ELSEVIER

Seminars in
NUCLEAR
MEDICINE

Diagnostic Accuracy Studies

Michaël Chassé, MD, PhD, FRCPC,[*,†] and Dean A. Fergusson, PhD, MHA[‡]

The medical community often assumes that the tests we use to diagnose various diseases are accurate, safe, and effective. However, the study designs traditionally used to determine whether such a diagnostic test is indeed accurate, safe, and effective are often at a higher risk of bias and are of lower methodological quality than those evaluating efficacy of therapeutic interventions. Several designs can be used to study diagnostic tests such as diagnostic accuracy cross-sectional studies, diagnostic accuracy case-control studies, and diagnostic accuracy comparative studies. Clinicians, researchers, and policy-makers may wish to consider moving toward higher quality study designs when studying new diagnostic modalities prior to their implementation in routine practice and diagnostic randomized trials are one such alternative.
Semin Nucl Med 49:87-93

Introduction

Randomized trials are considered the gold standard on which clinicians and policy-makers rely most to determine whether a medical intervention is effective. Scientific organizations place a higher level of quality and reliability on evidence coming from randomized trials and traditionally make stronger therapeutic recommendations when medical findings come from such studies.[1]

To decide whether a therapeutic intervention is warranted to treat a particular disease, clinicians must also rely on the accuracy of diagnostic information such as a detailed medical history, a thorough patient narrative, and a comprehensive medical examination. This information is often complemented by various diagnostic investigations such as laboratory tests, imaging procedures, and even tissue biopsies. Clinicians then select the appropriate therapeutic intervention based on the presumptive diagnosis to hopefully improve patient outcome. Clearly, if the diagnosis is wrong, even the most highly validated therapeutic intervention will likely not be beneficial and may even be harmful.

The medical community often assumes that the tests we use to diagnose various diseases are accurate, safe, and effective. However, surprisingly, the study designs used to determine whether such a diagnostic test is indeed accurate, safe, and effective are often at a higher risk of bias and are of lower methodological quality than those evaluating efficacy of therapeutic interventions. Reasons why are unclear. Unlike novel therapeutics, regulatory organizations do not require diagnostic tests to be validated with randomized trials. Moreover, given the higher amount of resources, number of patients, and methodological complexity associated with the conduct of randomized trials, researchers may be less compelled to select such a design if not absolutely required by the medical community.

Given the serious impact misdiagnoses can have on patient outcomes, and that most therapeutic interventions heavily rely on the accuracy of such diagnoses, clinicians, researchers, and policy-makers may wish to consider moving toward higher quality study designs when studying new diagnostic modalities prior to their implementation in routine practice. Diagnostic randomized trials are one such alternative to the more traditional diagnostic cohort studies that can significantly improve the quality and validity of the results obtained from a diagnostic study.[1,2] The use of this design may have the advantage of not only providing higher quality information at a lower risk of bias, but also by potentially providing additional information regarding their efficacy and safety in combination with treatment choices. Although they are not always possible to implement and have their limitations, more consideration should be given to a randomized study design when possible.

*Department of Medicine (Critical Care), University of Montreal Hospital, Montreal, Canada.
†University of Montreal Hospital Research Center (CRCHUM), Montreal, Canada.
‡Clinical Epidemiology Program, Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Ontario, Canada.
Address reprint requests to Michaël Chassé, Department of Medicine (Critical Care), University of Montreal Hospital, Montreal, Canada. E-mail: michael.chasse@umontreal.ca

https://doi.org/10.1053/j.semnuclmed.2018.11.005
0001-2998/

In the following sections, we will provide a brief overview of three different designs that can be used to study a new diagnostic test: (a) diagnostic accuracy cross-sectional (cohort) studies, (b) diagnostic accuracy case-control studies, and (c) diagnostic accuracy comparative studies. We will illustrate this review with a case example of a complex diagnostic situation (brainstem death) where several imaging modalities have been described to aid in the diagnosis and where an investigator may wish to identify the best imaging modality to help with clinical decision-making.

