

From the authors:

We thank very much I. Amirav and P.M. Boussuyt for their interest in our manuscripts [1, 2]. We agree that our two studies have some limitations, caused by the lack of a “gold standard” test for diagnosing primary ciliary dyskinesia (PCD), but we strongly disagree with their claim that we “did not notify the readers of these design deficiencies”. Instead, we had taken great care to highlight these uncertainties and risks of bias.

Our first manuscript evaluated the accuracy of different tests used to diagnose PCD: nasal nitric oxide, high-speed video microscopy analysis (HSVMA) and transmission electron microscopy (TEM), used alone or in combination, against an expert diagnostic consensus based on information from all available tests [1]. This was clearly described in the methods section; and several sensitivity analyses tested the robustness of our results by varying different parameters. Most notably, we began the discussion section by describing the lack of a single “gold reference standard” (second paragraph) as the major limitation of our study [1]. We highlighted that we used a surrogate reference standard, which was an expert multidisciplinary consensus, based on results from all available diagnostic tests. We also cautioned that, since each test contributed to the final diagnostic decision, our sensitivity and specificity estimates of the single tests might be overestimated.

Their correspondence is helpful as it highlights the challenges of investigating diagnostic accuracy for diseases where there is no “gold standard”. This is typical for many diseases including rare ones like PCD [3] and common ones like asthma. It complicates research, but should not impede it; or else we will never progress. Guidance recommends that in situations without a “gold standard” researchers can consider constructing a reference standard from multiple test results or use an imperfect reference [4]. In our clinical practice, a multidisciplinary panel of specialists considers the results of multiple tests to develop a consensus diagnosis of PCD, based on pre-determined rules. In the paper, we used this composite diagnostic outcome as the study’s reference standard. Pre-specified rules for deciding the composite diagnostic outcome make the method transparent and easy to use.

An alternative approach is to use an imperfect reference standard and then adjust the calculated sensitivity and specificity based on existing data about the imperfections [4]. For example, we could have used TEM alone as a single reference standard and calculated the accuracy of the other tests compared with TEM, taking into consideration that the latter has excellent specificity but limited sensitivity (70–90%). For completeness, we did this for TEM and for HSVMA (see online supplementary table S2 in [1]), but given that this was not our primary approach and that pre-existing data about the degree of imperfection were highly variable, we decided not to make any adjustments.

As discussed in our manuscript [1] and the accompanying editorial by HAARMAN and SCHMIDTS [5], the generalisability of our findings should be considered with caution. Accurate analysis of ciliary function by HSVMA and of ultrastructure by TEM depends on expertise. The UK PCD reference centres regularly audit each other’s analyses and discuss difficult cases [6]. We routinely analyse *de novo* cilia following culture of the original sample at the air–liquid interface allowing us to differentiate primary and secondary functional and structural defects [7]. These methods are technically demanding and not available in many centres. In centres without similar infrastructure subtle cases of PCD are more likely to be missed and secondary defects incorrectly attributed to PCD; even with these facilities, we may misdiagnose some patients. Introduction of new tests is likely to advance the accuracy of diagnostic decisions; we did not include data on genotype and immunofluorescence because these methods were introduced relatively recently and were not available during the study period.

The second manuscript that Amirav and Boussuyt refer to describes the PICADAR tool, a seven-item prediction rule aimed at identifying patients needing referral for diagnostic testing [2]. The tool uses information on clinical symptoms at presentation to predict the likelihood (or risk) of a final PCD diagnosis. As diagnostic outcome, we used the same composite reference standard as in the first study [1] based on results from all available diagnostic tests. A standardised clinical history was taken from all referred patients, prior to performing any of the diagnostic tests. As detailed in the manuscript, diagnosis was based on positive test results, but not on the clinical history. Symptoms were not part of the composite diagnosis, so there was no incorporation bias in this study. In addition to the published analyses, as strategy to assess potential model overfitting, we performed bootstrapping testing of the receiver operator characteristic (ROC)



curves of the derivation population, which indicated an expected shrinkage of <3% (data available from the authors). This suggests that there is no significant overestimation for the ROC curves as produced within the predictive logistic regression models (results not shown). We further published the external validation using an independent patient cohort [2], but pointed out that PICADAR should be further validated, and if necessary modified, in different study populations, in general respiratory clinics and different countries.

In summary, we agree that current methods for diagnosing PCD and for assessing diagnostic accuracy are imperfect and require further development and scrutiny. However, we confirm that our manuscripts, which were very transparent about diagnostic pathways, diagnostic and statistical methods, and their limitations, are an acceptable way forward while we establish better methods. The two publications are not intended as the definitive answer, but as significant steps in the plight to develop an evidence base to diagnostic testing. Waiting for a “gold standard” will not allow us to move forward in the near future.



@ERSpublications

Keeping up the momentum to develop an evidence base for the diagnosis of PCD

<http://ow.ly/o6zq300uhcw>

Jane S. Lucas^{1,2,3}, Borislav D. Dimitrov^{3,4}, Laura Behan^{1,2,3} and Claudia E. Kuehni⁵

¹Primary Ciliary Dyskinesia Centre, University Hospital Southampton NHS Foundation Trust, Southampton, UK.

²Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK. ³NIHR Southampton Respiratory Biomedical Research Unit, University of Southampton and University Hospital Southampton NHS Foundation Trust, Southampton, UK. ⁴Primary Care and Population Sciences, Faculty of Medicine, University of Southampton, Southampton, UK. ⁵Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland.

Correspondence: Jane S. Lucas, University of Southampton, Clinical and Experimental Sciences, Mailpoint 803, University Hospital Southampton, Tremona Road, Southampton, SO16 6YD, UK. E-mail: jlucas1@soton.ac.uk

Received: May 06 2016 | Accepted after revision: May 09 2016

Conflict of interest: None declared.

References

- 1 Jackson CL, Behan L, Collins SA, *et al.* Accuracy of diagnostic testing in primary ciliary dyskinesia. *Eur Respir J* 2016; 47: 837–848.
- 2 Behan L, Dimitrov BD, Kuehni CE, *et al.* PICADAR: a diagnostic predictive tool for primary ciliary dyskinesia. *Eur Respir J* 2016; 47: 1103–1112.
- 3 Lucas JS, Leigh MW. Diagnosis of primary ciliary dyskinesia: searching for a gold standard. *Eur Respir J* 2014; 44: 1418–1422.
- 4 Rutjes A, Reitsma J, Coomarasamy A, *et al.* Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technol Assess* 2007; 11: iii, ix–51.
- 5 Haarman EG, Schmidts M. Accuracy of diagnostic testing in primary ciliary dyskinesia: are we there yet? *Eur Respir J* 2016; 47: 699–701.
- 6 Lucas JS, Burgess A, Mitchison HM, *et al.* Diagnosis and management of primary ciliary dyskinesia. *Arch Dis Child* 2014; 99: 850–856.
- 7 Hirst RA, Jackson CL, Coles JL, *et al.* Culture of primary ciliary dyskinesia epithelial cells at air-liquid interface can alter ciliary phenotype but remains a robust and informative diagnostic aid. *PLoS One* 2014; 9: e89675.

Eur Respir J 2016; 48: 960–961 | DOI: 10.1183/13993003.00909-2016 | Copyright ©ERS 2016