Review Article

iMedPub Journals
www.imedpub.com

Journal of Pediatric Care
ISSN 2471-805X

2020
Vol.6 No.S1: 02

What Criteria should be met before a New Diagnostic Test is Deployed in Clinical Practice?

Marion Rowland*

Department of Pediatrics, UCD School of Medicine, University College Dublin, Catherine McAuley Research Centre, Dublin, D07 A8 NN, Ireland

***Corresponding author:** Marion Rowland, Department of Pediatrics, UCD School of Medicine, University College Dublin, Catherine McAuley Research Centre, Dublin, D07 A8 NN, Ireland, E-mail: marion.rowland@ucd.ie

Received date: November 23, 2020; **Accepted date:** December 07, 2020; **Published date:** December 14, 2020

Citation: Rowland M (2020) What Criteria Should be Met before a New Diagnostic Test is Deployed in Clinical Practice? J Pediatric Care Vol.6 No.S1:2

Abstract

Diagnostic tests are an important resource which facilitate clinical decision making with appropriate therapeutic options so that health outcomes for children are optimised. Diagnostic tests must be accurate (distinguish those with disease from the healthy population) and must have precision (that on average two measurements taken over a short period of time will be the same). However, there are no minimal performance criteria for the introduction of non-invasive tests into clinical practice. Despite efforts to improve the reporting of the evaluation of diagnostic tests much remains to be achieved. Many diagnostic tests are introduced into paediatric practice based on adult data and even when new non-invasive diagnostic instruments are evaluated in children the precision dimension is often poorly evaluated. Transient Elastography (TE) is a new non-invasive test for the diagnosis of liver disease in adults and children. We use our experience with TE to highlight the care which must be taken before adopting a new non-invasive test in paediatrics.

Keywords: Children; Transient elastography; Diagnostic accuracy; Repeatability

Abbreviations: TE: Transient Elastography; CF: Cystic Fibrosis

Introduction

Poorly performing diagnostic tests can negatively impact on patient safety and waste scarce healthcare resources [1]. Understanding how well a new test performs in different populations or different clinical situations is central to ensuring that we make the correct diagnosis and plan appropriate treatment for our young patients.

Early diagnosis of chronic liver diseases remains a challenge for the paediatric hepatologists. Despite significant investment in clinical diagnostics the search for the definitive non-invasive test which can distinguish early stage fibrosis or fatty infiltration from marked fibrosis or cirrhosis, remains elusive. While hepatologists have several investigations at their disposal each has limitations which must be carefully considered. Liver biopsy remains the gold standard for the diagnosis of liver disease and against which non-invasive tests are assessed. In addition to well described potential life-threatening complications associated with liver biopsy, sampling variation, when the disease process does not uniformly affect the liver, leads to uncertainty about the diagnosis or disease stage. Ultrasonography while inexpensive and widely available is a real-time examination which is operator dependant. CT scans and MRI have well described constraints for use in children, and the repeated radiation exposure of CT present significant risks which often outweigh the benefits of repeated examinations.

Literature Review

Recent advances in ultrasound-based measurement of liver stiffness have offered the possibility of a breakthrough in the early identification liver fibrosis. The most widely used of these is Transient Elastography (TE) (Fibroscan®, Echosens, Paris, France) which is a rapid non-invasive point of care test. Liver stiffness measurement serves as a surrogate marker of the degree of liver fibrosis and is now widely used to evaluate liver disease in adults and children in clinical settings and research studies.

What does the clinician want when using new or an established diagnostic test? In the first instance the test must distinguish those with the disease from those who do not have the disease (accuracy) and must provide the same results over a short period of time (precision).

The terminology to describe the performance of a clinical instrument can be confusing, and the literature abounds with a range of different terms to describe the two important facets of any diagnostic test or instrument [2-4]. When a new instrument or test is evaluated, it is important to determine both its diagnostic accuracy (that the instrument can clearly distinguish those with disease from the healthy population) and precision (that on average two or more measurements taken over a short period of time will be the same) Diagnostic accuracy is reported as sensitivity and specificity, area under the receiver operating

characteristic curve, likelihood ratios or predictive values. Measures of precision include reliability, repeatability, reproducibility, agreement or observer variation. Precision studies are usually conducted early in the introduction of a new test or technology, [2,3] and can be improved with training and good protocols, while some degree of clinical variability is inevitable.