Target Condition: Neurological Determination of Death

To retrieve a vital organ from a donor for the aim of transplantation, clinicians must be 100% certain that the donor is deceased. The diagnosis of neurological death (NDD) is the concept of irreversible loss of the capacity for consciousness combined with the irreversible loss of all brainstem functions including the capacity to breathe. When patients fulfill NDD criteria, they are legally declared "dead." Traditionally, NDD is a predominantly a clinical diagnosis, made by physical examination at the bedside.[3,4] There are however many situations where a complete and accurate clinical evaluation is impossible, and clinicians must order additional tests to confirm NDD. Various imaging modalities have been proposed to support clinicians in such situations. Radionuclide blood flow examinations including planar or SPECT, computed tomography angiography (CTA), and digital subtraction angiography are modalities that are often used by clinicians. Their use for confirming NDD raises diagnostic challenges such as in patients with clinical NDD but with negative ancillary testing (ie, with persistence of brain blood flow). When such a scenario occurs, it is sometimes unclear whether the flow observed is significant enough to translate into brain function or not. It is currently recommended that in this clinical context the selected modality should be able to demonstrate the presence or absence of brain blood flow within the cerebral hemispheres and in structures within the posterior fossa.[5] It is however also increasingly recognized that when these tests are applied to clinically confirmed neurologically deceased patients, a proportion of them demonstrate detectable brain blood flow, suggesting an imperfect correlation between the findings at physical examination and the findings of the confirmatory test. It is thus important to correctly assess these confirmatory tests as this could lead to a population of patients who would have received a NDD based on current clinical criteria but in whom residual blood flow actually remains (not deceased by actual diagnostic criteria).[6,7]

The Clinical Question

A critical care physician faces a situation where he cannot complete a complete clinical neurological determination of death because of significant facial trauma that limits the evaluation of brainstem function. He decides to order a CTA to check whether there is residual intracranial blood flow to confirm that the patient is neurologically deceased. The critical care fellow in charge of the unit that day however cites a case report[8] and a recent systematic review[9] stating that CTA may not be the appropriate test for this medical situation and suggests a radionuclide examination instead, citing a different review[10] and a recent study[11] that suggests better accuracy of the SPECT modality. The staff critical care physician is however well aware of the studies cited by the fellow and emphasizes that when compared to clinical evaluation, SPECT had a sensitivity of only 83% in that particular study (in 17% of NDD patients flow was demonstrated) and similarly has not been appropriately validated. They agree that both tests require additional validation to decide which one is the most appropriate in this setting. Although they face a complex clinical situation in which they will have to decide with only limited quality evidence available, they contemplate how a study to validate a diagnostic accuracy test to help with the clinical decision-making around neurological death could be performed.

Definition of the Research Question

In this clinical situation, the researcher will want to compare the diagnostic accuracy of CTA and the radionuclide study with a reference standard. The population of interest in the study should be as close as possible to the population on which the tests will be applied in practice. It must also include a mix of patients that do present the "disease of interest" (neurological death in this context) as well as patients who are not diseased but are close to the target population. One could therefore consider including patients with a severe brain injury who are at high risk of neurological death. Doing so, both the sensitivity (death by imaging criteria when the patient meets clinical NDD criteria) and specificity (alive by imaging criteria when the patient does not meet clinical NDD criteria) could be determined.

To evaluate the identified imaging modalities, one must carefully select the reference standard. The reference gold standard for neurological death is considered to be based on clinical examination performed at the bedside.[12] A patient would be considered neurologically deceased when presenting with (1) an established etiology capable of causing neurological death; (2) an absence of confounders that can mimic neurological death; (3) an absence of all brainstem reflexes; and (4) a positive apnea test.[3,4] For this validation study, we would therefore have to exclude patients with contraindications to CTA or SPECT, and because the reference standard for this study will be the clinical evaluation, any patient with a confounding factor precluding complete clinical neurological evaluation. The study would therefore compare the capacity of the test to correctly classify patients that have a clinical absence of brainstem function (the clinical definition of neurological death) as deceased, and patients with residual brainstem function as alive.

