Interpreting the evidence-base for making recommendations on medical tests

Gopalakrishna, G.

Citation for published version (APA):
Chapter One

Introduction
Medical tests are used in health care to provide information about the current or future status of a patient. Such tests include imaging tests, laboratory tests, physical examination and history taking or questionnaires. Tests are rarely used in isolation but typically form part of a test-treatment strategy; the results from testing guide, clinical management decisions around re-testing, follow up testing and treatment decisions.

Ideally, recommendations about testing should be based on their ability to improve patient outcome. Yet direct evidence of the impact of a test on patient outcomes is rare. For instance, the number of randomized trials that have evaluated test-treatment strategies is very small. Instead many evaluations of medical tests focus on a test’s accuracy: the ability of the test to correctly identify patients with the target condition. The results of such studies are typically reported as estimates of the test’s sensitivity and specificity.

Estimates of test accuracy are rarely sufficient to judge the health benefits from testing. A test with 99% test accuracy may not necessarily lead to improved patient outcomes if there is no effective treatment available, for example. Similarly, it is unclear to what extent a test with a specificity of 80% will generate an improvement in patient outcome, or harm. The effects of limited accuracy therefore have to be put in context, by offering a description of the consequences of misclassifications and correct classifications.

So far, the methods for developing recommendations about testing based on test accuracy are in their infancy. We know how to identify studies at risk of bias, and reasons for expressing concerns about the findings from test accuracy studies to a specific clinical question. It is less clear how we should put this all together: when and how can we recommend testing, based on test accuracy information only, and how should we qualify the strength of these recommendations. For this reason, guideline developers around the world seem to struggle when they face questions about the use of medical tests. These have to do with the methodology, but probably also with other factors. One way to facilitate the evaluation of medical tests, and the development of evidence-based guidelines, may be to link testing explicitly to management actions and downstream patient outcomes. This can be done by mapping out clinical pathways, also known as test-treatment pathways.

Mapping out such pathways describe the context in which testing may be used. This may help to define the possible downstream consequences of testing. These may be broadly defined as outcomes related to the test results or testing process such as direct health effects (i.e. related to target condition), emotional, social, cognitive, behavioral responses as a result of testing, legal or ethical effects of testing, and/or costs of testing.
Knowing these downstream outcomes and the pathway towards these outcomes can help in formulating more precisely the key questions, and types of evidence needed to make an evidence-based assessment if the new test or strategy is more meaningful than current practice. Several agencies such as the US Preventative Task Force (USPSTF), the Agency for Healthcare Research and Quality (AHRQ) and the Cochrane Diagnostic Test Accuracy (DTA) Systematic Reviews Handbook make reference to the development of such a pathway in the evaluation of a medical test. While there have been these efforts, the inclusion of evidence around patient outcomes both in the evidence surrounding medical tests and in the development of recommendations around medical tests is not yet mainstream.

The focus of our work reported in this thesis therefore has been to systematically look at the currently available approaches in evaluating and making recommendations around medical tests and to understand more precisely the reasons why test accuracy continues to be the dominant evidence base while evidence around downstream testing outcomes remains scarce. Guided by these findings, we proceed to propose strategies to address current gaps in currently available medical test evaluation methodologies.

Finally, we report on some initial user testing experiences on these approaches and make suggestions for future research in this area.

Chapter 2 lays the foundation of this thesis by demonstrating the challenges in medical test evidence appraisal as demonstrated in the conduct of a systematic review of diagnostic test accuracy studies. We present a systematic review of diagnostic accuracy studies of circulating antigen tests and urine reagent strips for detecting active schistosomiasis. This review underlines one of the main challenges in synthesizing medical test evidence, which is the variable and often poor study quality of accuracy studies, making the derivation of meaningful conclusions difficult.

Chapters 3 & 4 describe our work around understanding current methodologies that exist in evidence appraisal and making recommendations around medical tests as well the application of the GRADE for Diagnostics approach and limitations users might face while using this approach in medical test appraisal.
Chapter 3 is a systematic review of published systems for grading the evidence about medical tests. In the study reported in Chapter 4 we explore whether the dimensions used in GRADE, a widely used system to develop evidence-based recommendations in a systematic way, can be used to express the credibility, or certainty, of estimates of diagnostic test accuracy, generated by a study.

Chapter 5 presents the findings from a series of in-depth interviews with guideline developers in Europe aimed at better understanding the perspectives and beliefs among them on the challenges they face in evidence appraisal and guideline development around medical tests. We outline a number of solutions and key areas that were felt as obstacles to medical test guideline development.

Chapters 6 & 7 are concerned with the development of the clinical pathway as a tool to help systematic reviewers and guideline developers of medical test formulate specific key questions that can help guide them in their evidence selection with the intention of eventually making patient centred recommendations. These chapters focus on the importance of putting a medical test in the context of it’s use and makes suggestions on current challenges and areas needing improvement in the existing guidance.

Finally, in Chapter 8 we propose a step wise method that researchers of diagnostic test accuracy studies can employ in order to generate meaningful hypothesis prior to conducting diagnostic test accuracy studies. By doing so, we hope this can help researchers first consider the targeted purpose of the test in question prior to designing and conducting such a study.

We hope systematic review authors and guideline developers of medical tests as well as other healthcare decision makers who may need to make evidence-based recommendations on medical tests find the research reported in this thesis meaningful and useful in eventually making patient centred medical test recommendations.