Cyprus University of Technology

Faculty of Health Sciences

Cyprus International Institute for Environmental and Public Health

Advisor: Lecturer Stefania I. Papatheodorou

Development of Evidence Based Diagnostic Algorithm for Primary Ciliary Dyskinesia

PhD Thesis

Panayiotis Kouis

June 2017

Approval Page

PhD Dissertation

Development of Evidence Based Diagnostic Algorithm for Primary Ciliary Dyskinesia

Presented by

Panayiotis Kouis

Academic Supervisor _____

Stefania I. Papatheodorou

Committee member_____

Constantinos S. Pattichis

Committee member_____

Paris A. Skourides

Cyprus University of Technology

July 2017

Acknowledgments

This project was carried out within the framework of the PhD program in Environmental Health of Cyprus International Institute for Environmental and Public Health (Cyprus University of Technology) between September 2013 and July 2017. The work of this project was supported by FP7 EU 7th Framework Program EC-GA No. 305404 BESTCILIA and Cyprus University of Technology Starting Grant awarded to Dr Stefania Papatheodorou.

I would like to thank all the people who contributed to this study and especially:

- The academic supervisor of this work, Dr Stefania Papatheodorou for her supervision and guidance throughout the duration of the study.
- Professor Panayiotis K. Yiallouros, who had the original idea, co-supervised the work and secured most of the funding towards the implementation of the project.
- Professor John S. Evans who provided precious feedback and suggestions throughout the duration of the study.
- Dr Kyriacos Kyriacou, for his collaboration and important hands-on training and laboratory support he provided for the acquaintance with PCD diagnosting testing.
- Dr Nicos Middleton, for his precious comments from the beginning of this study.
- My fellow PhD candidates at Cyprus International Institute and especially Ms Despo Pampaka for the continuous encouragement and collaboration in the past 4 years.
- Finally, I owe special thanks to my family for their constant support. I would like to especially thank Zacharias for assisting me in the preparation of most of the original pictures presented in this study and Popi for her unceasing patience.

Abstract

Introduction

Primary Ciliary Dyskinesia (PCD) is a rare, genetically heterogeneous disorder which results from the dysfunction of small hair-like organelles, called motile cilia. Motile cilia project from the apical side of epithelial cells that line up the upper and lower airways (respiratory cilia) and can be found in a variety of other tissues. PCD patients usually suffer from recurrent respiratory infections which lead to chronic destructive airway disease characterized by progressive loss of lung function and structural damage of the airways (bronchiectasis).

Despite the fact that many of the manifestations of PCD present early in life, diagnosis is often delayed or missed completely, primarily due to the low specificity of some symptoms (e.g. cough, rhinorrhea), lack of awareness for PCD among clinicians and difficulties in the availability and interpretation of specialised diagnostic testing. Diagnostic testing for PCD usually involves at least three laboratory procedures: (a) nasal Nitric Oxide measurement, (b) assessment of ciliary motility and (c) examination of ciliary ultrastructure. Diagnostic testing for PCD is laborious and time consuming and many centers may lack access to necessary equipment or expertise to perform all required tests. As a result, different diagnostic algorithms for PCD diagnosis may be followed by different centers and this phenomenon is further influenced by the lack of knowledge regarding the diagnostic effectiveness and average cost of each test.

Aims

Towards further illuminating the decision making process for the establishment of the most efficacious diagnostic algorithm for PCD, we aimed first to characterize the diagnostic properties of the three main tests for PCD (nNO, TEM and HSVM) and second to evaluate different diagnostic algorithms in terms of overall health benefits for PCD patients and overall costs to the healthcare systems.

Methods

In separate systematic reviews and meta-analyses, all major electronic databases were searched from inception until 2016 using appropriate terms towards identifying eligible studies that reported estimates of diagnostic accuracy for TEM, nNO and HSVM as well as estimates of the prevalence of PCD in consecutive referrals of suspect cases. Eligible studies

included diagnostic information on PCD patients or PCD referrals that underwent a combination of diagnostic tests which included nNO, TEM, HSVM and genetic testing.

For the meta-analysis of nNO diagnostic accuracy, estimates of sensitivity and specificity of nNO measurement was calculated for each included study and a two-level mixed logistic regression model conditional on the sensitivity and the specificity of each study and a bivariate normal model for the sensitivity and specificity between studies were fitted. Summary receiver operating characteristic (HSROC) curves were drawn using the parameters of the fitted models separately depending on the breathing technique, Vellum Closure (VC) or non-Vellum Closure (non-VC), used for nNO measurement. For the meta-analysis of PCD prevalence in consecutive referrals of suspect cases and for the meta-analysis of TEM detection rate, a meta-analysis of proportions using a random effects model was performed. Meta-analysis of proportions allows the calculation of the pooled proportion across studies containing binomial data while random effects allow for each study to be assigned a weight which includes the within study variance and the between studies variance. Heterogeneity was assessed with the I^2 which describes the proportion of total variation in the effect estimate that results from the between-studies heterogeneity and ranges from 0 to 100%.

The evidence regarding the diagnostic properties of nNO and TEM as well as the evidence regarding the prevalence of PCD among suspect patients were combined along with diagnostic accuracy estimates for HSVM from individual studies to develop a probabilistic decision model that allowed the calculation of net sensitivity and specificity as well as the cost-effectiveness (CE) and incremental cost effectiveness for three diagnostic algorithms that were characterized by different combinations of nNO, TEM and HSVM. The evaluated combinations were (a) nNO+TEM in sequence, (b) nNO+HSVM in sequence and (c) nNO/HSVM in parallel followed, in cases with conflicting results, by confirmatory TEM (nNO/HSVM+TEM) and the model followed a hypothetical initial population of 1000 referrals (expected 320 PCD patients). Number of PCD patients identified, CE and ICE ratios were calculated using Monte Carlo analysis in ANALYTICA.

Results

PCD prevalence among referrals was 32% (95% CI: 25–39%, I^2 = 92%). TEM detection rate among PCD patients was 83% (95% CI: 75–90%, I^2 = 90%). Exclusion of studies reporting isolated inner dynein arm defects as PCD, reduced TEM detection rate and explained an important fraction of observed heterogeneity (74%, 95% CI: 66–83%, I^2 = 66%).

The overall sensitivity of nNO measured by VC techniques was 0.95 (95 % CI 0.91–0.97), while specificity was 0.94 (95 % CI 0.88–0.97). The positive likelihood ratio (LR+) of the test was 15.8 (95 % CI 8.1–30.6), whereas the negative likelihood ratio (LR-) was 0.06 (95 % CI 0.04–0.09). For non-VC techniques, the overall sensitivity of nNO measurement was 0.93 (95 % CI 0.89–0.96) whereas specificity was 0.95 (95 % CI 0.82–0.99). The LR+ of the test was 18.5 (95 % CI 4.6–73.8) whereas the LR- was 0.07 (95 % CI 0.04–0.12).

Regarding the probabilistic decision analysis model, out of 320 PCD patients, 311 were identified by nNO/HSVM+TEM, 274 with nNO+HSVM and 198 with nNO+TEM. The nNO/HSVM+TEM had the higher mean cost (€97K) followed by nNO+TEM (€56K) and nNO+HSVM (€39K). The nNO+HSVM algorithm dominated the nNO+TEM algorithm (less costly and more effective). The ICE ratio for nNO/HSVM+EM was €1600 per additional PCD patient identified.

Conclusions

Many centers for the diagnosis and treatment of PCD in the developed world follow different tests and a variety of algorithms for diagnosing PCD. In some low income countries, most likely, there is a complete lack of specialized diagnostic testing. The results of this PhD thesis suggest that diagnostic accuracy of nNO measurement both with VC and non-VC maneuvers is high and can be effectively employed in the clinical setting to detect PCD even in young children, thus potentiating early diagnosis. On the contrary, a significant percentage, at least as high as 26%, is missed by TEM and this limitation that should be accounted toward the development of an efficacious PCD diagnostic algorithm. The results of decision analysis approach employed in this study also suggest that a diagnostic algorithm which includes nNO during VC as a screening test followed by confirmatory HSVM identifies approximately 86% of PCD patients with a mean CER of 140€ per PCD case identified. The algorithm which maximizes the number of PCD patients identified involves parallel performance of nNO and HSVM as the first step, followed by TEM as a confirmatory test for the few cases where nNO and HSVM yield conflicting results, with a corresponding ICER of 1620€ per additional PCD patient identified. These findings can inform the dialogue about the development of evidence-based guidelines for PCD diagnostic testing and can illuminate discussions about how these guidelines can best be implemented across various healthcare systems.

Περίληψη

Εισαγωγή

Η Πρωτοπαθής Δυσκινησία των Κροσσών (ΠΔΚ) είναι μια σπάνια, γενετικά ετερογενής διαταραχή, η οποία προκύπτει από τη δυσλειτουργία των κινητών κροσσών. Οι κινητοί κροσοί αποτελούν μικρά τριχοειδή οργανιδία που προεκβάλουν από την κορυφαία μεμβράνη των επιθηλιακών κυττάρων που καλύπτουν τους ανώτερους και κατώτερους αεραγωγούς (αναπνευστικοί κροσσοί) αλλά οι δομές αυτές μπορεί να βρεθούν και σε άλλους ιστούς. Οι ασθενείς με ΠΔΚ υποφέρουν από επαναλαμβανόμενες λοιμώξεις του αναπνευστικού οι οποίες οδηγούν σε χρόνιες πνευμονοπάθειες που καταστρέφουν τους αεραγωγούς και χαρακτηρίζονται από προοδευτική απώλεια λειτουργίας του πνεύμονα και καταστροφή της δομής των αεραγωγών (βρογχιεκτασίες).

Παρόλο που η ασθένεια εκδηλώνεται σε νεαρή ηλικία, η διάγνωση της καθυστερεί ή απουσιάζει πλήρως, κυρίως λόγω των μη-ειδικών συμπτωμάτων (βήχας, ρινόρροια), έλλειψης γνώσης για την ΠΔΚ μεταξύ των ιατρών αλλά και λόγω της δυσκολίας στη διαθεσιμότητα και την ερμηνεία των ειδικών διαγνωστικών μεθόδων. Ο διαγνωστικός έλεγχος για ΠΔΚ συνήθως περιλαμβάνει τουλάχιστον 3 εργαστηριακές μεθόδους: α) μέτρηση ρινικού Μονοξειδίου του Αζώτου (nasal Nitric Oxide - nNO), β) εκτίμηση της κινητικότητας των κροσσών και γ) εξέταση της δομής των κροσσών. Ο διαγνωστικός έλεγχος για ΠΔΚ είναι επίπονος και χρονοβόρος ενώ πολλά κέντρα πιθανόν να μην διαθέτουν πρόσβαση σε εξειδικευμένο εξοπλισμό ή προσωπικό για την εκτέλεση όλων των απαιτούμενων ελέγχων. Σαν αποτέλεσμα, διαφορετικοί αλγόριθμοι μπορεί να χρησιμοποιούνται για τη διάγνωση της ΠΔΚ και αυτό το φαινόμενο επηρεάζεται περαιτέρω από την έλλειψη γνώσης σχετικά με την αποτελεσματικότητα και το μέσο οικονομικό κόστος της κάθε διαγνωστικής τεχνικής.

Στόχοι

Με ορίζοντα την βέλτιστη πληροφόρηση στη διαδικασία λήψης αποφάσεων για τον πιο αποτελεσματικό αλγόριθμο διάγνωσης για ΠΔΚ, ο στόχος της παρούσας εργασίας είναι, πρώτον να καθορίστουν οι ιδιότητες των τριών βασικών διαγνωστικών τεχνικών (nNO, ηλεκτρονική μικροσκοπία μετάδοσης - Transmission Electron Microscopy – TEM, βίντεομικροσκοπία υψηλής ταχύτητας – High Speed Video Microscopy - HSVM) και δεύτερον, η συνολική εκτίμηση διαφόρων διαγνωστικών αλγορίθμων ως προς το συνολικό

όφελος προς την υγεία των ασθενών με ΠΔΚ και του συνολικού οικονομικού κόστους στο σύστημα υγείας.

Μέθοδοι

Μέσω ξεχωριστών μελετών συστηματικής ανασκόπησης και μετα-ανάλυσης (systematic review and meta-analysis), όλες οι σημαντικές ηλεκτρονικές βάσεις δεδομένων έχουν ερευνηθεί μέχρι το 2016, με την χρήση κατάλληλων όρων για να αναγνωρίσουμε τα άρθρα που αναφέρουν στοιχεία σχετικά με την διαγνωστική ακρίβεια των μεθόδων nNO, TEM και HSVM, καθώς και στοιχεία της εκτίμησης της συχνότητας της ΠΔΚ σε πληθυσμούς ύποπτων ασθενών που έχουν παραπεμφθεί για διαγνωστικές εξετάσεις για ΠΔΚ. Οι επιλεγμένες μελέτες περιλάμβαναν διαγνωστικές πληροφορίες για τους ασθενείς με ΠΔΚ ή ύποπτους ασθενείς που έχουν υποστεί ένα συνδυασμό διαγνωστικών μεθόδων που περιλαμβάνουν nNO, TEM και HSVM και γενετικό έλεγχο.

Για την μετα-ανάλυση της διαγνωστικής ακρίβειας του nNO, έχουν υπολογιστεί η ευαισθησία και η ειδικότητα των μετρήσεων του nNO για όλα τα άρθρα που περιλήφθηκαν στη μελέτη καθώς και προσαρμόστηκαν ένα μοντέλο μικτής λογιστικής παλινδρομησης δύο επιπέδων βασισμένο στην ευασθησία και την ειδικότητα της κάθε μελέτης καθώς ένα διμεταβλητό μοντέλο κανονικής κατανομής για την ευαισθησία και την ειδικότητα μεταξύ των μελετών. Ιεραρχικές καμπύλες ROC (Summary receiver operating characteristic (HSROC) curves) αναπτύχθηκαν με βάση τις παραμέτρους των προσαρμοσμένων μοντέλων ξεχωριστά για την μέτρηση nNO με κλειστή την γλωττίδα (Vellum Closure - VC) και ξεχωριστά για την μετρήση του nNO με ανοικτή την γλωττίδα (non-Vellum Closure - non-VC). Για την μετα-ανάλυση της συχνότητας της ΠΔΚ σε πληθυσμούς ύποπτων ασθενών που έχουν παραπεμφθεί για εξέταση ΠΔΚ και για την μετα-ανάλυση του ποσοστού ανίχνευσης του TEM, χρησιμοποιήθηκε η μεθοδολογία της μετα-αναλυσης ποσοστού με τυχαίες επιδράσεις. Η μετα-ανάλυση ποσοστού επιτρέπει τον υπολογισμό του συγκεντρωτικού ποσοστού από όλες τις μελέτες που παριείχαν διωνυμικά δεδομένα ενώ η χρήση τυχαίων επιδράσεων επιτρέπει σε κάθε μελέτη να προσδιοριστεί με ένα συντελεστή που συμπεριλαμβάνει την διακύμανση εντός της μελέτης αλλά και την διακύμανση μεταξύ των μελετών. Η ετερογένεια εκτιμήθηκε με την παράμετρο I^2 η οποία περιγράφει το ποσοστό της συνολικής παραλλαγής στην εκτίμηση του αποτελέσματος από την ετερογένεια μεταξύ των μελετών και το εύρος των τιμών που μπορεί να πάρει, κυμαίνεται μεταξύ 0 μέχρι 100%.

Τα στοιχεία σχετικά με τις διαγνωστικές ιδιότητες του nNO και TEM και τα στοιχεία σχετικά με την επικράτηση της ΠΔΚ μεταξύ ύποπτων ασθενών συνδυάστηκαν με τις εκτιμήσεις για την διαγνωστική ακρίβεια του HSVM από μεμονωμένες μελέτες κατά την ανάπτυξη ενός πιθανολογικού μοντελου ανάλυσης το οποίο επιτρέπει να υπολογιστεί η συνολική ευαισθησία και ειδικότητα του εκάστοτε διαγνωστικού αλγορίθμου, καθώς και η σχέση κόστους-αποτελεσματικότητας και η αυξανόμενη σχέση κόστους-αποτελεσματικότητας για τους τρεις διαγνωστικούς αλγόριθμους οι οποίοι χαρακτηρίζονται από διάφορους συνδυασμούς του nNO, TEM και HSVM. Οι αξιολογούμενοι συνδυασμοί ήταν α) nNO+TEM σε ακολουθία, β) nNO+ HSVM σε ακολουθία και γ) παράλληλη διενέργεια nNO/HSVM ακολουθούμενο από TEM ως επιβεβαιωτική εξέταση στις περιπτώσεις αλληλοσυγκρουόμενων αποτελεσμάτων (nNO/HSVM+TEM). Το πιθανολογικό μοντέλο υπολογίζονταν με αρχικό υποθετικό πληθυσμό 1000 παραπομπών ύποπτων ασθενών με αναμενόμενη μέση συχνότητα ΠΔΚ ίση με 320 ασθενείς στις 1000 παραπομπές). Τα αποτελέσματα του μοντέλου περιλάμβαναν τον εκτιμόμενο αριθμό των ασθενών με ΠΔΚ από κάθε διαγνωστικό αλγόριθμο, τον λόγο κόστους-αποτελεσματικότητας (Cost-Effectiveness Ratio – CER) και τον λόγο αυξανόμενου κόστους-αποτελεσματικότητας (incremental Cost-Effectiveness Ratio - ICER) για κάθε διαγνωστικό αλγόριθμο. Τα αποτελέσματα υπολογίστηκαν χρησιμοποιώντας τη προσομοίωση Monte Carlo στο πρόγραμμα ANALYTICA.