Finally, depending on the selected design, the outcome of interest may be limited to diagnostic accuracy measures only, but could be expanded to clinical outcomes such as number of patients who donated organs or time from declaration to organ donation if a therapeutic intervention is paired with the result of the diagnostic test.

Types of Designs for Diagnostic Test Studies

Diagnostic tests all have an overarching objective to correctly classify a patient as having a disease or not.[13,14] Depending on the chosen design when studying a diagnostic test, if the result of the test is paired with a therapeutic intervention, it may sometimes be possible to measure, in addition to standard accuracy measures, meaningful outcomes that result from the clinical decisions derived from the diagnostic test.[2] An ideal diagnostic accuracy test should not result in any false-positive (patient wrongly diagnosed with the disease) or false-negative (patient wrongly diagnosed without the disease) results. The test should ideally be quickly performed, safe, readily available and accessible, noninvasive, not expensive, not susceptible to external factors that may affect test results, and standardized.[15] Diagnostic tests are rarely perfect, and the clinicians must accept a certain error rate in their results. It is thus important to study their properties carefully to have a sense of their strengths and limitations, a good understanding of their diagnostic accuracy and, ideally, the clinical impact of using this new test.

A properly conducted diagnostic test study will provide all the required information to decide whether this test should be used in the clinical setting. Patients must be selected using a consecutive or random sample of the target patient population and inappropriate exclusions should be avoided. The studied test should be interpreted without knowledge of the results of the reference standard and if a threshold is used for a reference standard, it should be prespecified. The reference standard used in such studies must classify correctly the target condition (disease) and be interpreted without knowledge of the results of the studied test. Finally, the time interval between each test should be as short as possible to avoid the clinical condition of the study subject to change, all patients must be assessed using the same reference standard and all patients should be analyzed.[16]

Diagnostic Accuracy Cross-sectional Studies

The diagnostic accuracy cross-sectional study (often referred as cohort study) involves comparing at least one diagnostic test to a comparator that is considered the reference standard for the target condition to be studied.[14,17] In this design, all the patients enrolled have the same set of inclusion and exclusion criteria, and they all are exposed to the same studied test and reference standard (Fig. 1). The statistics of the tests (sensitivity, specificity, positive and negative predictive values, and likelihood ratios) are then calculated and used to estimate the accuracy of the test.

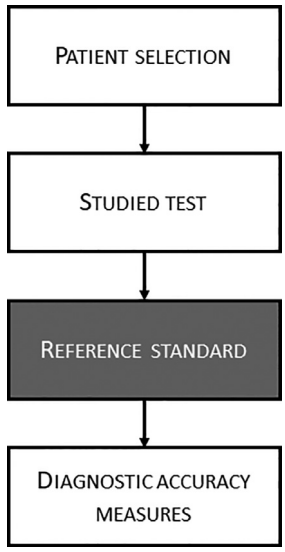

Figure 1 Diagnostic accuracy cross-sectional studies.

In our clinical scenario, to validate whether SPECT is accurate compared to the clinical brainstem function evaluation, one could design a study that includes patients with a brain injury severe enough to lead to brainstem death and exclude all patients with confounding factors that preclude a good clinical evaluation (such as the use of sedatives and paralytics, or the inability to complete the clinical evaluation). Identical inclusion and exclusion criteria should be used to enroll all patients in the study. These criteria should ensure that a mix of patients that have and do not have the target condition (absence of brainstem function) are enrolled and they should enroll patients with characteristics as close as possible to the target population to be studied.

All enrolled patients would then undergo the reference standard (complete clinical evaluation) to obtain their brainstem death status (deceased or not). As soon as possible after the clinical evaluation, to avoid any change in the clinical condition of the patient, scintigraphy would then be performed and interpreted blindly from the clinical evaluation by the nuclear medicine specialist. The results from this interpretation would then be compared with the reference standard and diagnostic test statistics calculated.