There is an extensive literature on diagnostic accuracy studies of TE in children with a variety of underlying liver diseases. However the degree of heterogeneity of studies in terms of age of children, probe size, underlying liver disease, the cut-points used to determine the degree of liver fibrosis and how it is assessed (METAVIR score or Ishak Score) makes it difficult to draw summary conclusions about the accuracy of TE using systematic review or meta-analysis techniques. Possibly, the most informative studies of TE in children was performed by Lee et al. [5,6] who used a 2 study approach, with liver biopsy as the gold standard, to determine optimal cut-points of liver stiffness measurement which would discriminate advanced liver fibrosis (Metavir F3-F4) and cirrhosis (F4). In the first study they demonstrated that an optimal TE cut-point for F3-F4 was >8.6 kPa ((sensitivity 79.4% (95% CI 62.1-91.3) specificity 82.5% (95% CI 70.9-90.9)) while F4 the cut-point was 11.5 kPa ((sensitivity 83.3% (95% CI 58.6-96.4) specificity 83.5% (95% CI 73.5-90.9)) [5]. In the second study they validated the cut-points outlined above in a further cohort of children who also had liver biopsy for the diagnosis of different underlying liver diseases [6]. The sensitivity and specificity declined for both cut-points. For the cut-point of 8.6 kPa the sensitivity was 70.8% (95% CI 55.9-83.0), specificity 65.6% (95% CI 56.4-73.9) while for the cut-point of 11.5 kPa the sensitivity was 78.9% (95% CI 54.4-93.9) and specificity 74.8% (95% CI 67.1-81.5) [6]. The results of this carefully conducted study of TE, raises questions about the clinical deployment of TE for the diagnosis or monitoring liver disease in children. The authors considered a wide range of potential biases in their study including disease heterogeneity, selection bias as not all consecutive patients were enrolled, the exclusion of 17% of participants due to invalid TE measurements, and that they had not been in a position to follow the manufacturers new guidelines on fasting which had been extended to three hours. However, the authors provide no data on the precision of TE in their centre [5,6].

In our experience with TE as a research instrument we demonstrated that TE lacked precision in healthy children [7]. To facilitate early diagnosis and monitor liver disease progression in Cystic Fibrosis (CF) we first sought to confirm that any change in TE measurements over time could be attributed to a change in the underlying severity of liver disease in persons with CF. We assumed when planning our investigation that this would be a straightforward preliminary exercise in a much larger programme of research. However, be found that TE lacked precision in healthy volunteers. There was a difference of >1 kPa between measurements in 61/235 (25.9%) children who had measurements performed at least 24 hours apart. Using the 95% Limits of Agreement [8,9]. We reported that the range within which 95% of the differences between two measurements lay was -0.8 to +0.76 kPa [7]. Despite careful evaluation of our protocols and completing two measurements in over 250 healthy volunteers we failed to demonstrate sufficient precision to introduce TE, as a research instrument, to our national prospective follow-up study of CFLD. We only included children over 7 years of age, and used a portable TE machine with a M probe in all children [7]. We did not require children to fast for 3 hours according to revised recommendations by the manufacturer. The research evidence in children does not support the recommendation of prolonged fasting to improve the precision of TE [6,10].

What do previously published studies tell us in terms of the validation of TE as a diagnostic tool and specifically about precision? In children, despite the widespread reporting of the diagnostic accuracy of TE in a variety of pediatric liver diseases [11,12] only two previous studies examined precision. The findings in these studies, which are consistent with our study indicate that precision is poor unless ultrasound guidance [13] or a marked point on the skin [10] to position the TE probe was used. In adults only two studies examined precision, in the first the number of research participants with repeat examinations was small (n=15) [14] while the second identified that the performance of TE was poor when in those with milder forms of liver disease [15] More importantly other adult studies have suggested that variability in repeat TE measurements could result in a change in liver disease classification without any change in underlying pathology [16-18] There is, however, an extensive adult literature on instrument precision (10 valid measurements and Median/IQR ratios of <30%) but this should not be confused with observer or participant variation [19].