Αποτελέσματα

Η μέση συχνότητα της ΠΔΚ μεταξύ των παραπομπών ήταν 32% (95% Διάστημα Εμπιστοσύνης (ΔΕ): 25–39%, I^2 = 92%). Το ποσοστό ανίχνευσης με TEM μεταξύ των ασθενών με ΠΔΚ ήταν 83% (95% ΔΕ: 75–90%, I^2 = 90%). Ο εξαίρεση των μελετών που αναφέρουν μεμονωμένους ελαττωματικούς βραχίονες δυνείνης ως ΠΔΚ, μείωσε το ποσοστό ανίχνευσης με TEM και εξηγεί το ποσοστό της παρατηρούμενης ετερογένειας (74%, 95% ΔΕ: 66–83%, I^2 = 66%).

Η συνολική ευαισθησία του nNO κατά την μέτρηση με VC ήταν 0.95 (95 % ΔΕ: 0.91–0.97), ενώ η ειδικότητα ήταν 0.94 (95 % ΔΕ: 0.88–0.97). Ο θετικός λόγος πρόβλεψης (Positive Likelihood Ratio - LR+) ήταν 15.8 (95 % ΔΕ: 8.1–30.6), ενώ ο αρνητικός λόγος πρόβλεψης (Negative Likelihood Ratio - LR-) ήταν 0.06 (95 % ΔΕ: 0.04–0.09). Για τις μετρήσεις non-VC, η συνολική ευαισθησία του nNO ήταν 0.93 (95 % CI 0.89–0.96), ενώ η ειδικότητα ήταν

0.95 (95 % CI 0.82–0.99). Ο θετικός λόγος LR+ ήταν 18.5 (95 % CI 4.6–73.8), ενώ ο αρνητικός λόγος LR- ήταν 0.07 (95 % CI 0.04–0.12).

Σχετικά με τα αποτελέσματα του πιθανολογικού μοντέλου, από τους 320 αναμενόμενους ασθενείς με ΠΔΚ, 311 ανιχνεύτηκαν με nNO/HSVM+TEM, 274 με nNO+HSVM και 198 με nNO+TEM. Ο συνδυασμός nNO/HSVM+TEM είχε το πιο ψηλό μέσο κόστος (€97K) ακολουθούμενο από το συνδυασμό nNO+TEM (€56K) και τον συνδυασμό nNO+HSVM (€39K). Ο αλγόριθμος nNO+HSVM επικράτησε του αλγόριθμου nNO+TEM (χαμηλότερο κόστος και πιο αποτελεσματικό). Ο λόγος ICER για το nNO/HSVM+EM ήταν €1620 για κάθε επιπλέον διαγνωσμένο ασθενή με ΠΔΚ.

Συμπεράσματα

Πολλά από τα κέντρα που ασχολούνται με τη διάγνωση και τη θεραπεία της ΠΔΚ στις ανεπτυγμένες χώρες, ακολουθούν διαφορετικές τεχνικές και ποικίλους αλγόριθμους για τη διάγνωση της ασθένειας. Σε χώρες με χαμηλό εισόδημα, πολύ πιθανό να υπάρχει πλήρης έλλειψη ειδικών διαγνωστικών μέσων. Τα αποτελέσματα αυτής της διδακτορικής διατριβής εισηγούνται ότι η διαγνωστική ακρίβεια των μετρήσεων nNO και με τους δύο δυνατούς αναπνευστικούς ελιγμούς (κλειστή/ανοικτή γλωττίδα) είναι υψηλή και μπορεί να υιοθετηθεί αποτελεσματικά στην διαδικασία κλινικού εντοπισμού της ΠΔΚ ακόμα και σε παιδιά, με στόχο την πρώιμη διάγνωση. Αντίθετα, ένα σημαντικό ποσοστό ασθενών, τουλάχιστον της τάξης του 26% δεν μπορεί να διαγνωστεί με TEM και αυτός ο περιορισμός θα πρέπει να αποτελέσει βάση για την ανάπτυξη αποτελεσματικών διαγνωστικών αλγόριθμων για ΠΔΚ. Τα αποτελέσματα αυτής της μελέτης εισηγούνται επίσης πως ο διαγνωστικός αλγόριθμος ο οποίος περιλαμβάνει το nNO με VC ως πρώτο βήμα ακολουθούμενο από HSVM για επιβεβαίωση, ανιχνεύει περίπου το 86% των ασθενών με ΠΔΚ με μέση τιμή λόγου κόστους-αποτελεσματικότατας (CER) 140€ για κάθε διαγνωσμένο ασθενή με ΠΔΚ. Ο αλγόριθμος που μεγιστοποιεί τον αριθμό των ασθενών με ΠΔΚ περιλαμβάνει παράλληλη εξέταση με nNO και HSVM ως πρώτο βήμα, ακολουθούμενο από TEM ως επιβεβαιωτικό διαγνωστικό τεστ για τις λίγες περιπτώσεις όπου το nNO και το HSVM παρουσιάζουν αντικρουόμενα αποτελέσματα, με αντίστοιχο λόγο αυξανόμενου κόστους-αποτελεσματικότητας (ICER) ίσο με 1620€ για κάθε επιπλέον ασθενή διαγνωσμένο με ΠΔΚ. Τα ευρήματα αυτά επισημάνουν την ανάγκη για ανάπτυξη οδηγιών για τον διαγνωστικό έλεγχο των ασθενών με ΠΔΚ βασισμένες σε επιστημονικά τεκμήρια και παρέχουν νέες πληροφορίες στον εν εξελίξει

διάλογο για το πως οι οδηγίες αυτές μπορούν να εκτελεστούν από τα διάφορα συστήματα υγείας.

Table of Contents:

List of Figures

Chapter 2:

Chapter 3:

Chapter 4:

Chapter 5:

Chapter 6:

Chapter 7:

List of Tables

Chapter 2:

Chapter 3:

Chapter 5:

Chapter 6:

Chapter 7:

Chapter 1: Introduction

Motivation

Primary Ciliary Dyskinesia (PCD) is a rare, genetically heterogeneous disorder that affects one in approximately 15 000 live births [1] and results from the dysfunction of small hair-like organelles, called motile cilia. Motile cilia can be found in a number of human tissues including the upper and lower airways epithelium (respiratory cilia), the lining of brain ventricles (brain ependymal cilia), the lining of fallopian tubes and uterus (oviduct cilia) and the sperm tail (single flagellum) [2]. The active and coordinated beating of respiratory cilia is the driving factor behind the process of mucociliary clearance, a critical mechanical defense mechanism of the respiratory system responsible for the removal of inhaled pathogens and other hazardous substances as well as cell debris from the upper and lower airways [2].

PCD patients usually suffer from recurrent respiratory infections which lead to chronic destructive airway disease characterized by progressive loss of lung function and structural damage of the airways (bronchiectasis), lifetime rhinorrhea and recurrent acute sinus and ear infections [3]. Furthermore, PCD patients may also present with situs abnormalities, as the motile cilium in the embryonic node, which determines the organization of organ placement in the body during embryogenesis, could also be affected, thus resulting to a random organ placement and laterality [4]. In fact, part of PCD clinical spectrum is known as Kartagener Syndrome, defined as the presence of the clinical triad: Sinusitis, Bronchiectasis and Situs Inversus (mirror image in the placement of visceral organs). However this definition fails to cover the whole spectrum of PCD patients since about half of PCD patients do not present with Situs Inversus [5]. A small fraction of PCD patients may present with heterotaxy (situs ambiguous) accompanied with congenital cardiovascular abnormalities [6]. Other clinical manifestations that may lead to the consideration of PCD are a history of unexplained

neonatal respiratory distress syndrome, nasal polyps, family history of PCD, male infertility and chronic productive cough in the absence of more common causes of chronic lung disease [7]. A recent systematic review and meta-analysis summarized the published evidence of clinical manifestations of PCD and reported the pooled prevalence of each symptom. Although considerable heterogeneity was found between the assessed studies, chronic cough and sputum production was found to be the most prevalent symptoms with a mean prevalence of 88% and 89% respectively. Chronic rhinorrhea and otitis media (with or without effusion) were the most frequent upper respiratory symptoms with a reported mean prevalence of 75% and 74% respectively. The mean prevalence of other upper respiratory manifestations such as sinusitis was found to be 69%, while for nasal polyps the mean prevalence was 19%. The mean prevalence of a history of lower respiratory infections, including pneumonia, was 72% and mean prevalence of development of bronchiectasis was 56%. The prevalence of situs abnormalities was 49% while the mean prevalence of congenital cardiovascular abnormalities was 5%. In studies that evaluated infertility in adults, 100% of males were found to be infertile as well as 58% of females [8]. However, older studies reported infertility in approximately only 50% of male patients [9, 10].

Despite the fact that many of the manifestations of PCD present early in life, diagnosis is often delayed or missed completely, primarily due to the low specificity of some symptoms (e.g. cough, rhinorrhea) and lack of awareness for PCD among clinicians [5]. In Europe, as indicated by a recent survey of 223 centers from 26 countries, the median age at PCD diagnosis was 5.3 years. However in patients with situs inversus, the median age of diagnosis was 3.8 years compared with a median age of 5.8 years in patients with situs solitus [5]. In addition to lack of awareness, difficulties in establishing PCD diagnosis, both due to lack of equipment and or lack of expertise, might further contribute towards missing the diagnosis or diagnosing patients at an older age [11].

Diagnostic testing for PCD usually involves at least three laboratory techniques/procedures: (a) nasal Nitric Oxide measurement, (b) assessment of ciliary motility and (c) examination of cilia ultrastructure in ciliated epithelial cells obtained via a biopsy of the epithelium of the nasal passages or the bronchi [12]. More recently, additional diagnostics tests have been developed for PCD such as the targeted genetic screening [13] and immunofluorescence analysis [14] but to date only a small number of specialized centers have incorporated them as part of the routine diagnostic procedures for PCD [13]. A recent survey undertaken by the European Respiratory Society (ERS) in 2012 confirmed that most European countries were not offering a centralized service for PCD diagnosis and revealed that substantial variability existed in the availability of the diagnostic tests at each center [15]. Diagnostic testing for PCD is laborious and time consuming and many centers may lack access to necessary equipment or expertise to perform all required tests. As a result, different diagnostic algorithms for PCD diagnosis may be followed by different centers [16]. This variability is further influenced by the lack of knowledge regarding the diagnostic effectiveness and average cost of each test. Characteristically, the Standardized Operational Procedures (SOPs) for PCD diagnostic testing developed recently by the FP7 project BESTCILIA [17] and the recent ERS guidelines for PCD diagnosis [18] are primarily based on the experience of PCD specialists rather than evidence-based estimates regarding the diagnostic efficacy of each test and do not account for the cost-effectiveness of implementing different diagnostic algorithms in clinical practice.

Aims

Towards further illuminating the decision making process for the establishment of the most efficacious diagnostic algorithm for PCD, we aimed first to characterize the diagnostic properties of the three main tests for PCD (nNO, TEM and HSVM) and second to evaluate

different diagnostic algorithms in terms of overall health benefits for PCD patients and overall costs to the healthcare systems. The first part of the study involved the systematic review of the literature and development of summary estimates of diagnostic efficacy for the three tests with the use of meta-analytic approaches while the second part included the comparison of the different diagnostic algorithms through a probabilistic decision tree model.

In summary, this PhD thesis focused on the:

i. Description of PCD through an extensive literature review.

ii. Systematic review and meta-analysis of the prevalence of PCD in consecutive referrals of suspect cases.

iii. Systematic review and diagnostic accuracy meta-analysis of nNO for establishing PCD diagnosis.

iv. Systematic review and meta-analysis of the TEM detection rate in patients suspected for PCD.

v. Cost effectiveness analysis of different diagnostic algorithms for Primary Ciliary Dyskinesia.

Thesis Overview

This study was performed in Cyprus International Institute for Environmental and Public Health, Cyprus University of Technology during the period September 2013 – May 2017. Main advisor of this work was Dr Stefania Papatheodorou (Cyprus University of Technology) and co-advisors were Professor Panayiotis Yiallouros (University of Cyprus) and Professor John S. Evans (Harvard School of Public Health). Members of the Advisory

Committee were also Dr Nicos Middleton (Cyprus University of Technology) and Professor Kyriacos Kyriacou (Cyprus Institute of Neurology and Genetics).

Chapter 1 describes the motivation for this work and presents the main aims of this study while Chapter 2 presents an extensive literature review of PCD. The literature review covers all major aspects of PCD including the pathophysiology, genetic background, epidemiology, diagnostic testing, clinical picture, clinical management and disease burden in PCD.

Chapter 3 describes the basic concepts of key methodologies used throughout this study, namely the performance of a systematic review and meta-analysis towards summarizing the published evidence about a scientific question. Particular focus is given in describing the theoretical background of specific meta-analytic approaches such as meta-analysis of proportions and diagnostic accuracy meta-analysis as well as development of Hierarchical Summary Receiver Operating Curves (HSROC)

Chapter 4 provides a brief overview regarding the use of decision trees and economic evaluation towards informing evidence-based decision making in healthcare. A more detailed description of the methodology underlying the performance of Cost Effectiveness Analysis, the most widely used method for economic evaluation in healthcare, is provided and includes sections on how to value costs and health effects and how to handle uncertainty in this type of analysis.

Chapter 5 presents the systematic review and meta-analysis regarding the diagnostic accuracy of nNO measurements towards establishing diagnosis PCD. The meta-analysis methodology is presented in detail and the different breathing manoeuvres that can be used during the measurement of nNO are discussed and a separate analysis for each manoeuvre used was performed. The work presented in Chapter 4 has been published in *BMC Pulmonary Medicine* as a research manuscript titled "Diagnostic accuracy of nasal nitric oxide for establishing diagnosis of primary ciliary dyskinesia: a meta-analysis" [19].

Chapter 6 presents the systematic review and meta-analysis for both the prevalence of PCD in cohorts of consecutive patients referred for specialized testing and the detection rate of TEM. This chapter

reports an evidence-based estimate of PCD prevalence in patients with relevant symptoms and discusses the limitations of ciliary ultrastructural assessment using TEM for PCD diagnosis. The work presented in Chapter 5 has been published in *Pediatric Research* as a research manuscript titled "Prevalence of primary ciliary dyskinesia in consecutive referrals of suspect cases and the transmission electron microscopy detection rate: a systematic review and meta-analysis" [20].

Chapter 7 describes the probabilistic decision tree model developed during this work and presents the methodology and results of the cost-effectiveness analysis for three different diagnostic algorithms for PCD.

Finally, in Chapter 8, the main conclusions and limitations of this body of work in light of previous research findings and the implications for future research are discussed.

Chapter 2: Literature Review

Cilia

Cilia, as well as the structurally and functionally similar flagella, are evolutionary conserved organelles and were first developed to provide motility in unicellular organisms [21]. Several model organisms such as the *Paramecium* and *Chlamydomonas reinhardtii* were employed to understand the basic structure of these organelles [22] while previous studies have provided additional insight into the genome and proteome of human cilia [23]. Although historically thought to primarily serve motility in water or transport of fluids above a mucosal surface, cilia are now known to also serve as sensory organelles [24]. This functional distinction formed the basis for the characteristic terms used for human cilia, "motile cilia" and "primary cilia". Motile cilia can be found on the apical side of several human epithelial tissues including the upper and lower airways epithelium (respiratory cilia), the lining of brain ventricles (brain ependymal cilia), the lining of fallopian tubes and uterus (oviduct cilia) and the sperm tail (single flagellum) while primary cilia are known to exist in kidney cells (renal cilia), in photoreceptor cells of the eye (photoreceptor cilia) and in almost all human tissues [21, 25]. Primary cilia are borne as solitary attachments (monocilia) are immotile and project from the basal body, which is a specialized centriole attached on the cell surface. Similarly, motile cilia project from basal bodies but are found in numerous ciliary bundles [24]. A highly organized array of microtubules composed of tubulin monomers organized in protofilaments makes up the organelle's cytoskeleton but the pattern of this microtubule arrangement differs between primary and motile cilia.

A "9+2" arrangement characterizes motile cilia with 9 pairs of outer microtubules, often called peripheral doublets, where several dynein arms (DA) are docked, and a single pair of microtubules in the center of the cilium. The central pair is linked with the outer doublets via

multiprotein complexes called radial spokes, while the outer doublets are linked between them through other multiprotein complexes termed nexin links [21, 24] (Figure1B). These protein complexes facilitate stability and movement while other proteins involved in intraflagellar transport are essential in ciliogenesis and the maintenance of the cilium, as no protein synthesis is present within the cilium axonemal shaft [25].

Primary cilia are characterized by a "9+0" arrangement of microtubules with the same pattern at the periphery as the motile cilia but with absence of the central pair (Figure 1A). DA are also absent in primary cilia, although present in motile cilia, as DA with their adenine triphosphatase (ATPase) activity act as molecular motors and are essential for ciliary movement. A third category of cilia are called nodal cilia or embryonic cilia. They are solitary but although they have a "9+0" arrangement, they do contain DA and are motile [24].