Strengths and Limitations

The diagnostic accuracy cross-sectional study design will allow us to accurately calculate the diagnostic accuracy of a new test when compared to a known accurate reference standard. When conducted properly, the estimates obtained from this study are at low risk of bias and can provide useful information about the accuracy and safety of a test compared to current standards.[17] However, it is impossible when using this design to compare the accuracy of the studied test to a different test that was not assessed in the same study. When doing so, the included patients are likely different from the

target population, and the tests were applied in different clinical settings or time frames. It is thus often hard, if not impossible, to compare the accuracy of two tests if not compared directly, similarly as it may be difficult to compare the accuracy of two drugs that have only been compared to placebo, but not to each other directly. It is also possible in some circumstance that the new test is more accurate than the reference standard (more sensitive or specific). When this is the case, it can be difficult or impossible to know for sure if the reference or the new test provided the correct diagnosis.

Diagnostic Accuracy Case-Control Studies

It is sometimes hard to find a unique set of criteria that will allow enrollment of patients in diagnostic test studies. Sometimes, researchers will use two sets of criteria to conduct a diagnostic accuracy study, one set of criteria to identify known cases, and one set of criteria to identify healthy controls. This design is often referred to as a diagnostic accuracy case-control study (Fig. 2).[14,17] Both group of patients will undergo the studies test and the reference standard. The group of known cases will be used to calculate sensitivity, and the group of healthy control (or negative cases) will be used to compute specificity and the results reported as one cohort. Although it may sound advantageous for research feasibility reasons, this design is at higher risk of bias than a diagnostic accuracy cross-sectional study or a comparative study.[17,18]

When applied to our clinical scenario, the researchers may decide to conduct a study that will enroll patients that have been already declared neurologically deceased by clinical criteria, then enroll patients who are comatose but are known to still have residual brainstem function in the control group. Scintigraphy would then be applied to all enrolled patients and accuracy statistics calculated. It is thus obvious that in this context, the patients with "borderline" criteria for neurological death would not be included in that study. It would therefore be unclear how the test would perform in this population, limiting its generalizability in clinical practice where these tests are mostly applied in such "borderline cases."

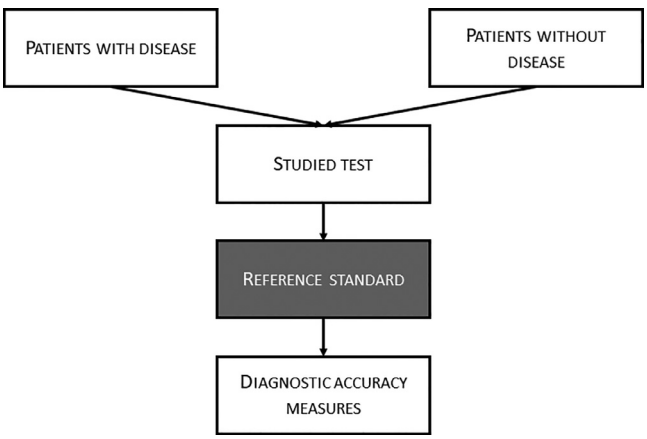

Figure 2 Diagnostic accuracy case-control studies.

Strengths and Limitations

This design is very similar to the diagnostic cross-sectional study except for patient selection. The decision to enroll patients with different inclusion criteria may simplify patient identification and optimize the number of patients to be enrolled by ensuring an equal mix of diseased and nondiseased patients. Doing so will however have important effects on the interpretation of the results. When designing a diagnostic accuracy study, researchers need to ensure to select a population that is representative of the population the test will be applied to in the clinical setting. Such is the case when using a unique set of inclusion criteria that represents the target population to obtain a mix of deceased and not deceased patients at different stages of the disease of interest, thus giving information about the accuracy of the test for a larger scope of patients. When using a case-control design, the researcher will select two different populations, thus restricting the scope of diseases included in the study. This will result in inflated sensitivity and specificity measures and lead to overestimation of the test accuracy. The additional challenges of selecting an appropriate unique set of inclusion criteria (and thus avoiding this design) may usually be worth the effort to ensure the most accurate and clinically meaningful results.