In clinical practice measurement variation is inevitable but the degree of variation that can be deemed acceptable is determined by what constitutes a clinically important difference between measurements, and should be determined before a study of precision commences. It has been suggested that in our study the differences between repeat measurements are small and therefore not clinically important. Normal TE values in healthy children range between 2.45 kPa to 5.56 kPa [20]. The differences between the first and second measurement in our study was greater than 1 kPa in 25% of healthy children. While these differences appear inconsequential, they could result in a change in liver disease classification based on TE measurements without any change in underlying pathology, as noted previously in adult studies [16-18]. In the small number of children included in this study, who had Clinically Significant Liver Disease (CFLD) the difference between repeated measurements were much greater over a very short period of time. This random variation in paired measurements underlines the lack of precision, which may explain the differences reported by Lee et al. in their validation cohort [5,6].

There are many statistical techniques to compare the precision of diagnostic instruments. Currently Bland and Altman limits of agreement is the method most widely used and is based on examining the spread or standard deviation of the paired differences between two or more measurements. A histogram is first used to examine the distribution of the paired differences. A wide distribution signifies a wide standard deviation which demonstrates lack of agreement between observers or tests [8,9]. We demonstrated that while there was

no statistically significant difference in the paired means of 2 TE measurement in 235 healthy volunteers (mean difference -0.044 kPa p=NS), and the distribution of paired differences followed a normal distribution, the standard deviation was wide SD=0.414. When we plotted the difference of TE measurements against the mean of paired measurements the wide scatter of paired measurements is apparent [7]. The range of minus 0.85 kPa to plus 0.76 kPa is a clinically important difference when the normal range of TE measurements in healthy children is 2.45-5.56 kPa. Further research is required to improve the precision of TE in children.

All aspects of diagnostic validation studies must be conducted with children. The prevalence of the disease determines the accuracy of the test, and therefore diagnostic accuracy in children cannot be inferred from adult studies. Diagnostic tests are used in different situations and for different reasons. For example, screening test have different requirements from tests used in tertiary referral centres. The degree of diagnostic accuracy required depends on the setting, the confirmatory test and the therapeutic options available for serious diseases. However, all tests require precision. Currently there are no minimal performance criteria for the introduction of non-invasive tests into clinical practice [21,22] and there is significant evidence of over interpretation of results of diagnostic accuracy studies without due consideration of the risk of bias in most studies [1]. Despite efforts to improve the validation and reporting of the evaluation of new diagnostic tests, much remains to be done [23].

Conclusion

While this review has focused on liver disease there are both invasive and non-invasive diagnostic tests and technologies in other areas of clinical practice which are reported in such a way as to lead to unjustified optimism about the test performance. Developing new diagnostic instruments is expensive and time consuming, and many will be less than optimal. However, clinicians must be willing to critically evaluate new technologies, ensuring that the reported metrics for accuracy and precision are achieved in clinical practice. Diagnostic instruments will have limitations, but we must understand how these limitations impact on clinical decision making. We must strive for the best diagnostic tools for our patients, and we must ensure that clinical decision making does not put patient safety at risk, or waste scarce healthcare resources.

Funding Source

The Health Research Board (Grant HRB-PHS-662). The funder had no role in the study or in the preparation of this review.