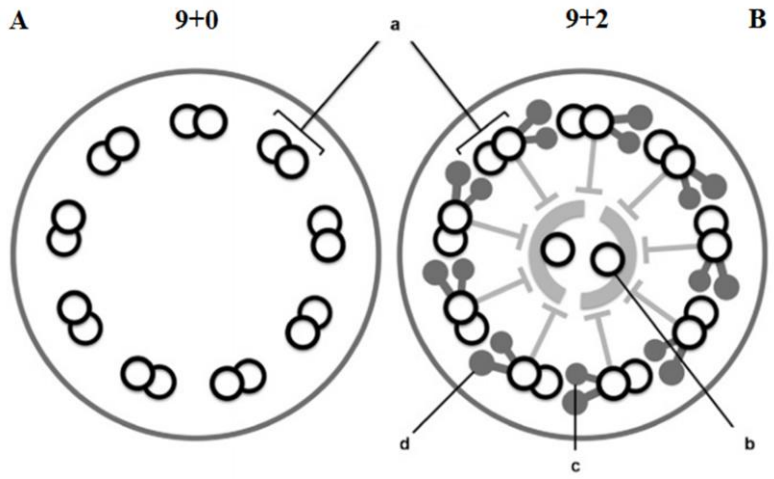


Figure 2.1: *The axonemal ultrastructure of primary (A) and motile cilia (B). Both types express 9 peripheral microtubule doubles (a) and a central pair of microtubules (b) while primary cilia lack the central pair. On the peripheral doublets of motile cilia inner (c) and outer (d) dynein arms are docked. Adapted from Takeda S 2012*

Cilia in health and disease

Primary, "9+0" cilia are equipped with a number of receptors and ion channels and are thought to primarily act as sensory organelles (antennae) with important roles in both

development and homeostasis [26]. Examples of receptors that are present on primary cilia include the polykystin 1 receptor, the serotonin receptor 5, the melanin-concentrating hormone receptor 1 and important components of the Hedgehog and non-canonical Wnt/planar cell polarity signaling pathways [27]. The importance of cilia in homeostasis is highlighted by the fact that composition and flow rate of urine in nephrons is monitored by primary cilia that are present on the epithelium of renal cilia and act as mechanosensors. By bending upon fluid shear stress, a signaling cascade which involves conformational changes to integral membrane proteins such as polykystin 1, activates Ca^{2+} channels (such as polykystin 2) and the resulting influx of extracellular Ca^{2+} influx in the cilium eventually modifies gene expression, growth, differentiation and other important cellular functions. Mutations in polykystin 1 and polykystin 2 are known to cause Polycystic Kidney Disease, an autosomal dominant disorder characterized by formation of fluid-filled cysts in the kidneys, which eventually leads to renal failure [28]. Furthermore the primary cilium appears to be involved in cell cycle control as its disassembly acts as prerequisite for mitosis probably through the interaction with mitotic kinases such as Aurora and NIMA-related kinases [29, 30] and defective primary cilia have been associated with tumor formation as a result of abnormal mitogenic signaling [31].

Many developmental signaling pathways include components located on the cytoplasm or membrane of the primary cilium such as proteins Smo, Sufu and Gli that are part of the Hedgehog signaling network [32]. The Hedgehog network is involved in embryonic development and plays a crucial role in left-right asymmetry and limb and heart development as well as cell proliferation and differentiation of neural tissues [33, 34]. Similarly, the non-canonical Wnt/planar cell polarity pathway operates through the primary cilium localized membrane protein Van Gogh-like 2 (Vangl2) to control cytoskeletal changes, cell adhesion, cell migration, planar polarity and apical–basal polarity in epithelial tissues [32]. Genetic

defects that affect components of specific signaling cascades that are localized on primary cilia or basal bodies result in a class of serious genetic disorders termed ciliopathies. Ciliopathies are further distinguished in a number of syndromic diseases such as Polycystic Kidney Disease, Bardet-Biedl syndrome, Joubert syndrome, Meckel syndrome, Senior-Løken syndrome, Jeune syndrome, Nephrophthisis, Ellis van Creveld syndrome and Alstrom Syndrome [35]. These syndromes are characterized by broad phenotypic manifestations demonstrating the primary cilium extensive tissue and cellular distribution. However, given that all syndromes result from primary cilia specific abnormalities, there is considerable genetic and clinical overlap [36]. The major overlapping clinical characteristics are renal and/or hepatobiliary disease, laterality defects, polydactyly, agenesis of corpus callosum, cognitive impairment, degeneration of the retina, skeletal defects and encephalocoele [37].

Motile, "9+2" cilia primarily serve as fluid propulsion organelles located on the apical side of epithelial tissues in the human body but it is possible to maintain a sensory role as well. It has been shown that respiratory cilia motility parameters such as ciliary beat frequency (CBF) and waveform are controlled through Ca^{2+}, cyclic adenosine monophosphate (cAMP) and cyclic guanine monophosphate (cGMP) [38]. Similarly, recent evidence from oviduct cilia demonstrated that polykystin 1 and polycystin 2 as well as progesterone receptors are present on ciliary membrane and their regulation is fine-tuned by the menstrual cycle, demonstrating that signaling through motile cilia receptors facilitates the ciliary motility and transport of the ovulated oocyte [32]. Furthermore an increasing number of other signal receptors have been found to localize on motile cilia although their effects appear not to be related to CBF and the beating waveform thus indicating a signaling role of motile cilia independent of the regulation of ciliary motility. Among these receptors are angiopoietin receptors that localize on oviduct cilia [39], Vangl2 receptors and fibroblast growth factor receptors that have been found to localize on respiratory cilia [40, 41].

Interestingly motile monocilia (nodal cilia) do have an essential role in the development of the human fetus by rotating in a clockwise fashion and thus producing a leftward nodal flow which results in normal situs anatomy, termed Situs Solitus. Malfunction of nodal cilia has been found to cause left-right patterning asymmetry in organ placement and orientation (Figure 2). This malfunction and resulting asymmetry could lead to complete (Situs Inversus Totalis) or partial reversal (Partial Situs Inversus) of the major visceral organs or even to an erratic distribution of organs which is termed Situs Ambiguous [42]. There are two prevailing hypotheses regarding the role of nodal cilia in left-right asymmetry. The first one proposes that nodal cilia transport an extracellular morphogen (such as Hedgehog lipoproteins, [33] towards the left side thus triggering an asymmetrical laterality signaling cascade of morphogenesis. The second hypothesis proposes that the motile nodal cilia located at the center of the embryonic node, generate a leftward motion of extracellular fluid, which is picked up by sensory cilia at the periphery and thus initiate the signaling cascade of morphogenesis [21, 43].

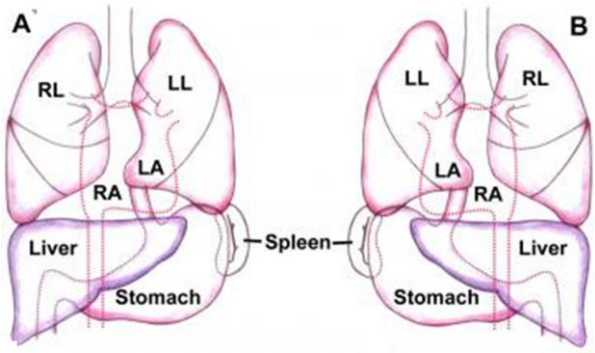


Figure 2.2: *Organ placement in Situs Solitus (A) and Situs Inversus (B). RL: Right Lung, LL: Left Lung, RA: Right heart Atrium, LA: Left heart Atrium. Adapted from Wilhem A 2015 http://emedicine.medscape.com/article/413679-overview*

Mucociliary Clearance

Humans inhale approximately 1000-21000 liters of air per day depending on age, body size and physical activity with the average lung ventilation rate for an adult in resting state being

approximately 10000 liters per day [44]. As a result the extensive respiratory epithelium which lines the conductive airways (nose, pharynx, trachea, bronchi and bronchioles) and the alveoli is constantly exposed to a large burden of organic and inorganic pollutants, airborne pathogens and viral agents. Towards removing these potentially dangerous materials, the respiratory system relies on an extensive array of innate defense mechanisms. Firstly, through filtration large particles with a diameter greater than 1μm deposit in the nasopharynx and tonsilar regions but smaller particles and bacteria do reach the tracheobronchial and alveolar region of the lung [45, 46]. The first line of defense against such pollutants, bacteria and viruses is Mucociliary Clearance (MCC) while other innate defense mechanisms of the lung include cough and soluble immunity components such as the complement proteins and pulmonary surfactant and alveolar macrophages in the terminal bronchioles that phagocytose small particles and bacteria and transport them to the local lung associated lymph nodes [46, 47]. MCC serves a threefold purpose; a) it acts as a mechanical barrier by trapping any inhaled material (as well as endogenous cell debris) in the mucus liquid which covers the airway epithelium and eventually clearing it from the respiratory tract through the constant beating of respiratory cilia (mucociliary escalator), b) the airway mucus itself has antioxidant properties and provides a biological barrier for microorganisms as macrophages and neutrophils, as well as c) other substances (e.g. lysozyme) with antimicrobial properties are present in the mucous [48]. In cases where MCC is not adequate or fails completely, coughing serves as a back-up system to shift mucus towards the pharynx [48].

Ciliated cells in the respiratory tract have approximately 200 cilia per cell, each with an axonemal diameter of 250 nm and an axonemal length of approximately 6 μm. There are about 10^9 cilia per cm^2 of ciliated airway epithelium [49]. Respiratory cilia lie within the so-called periciliary liquid which lies below the mucus blanket and is considerably less viscous than the upper layer thus allowing cilia to beat rapidly. Average (CBF), when measurements

are made ex-vivo in 37°C, ranges between 10 and 18 Hz [50]. In order to propel mucus (and trapped material), cilia need to beat in a specific asymmetric beating pattern which includes both an effective and a recovery stroke. During the effective stroke the cilia beat in a plane perpendicular to the cell surface thus generating a mucus flow in the same direction as its motion. On the contrary, during the recovery stroke the cilia bend sideways (parallel to the cell surface) thus generating a weaker backward flow as it returns to its original configuration [51] (Figure 3). During the effective stroke the cilia tips may also engage the mucus layer above the periciliary liquid, further enhancing the flow of mucus towards the beating direction [48].

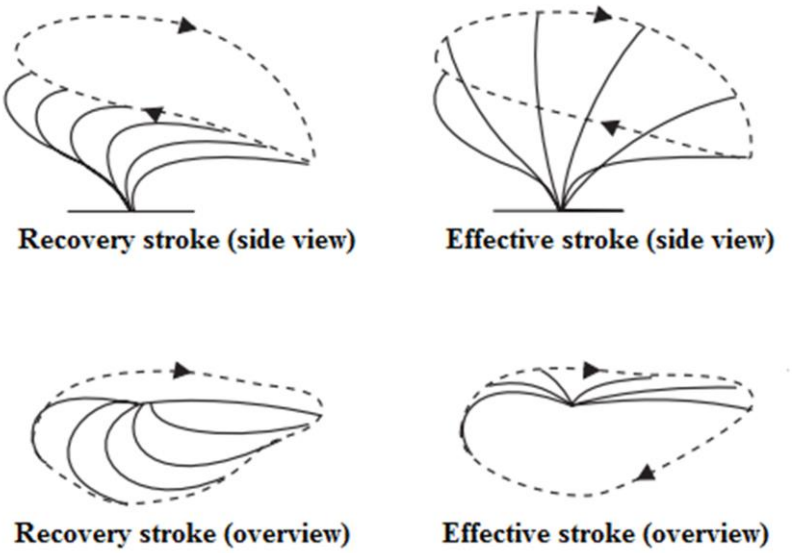


Figure 2.3: *Schematic representation of respiratory cilia beat pattern. Adapted from Hoodmeyers et al* [48]

The beating of cilia is powered by ATP-dependent reactions that allow the sliding movement of microtubular doublets through the active binding of the DA of the first doublet in successive binding sites along its neighboring doublet [2]. In health, the anchoring of basal bodies is oriented as in neighboring cilia resulting in beating strokes that have the same direction [52] and enhances effective mucus transport. Interestingly, although collective beating of cilia is highly coordinated, it is not entirely synchronous but is rather organized in propagating metachronal waves. Metachronal waves result from cilia performing similar

beating patterns but with a small phase difference in respect to cilia from neighboring segments of the epithelium [51].

Primary Ciliary Dyskinesia

The importance of MCC in lung defense is highlighted by pathologic conditions that are characterized by disruption of the mucociliary escalator. Such conditions are Cystic Fibrosis (CF) and PCD. Cystic Fibrosis, the more well-known of the two, is an autosomal recessive disease resulting from mutations located in the Cystic Fibrosis Transmembrane conductance Regulator (*CFTR*) gene and is characterized by abnormal transport of chloride ions across epithelial surfaces. Abnormal ion transport across the lung epithelium in CF is followed by dehydration of the mucus layer, depletion of the periciliary layer and abnormal epithelial fluid transport in the lung. The mucus in CF patients is characterized by increased viscosity and is usually described as thick or sticky and although respiratory cilia are functioning, mucociliary clearance is significantly impaired, as well as the effectiveness of coughing to clear up material from the airways [53, 54]. PCD is a rare, genetically heterogeneous motile ciliopathy that affects one in approximately 15 000 live births [1]. It is caused by dysfunctional motile cilia that are characterized by either complete immotility or reduced cilia beat frequency (CBF) and/or abnormal ciliary beat pattern (CBP). Abnormal cilia motility results in disrupted mucociliary clearance. The removal of inhaled pathogens and other hazardous substances from the upper and lower airways of the lungs fails and patients suffer from recurrent respiratory infections [3]. In addition other organ systems such as the cardiovascular (congenital heart defects) and the reproductive system (infertility) can be affected while almost half of the PCD patients are diagnosed with situs abnormalities (Situs Inversus, partial Situs Inversus or Situs Ambiguous) [1].

A patient with the characteristic triad of PCD symptoms (bronchiectasis, chronic sinusitis and Situs Inversus) was first described by Siewert in 1904 [55] but the condition was subsequently classified as a distinct congenital disorder by Manes Kartagener, a Polish/Swiss pulmonologist, in 1933 [56] and was thereafter referred as Kartagener's Syndrome for many years. The link between Kartagener's syndrome and cilia abnormalities was discovered by the Swedish physician Bjorn Afzelius in 1975, who reported absence of dynein arms in electron micrographs of immotile cilia and sperm flagella [57, 58] and for many years the condition was also termed as 'immotile cilia syndrome'. However, following reports of a subset of Kartagener's syndrome patients with dyskinetic or asynchronous cilia, rather than completely immotile and patients with immotile cilia but no Situs Inversus [59], favored the adoption of the more inclusive term Primary Ciliary Dyskinesia (PCD). PCD patients with Situs Inversus may still be referred as Kartagener's syndrome patients but it is now accepted that Kartagener's Syndrome is a subset of PCD.

Clinical Features of Primary Ciliary Dyskinesia

PCD patients usually suffer from recurrent respiratory infections which lead to chronic destructive airway disease characterized by progressive loss of lung function and structural damage of the airways (bronchiectasis), lifetime rhinorrhea and recurrent acute sinus and ear infections [3]. About half of PCD patients also present with Situs Inversus as a result of a dysfunctional cilium in the embryonic node, which determines the organization of organ placement in the body during embryogenesis [4, 60]. A small fraction of PCD patients may present with heterotaxy (situs ambiguous) accompanied with congenital cardiovascular abnormalities [6]. Other clinical manifestations that may lead to the consideration of PCD are a history of unexplained neonatal respiratory distress syndrome, nasal polyps, family history

of PCD, male infertility and chronic productive cough in the absence of more common causes of chronic lung disease [7]. A recent systematic review and meta-analysis summarized the published evidence of clinical manifestations of PCD and reported the pooled prevalence of each symptom. Although considerable heterogeneity was found between the assessed studies, chronic cough and sputum production was found to be the most prevalent symptoms with a mean prevalence of 88% and 89% respectively. Chronic rhinorrhea and otitis media (with or without effusion) were the most frequent upper respiratory symptoms with a reported mean prevalence of 75% and 74% respectively. The mean prevalence of other upper respiratory manifestations such as sinusitis was found to be 69%, while for nasal polyps the mean prevalence was 19%. The mean prevalence of a history of lower respiratory infections, including pneumonia, was 72% and mean prevalence of development of bronchiectasis was 56%. The prevalence of situs abnormalities was 49% while the mean prevalence of congenital cardiovascular abnormalities was 5%. For studies that evaluated infertility in adults, 100% of males were found to be infertile as well as 58% of females [8] although older studies reported infertility in approximately only 50% of male patients [9, 10].