Diagnostic Accuracy Comparative Studies

Comparative diagnostic accuracy studies can, in addition to measuring the accuracy of a new diagnostic test, provide comparative accuracy between two tests, and can allow for the measurement of meaningful clinical outcomes.

There are three main design variations that will allow direct comparisons of diagnostic tests.[2,17] All designs use only one set of inclusion criteria to enroll patients. In the first situation, the included patients will all undergo the two tests to be compared, as well as the accepted reference standard (Fig. 3). It will then be possible to obtain accuracy statistics for each test by comparing them to the reference standard, as well as allow a direct comparison of the two tests. For this design to be feasible, the conduct of each individual test must not affect the course of the target disease, and the result of one of the tests must not affect the performance of the other.[16] This may happen for example when the two tests require the use of a contrast agent that will then affect the interpretation of the other test. This design is also more demanding for the patient as the patient may be exposed to several tests for the same target condition, exposing them to additional potential risks of adverse events. Also, the patients willing to consent and be enrolled in such a study may be different from those who decline, thus changing the characteristics of the enrolled patients and making the enrolled cohort of patients different from the intended target population.

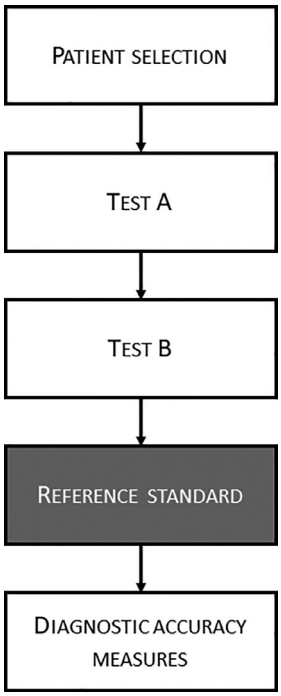

Figure 3 Nonrandomized comparative diagnostic accuracy study.

An alternative design involves randomizing the included patients into groups (Fig. 4). Each group will then be exposed to only one test and be compared to the reference standard. The interpretation of each test and the reference standard should be blinded to the study group like in standard randomized control trials. If done properly, this design is at low risk of bias. It also allows for the measure of patient outcomes by blinding the clinician to the study group of each patient and providing only the result of the studied test. A therapeutic intervention can then be applied based depending on the results of the test and clinical outcome measured. This design is however less statistically powerful as more patients need to be enrolled to show similar accuracy statistics between two proposed tests. In addition, it does not allow for pairwise comparisons between the two studied tests as no patients received both tests.

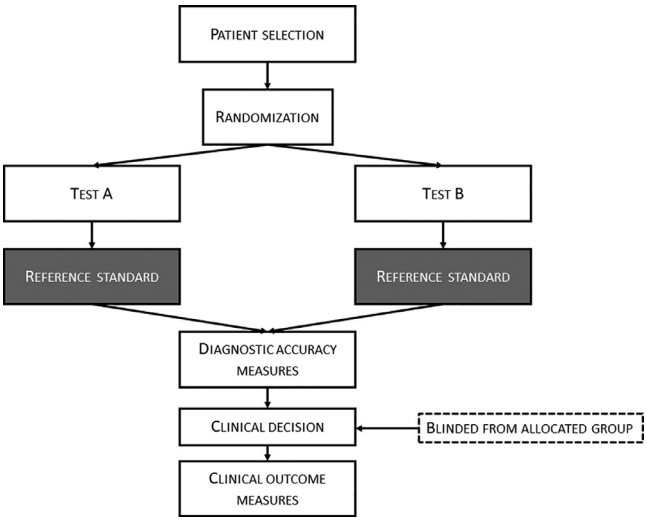

Figure 4 Randomized diagnostic accuracy study—option 1.

