Clinica Chimica Acta 427 (2014) 49–57

Contents lists available at ScienceDirect

Clinica Chimica Acta

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

Special report

From biomarkers to medical tests: The changing landscape of test evaluation

Andrea R. Horvath [a,b,*], Sarah J. Lord [b,c,l], Andrew StJohn [d], Sverre Sandberg [e], Christa M. Cobbaert [f], Stefan Lorenz [g], Phillip J. Monaghan [h], Wilma D.J. Verhagen-Kamerbeek [i], Christoph Ebert [j], Patrick M.M. Bossuyt [k],

For the Test Evaluation Working Group of the European Federation of Clinical Chemistry Laboratory Medicine

[a] SEALS Department of Clinical Chemistry, Prince of Wales Hospital and School of Medical Sciences, University of New South Wales, Australia
[b] Screening and Test Evaluation Program, School of Public Health, University of Sydney, Australia
[c] National Health and Medical Research Council (NHMRC) Clinical Trials Centre, University of Sydney, Australia
[d] ARC Consulting, Perth, Australia
[e] Laboratory of Clinical Biochemistry, Haukeland University Hospital, and The Norwegian Quality Improvement of Primary Care Laboratories (NOKLUS), Department of Public Health and Primary Health Care, University of Bergen, Norway
[f] Department of Clinical Chemistry and Laboratory Medicine, Leiden University Medical Center, The Netherlands
[g] Abbott Diagnostics, Wiesbaden, Germany
[h] Department of Clinical Biochemistry, The Christie Hospital NHS Foundation Trust, Manchester, UK
[i] Medical and Scientific Affairs, Roche Diagnostics International Ltd., Rotkreuz, Switzerland
[j] Medical and Scientific Affairs, Roche Diagnostics GmbH, Penzberg, Germany
[k] Department of Clinical Epidemiology, Biostatistics & Bioinformatics, Academic Medical Center, University of Amsterdam, The Netherlands
[l] School of Medicine, Sydney, The University of Notre Dame Australia, Australia

ARTICLE INFO

Article history:
Received 22 April 2013
Received in revised form 15 September 2013
Accepted 15 September 2013
Available online 27 September 2013

Keywords:
Biomarker
Clinical effectiveness
Evidence-based laboratory medicine
Medical test

ABSTRACT

Regulators and healthcare payers are increasingly demanding evidence that biomarkers deliver patient benefits to justify their use in clinical practice. Laboratory professionals need to be familiar with these evidence requirements to better engage in biomarker research and decisions about their appropriate use.

This paper by a multidisciplinary group of the European Federation of Clinical Chemistry and Laboratory Medicine describes the pathway of a laboratory assay measuring a biomarker to becoming a medically useful test. We define the key terms, principles and components of the test evaluation process. Unlike previously described linearly staged models, we illustrate how the essential components of analytical and clinical performances, clinical and cost-effectiveness and the broader impact of testing assemble in a dynamic cycle. We highlight the importance of defining clinical goals and how the intended application of the biomarker in the clinical pathway should drive each component of test evaluation. This approach emphasizes the interaction of the different components, and that clinical effectiveness data should be fed back to refine analytical and clinical performances to achieve improved outcomes.

The framework aims to support the understanding of key stakeholders. The laboratory profession needs to strengthen collaboration with industry and experts in evidence-based medicine, regulatory bodies and policy makers for better decisions about the use of new and existing medical tests.

Abbreviations: ACR, albumin:creatinine ratio; ACS, acute coronary syndrome; AMI, acute myocardial infarction; CRP, C-reactive protein; BNP, B-type natriuretic peptide; CE, Conformité Européenne; CK-MB, creatine kinase MB isoform; cTn, cardiac Troponin; EFLM, European Federation of Clinical Chemistry and Laboratory Medicine (formerly abbreviated as EFCC); EU, European Union; FDA, Food and Drug Administration; HbA_{1c}, Hemoglobin A_{1c}; hs-cTn, high sensitivity cardiac Troponin; IFCC, International Federation of Clinical Chemistry and Laboratory Medicine; INR, international normalized ratio; IVD, in vitro diagnostics; NGSP, National Glycohemoglobin Standardization Program; NHS, National Health Service in the United Kingdom; NICE, National Institute for Health and Care Excellence; PoCT, point of care testing; RCT, randomized controlled trial.

* Corresponding author at: SEALS North, Department of Clinical Chemistry Level 4, Campus Centre, Prince of Wales Hospital Barker Street, Randwick NSW 2031, Sydney, Australia. Tel.: +61 404 027 843.
E-mail address: rita.horvath@sesiahs.health.nsw.gov.au (A.R. Horvath).

0009-8981/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cca.2013.09.018

1. Introduction

There is an increasing awareness that the introduction of costly new medical interventions, including medical tests, can only be justified if they deliver proportionate benefits to patients. Increased public, media and political awareness of the harms from medical tests has come from debates about the potential for over-diagnosis in asymptomatic patients [1] and concerns about the harms of direct-to-consumer testing [2]. Weaknesses of the current systems to assure the quality and clinical utility of in vitro medical devices (IVDs) have been pointed out [3]. Amidst all this, the regulatory environment for therapeutic and diagnostic technologies is changing rapidly. Revisions of the European directives on medical [4] and in vitro diagnostic devices [5] are being prepared in parallel. The ability of novel medical tests to improve health outcomes is also becoming more central in discussions about their market entry and reimbursement. The increasing requirements for clinical benefits and patient safety mirror public and political pressures for more transparency. These processes are affecting the way novel medical tests and biomarkers are being developed, and are likely to reshape the landscape of medical test evaluation. Laboratory professionals need to be familiar with these evidence requirements to better engage in biomarker research and in clinical and policy decisions about the appropriate use of laboratory tests.