PCD has been shown to greatly affect lung function. Decline in spirometric lung function parameters such as Forced Vital Capacity (FVC) and Forced Expiratory Volume during the first second (FEV1) has been found to correlate with age [61] but also with gender [62]. Earlier diagnosis is known to be associated with better FVC and FEV1 and this was evident from small scale studies comparing PCD patients diagnosed before vs during adulthood [63, 64]. A large study that assessed the relationship between age of diagnosis and lung function expressed as percentage (%) of predicted, reported a mean annual decline of lung function across a cohort of 74 PCD patients equal to 0.8% per year [61]. Recently, female patients were found to have lower baseline lung function as well as a greater decline in FEV1 compared to male patients [62]. Sexual differences in disease severity is understudied and not

well understood in PCD but is a well-known feature of CF patients. Female CF patients do worse than males possibly due to the interaction between levels of female hormones (estrogens) and infection susceptibility [65]. Estradiol has been associated with the presence of more aggressive strains of colonizing bacteria in the lung [66] while progesterone is known to have cilioinhibitory effects [67]. Similar mechanisms may explain the possible presence of sexual differences in disease outcomes (including lung function) among PCD patients.

Genetics of Primary Ciliary Dyskinesia

Primary Ciliary Dyskinesia is the result of dysfunctional proteins that either make up the axonemal structure of motile cilia or are involved in particle trafficking across the cilium [24]. Proteomic analysis of isolated human motile cilia from respiratory epithelial revealed that they are made up by >250 proteins [23]. Consequently, it is not surprising that PCD is a genetically heterogeneous disease with 37 genes reported to date to cause PCD.

The cilia are evolutionary conserved structures. Most studies regarding the composition and structure of motile cilia and dynein arms in particular has been mostly studied in *Chlamydomonas reinhardtii*. *Chalmydomonas* is a unicellular biflagellate aquatic organism (alga) which allowed the easy generation and detection of immotile mutants and subsequent morphological and biochemical analysis [22]. The first gene that was found to cause PCD, *DNAI1*, was first described in 1999 and was discovered through a candidate gene approach which relied on the presence of human orthologs in Chlamydomonas (specifically intermediate chain 78) that caused a flagellar dynein arm defect. *DNAI1* in humans, and its ortholog *IC78* in *Chlamydomonas,* code for a dynein axonemal intermediate chain [68] which is part of the outer dynein arm. The dynein arms are multiprotein complexes that are formed

by polypeptides of different size and that are characteristically called heavy, intermediate and light (polypeptide) chains. Based on Chlamydomonas studies, the structure of dynein arms is largely understood and it is now known that outer arms are made up from three heavy chains, two intermediate chains and eight light chains while the more variable inner dynein arm is comprised of at least seven isoforms of eight heavy chains and three intermediate and three light chains [21]. The most commonly mutated gene in PCD, *DNAH5*, was first described in 2005 by Olbrich et al using homozygosity mapping. *DNAH5* (OMIM: 603335) mutations result in mislocalization of DNAH5 protein (a heavy chain protein), abnormal outer dynein arms and immotile cilia beat pattern [69]. *DNAH5* mutations are considered the most common in PCD following large scale studies that demonstrated that mutations in *DNAH5* were found in 28% of a total of 134 patients with PCD from 109 unrelated families [70]. Another gene that is commonly mutated in PCD is *DNAH11* (OMIM: 603339) gene which also codes for a heavy chain protein in outer dynein arms and is the most frequently mutated gene in PCD patients in which dynein arms appear normal in TEM micrographs [24]. DNAH11 protein localizes only on the proximal region of respiratory cilia and its absence or truncation results in hyperkinetic but characteristically stiff cilia beating. This phenotype demonstrates the importance of DNAH11 protein driven bending of the proximal axonemal region during both effective and recovery strokes [71, 72].

However, causative mutations for PCD are not restricted to genes coding for dynein arm components but also expand to genes coding for radial spokes, nexin links, proteins that make up the central pair apparatus and cytoplasmic dynein arm pre-assembly or scaffold proteins. A characteristic example of a genetic defects resulting in radial spoke abnormalities are biallelic mutations in genes *RSPH9* (OMIM: 612648) and *RSPH4a* (OMIM: 612647). Radial spokes are T-shaped structures spaced along the ciliary axoneme in regular intervals and are composed of a multiprotein “stalk” and “head” component (Figure 4). The “stalk” component

of radial spoke anchors on the outer microtubular doublets while the "head" component is attached to the central pair apparatus and evidence from *Chlamydomonas* indicate that they provide a mechanosignaling link between the central pair apparatus and inner dynein arm activity to fine tune beating velocity and beating waveform of the cilium [73, 74]. RSPH9 and RSPH4a are radial spoke "head" proteins that when defective, cause the cilia to exhibit an abnormal rotational beating pattern that resembles the beating pattern of "9+0" nodal cilia. This finding confirmed the previous observations from *Chlamydomonas* that radial spoke "head" and central pair apparatus interaction is essential for the establishment of the characteristic waveform of "9+2" motile cilia [75].

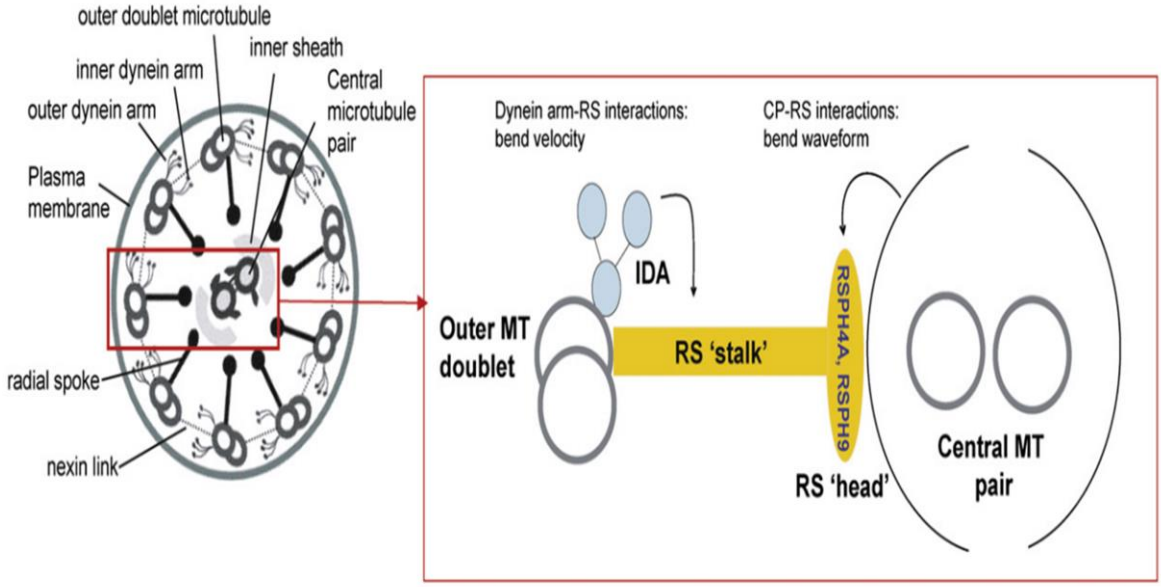


Figure 2.4: *Localization of RSPH9 and RSPH4a. Adapted from Castleman et al 2010* [75]

Another category of known genetic defects regard to nexin link defects were firstly described as causative for PCD in 2015. Biallelic mutations in *GAS8* (OMIM: 616726) gene which codes for a protein that spans across the whole nexin-dynein regulatory complex (for simplicity, usually termed as nexin link) have been found to be associated with absence of observable nexin links in TEM micrographs [76]. Nexin links, through binding CCDC39 and

CCDC40 heterodimers, attach to the A-tubule of the outer doublet and extend towards the B-tubule of the adjacent outer doublet [77]. It is believed that nexin links are involved in signal transduction but also that nexin links have an important role in regulating the microtubular doublet sliding which drives the cilia bending through either possible elastic properties of the subunits that make up the nexin link or a "release and re-attach" mechanism following the sliding movement of the microtubules [77]. Mutations in one of the largest genes found to be causative for PCD, 86 exons wide *HYDIN* (OMIM: 610812), are characteristic of genetic defects that involve directly the central pair apparatus. HYDIN protein is part of the C2b projection of the apparatus, and when defective, results in very subtle beating defects and considerable less frequency of situs abnormalities in a similar fashion as radial spokes defects [78] (Figure 5).

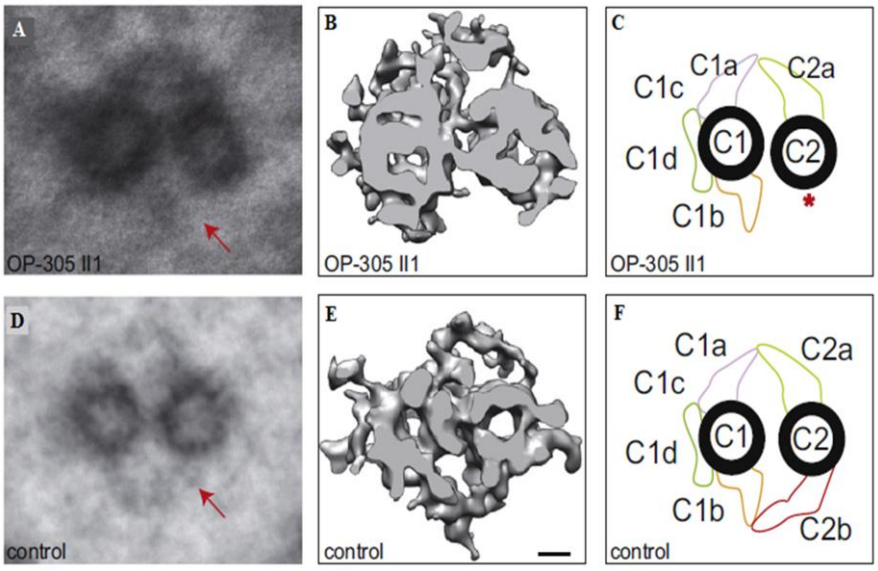


Figure 2.5: *A: TEM micrograph of a PCD patient with HYDIN mutations, B: Electron tomogram identified lack of C2b projection, C: Schematic diagram of central pair apparatus in PCD patient with HYDIN mutations. D,E,F correspond to TEM microgram, electron tomogram and scematic diagram of a healthy control. Adapted from Olbrich et al 2012*[78].

Genetic defects involving genes such as LRRC6 (OMIM: 614930), DNAAF1 (OMIM: 613190) and others that take part in the cytoplasmic pre-assembly and transport of ciliary axonemal proteins (i.e. DNAI1 or DNAH5) usually result in absence of these proteins from

the whole ciliary axoneme and frequently mislocalized staining of the axonemal proteins (i.e.DNAI1 or DNAH5) in the cytoplasm instead of the axoneme during immunofluorescence analysis [79, 80] (Figure 6).

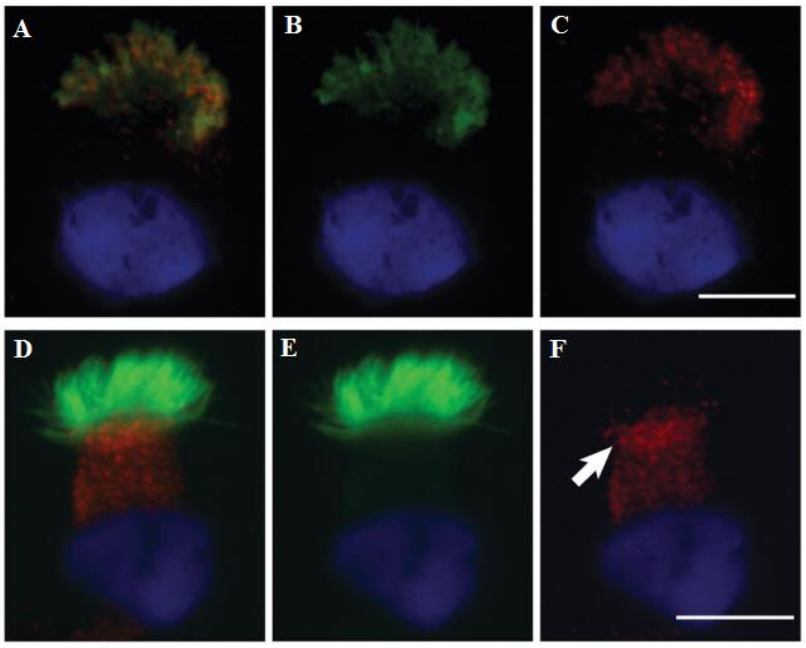


Figure 2.6: *Immunofluorescent staining in ciliated cells from healthy control (A,B,C) and PCD patients with biallelic mutations in LRRC6 (D,E,F). A: Staining overlay in healthy control, B: Tubulin staining in healthy control, C: DNAI1 staining in healthy control. D: Staining overlay in PCD, E: Tubulin staining in PCD, F: DNAI1 staining in PCD, DNAI1 stays trapped in the cytoplasm due to defective pre-assembly factor LRRC6. Adapted from Horani et al 2013*[80].

Interestingly, two additional genes, CCNO and MCIDAS also described as PCD related genes, have been shown not to cause immotility or dyskinesia of motile cilia but rather complete absence of these organelles from the airway epithelium. Mutations in these genes result in defective multiciliated cell differentiation and eventually reduced generation of multiple motile cilia. In patients with *CCNO* (OMIM: 607752) or *MCIDAS* (OMIM:614086) pathogenic mutations, severely reduced numbers of both motile cilia and basal bodies were observed in TEM studies indicating defective centriole formation which is an essential organelle for generation of multiple motile cilia. *MCIDAS* codes for Multicilin protein whereas *CCNO* codes for Cyclin O protein, both of which are part of the same pathway that

regulates assembly and maturation of centrioles. Multicilin acts upstream of Cyclin O but also regulates the expression of *FOXJ1* (OMIM 602291) which is known to be a prerequisite for centriole docking on the apical membrane of cells and expression of motile cilia proteins (e.g. DNAH5) [81, 82]. A comprehensive list of all genes that have been found to be causative for PCD (including CCNO and MCIDAS) until April 2017 is provided in Table 1.

Although many genes have been found to be causative for PCD, it is considered a monogenic disease that is usually passed down from one generation to the next through autosomal recessive mode of inheritance (two copies of one single defective gene are required to cause the disease). However there are few reports that highlight the probability of alternative modes of inheritance in PCD. More specifically an X-linked mode of inheritance was demonstrated in males carrying *PIH1D3* mutations [83] as well as in syndromic forms of PCD in males carrying *RPGR* and *OFD1* mutations. On the contrary, there is no unequivocal evidence of a digenic mode of inheritance in PCD (heterozygous mutations in two different PCD genes) as only one study reported ciliary abnormalities in patients with heterozygous mutations in candidate PCD gene *DNAH6* and known PCD gene *DNAH5* [84]. In addition there is no additional evidence of oligogenic or polygenic inheritance in PCD. In general, it is considered unknown what is the effect of multiple heterozygous mutations or single nucleotide polymorphisms (SNPs) in known PCD genes or in genes with unknown penetrance. Up to date, an important fraction of PCD patients lack a genetic diagnosis and the possibility of a polygenic effect of SNPs across multiple genes cannot be ruled out. Nevertheless, the introduction of whole exome sequencing and developments in understanding of disease pathophysiology has allowed clinicians to characterize a very high percentage of PCD subtypes [85]. Taking into account the early development of gene therapy approaches for PCD [86-88] along with the characterization of the causative genetic defect in each PCD patient, it is possible that in the future, PCD patients or at least a subset of them

with specific genetic defects could be the recipient of novel personalized medicine treatments based on genetic editing or gene silencing techniques [89]. Gene editing techniques may include zinc finger nucleases, TALENs or CRISPR-Cas9 [90] while gene silencing techniques may additionally include antisense oligonucleotides, siRNAs or miRNAs [91]. Furthermore, the greater characterization of the genetic defects as well as the improved description of the phenotypic expression of the disease may allow for the classification of patients in terms of risk of a more severe disease progression as early as with the completion of diagnostic tests. This knowledge on genotype-phenotype correlation will allow clinicians to more closely monitor and more aggressively treat the most susceptible PCD patients thus reducing complications in later life as it has been shown in CF [92] and as it has been suggested in PCD [93, 94].