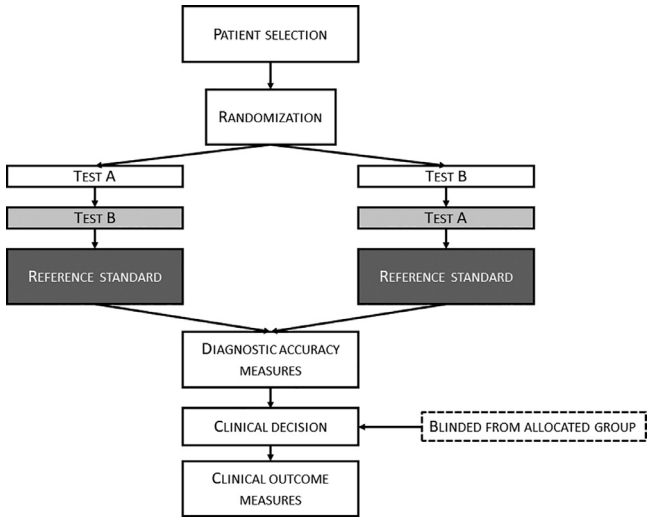

Figure 5 Randomized diagnostic accuracy study—option 2.

The last two comparative designs discussed can be merged into a single paradigm (Fig. 5). After inclusion, the patients are randomized into two different groups. All the studied tests and reference test are conducted in the two groups, blinded to each other. The clinician is also blinded to the study group and will obtain the result only from one test, depending on the randomization group. This design allows one to measure the accuracy of each test compared to the recommended reference, it will provide information regarding the comparative accuracy of the tests between each other, and when paired to a therapeutic intervention, enable the evaluation of patient outcome measures.

For the last two designs, each test is usually compared to a reference standard because most of the time, it is not known if a new test is as accurate as the reference standard. It is however in theory possible to randomize patients to one group where only the new test will be performed, and a control group where only the reference standard will be applied. However, for this to be ethically sound, the accuracy of the new test compared to the reference must be known and significant clinical equipoise must remain regarding the added clinical benefits of the new test compared to the reference.[1] If this is not the case, the patients could be placed in significant risk due to potential diagnostic inaccuracies of the new test. In summary, the randomized diagnostic studies are appropriate to assess simultaneously diagnostic accuracy of each test compared to the reference standard and the comparative accuracy of the studied tests. They also have the added benefit of allowing the measurement of patient outcomes.

Assessment of Clinical Outcomes

It is often assumed that if a test is as accurate as a reference standard, then that this test will provide similar information to the clinician and that the clinical outcome will also be similar. Each test may however have a different safety profile, they may provide additional clinically meaningful information or

provide it differently, or they may affect the clinical decision-making in ways not always easy to predict. By randomizing patients in two or more groups, and by blinding the clinicians to the study group (but not the result of the diagnostic test), it becomes possible to measure patient clinical outcomes. Randomization will distribute unmeasured confounding factors between groups, hopefully isolating the effect of the diagnostic test on the decision-making and clinical outcome, like in standard intervention studies. When possible, blinding will also greatly reduce the risk of biases associated with the clinician knowing the study group. A new test may, for example, be more sensitive for a target condition. Although this may seem desirable at first sight, this hypothetical test may only detect the disease at stages where it has no clinical consequences. The clinicians may decide to treat these preclinical conditions. One can imagine circumstances where a patient could be exposed to a potentially unnecessary treatment (which may have some risks), with no outcome improvement, thus increasing risks to the patient and costs to the health system. For example, a current controversy exists around this issue in the realm of imaging of pulmonary emboli. While tomographic SPECT imaging is more accurate and will detect additional and smaller emboli than planar imaging, there is a real concern that treating patients with more trivial disease will expose them to the risks of anticoagulation without any real clinical benefits.[19,20]

For our clinical scenario, in addition to the diagnostic accuracy of scintigraphy compared to the clinical brainstem evaluation, the clinicians were also interested in the diagnostic accuracy of CTA. Because neither scintigraphy nor CTA have been appropriately validated and considered suitable enough to be used as a reference standard, a comparative design could be selected. If feasible, comatose patients with no factors limiting a complete clinical brainstem evaluation could be enrolled and scintigraphy, CTA, and a clinical brainstem evaluation could be performed consecutively. This would allow the assessment of the comparative accuracy of both tests between each other and compared to the reference standard. One must be careful such that the conduct of the two tests should not be too much of burden for such patients. The included patients could then be randomized into two groups. One group would be assessed using CTA and the other group using scintigraphy, and all patients would undergo a formal clinical brainstem evaluation. In addition to standard accuracy statistics, the clinician could then be informed of only the result of the CTA or the result of scintigraphy, while kept blinded of the study group, and then use these results for clinical decision-making. In this example, the number of patients who undergo organ donation, the time from patient identification to organ donation, or clinician and family satisfaction could be compared between each group to provide meaningful clinical outcome, in addition to determination of the accuracy of the tests.

