.

4. Evidence base for laboratory tests

It is important to consider what evidence underpins guidelines on test selection. Many schemes for the evaluation of medical tests have been proposed. In the last few decades the introduction of new medical tests into practice, has been compared to the introduction of a new drug. The latter has a well-defined hierarchical model of phased evaluation and proposals to adopt a similar approach for the introduction of a new test have also been made [10–12]. This is keeping with a paradigm shift in the way undertaking medical tests are rightly regarded as equal to any other medical intervention.

4.1. Ideal evaluation of a medical test

Lijmer et al. undertook a systematic review of 19 different published models for evaluation of medical tests [10]. The models reviewed had striking similarities and in general included:

- an early phase where tests were developed;
- a diagnostic accuracy phase where the test was evaluated in a variety of settings, including healthy volunteers, diseased population, population similar to the intended use and against a reference comparator;
- a clinical effectiveness phase, which included some measure of diagnostic thinking efficiency (the degree to which decision making was altered as a result of the test) and randomised controlled trials to assess effect on health outcomes;
- some models additionally assessed cost effectiveness and other outcomes e.g. secondary changes in practice.

In a departure from drug evaluations, the proposal for phased evaluations of medical tests describes a cyclical rather than a linear unidirectional course. Such a cyclical process has been described by the Centre for Disease Control for evaluation of genetic tests, further developed by the Evaluation of Genomic Applications in Practice and Prevention (EGAPP) working group [11]. The approach asks 44 targeted questions which comprehensively review the new genetic test. These are broadly separated into 4 domains:

- Analytical validity: how accurately the test measures the genotype of interest
- Clinical validity: how consistently and accurately the test detects or predicts the intermediate or final outcomes of interest
- Clinical utility: how likely the test is to significantly improve patient outcomes
- Ethical, legal and social implications: identification of such issues and any safeguards in place.

The Global Harmonization Task Force (GHTF) for the scientific validity, determination and performance evaluation of in vitro diagnostic medical devices also published a cyclical phased approach with similar domains [12].

4.2. Points to consider in existing studies evaluating medical tests

The above approaches highlight that a focused evaluation of medical tests on diagnostic accuracy studies alone is not sufficiently robust, but despite this, it may not be feasible or appropriate to undertake a thorough evaluation of all the domains. However, some evidence of clinical utility, either directly through randomised studies (see below) or indirectly through the use of decision analysis models, is imperative.

- Tests form part of a clinical/diagnostic pathway or testing strategy: In reality the evaluation of medical tests is even more complex as tests are seldom undertaken in isolation and the vast majority do not usually have a direct uninterrupted effect on patient outcome [5]. The decision to utilise a test is usually part of a clinical pathway and interpretation of the test result will be combined with downstream outcomes. The evaluation of a test will be pathway specific and care