Table 2.1: A summary of PCD-related genes

#	Gene	Chromosomal Location	Protein localisation/function	Ultrastructural defect	OMIM	References
1	DNAH5	5p15.2	ODA - Heavy Chain	ODA defect	603335	[69]
2	DNAI1	9p13.3	ODA - Intermediate Chain	ODA defect	604366	[68]
3	DNAI2	17q25.1	ODA - Intermediate Chain	ODA defect	605483	[95]
4	DNAL1	14q24.3	ODA - Light Chain	ODA defect	610062	[96]
5	TXNDC3	7p14.1	ODA - Light Chain	ODA defect	607421	[97]
6	ARMC4	10p12.1	ODA – Docking Complex	ODA defect	615408	[98]
7	DNAAF1	16q24.1	Cytoplasmic, axonemal assembly	ODA+IDA defect	613190	[99]

8	DNAAF2	14q21.3	Cytoplasmic, axonemal assembly	ODA+IDA defect	612517	[100]
9	DNAAF3	19q13.42	Cytoplasmic, axonemal assembly	ODA+IDA defect	614566	[101]
10	HEATR2	7p22.3	Cytoplasmic, axonemal assembly	ODA+IDA defect	614864	[102]
11	LRRC6	8q24.22	Cytoplasmic, axonemal assembly	ODA+IDA defect	614930	[79]
12	DYX1C1	15q21.3	Cytoplasmic, axonemal assembly	ODA+IDA defect	608706	[103]
13	ZMYND10	3p21.31	Cytoplasmic, axonemal assembly	ODA+IDA defect	607070	[104]
14	SPAG1	8q22.2	Cytoplasmic, axonemal assembly	ODA+IDA defect	603395	[105]
15	C21orf59	21q22.11	Cytoplasmic, axonemal assembly	ODA+IDA defect	615494	[106]
16	CCDC39	3q26.33	Nexin Dynein Regulatory Complex	AD + IDA defect	613798	[107]
17	CCDC40	17q25.3	Nexin Dynein Regulatory Complex	AD + IDA defect	613799	[108]
18	CCDC114	19q13.33	ODA – Docking Complex	ODA defect	615038	[109]
19	CCDC164	2p23.3	Nexin Dynein Regulatory Complex	Nexin link defect	615288	[110]
20	CCDC65	12q13.12	Nexin Dynein Regulatory	Nexin link defect	611088	[106]

			Complex			
21	GAS8	16q24.3	Nexin Dynein Regulatory Complex	Nexin link defect	616726	[76]
22	CCDC151	19p13.2	ODA – Docking complex	ODA defect	615956	[111]
23	CCDC103	17q21.31	Cytoplasmic, axonemal assembly	ODA + IDA defect	614677	[112]
24	HYDIN	16q22.2	Central Pair	CP defect	610812	[78]
25	DNAH11	7p15.3	ODA – Heavy Chain	Normal Ultrastructure	603339	[72]
26	RSPH1	21q22.3	Radial Spokes	CP + RS + AD defect	609314	[113]
27	RSPH4A	6q22.1	Radial Spokes	CP + RS + AD defect	612647	[75]
28	RSPH9	6p21.1	Radial Spokes	CP + RS + AD defect	612648	[114]
29	RPGR	Xp11.4	Cytoplasmic	PCD accompanied by X-linked Retinitis Pigmentosa	312610	[115]
30	OFD1	Xp22.2	Cytoplasmic	PCD accompanied by X-linked mental retardation	300170	[116]
31	CCNO	5q11.2	Cytoplasmic, basal body migration	Absence/reduced cilia	607752	[81]
32			Cytoplasmic, basal body	Absence/reduced		[82]

	MCIDAS	5q11.2	migration	cilia	614086	
33	DNAJB13	11q13.4	Radial spokes	CP defect	610263	[117]
34	RSPH3	6q25.3	Radial spokes	CP + RS + AD defect	616481	[118]
35	TTC25	17q21.2	ODA – Docking complex	ODA defect	617095	[119]
36	DYNC2H1	11q22.3	Ciliary intraflagellar transport	Unknown	603297	[120]
37	PIH1D3	Xq22.3	Cytoplasmic,axonemal assembly	ODA+IDA defect	300933	[83]

OMIM: Online Mendelian Inheritance in Man Database, ODA: Outer Dynein Arm defect, IDA: Inner Dynein Arm Defect, AD: Axonemal Disorganization, CP: Central Pair, RS: Radial Spokes,

Animal Studies in Primary Ciliary Dyskinesia

A number of different animal models have been used in previous decades for the study of ciliary development, ciliary function and ciliary abnormalities. Given that the ciliary axonemal structure has been highly conserved throughout evolution [121] employed animal models ranged from unicellular eukaryotes like *Chlamydomonas reinhardtii* [122] and *Tetrahymena Thermophila* [123] to amphibian species like the embryos of *Xenopus laevis* [124], fish species like *Medaka* [125] and *Zebrafish* [126] to mammals like mice [127], and rats [128]. Several developments in the understanding of ciliary structure and function as well discoveries in PCD genetics and PCD pathophysiology is the result of studies performed in different animal models [121]. In the past, comparative genomic approach that compared the proteins found in non-flagellar and non-ciliated organisms like *Arabidopsis, Saccharomyces* and *C. elegans* with the proteome of ciliated organisms such as *Chlamydomonas, zebrafish* and humans led to the identification of 200 genes that constitute primary targets for genes involved in PCD [129]. Furthermore, a great amount of information regarding the waveform of ciliary beating as well as regulation ciliary and flagellar activity has been generated

through studies performed in *Chlamydomonas* [130-132] while several studies have studied the regulation of ciliary beat frequency in mammals [133-135]. Lastly, studies performed in Xenopus embryos have provided crucial insights in the development of planar cell polarity and directional ciliary beating [136] as well as determination of left-right axis during embryogenesis [137]. In summary, up to date, most animal studies have provided evidence regarding the biology of cilium development and function and pathophysiology of ciliopathies. However, with reducing gaps in knowledge about cilium biology and pathophysiology of disease, focus is expected to turn towards employing animal models in the quest for effective treatment approaches in PCD and other ciliopathies. In recent years a number of animal studies have been published demonstrating the effect of novel treatment approaches on ciliary structure and function [87, 138-140]. More specifically, the study by Ostrowski et al demonstrated that gene transfer to undifferentiated cultures of mutant Dnaic1-/- mouse cell through a lentiviral vector restored Dnaic1 expression and ciliary motility. The same study however, demonstrated low levels of efficient gene transfer to the nasal epithelium due to the presence of severe rhinitis in mutant mice [87].

In general, animal models will remain a precious tool in the understanding of PCD and other ciliopathies and it is expected that animal models will play a key role in the development of novel treatments for this class of diseases. Nevertheless animal models are just a proxy of human disease and several limitations in their applicability or difficulties in their use may apply. Ciliary abnormalities may primarily affect different organ systems in different animal models and in humans (e.g. in adult *zebrafish* the primary site of motile cilia involvement is the nephron, in *Xenopus laevis* is the skin while in mice and animals is the respiratory system [121]. Furthermore, although abnormal ciliary motility can be reliably reproduced in animal models, additional manifestations may be observed compared to humans such as increased frequency of hydrocephalus (likely attributed to anatomical differences of brain ventricles

between mice and humans) and increased frequency of cardiac abnormalities in mice and rats [141, 142]. These additional manifestations may significantly impede the long-term observation of disease progression or treatment results in adult mice and rats in future whole animal testing of candidate compounds [127].

Diagnosting Testing for Primary Ciliary Dyskinesia

The high genetic heterogeneity and phenotypic variability observed in PCD along with the presence of non-specific respiratory symptoms (e.g rhinorrhea, cough) make the establishment of PCD diagnosis a challenging task [11]. Although a number of different diagnostic tests have been developed to date and are in use in PCD referral centers across the world, no single test is considered as the gold standard for diagnosis, as none has been found to be 100% sensitive and 100% specific [16]. The tests developed for PCD diagnosis include the indirect in vivo assessment of mucociliary clearance in the nose (saccharine test), the measurement of nasal Nitric Oxide (nNO), the assessment of ciliary motility using simple Light Microscopy or High Speed Video Microscopy (HSVM), the assessment of ciliary ultrastructure using transmission electron microscopy (TEM), immunofluorescence analysis, electron tomography of ciliary ultrastructure, radiolabeled mucociliary clearance and genetic testing [1, 18]. Among these, the measurement of nNO, the examination of cilia ultrastructure with TEM and the evaluation of ciliary motility with HSVM are the more established, better validated and more frequently used for PCD diagnosis [15, 143]. A detailed description of nNO, HSVM and TEM is provided in the next sections.

The saccharine test involves the placement of a small (1mm wide) saccharine tube approximately 1 cm from the nasal inferior turbinate and the time period required for the patient to detect sweet taste is reported. A saccharine test with a mean transport time greater

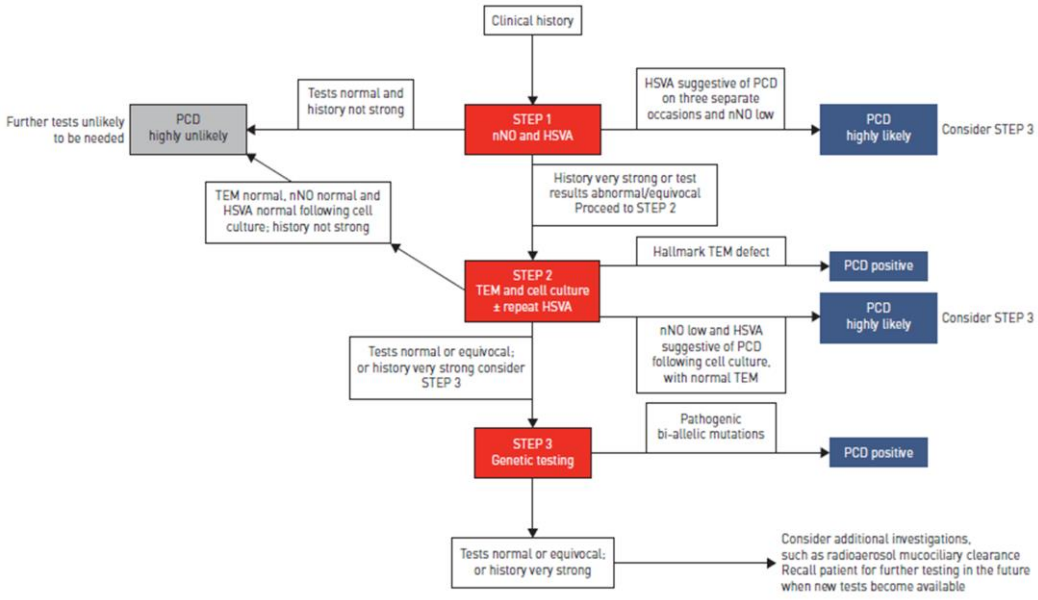
than 60 min is considered abnormal and indicative of a mucociliary disorder [144, 145]. Although saccharine test is cheap and easy to perform it suffers from many limitations and its use as a diagnostic test has now been abandoned. The main limitations were the inability to use the test in non-cooperative patients such as young children and the requirement from the subjects, not to sneeze, sniff or cough during the performance of the test, a requirement that patients suffering from respiratory symptoms couldn't easily adhere to [146].

Immunofluorescence analysis allows the visualization of the presence and localization of different proteins across the ciliary axoneme with the use of protein specific antibodies. The antibodies that also carry fluorescent tags can be visualized using fluorescent (or confocal) microscopy. Absence of mislocalization of staining signifies absence, truncation or mislocalization of the specific ciliary protein and thus can be used as a diagnostic test [14]. However, given that it is a new diagnostic technique for PCD (first described in 2005) and for many years it was used only for research purposes such as gene discovery and understanding of the effects of mutations at the protein level [147], still only few centers possess the expertise to carry out and interpret the findings of this technique which is generally considered as an auxiliary test when results from other tests are contradictory or equivocal [146].

Lastly, the discovery of the underlying genetic defect in many PCD variants has allowed the use of genetic testing as a reliable diagnostic tool with the use of either Sanger sequencing or novel Next Generation Sequencing (NGS) technologies. Through the option of whole exome sequencing or through the development of PCD specific gene panels (kits) for targeted genetic screening, NGS can be used as a stand-alone diagnostic test [13]. In contrast, genetic testing using Sanger sequencing is considerably ineffective due to the high number of PCD-related genes and the large size of these genes and as a result, a priori performance of HSVM and TEM, along with the underlying information on genotype-phenotype correlations, is

usually used to inform about the most likely causative gene to be sequenced. Genetics are only utilized as a diagnostic test in few centers in Europe and the current ERS guidelines suggest that genetic testing can be used as an additional method to further characterize the underlying defect, following abnormal results of other tests such as HSVM and TEM or as a diagnostic method when there is no availability of other diagnostic tests [148].

In general, most PCD referral centers rely on diagnostic algorithms that mainly include a combination of HSVM, TEM and nNO. Slightly different diagnostics algorithms for PCD have been suggested in recent years by the ERS [18] (Figure 7) and the FP7 BESTCILIA project (Figure 8) [17]. Both require the parallel performance of nNO and HSVM as a first step and performance of TEM as a second step if nNO and HSVM results are equivocal or even normal but the referred patient has a clinical history that is strongly suggestive of PCD [18].



The BESTCILIA algorithm suggests that PCD can be confirmed following abnormal HSVM and abnormal nNO and that repeat testing of HSVM is required to be abnormal on 3 occasions for a positive diagnosis only in the case that HSVM is the only abnormal test. The

main differences between the two algorithms lie in the requirement in the ERS guidelines for confirmation of PCD using genetic testing following abnormal nNO and abnormal HSVM and the requirement of 3 separate abnormal HSVM test results at the first step together with an abnormal nNO result.

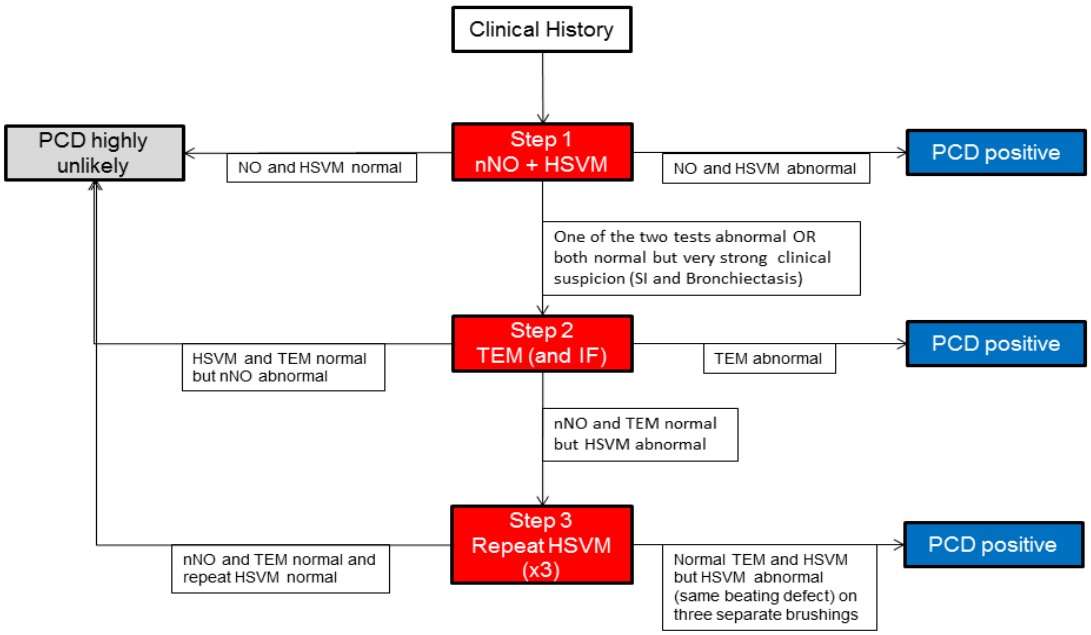


Figure 2.8: *Diagnostic algorithm for PCD as suggested by the BESTCILIA project.*

High Speed Video Microscopy (HSVM) – Cilia Beat Frequency and Cilia Beat Pattern

The functional analysis of respiratory cilia motility is performed as soon as the sample of nasal brushing or bronchial biopsy is obtained and while the cells are kept alive in a cell culture medium. The beat frequency and beat pattern of cilia are usually evaluated in a quantitative and qualitative manner respectively but additional, composite metrics like the immotility index and the percentage of dyskinetic edges have been suggested as well [149]. Overall a CBF>11 Hz (beats per second) is considered normal although a subset of patients (carriers of biallelic mutations at *DNAH11* gene) present with a hyperkinetic but also

dyskinetic beat pattern [11]. Normal cilia beat in a characteristic wavelike motion with a forward power stroke followed by a backward recovery stroke that does not sweep to the side, as opposed to dyskinetic or completely immotile cilia that have been described in PCD [50]. Dyskinetic cilia could be characterized by circular or stiff beating with reduced bending. Evaluation of CBF and CBP is possible with the use of an HSVM system which includes the light microscope, a high speed camera and relevant image analysis software. The system allows cilia movement to be recorded and played back at a slower rate thus decreasing the possibility of errors. During the playback, the user identifies healthy looking ciliated strips (continuous ciliated epithelium sections without large projections and absence of mucous and debris) and carefully evaluates the CBP. This evaluation involves the assessment of beat amplitude, beat direction and synchronization among cilia. Usually the user, observes and records the cilia from the side (Figure 9, A) and from above (Figure 9, B) in order to correctly assess these parameters.