References

1. McGrath TA, McInnes MDF, Van Es N, Leeflang MMG, Korevaar DA, et al. (2017) Overinterpretation of research findings: Evidence of "Spin" in systematic reviews of diagnostic accuracy studies. Clin Chem 63: 1353-1362.
2. McAlinden C, Khadka J, Pesudovs K (2015) Precision (repeatability and reproducibility) studies and sample-size calculation. J Cataract Refract Surg 41:2598-2604.
3. Lord SJ, Staub LP, Bossuyt PM, Irwig LM (2011) Target practice: Choosing target conditions for test accuracy studies that are relevant to clinical practice. BMJ 343: d4684.
4. Harris EF, Smith RN (2009) Accounting for measurement error: A critical but often overlooked process. Arch Oral Biol 54: S107-117.
5. Lee CK, Perez-Atayde AR, Mitchell PD, Raza R, Afdhal NH, et al. (2013) Serum biomarkers and transient elastography as predictors of advanced liver fibrosis in a United States cohort: The Boston children's hospital experience. J Pediatr 163: 1058-1064.
6. Lee CK, Mitchell PD, Raza R, Harney S, Wiggins SM, et al. (2018) Validation of transient elastography cut points to assess advanced liver fibrosis in children and young adults: The Boston children's hospital experience. J Pediatr 198: 84-89.
7. Rowland M, McGee A, Broderick A, Drumm B, Connolly L, et al. (2020) Repeatability of transient elastography in children. Pediatr Res 88: 587-592.
8. Bland JM, Altman DG (1999) Measuring agreement in method comparison studies. Stat Methods Med Res 8: 135-160.
9. Altman DG (2009) Assessing new methods of clinical measurement. Br J Gen Pract 59: 399-400.
10. Goldschmidt I, Streckenbach C, Dingemann C, Pfister ED, Nanni AD, et al. (2013) Application and limitations of transient liver elastography in children. J Pediatr Gastroenterol Nutr 57(1): 109-13.
11. Hwang JY, Yoon HM, Kim JR, Lee JS, Jung AY, et al. (2018) Diagnostic performance of transient elastography for liver fibrosis in children: A systematic review and meta-analysis. AJR Am J Roentgenol 211: W257-W66.
12. Lam S, Nettel-Aguirre A, Van Biervliet S, Roeb E, Sadler MD, et al. (2019) Transient elastography in the evaluation of cystic fibrosis-associated liver disease: Systematic review and meta-analysis. J Can Assoc Gastroenterol 2: 71-80.
13. Nobili V, Vizzutti F, Arena U, Abraldes JG, Marra F, et al. (2008) Accuracy and reproducibility of transient elastography for the diagnosis of fibrosis in pediatric nonalcoholic steatohepatitis. Hepatology 48: 442-448.
14. Sandrin L, Fourquet B, Hasquenoph JM, Yon S, Fournier C, et al. (2003) Transient elastography: A new noninvasive method for assessment of hepatic fibrosis. Ultrasound Med Biol 29: 1705-1713.
15. Fraquelli M, Rigamonti C, Casazza G, Conte D, Donato MF, et al. (2007) Reproducibility of transient elastography in the evaluation of liver fibrosis in patients with chronic liver disease. Gut 56: 968-973.
16. Nascimbeni F, Lebray P, Fedchuk L, Oliveira CP, Silva MRA, et al. (2015) Significant variations in elastometry measurements made within short-term in patients with chronic liver diseases. Clin Gastroenterol Hepatol 13(4): 763-771.
17. Vignier N, Esmat G, Elsharkawy A, Hassany M, Bonnard P, et al. (2011) Reproducibility of liver stiffness measurements in hepatitis C virus (HCV)-infected patients in Egypt. J Viral Hepat 18: e358-365.
18. Vuppalanchi R, Sanyal AJ (2015) Myths and mysteries about staging hepatic fibrosis by fibroscan. Clin Gastroenterol Hepatol 13: 780-782.

19. Boursier J, Zarski JP, de Ledinghen V, Rousselet MC, Sturm N, et al. (2013) Determination of reliability criteria for liver stiffness evaluation by transient elastography. Hepatology 57: 1182-1191.

20. Li DK, Rehan Khan M, Wang Z, Chongsrisawat V, Swangsak P, et al. (2020) Normal liver stiffness and influencing factors in healthy children: An individual participant data meta-analysis. Liver Int 40: 2602-2611.

21. Majumdar A, Campos S, Gurusamy K, Pinzani M, Tsochatzis EA (2020) Defining the minimum acceptable diagnostic accuracy of non-invasive fibrosis testing in cirrhosis: A decision analytic modeling study. Hepatology 71: 627-642.

22. Steadman R, Myers RP, Leggett L, Lorenzetti D, Noseworthy T, et al. (2013) A health technology assessment of transient elastography in adult liver disease. Can J Gastroenterol 27: 149-158.

23. Korevaar DA, Wang J, van Enst WA, Leeflang MM, Hooft L, et al. (2015) Reporting diagnostic accuracy studies: Some improvements after 10 years of stard. Radiology 274: 781-789.