Discussion

To make the best medical decisions for their patients, clinicians rely heavily on the interpretation of diagnostic test studies. Unfortunately, the quality requirements of diagnostic studies are different than for medical interventions. Yet, the consequences of incorrectly interpreting the results of such tests have the potential to greatly affect patient care as it affects the medical diagnosis itself, and thus also affects all the therapeutic decisions that inevitably follow. We described a number of potential diagnostic accuracy study designs, each having strengths and weaknesses. We however argue that in most circumstances, randomized designs will provide all the required information to first assess the accuracy of a test, with the added benefit of reduced risk of biases, and can provide additional valuable information regarding patient outcomes.

Although it is possible to obtain a high-quality diagnostic cross-sectional study by selecting a representative sample of the target population using a unique set of inclusion criteria, by keeping the interpretation of the studied test and reference standard blinded to each other, and by following strict study flow and procedure, it is very hard to directly compare the accuracy of two tests if they were not assessed in the same study and it is nearly impossible to assess patient clinical outcomes. By including randomization in the study design, it becomes possible to directly measure clinical outcomes in addition to diagnostic accuracy. This enables the medical community to confirm if in addition to a comparable or better accuracy, these results translate into improved clinical outcome. It also becomes easier to compare different randomized studies one to each other by contrasting observed clinical outcomes of each trial.

It is important to note that the conduct of a diagnostic accuracy study that meets all quality criteria (including appropriate patient selection and study flow, and blinding) can be complex. The added complexity of adding randomization to this design may not increase significantly its complexity while significantly increasing its expected output. Randomization will inevitably reduce the power of the study and require an increased sample size to detect clinically meaningful outcomes. The study will therefore be more expensive and require additional data management. To improve efficiency, it has recently been suggested that rather than randomizing patients between two test groups, all patients would undergo both tests of interest. Patient would then undergo treatment A if both test results were positive and treatment B if both test results were negative. For pairs where the two tests would disagree, treatment assignment would be decided by randomization between the two treatment arms.[21]

The added benefits of conducting a randomized diagnostic test study may however be worth these limitations and they should be considered as the first design to consider when planning to study a new diagnostic test to be used in clinical practice.

Conclusion

There are a high number of new diagnostic tests made available each year, with most of these tests being associated with

higher costs, for unclear improved patient outcome. Traditional diagnostic accuracy case-control or cross-sectional study designs can be at high risk of bias (for the former) or at risk to provide a limited amount of clinically meaningful information (for the latter). It is critical that diagnostic test be considered like any other interventions. They should be studied using the same standards as new therapeutic interventions that not only ensure that the new test is "as accurate as" the former test, but also that its use is associated with improved patient outcomes. Randomized diagnostic test studies should be considered as the primary design to study new tests to ensure the best and appropriate care for our patients.