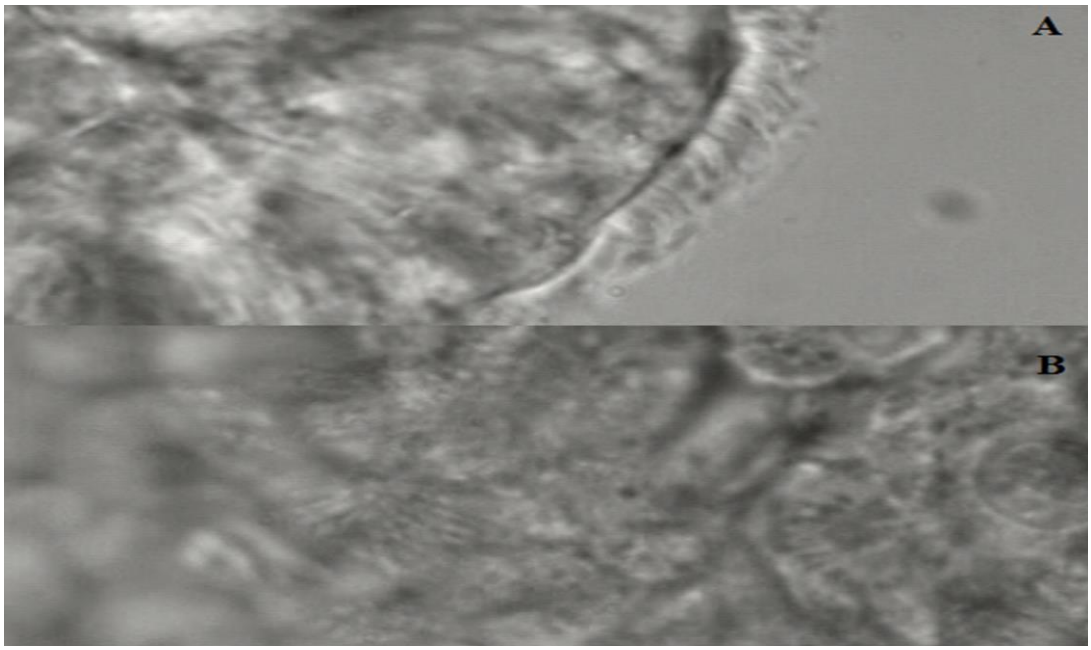


Figure 2.9: *Video recording snapshot during HSVM (SAVA system). A: Planar (side) view, B: Overview*

As opposed to CBP evaluation, newer applications have almost completely automated the measurement of ciliary beat frequency and usually also provide an array of additional metrics such as mean values and variance estimates. Several automated HSVM systems have been developed by academic centers and most of them use a combination of digital image processing techniques such as waveform analysis [150] and Fast Fourier Transformation [150-152]. The Sisson-Ammons Video Analysis (SAVA) system [150] is commercially available and is used by a large number of PCD diagnostic centers. The majority of HSVM systems require the user to observe the moving microscope image and select a region of beating cilia, followed by the automatic calculation of CBF for that specific region (often called Region of Interest-ROI). The SAVA system, provides in addition to the user the ability to just record the CBF from the whole field of the moving image (Whole Field Analysis-WFA) thus removing any potential bias that could result from the selection by the user of the ROI's [150]. Overall the measurement of CBF is now considered a straightforward and automated procedure with little room for subjectivity, while the evaluation of CBP still remains a challenge and is largely dependent to the observer's experience and subjectivity [153].

Transmission Electron Microscopy

As soon as an adequate sample of respiratory epithelial cells is obtained either by nasal or bronchial brushing, analysis of cilia ultrastructure is possible with the use of Transmission Electron Microscopy (TEM). The normal characteristic, 9+2 configuration (9 pairs of microtubules at the periphery and a central pair in the middle) of healthy cilia, can be visualized in cross-sections of the cilia axoneme [24]. Analysis of cilia ultrastructure in PCD suspect patients includes not only the assessment of the 9+2 configuration but also the careful

examination of dynein arms (protein complexes that consist of light, intermediate and heavy chains) that are attached on the peripheral tubules. Defective (absent or short) dynein arms are a common structural abnormality that is recovered in PCD patients. Central pair and or peripheral tubules abnormalities can also be detected in PCD patients, as well as microtubular disorganization and problems with the orientation of the cilia [154]. Usually ultrastructural defects of the axoneme are categorized as outer dynein arm (ODA) defect (short/absent), inner dynein arm (IDA) defect (short/absent) combined ODA and IDA (ODA/IDA) defect, central pair (CP) defect (degradation, absence), microtubular disorganization (MTD) defect and orientation defect [50]. For many years in the past, cilia ultrastructural analysis with TEM was considered as the gold standard for the diagnosis of PCD but recent advances in the genetics of PCD have demonstrated that an important subset of patients may be missed by this technique primarily due to the limitations of the existing technology of electron microscopy [11]. Characteristic examples of ultrastructural defects that have been implicated in PCD are presented in Figure 10.

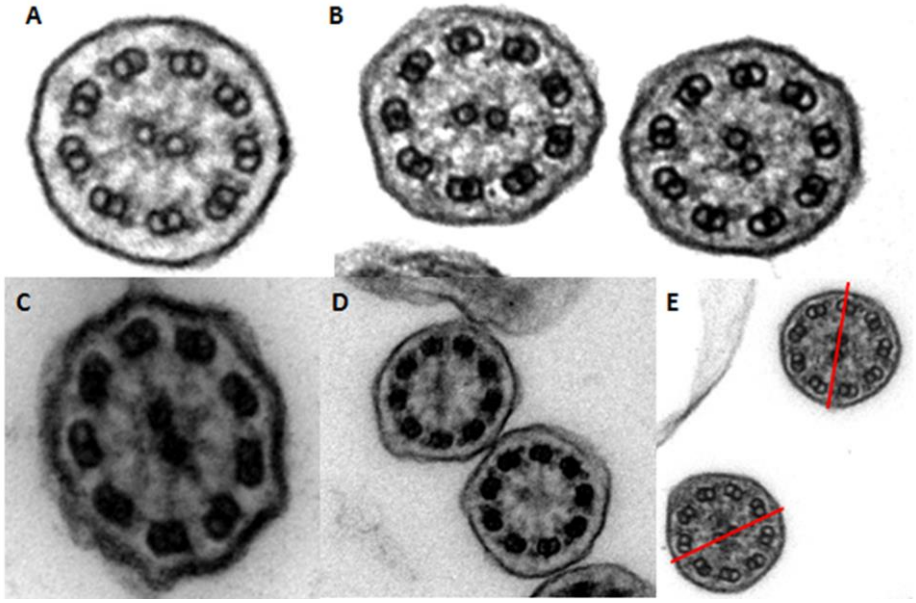


Figure 2.10: *TEM micrographs of ciliary axonemal cross-sections. A: Normal Ultrastructure, B: Absent ODA, C: Absent ODA+IDA, D: Central pair defect, E: Orientation defect*

Nasal Nitric Oxide

Low levels of nasal nitric oxide (nNO) in PCD patients were first described by Lundberd [155] more than 20 years ago, a finding that has been verified by several researchers both in case control and prospective cohort studies [156]. NO, a small free radical molecule, is well known as a hazardous environmental pollutant in environmental sciences but also as a very important signaling molecule in human physiology. Better studied for its role in control of blood flow and neurotransmission [157], NO is also present in the airways with the primary site of production being the paranasal sinuses [158]. In the respiratory system directly or through its metabolites, NO exerts a variety of biological roles including assisting bronchodilation and restriction of airway hyperresponsiveness, facilitation of pulmonary blood circulation and modulation of airway secretions [157]. In addition, NO has a bacteriostatic, anti-biofilm action and a role in modulation of CBF [159]. Measurement of NO, either of nasal origin (from the nares) or exhaled from the lower airways (from the mouth) can be achieved in the clinical setting by equipment using the chemiluminescence or the electrochemical methods [159]. A recent review, which synthesized the available published evidence until 2014, reported a mean difference of 231nl/min between PCD patients versus healthy controls and a mean difference of 114 nl/min between PCD patients versus Cystic Fibrosis patients [156]. While ciliary ultrastructural and functional assessments require significant investment in terms of equipment, time and expertise, measurement of nNO is fast and easy to perform and could serve as a first line screening test in PCD. A screening test could be potentially useful in clinical practice as the main PCD manifestations from the upper and lower airways are quite common in the pediatric population and only a small fraction are caused by PCD [159]. As a result measurement of NO could allow for the earlier identification of PCD patients and reduce the burden of unnecessary nasal brushings and sophisticated diagnostic testing in non-PCD subjects that present with PCD-like

symptoms. On the other hand, patients missed by nNO screening will not undergo additional diagnostic testing and could potentially be missed completely by such a diagnostic algorithm.

Regardless of the value of nNO measurement as a screening test, neither the etiology of its low levels in PCD nor the biological mechanism relating NO to ciliary function has been fully elucidated. The prevailing hypotheses suggest that the low nNO levels in PCD could be explained by either increased breakdown of NO, decreased synthesis of NO or entrapment of NO in, blocked by mucus, paranasal sinuses [159]. The hypothesis of increased NO breakdown is mainly supported by data showing increased levels of oxidative stress markers in PCD patients versus controls in combination with findings that suggest that these oxidative markers are coupled with the presence of peroxynitrite and nitrogen dioxide that are known derivatives of NO breakdown [158]. However these results were not replicated in an independent case-control study [160]. The second hypothesis supports that mucus leads to entrapment of NO in the paranasal sinuses or that the paranasal sinuses are not adequately developed in PCD patients. A recent study has shown that there is a significant correlation between the level of sinus aplasia and nNO values in an Italian cohort of PCD patients [161], giving some support to this hypothesis as opposed to the entrapment of NO in the paranasal sinuses which is quite unlikely as methods that increase sinuses ventilation did not have significant effect on nNO in PCD patients [162]. The last hypothesis suggests that there is reduced biosynthesis of NO due to either decreased expression of NO synthetase (NOS) isoenzymes, mechanochemical uncoupling or lack of NOS substrate L-Arginine. One recent study provided evidence of reduced expression of NOS in PCD [163] but more studies are needed whereas the possibility of genetic linkage with NOS genes' polymorphisms is low, given that PCD is a polygenic disease [159]. On the other hand, the reduced biosynthesis of NO could be the result of defective mechanochemical coupling of cilia with NOS in PCD in a similar fashion as dystrophin gene mutations lead to uncoupling on neuronal NOS from the

contractile apparatus and this also results in lower NO levels in serum [164]. With regards to the availability of L-Arginine, currently there are no data comparing levels of L-Arginine in PCD patients and healthy controls. However, some studies have demonstrated that intravenous and nebulized administration of L-Arginine resulted in an increase in CBF and nasal NO values [165, 166].

Epidemiology of Primary Ciliary Dyskinesia

The lacks of awareness for PCD among practicing clinicians as well as the genetic heterogeneity characterizing this disease and the difficulties in performing and interpreting the diagnostic tests for PCD, result in under-diagnosis and misclassification of PCD patients [5]. The true prevalence of the disease among live births is not known and only few studies have attempted to calculate measures of PCD frequency and describe its epidemiology. Based on radiological findings (Situs Inversus combined with Bronchiectasis) from plain chest X-rays (CXR's) in a large Norwegian population, Torghensen reported the prevalence of Kartagener's Syndrome to be close to 1:40000 [167]. This number is considered to be a significant underestimation as bronchiectasis may not be evident in plain CXR's, may have not yet developed in young patients whereas Kartagener's syndrome is present in about half of PCD affected individuals [5]. More recent studies have reported a prevalence of 1:10000 in Sweden [168] and approximately 1:4000 among atomic bomb survivors in Japan [169]. An estimate of 1:15000 is usually used in published reports although it is believed to still be an underrepresentation of the real burden of PCD [24, 145]. Furthermore, the most recent European survey involving 223 PCD centers from 26 countries reported that small countries with centralized national reference centers for PCD exhibited the highest prevalence of PCD among the pediatric population (5-14 years old) in the continent. With a frequency of

1:11000 Cyprus had the highest prevalence of PCD among children followed by Denmark and Switzerland with a prevalence of 1:20000. The same survey however identified large differences among participating European centers in terms of diagnostic approach and patient reporting, indicating the difficulty of pinpointing the actual true incidence and prevalence of PCD across Europe [5]. Discrepancies in availability, performance and interpretation of diagnostic tests will continue to contribute to the large differences in the prevalence estimates among different countries and will continue to play an important part in discrepancies regarding the age of diagnosis among different countries or regions. Per capita national health expenditure and the presence of a centralized PCD diagnostic center supported by a referral network involving pediatricians, pulmonologists, otorhinolaryngologists and cardiologists have been reported to influence prevalence of diagnosed PCD cases [15].

Management and Treatment of Primary Ciliary Dyskinesia

Currently there are no evidence-based management guidelines for PCD.As an orphan disease very few short-term and no long-term randomized clinical trials have been performed in PCD patients. As a result most treatment protocols for PCD are largely extrapolated from the CF literature and empirical evidence [143]. The current approaches focus on facilitating the removal of secretions from the lung using physiotherapy and airway clearance techniques and on infection control by prescribing antibiotics to treat breakthrough infections or long term antibiotics as a prophylactic measure for incident respiratory infections.

Given that the pathophysiological mechanism that results in recurrent respiratory infections in PCD is the impairment of mucociliary clearance, airway clearance techniques to facilitate secretions removal from the lung have a central role in PCD management [170]. Airway clearance techniques involve forced cough and several breathing control and assisted

expulsion techniques that can be used alone or in combination to each other in order to clear mucus from the lung. Such physiotherapy techniques are Chest Percussion, Vibrations, Postural Drainage, Forced Exhalation Techniques and Autogenic Drainage. Other breathing techniques that require use of accessory mechanical equipment are Positive Expiratory Pressure, Oscillating Positive Expiratory Pressure and High Frequency Chest Wall Compression [171]. In addition, all forms of exercise which induce increase in ventilation rate and increased bronchodilation are considered to have a beneficial effect in mucus clearance and may be used in addition to standard airway clearance techniques. Notably, performance of exercise prior physiotherapy may further increase mucus clearance during the airway clearance techniques [172]. An array of mucolytic agents such as recombinant human deoxyribonuclease I (rhDNase) and N-acetylcysteine and hyposmolar agents such as hypertonic saline, mannitol and Uridine 50 Triphosphate are commonly used in CF management to facilitate mucus clearance but their effectiveness in PCD still remains unclear [170]. The results of the first clinical trial that evaluated the use of hypertonic saline in PCD patients were published in 2017 and results were mostly negative as no statistically significant differences in health-related quality of life parameters (main outcome) were demonstrated [173].

Antibiotic administration is used aggressively to treat or prevent recurrence of respiratory tract infections and should be combined with routine microbial surveillance through sputum cultures or oropharyngeal cultures for very young patients. Commonly cultured pathogens are Haemophilus Influenzae, Staphylococcus Aureus and Streptococcus pneumonia but Pseudomonas Aeruginosa species have also been reported but mostly in adults [143]. Administration of oral antibiotics in high doses is suggested as a first response to worsening respiratory symptoms, deterioration in lung function or positive sputum culture. In case of no response to oral antibiotics, intravenous antibiotics can be administered. Patients with

Pseudomonas Aeruginosa positive cultures are treated with three-month eradication protocols (as in CF) which usually include oral and inhaled antibiotics [146].

Other prescribed medications for PCD patients may include inhaled or systemic corticosteroids but evidence regarding the efficacy of these approaches is still lacking. Similarly, evidence regarding the performance of surgical procedures in PCD is still quite poor. Surgical resection of lung segments (lobectomy) should only be considered with caution, in selected cases of localized bronchiectasis and when other approaches were not efficient [143]. Nevertheless, this approach remains controversial and to date there are only two reports in the literature describing the clinical course in PCD patients after lobectomy with conflicting results [63, 174]. The first study by Smit et al compared lobectomised adult PCD patients (n=13, age range: 32-61 years) to non-lobectomised adult PCD patients (n=8, age range: 24-66 years) and did not find any significant differences in the prevalence of respiratory symptoms between the two groups. Despite this, 85% of the lobectomised patients subjectively experienced the operation as beneficial [174]. In a more recent study, Yiallouros et al. compared 5 lobectomised PCD patients (lobectomy performed prior to PCD diagnosis, age range: 37 -49 years) with 7 non-lobectomised PCD patients (age range: 30-64 years) and reported that patients with lobectomy had a more severe clinical picture at time of diagnosis and consistently lower lung function across time compared to the non-lobectomised PCD patients [63]. Both reports were single-center studies including small number of patients and generalizability of their results should be avoided. An international multi-center study is currently under way with the aim to recruit a larger number of PCD patients and synthesize information from several PCD specialized centers across the world in order to provide data on the prevalence of lobectomy in PCD patients, describe the course of the disease after lobectomy and determine factors associated with poorer outcomes in lobectomised PCD patients [175]. Patients with end stage lung disease could be referred for lung transplantation,

although there are no specific transplantation referral criteria. A number of case reports or case series have described successful lung transplantations [176, 177] or living donor lobar transplantations [178] in PCD but long term survival of these patients remains unknown.

Quality of Life in Primary Ciliary Dyskinesia

Patients affected by PCD are considered to have a "near normal" life expectancy [179], although no data regarding average life expectancy have been published to date [180]. However, progression of bronchiectasis with ascending age could be associated with a significant reduction in lung function and possible complications such as respiratory failure. In reality, evidence from many longitudinal studies and registries indicates that PCD patients may exhibit a wide spectrum of lung disease severity that could be associated with normal lifespan or could result in early death or need for lung transplantation in early adulthood [180]. During their lifetime, PCD patients do suffer from progressive deterioration of lung function compared to their healthy peers, as longitudinal lung function studies have highlighted [61, 64, 181]. Reduction of lung function has been strongly associated with earlier death in large population cohorts [182, 183] and poorer Health Related Quality of Life (HRQoL) in Cystic Fibrosis (CF) [184, 185], a disease which is characterized by similar but more profound disease manifestations than PCD. This relationship of lung function with HRQoL has also been demonstrated in Chronic Obstructive Pulmonary Disease (COPD) ([186, 187].