Over the past decade landmark advances have been made to define the types of evidence required to evaluate medical tests and distinguish between the different phases of test evaluation from discovery to assessment of cost-effectiveness [6–10]. There is less guidance, however, about the most efficient approaches to produce this evidence and judging whether it is adequate for proving the clinical effectiveness of biomarkers.

The evaluation of medical tests differs from comparable processes for therapeutic interventions. One of the most important differences is that medical testing rarely improves health outcomes directly. Testing is usually part of a more complex clinical pathway where test results guide treatment decisions, which include a variety of medical actions and processes. All of these shape the final health outcomes for the patients tested. Test evaluation therefore requires the consideration of all the consequences of clinical management decisions that are guided by the test results. An understanding of these more complex concepts for test evaluation is becoming essential for informed decision making by all potential stakeholders.

To help address these issues, the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) has formed a Working Group on Test Evaluation which consists of laboratory professionals, clinical epidemiologists, health technology assessment experts and representatives of the in vitro diagnostics industry. The primary purpose of this working group is to provide key stakeholders, i.e. laboratory professionals, clinicians, researchers, manufacturers, policy makers and purchasers, with guidance and practical tools for assessing the clinical benefits of in vitro medical tests. In this introductory paper, the working group outlines the key principles and defines some of the key components of contemporary approaches to test evaluation, such as analytical performance, clinical performance, clinical effectiveness, cost-effectiveness, and the broader impact of testing on social, psychological, legal, ethical, societal, organizational and other consequences. We additionally present a framework for the evaluation of medical tests that integrates these components into a dynamic process. We illustrate the key principles and components with examples from the literature on Hemoglobin A_{1c} (HbA_{1c}) and cardiac markers, including cardiac Troponins (cTn) and B-type natriuretic peptides (BNP).

2. Key definitions and principles

There is no international consensus on the terminology related to test evaluation and numerous definitions exist in the literature. Table 1 lists a number of alternative terms and illustrates the proposed definitions with examples.

Under the general umbrella term of *medical tests*, which encompasses tests from all clinical disciplines, specialties, or types (laboratory, histopathology, imaging, and others), we define and focus this paper on in vitro medical tests only; yet the key messages outlined here can also be adapted to any other forms of medical tests. We distinguish in vitro medical or laboratory assays and measurement procedures [11] from biomarkers that are measured by these assays [12].

From the regulatory and medical laboratory perspective, *test evaluation* refers to a set of processes which start when a laboratory assay capable of measuring a biomarker with potential application in clinical care becomes available. Ideally, and before the test evaluation process starts, the potential purpose of the new marker is defined to address an unmet clinical need. Laboratory assays can then be developed to measure the marker with this purpose in mind. For example, in the field of cardiac biomarkers, CK-MB and Troponins are considered as tests of myocardial damage, but clinicians have long been waiting for non-invasive markers that can predict myocardial infarction before cell damage happens. An early research finding revealed that endothelial cells are shed from coronary arteries several days to weeks before heart attack. This finding led to the development of a method to measure circulating endothelial cells [13]. Translation of such primary findings usually starts with proof-of-concept studies which explore the association of the disease or condition with the new potential biomarker, usually in diseased versus control patients. Such 'case–control' designs tend to overestimate the clinical performance of a diagnostic assay, as they are designed to test proof of concept. Additional study designs are required for other phases in the medical test evaluation process.

The key principle of medical test evaluation is the fundamental premise that the introduction of any new test should eventually improve health outcomes, or provide other benefits, e.g. reduce costs, or simplify health care delivery without compromising the well-being of patients. Therefore evaluation begins with defining the potential health outcomes (benefits and harms) of introducing the test. Health outcomes should include consequences most relevant to patients. As discussed by Porter, these include survival, sustaining health, achieving recovery, improving functioning and reducing complications [14]. It also includes process outcomes such as reducing delays in time to diagnosis which also have direct patient benefits for reducing anxiety and improving treatment outcomes [14].

New tests should provide added benefit for patients or society over currently existing clinical pathways. *Clinical pathways* (also termed as clinical care pathways or test-treatment pathways) describe the typical processes of care for managing a specific condition in a specific group of patients [15], and provide a map that links testing to health and other outcomes (Table 1). The clinical pathway therefore plays a central role in the test evaluation process. Its description can be supported by information in well accepted best practice guidelines. For example, the National Institute for Health and Care Excellence (NICE) provides interactive clinical pathways supported by existing evidence-based guideline recommendations and tools for implementation (http://pathways.nice.org.uk/). For the assessment and management of suspected acute coronary syndrome the pathway shows the use of ECG and describes the timing and role of Troponin T or Troponin I tests and how they should inform subsequent clinical decisions for management based on test results (Table 1).

Since tests usually do not affect health outcomes directly, one has to define, right at the beginning of the evaluation process, the *purpose* and *role* of the medical test in the clinical pathway and the relevant patient population for each testing application. *Test purpose* describes the intended clinical application of the test and how the test information will be used to improve clinical management in practice (Table 1). Medical tests can be used for diagnosis and prognosis, but also for monitoring, early detection, screening, risk classification, treatment selection, surveillance after treatment, and many more. Within these applications,