References

1. Rodger M, Ramsay T, Fergusson D: Diagnostic randomized controlled trials: The final frontier. Trials 13:137, 2012. https://doi.org/10.1186/1745-6215-13-137
2. Bossuyt PM, Lijmer JG, Mol BW: Randomised comparisons of medical tests: Sometimes invalid, not always efficient. Lancet 356:1844-1847, 2000. https://doi.org/10.1016/S0140-6736(00)03246-3
3. Shemie SD, Doig C, Dickens B, et al: Severe brain injury to neurological determination of death: Canadian forum recommendations. CMAJ 174: S1-S13, 2006. https://doi.org/10.1503/cmaj.045142
4. Gardiner D, Shemie S, Manara A, et al: International perspective on the diagnosis of death. Br J Anaesth 108 (suppl 1):i14-i28, 2012. https://doi.org/10.1093/bja/aer397
5. Shemie SD, Lee D, Sharpe M, et al: Brain blood flow in the neurological determination of death: Canadian expert report. Can J Neurol Sci 35:140-145, 2008. http://www.ncbi.nlm.nih.gov/pubmed/18574925. Accessed February 4, 2013
6. Roberts DJ, MacCulloch KA, Versnick EJ, et al: Should ancillary brain blood flow analyses play a larger role in the neurological determination of death? Can J Anaesth 57:927-935, 2010. https://doi.org/10.1007/s12630-010-9359-4
7. Lessard MR, Brochu JG: Challenges in diagnosing brain death. Can J Anaesth 57:882-887, 2010. https://doi.org/10.1007/s12630-010-9361-x
8. Greer DM, Strozyk D, Schwamm LH: False positive CT angiography in brain death. Neurocrit Care 11:272-275, 2009. https://doi.org/10.1007/s12028-009-9220-1
9. Taylor T, Dineen RA, Gardiner DC, et al: Computed tomography (CT) angiography for confirmation of the clinical diagnosis of brain death. Cochrane Database Syst Rev 3:CD009694, 2014. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=24683063
10. Joffe AR, Lequier L, Cave D: Specificity of radionuclide brain blood flow testing in brain death: Case report and review. J Intensive Care Med 25:53-64, 2010. http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=20095080
11. Suarez-Kelly LP, Patel DA, Britt PM, et al: Dead or alive? New confirmatory test using quantitative analysis of computed tomographic angiography. J Trauma Acute Care Surg 79:995-1003, 2015. https://doi.org/10.1097/TA.0000000000000831. discussion 1003
12. Heran MKS, Heran NS, Shemie SD: A review of ancillary tests in evaluating brain death. Can J Neurol Sci 35:409-419, 2008. http://www.ncbi.nlm.nih.gov/pubmed/18973057. Accessed February 4, 2013
13. Bossuyt PMM: Interpreting diagnostic test accuracy studies. Semin Hematol 45:189-195, 2008. https://doi.org/10.1053/j.seminhematol.2008.04.001
14. Bossuyt P, Leeflang M: The Cochrane Collaboration. Developing criteria for including studies. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.4 [Updated September 2008]. The Cochrane Collaboration, 1-7, 2008
15. Chassé M, Glen P, Doyle M-A, et al: Ancillary testing for diagnosis of brain death: A protocol for a systematic review and meta-analysis. Syst Rev 2:100, 2013. https://doi.org/10.1186/2046-4053-2-100
16. Whiting PF, Rutjes AWS, Westwood ME, et al: QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 155:529-536, 2011. https://doi.org/10.1059/0003-4819-155-8-201110180-00009
17. Leeflang MM, Davenport CF, Takwoingi YDJ: Eligibility designs. Lesson 2.2: Cochrane Collaboration DTA Online Learning Materials. 2014 The Cochrane Collaboration. http://training.cochrane.org. Published Accessed August 1, 2018
18. Whiting PF, Rutjes AWS, Westwood ME, et al: Quadas-2: A Quality Assessment Tool for Diagnostic Accuracy Studies. University of Bristol, http://www.bris.ac.uk/quadas/resources/ Published 2018
19. Le Gal G, Righini M, Parent F, et al: Diagnosis and management of subsegmental pulmonary embolism. J Thromb Haemost 4:724-731, 2006. https://doi.org/10.1111/j.1538-7836.2006.01819.x
20. Yoo HHB, Queluz THAT, El Dib R: Anticoagulant treatment for subsegmental pulmonary embolism. Cochrane Database Syst Rev 2016: CD010222. https://doi.org/10.1002/14651858.CD010222.pub3
21. Lu B, Gatsonis C: Efficiency of study designs in diagnostic randomized clinical trials. Stat Med 32:1451-1466, 2013. https://doi.org/10.1002/sim.5655