Symptoms affecting the upper and lower airways also contribute to the morbidity burden of PCD as chronic rhinosinusitis, nasal polyps and chronic wet cough are hallmarks of the disease. All of these have been found to affect various aspects of the patients' life and mainly the psychosocial component of HRQoL [188, 189]. In a similar fashion mucus plugging of

the middle ear results in frequent ear infections and subsequently insertion of ventilation tubes or use of hearing aids, both with important social implications for the individual [190]. A systematic review summarized all available studies on the effect of PCD in HRQoL. In general, patients reported that it was difficult to keep up with their peers, they were feeling easily tired while in terms of social impact, patients reported feeling embarrassed, isolated and very frequently tried to conceal diagnosis or symptoms. PCD patients also reported feelings of anxiety about getting sick or about their future health [191]. A common conclusion of all studies was that PCD has a severe impact on quality of life and that earlier diagnosis and hence earlier treatment significantly improves the impact of the condition [191-193]. However, most of these studies relied on HRQoL questionnaires for other respiratory diseases such as the St George's Respiratory or generic health related quality of life questionnaires such as the Medical Outcomes Study Short Form-36 questionnaire. Since 2016, a PCD-specific HRQoL questionnaire has been developed for all age groups [194, 195] and has already been translated in different languages and is currently in use in a number of centers. Although HRQoL in PCD has been studied to some extent and a PCD-specific HRQoL questionnaire has been developed recently, the health utility (or health preferences) of PCD patients has not yet been described. As a result the number of Quality Adjusted Life Years (QALYs) lost in PCD cannot be estimated. Health preferences' studies aim to describe not the actual symptoms (or feelings) that patients may experience (as in traditional HRQoL) but rather aim to describe the values that patients attach to their overall health status. Usually health utility scores are combined in a single value between zero and one. One representing perfect health and zero representing death [196]. Although health utility estimates for PCD are lacking, evidence from other diseases that share common symptoms with PCD do provide an estimation of what the expected health utility could be in PCD. Characteristically, in a random sample of adults, participating in the Beaver Dam Health Outcomes Study on

Chronic Bronchitis and Chronic Sinusitis, conditions that share common respiratory symptomatology with PCD, exhibited an age adjusted Quality of Wellbeing Index score of 67% and 72% respectively and an age adjusted health utility score (measured by Time Trade off Index) score of 72.4% and 87.4% respectively. Estimates of health utilities for CF using the EQ-5D questionnaire [197] have been reported to be 0.85 (95% CI: 0.80-0.89), 0.79 (95%CI: 0.67-0.91) and 0.60 (95%CI: 0.44-0.76) depending on the presence of no, mild or severe pulmonary exacerbations [198].

Chapter 3: Methodology - Systematic Review and Meta-analysis

Systematic Review

During the last three decades, the rise in the universal requirement for evidence based practice in healthcare has led to the development of several methods in order to inform healthcare practitioners about the most current evidence that is available from scientific literature towards supporting decision making [199]. The most well-known and more frequently used method to summarize available evidence for a particular topic is the performance of a systematic literature review. Unlike other type of reviews such as standard literature or state-of the art reviews, a systematic review aims to systematically search for, appraise and synthesize all available research evidence usually through adhering to specific guidelines [200]. In more detail, a systematic review uses predefined eligibility (or inclusion and exclusion) criteria to identify relevant studies that provide evidence to answer a specific scientific question while using systematic methods to minimize different types of potential biases [201]. Such methods include the detailed description of the methodology (search strategy) used towards allowing reproducibility by other researchers, systematic search across several databases to identify studies that meet the predefined eligibility criteria, an assessment of the validity of the findings and the systematic synthesis and presentation of the characteristics and findings of the included studies [202]. A systematic review may or may not be accompanied by a meta-analysis. Meta-analysis refers to the statistical approach that allows the integration of results from several (two or more) independent studies to one summary estimate [203]. Meta-analyses allow for increased power and precision as a result of greater sample size compared to individual studies and provide the opportunity to assess the consistency or differences of findings across studies. In addition, meta-analyses allow researchers to answer additional questions and develop new hypotheses to explain differences between the included studies [202].

Eligibility criteria and search strategy

The eligibility criteria for studies included in a systematic review are usually specific characteristics of the individual studies and their careful selection constitute one of the most important factors towards integrating published evidence on a specific topic. These criteria should be broad enough to allow the review to include studies with some diversity but narrow enough to ensure that the review is within scope of the research question and able to provide a meaningful answer [202]. Despite the fact the each systematic review is trying to answer a different scientific question and eligibility criteria are expected to be unique, these typically fall within one or more of the following categories: (i) Study participants, (ii) intervention evaluated, (iii) outcome(s) reported (iv) time period and (v) methodological quality [204]. Some examples of factors that a reviewer usually considers during the development of eligibility criteria are described in Table 3.1 [202].

Table 3.1: Factors to consider towards developing eligibility criteria for included studies

Category	**Examples of factors to consider**
Study participants	What was the definition of disease (e.g. diagnostic criteria)? Are there any relevant demographic factors (e.g. age or gender)? What was the setting (e.g. hospital or community based)?
Intervention Evaluated	What are the experimental and comparator interventions of interest? Are there variations in the intervention implementation regarding e.g. equipment, dosage or mode of delivery?
Outcome Reported	Primary or secondary outcomes? Quantitative or qualitative outcomes? What is the type and timing of outcome measurements?
Methodological Quality	Was the study a double blind RCT? Were the participants/interventions/outcomes described clearly?
Time Period	Only contemporary studies should be included (e.g. last 10 years)? May limit the number of eligible studies however it may be appropriate on the basis of a time point that a new intervention was introduced.

Source: [202]

Furthermore, the reviewer may apply another eligibility criterion, that of linguistic range (e.g. excluding studies that have not been published in the English language) although it is known to limit the scope of the review and additionally introduce publication bias in the results. Nevertheless, it is not an uncommon phenomenon to limit included studies to those published in English for practical reasons as the vast majority of scientific literature is published in the English language, but in fields with a lot of non-English literature this restriction should be avoided [204].

The bibliographic search for a systematic review should rely primarily on MEDLINE (PubMed) and Cochrane Central Register of Controlled Trials (where applicable) but also in other electronic databases such as EMBASE, Scopus, Google Scholar or Web of Science [205]. PubMed utilizes Medical Subject Heading (MeSH) terms which constitute a comprehensive vocabulary that facilitate indexing of journal articles and books across MEDLINE. Search strategies may utilize MeSH terms to focus the search to only appropriately indexed articles or can make use of keywords. Keywords can be searched in titles, abstracts or article text, depending on the electronic database and settings. Keywords or MeSH terms can be combined with Boolean operators such as AND, OR and NOT to create search algorithms that allow for efficient article retrieval [206]. Bibliographic search is completed with the identification of several electronic records that the reviewer will examine carefully towards deciding which of the identified records should be part of the systematic review [202, 207]. Apart from the applicability of eligibility criteria, the overall quality of the included studies should be evaluated and taken into account during the performance of a systematic review. Several individual studies of healthcare interventions as well as several individual observational studies may be characterized by significant weaknesses in study design and/or study resulting in significant risk of bias in their results. The underlying weaknesses in the included studies may lead to bias (overestimation or underestimation) in

the results of the systematic review and meta-analysis [208]. Table 3.2 summarizes common sources of bias in systematic reviews of observational studies or clinical trials.

Table 3.2: Different sources of bias in RCT and observational studies

Source of Bias	Application in clinical trial studies	Application in observational studies	How to avoid it
Selection bias	Differences in baseline characteristics of compared groups in clinical trials	Study population is not a random selection from the target population for which results are supposed to apply	Randomization*
Performance bias	Differences between groups in the care provided, or in exposure to factors other than the interventions of interest	n/a	Blinding**
Detection bias	Differences in how outcomes are determined	Differences in how outcomes are determined	Blinding
Attrition bias	Differences between groups in withdrawals from a study (loss to follow up)	Loss to follow up in cohort studies	Motivating, case management, incentives
Reporting bias	Differences between reported and unreported findings in published reports	Differences between reported and unreported findings in published reports	Predefined set of outcomes reported in study protocol
Publication bias	The *publication* or *non-publication* of research findings, depending on results' nature and direction	The *publication* or *non-publication* of research findings, depending on results' nature and direction	Avoid outcome related search terms, search trial registries, updating systematic reviews
Recall bias	n/a	Imprecise answers to questions about past events	Verify information given with a reliable third party or medical record
Measurement bias	Measurement errors arise from imprecision of the instruments, measurement procedure, or human investigator.	Measurement errors arise from imprecision of the instruments, measurement procedure, or human investigator.	Precise measurements performed in validation study in a subset of study participants
Confounding	Effects of exposure under study on a given outcome are mixed with the effects of additional factors.	Effects of exposure under study on a given outcome are mixed with the effects of additional factors.	Inclusion restricted by confounding variables(e.g. age), stratification, multivariate analysis

*Randomization: Randomly allocate interventions (or placebo) to participants **Blinding: Clinician is unaware of who received intervention or placebo Sources: [209-213]

Assessment of Methodological Quality

Assessment of methodological quality of included studies for the evaluation of the risk of bias is considered a necessary component of any systematic review and several tools have developed to facilitate reviewers to assess quality and bias for different types of studies (randomized, non-randomized, clinical trials, observational epidemiological studies) [214]. Depending on the type of systematic review and the type of included studies several tools in the form of simple checklists, summary judgment checklists and scales, have been suggested to assess methodological quality of included studies [215]. Randomized control trials (RCT) are usually evaluated using the Cochrane Collaboration's tool [207], non-randomized intervention studies are evaluated using the Methodological Index for Non-Randomized Studies (MINORS) [216], observational case-control and cohort studies are evaluated using the Newcastle Ottawa Scale [217], diagnostic accuracy studies are usually evaluated with the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool [218, 219] and animal studies are evaluated using the SYstematic Review Centre for Laboratory animal Experimentation (SYRCLE) risk of bias tool [220].

Reporting in Systematic Reviews

An important distinction exists between the quality assessment of the actual design conduct and analysis of the primary studies and quality of reporting in the primary studies. The first may refers to the propensity of the presence of biases in the primary studies and the second refers to whether a primary studies report their results properly and in a proper order. Accurate and extensive reporting is encouraged by a number of consensus statements such as the STROBE statement for observational studies [221], the CONSORT statement for non-randomized control trials [222], the STARD statement for diagnostic accuracy studies [223] and the QUOROM statement for systematic reviews [224]. The improved reporting resulting

from such statements/guidelines can ensure studies make available all necessary information to the reader and facilitate the quality assessment of the study contents. In the case of a systematic review that is accompanied by a meta-analysis the use of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement is required by many scientific journals. The PRISMA statement provides a well-defined framework for reporting and presentation of systematic review search results and it consists of a 27-item checklist, which includes instructions to reviewers regarding several aspects of the systematic review manuscript. The PRISMA statement also includes a flow diagram template that allows the author to demonstrate the review process and how the identified records have been screened and how the eligibility criteria were applied [225]. The PRISMA diagram is displayed in figure 3.1.

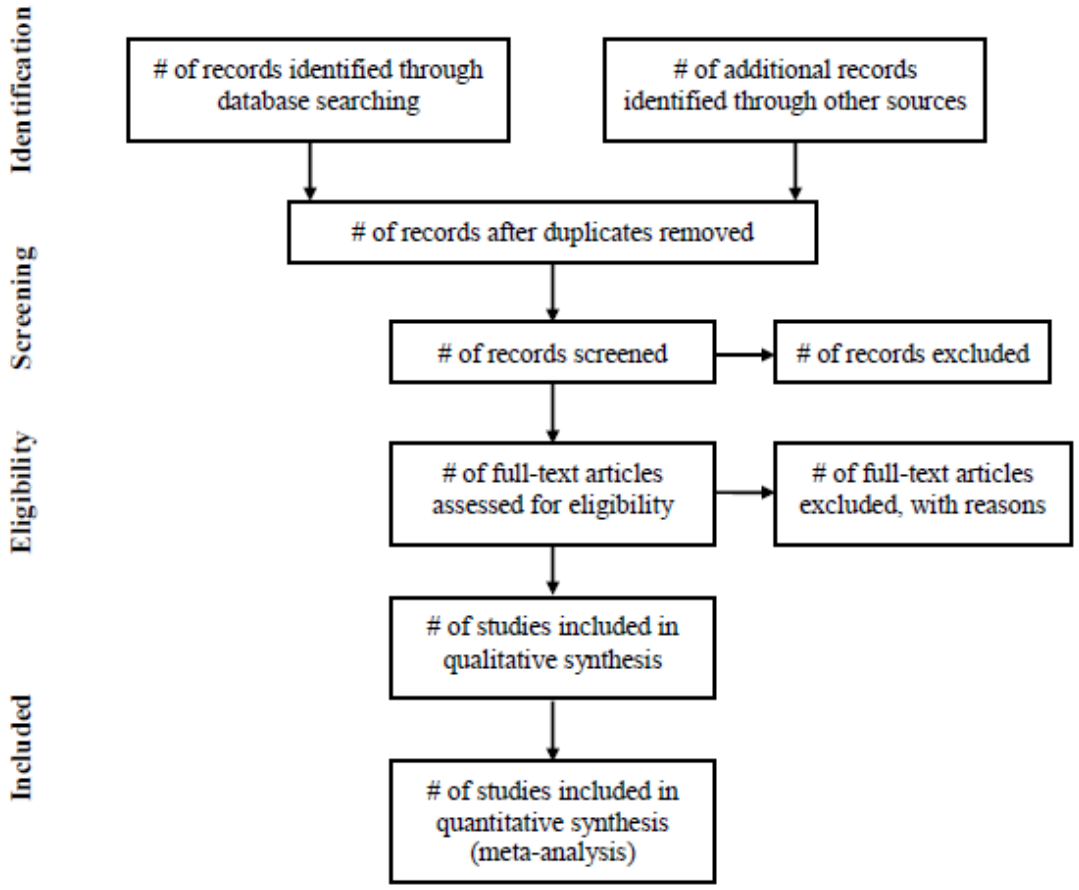


Figure 3.1: *PRISMA statement flow diagram Adapted from* Moher D et al [225].

Meta-analysis

Given that the numerical pooling of the results of the individual studies is merited, statistical methods they have been used widely in the last two decades following the increase in research findings and the requirement for evidence based decision making in healthcare [203]. In general a meta-analysis can be described as a sequence of two steps which involve, as a first step, the extraction of appropriate data from the individual studies that have addressed the specific research question in the past and that have been identified through systematic review and in a second step, the pooling of the effect estimates from each study into one single measure (e.g. Odds ratios, Relative Risks or Hazard Ratio for dichotomous data or mean differences for continuous data) taking into account the precision in the study estimate and weight of each study [226]. The weights are chosen to reflect the amount of information that can be found within a study and the resulting averaging estimate is considered a weighted average of the individual studies effects. The Confidence Intervals (CI) around this weighted average represents the precision (or alternatively the uncertainty) in the meta-analysis estimate and the p value that accompanies the effect and CI represents the statistical significance of the hypothesis that there is an effect versus the alternative hypothesis that there is no effect (null hypothesis) [202].

Meta-analyses can be represented graphically with the use of forest plots. These plots are useful in providing the reader with an overview of data that were used for the meta-analysis, the heterogeneity that exists between included studies and an estimate of the synthesized results. Forest plots are usually accompanied by the name of the first author and year of publication or other reference type of each included study but may also be accompanied by a table that presents the data that are synthesized. Different forest plots can be created depending of the type of studies synthesized in the meta-analysis. A typical forest plot in meta-analysis of observational studies is presented in Figure 3.2 which is adopted from a

meta-analysis of observational studies about the association of 25-hydroxyvitamin D with different types of cancers, namely colorectal, breast and prostate cancer [227].

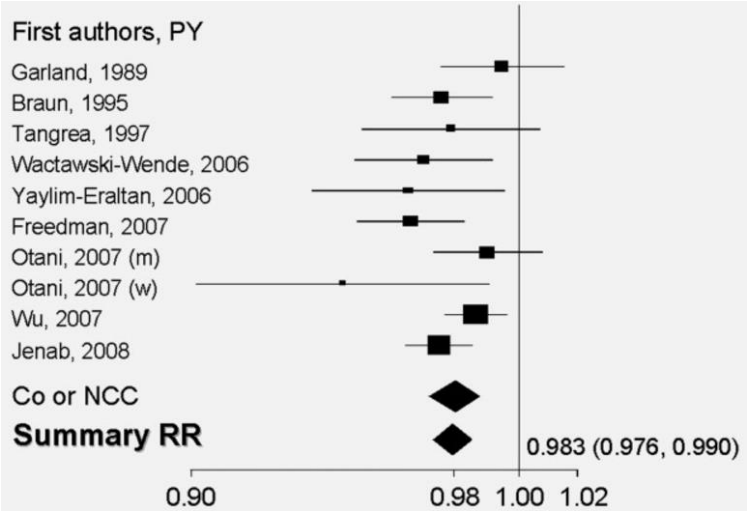


Figure 3.2: *Forest plot presenting the relative risk for colorectal cancer associated with 1 ng/ml increase in serum level of 25-hydroxyvitamin D. PY refers to publication year, RR refers to Relative Risk, Co refers to cohort study and NCC refers to case-control study nested within a cohort study. Adapted from Gandini et al* [227]

The vertical line in the center indicates the point of no effect while the effect of each study is presented by a square. The null value in the case of relative measures of association such as Relative Risk (RR) or Odds Ratio (OR) or Hazard Ratio (HR), is 1 meaning that there is no difference in the proportion of events between the two groups. The size of the square is proportional to each individual study weight and the vertical lines represent the CI of the effect of each individual study. The diamond at the bottom of the forest plot provides the pooled estimate of the RR. The horizontal extent of the diamond represents the CI of the pooled RR.

A slightly different kind of forest plot is presented in figure 3.3. This forest plot resulted from a meta-analysis performed in order to assess whether lithium, a known mood stabilizer medication has a specific preventive effect for suicide and self-harm in people with unipolar and bipolar mood disorders. This forest plot displays information of different comparisons,

namely the meta-analysis of RCT for the effect of lithium versus amitriptyline, versus carbamazepine, versus lamotrigine, versus olanzapine and versus placebo [228].

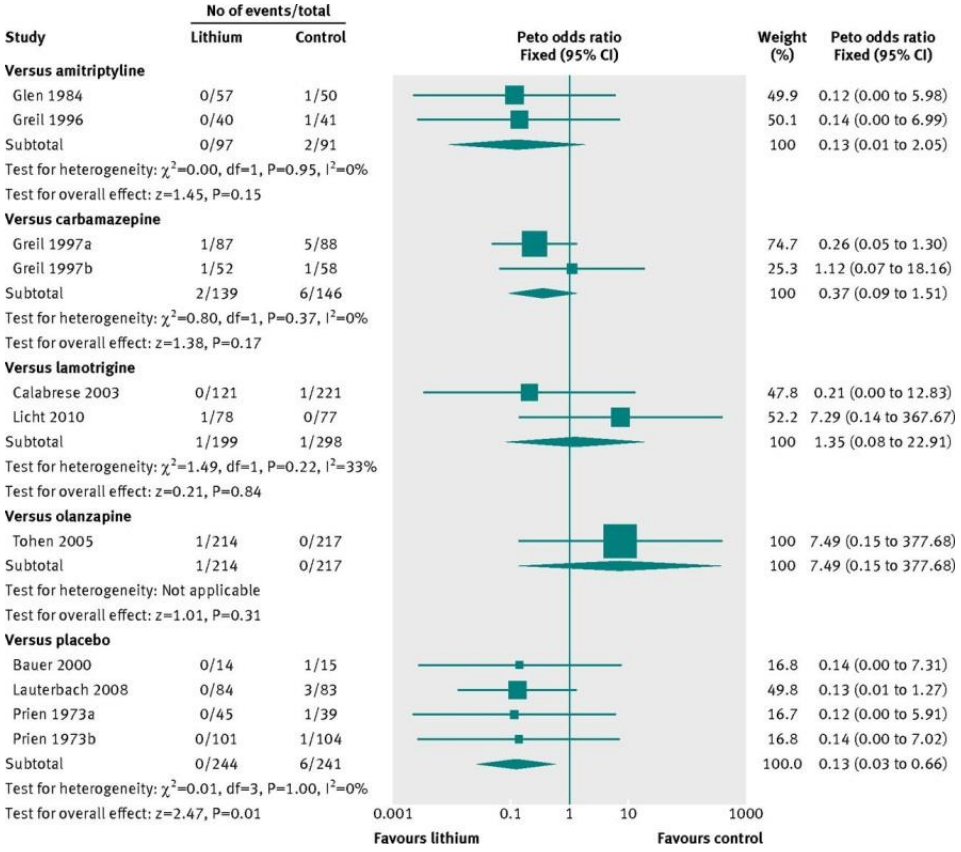


Figure 3.3: *Forest plot showing meta-analysis of suicides in randomized trials comparing lithium with placebo or with active comparators. Adapted from Cipriani A et al 2013* [228]

This forest plot shares many similarities with the forest plot of Figure 3.2 such as the vertical line, alternative size of squares but it differs in that it presents additional information in addition to author name and publication year. It features a table with individual studies outcomes which are either events of suicide or self-harm and it also features information about the heterogeneity between the summarized studies. As a dichotomous outcome "events" are presented numerically as number of events/total participants with a resulting OR.

Figure 3.4 presents a forest plot from a meta-analysis of RCTs of rhythmic cueing to improve walking speed in people with Parkinson's disease [229]. In this case, the outcome of RCTs

was not a relative difference in the proportion of events between the two groups but rather the absolute difference between the two groups in the outcome of interest (walking speed) which is a continuous outcome measure. The vertical line represent the point of no effect with the null hypothesis tested is that there is no difference in absolute value of the outcome between the group receiving the intervention and the control group (null=0).

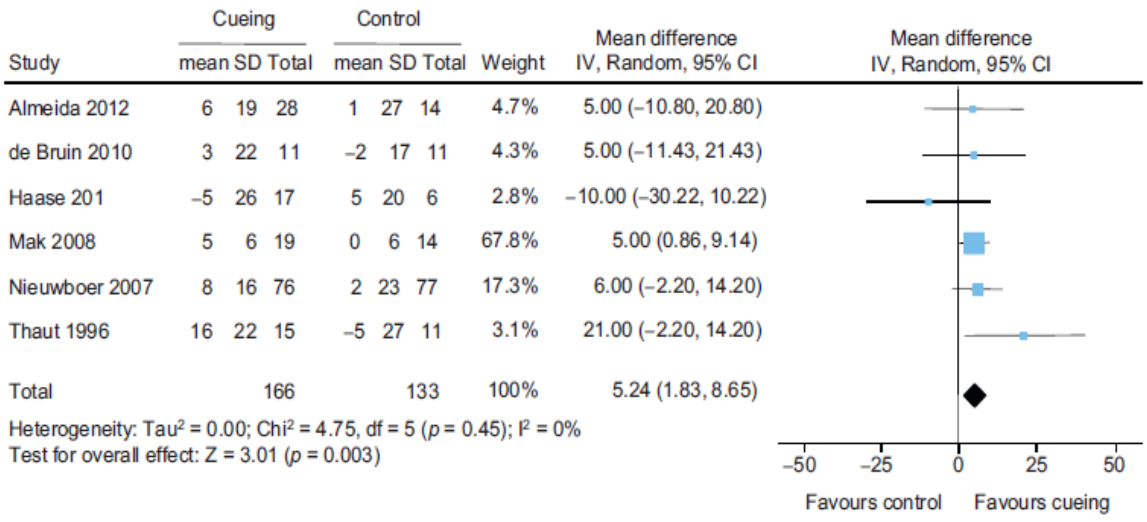


Figure 3.4: *Forest plot of a meta-analysis of trials of the effect of cueing versus no cueing on gait speed (cm/s) in people with Parkinson's disease. A negative mean gait speed means a slower overall gait speed. Adapted from Tomlinson et al* [229]

Figure 3.5 presents a forest plot from a meta-analysis summarising studies assessing the prevalence of depression in chronic obstructive pulmonary disease (COPD) [230]. In this example, the individual estimates of prevalence are only accompanied by the study reference and no numerical information is displayed about the number of COPD patients with depression in each study as well as no numerical information of sample size of each study. In forest plots of prevalence meta-analyses, estimates are plotted over the x axis which extends from zero (0) to one (1) with 0 demonstrating no member of the study population having the condition and 1 demonstrating that all members of the study population having the condition [231]. A vertical line is not used in this kind of forest plot as there is no point that can be considered of "no effect".

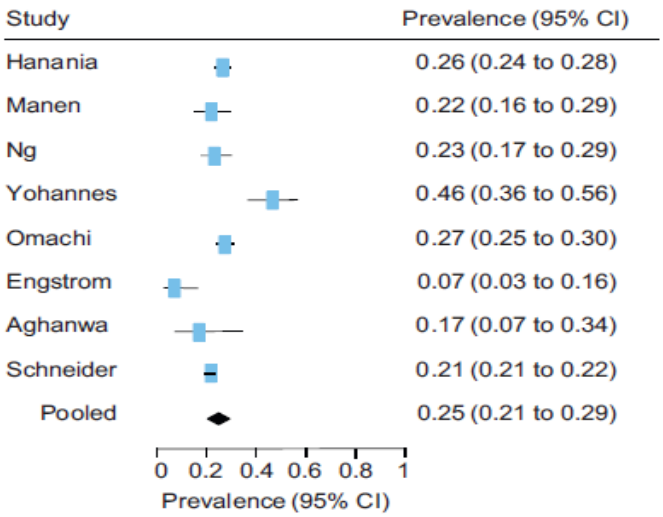


Figure 3.5: *Forest plot of eight studies of the prevalence of depression among people with chronic obstructive pulmonary disease. Adapted from Zwang MWB et al* [230]

Limitations of Meta-analysis

Although meta-analysis remains a valuable tool in the hands of clinicians, researchers and regulators, it is not without flaws. As a separate study design, meta-analyses may also be characterized by several disadvantages, most of which originate from the methodological quality of the included studies and as a result special care is required prior relying on the generated output of the meta-analysis.

The high risk of introducing publication bias in the meta-analysis output is one of the primary disadvantages of this study design. Publication bias refers to the phenomenon that negative or statistically non-significant results may remain unpublished and since a meta-analysis study is usually restricted to the published literature, inevitably, the validity of the results may be compromised [232]. However, several statistical methods have been developed to detect and adjust for the presence of publication bias in a meta-analysis [233]. Furthermore the conclusions generated by a meta-analysis are heavily influenced by the quality of the studies selected for the estimation of the pooled effect [234]. For this reason, the quality assessment of the included studies during the systematic review process is an important step that should

precede the statistical analysis [235]. Another source of bias in the meta-analysis study design is the "small study effect" which refers to the phenomenon that included studies with small sample size report systematically different effect estimates from included studies with larger sample size [236]. It has been reported that small sample studies tend to usually demonstrate larger effect estimates than larger studies and this has been attributed to the fact that smaller studies are more easily affected by publication bias or by lower methodological quality [237]. It is of note however, that heterogeneity in the effect estimates of small and large studies may refer to actual clinical heterogeneity originating from the more detailed selection process and experimental design that smaller studies may employ in contrast to the less rigorous process that may characterize larger studies [235]. Towards identifying the "small study effect", the effect estimates between fixed and random effects can be compared and in the absence of substantial differences, the "small study effect" can be considered minimal [235].

Statistical methods in Meta-analysis

In general, all statistical methods commonly used in meta-analysis first calculate the observed intervention effect for each included study followed by the calculation of a weighted average of these effects. A weighted average is calculated as follows:

$$Weighted\ Average = \frac{Sum\ (Effect * Weight)}{Sum\ of\ Weights}$$

The Effect could be the mean differences for continuous outcomes, HR for survival data or RR and OR for dichotomous outcomes. The weights are selected to reflect the amount of information contained in each study and the bigger the weight of a specific study, the more it will influence the resulting weighted average [202, 238].

Nevertheless, a number of different methods have been developed to analyze data in meta-analyses depending on whether dichotomous or continuous data are used and depending on

how weights are assigned to each study. The broad classification of meta-analysis methods is whether they rely on fixed or random effects models. Fixed effects methods assume that the true effect size that individual studies have investigated is the same across all studies. Under this assumption, it is expected that the only source of error in the estimate of the pooled effect is chance, which is reflected in the random (standard) error in the estimates of the individual studies (within study error). Consequently, in studies with large sample size, standard error becomes smaller as well as the error around the pooled estimate in meta-analyses of such studies [239]. The most common and simpler fixed effect method to pool both dichotomous and continuous data is the inverse variance method. The pooled inverse variance effect size (denoted θ_{IV}) is calculated by the general formula:

$$\theta_{IV} = \frac{\sum w_i \theta_i}{\sum w_i}$$

Where θ_i is the effect estimate from the *i*th individual study and w*i* is the weight of the *i*th individual study. The weight of each individual study is calculated as the reciprocal of the squared standard errors (SE) [240]:

$$wi = \frac{1}{SE\ (\theta_i)^2}$$

With the inverse variance method, studies with smaller SE are assigned greater weight compared to studies with greater SE. Another group of fixed effects methods that are commonly used to calculate pooled estimates in meta-analyses of dichotomous data are the Mantel-Haenszel methods [241]. Mantel Haenszel methods use different approaches to assign weights to individual studies depending on the type of dichotomous outcome that is evaluated (e.g. risk ratio, odds ratio, risk difference) and are more robust compared to the inverse variance method when data (such as event rates) are small [240].

On the contrary to fixed effects methods, random effects methods assume that there is no one true effect size but rather a distribution of true effect sizes and therefore the pooled meta-analysis estimates not a common true effect but rather the mean of the distribution of pooled effects. The error in the pooled estimate incorporates the between studies' variance (between study error) and the within study error. The DerSimonian and Laird method is the most commonly used random effects method and is also based in the inverse variance method [242] and always results in more conservative estimates (i.e. wider confidence intervals) compared with the fixed effects methods [240] (except when between study variance is equal to zero). The difference between fixed and random effects is presented graphically in Figures 3.6 and 3.7.

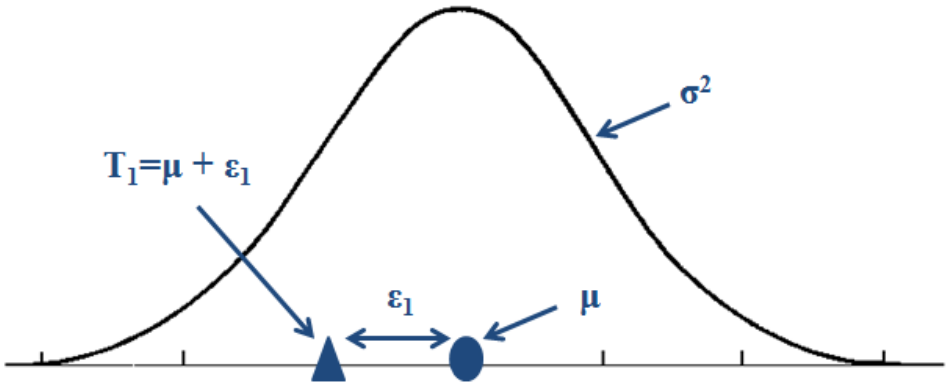


Figure 3.6: Fixed effects schematic. T1 represents the individual study effect which is determined by the true effect (μ) and the within study error (ε_1). Adapted from Borenstein M et al [239].

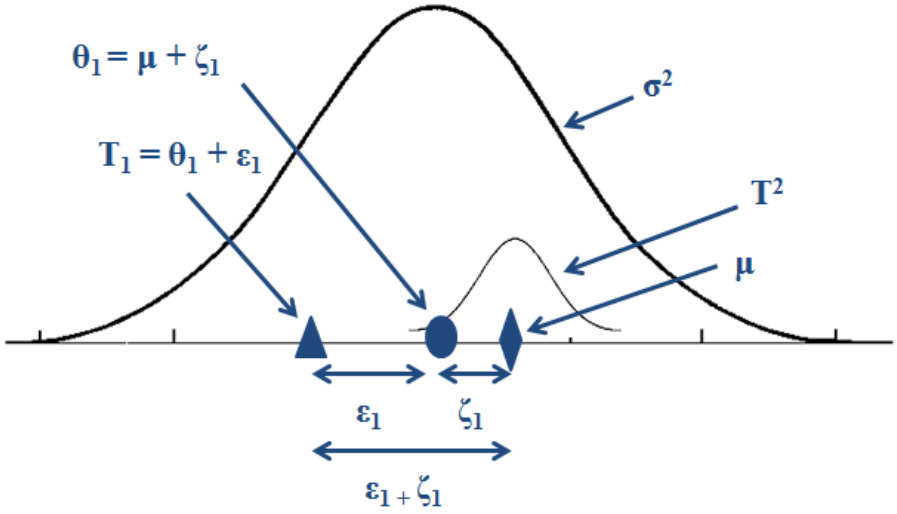


Figure 3.7: Random effects schematic. T1 represents the observed effect from the individual study which is determined by the true effect and (θ_1) and the within study error (ε_1). The true effect (θ_1) is determined by the mean of all true effects (μ) and the between study error (ζ_1). Adapted from Borenstein M et al [239].

Statistical methods in assessing Heterogeneity

Effect estimates are expected to vary between synthesized individual studies to, either due to random sampling error or due to between study differences such as variations in study design, statistical analysis, differences in study population and more. Any types of differences between the different included studies in a meta-analysis are described with the term "Heterogeneity" and could be clinical, methodological or statistical in origin [243]. Clinical heterogeneity refers to differences in the characteristics of the studied population, the evaluated interventions or the outcomes reported while methodological heterogeneity refers to differences in study design and risk of bias. Both methodological and clinical heterogeneity, give rise to statistical heterogeneity which refers to large differences in the effect estimate of individual studies that cannot be explained by chance alone [202].

The heterogeneity in meta-analysis studies is usually assessed with the Cohran Q test or the I^2 statistic. The Q test, has been originally developed by Cohran in 1954 [244], has been used

extensively to assess the presence or absence of statistically significant heterogeneity in meta-analysis studies is calculated by the following formula:

$$Q = \sum_{i=1}^{k} w_i (x_i - \overline{x}_w)^2$$

Where k is the number of included studies, w_i is the weight of each study calculated as the reciprocal of the squared standard error, x_i is the effect estimate of study i and $\overline{x}_w$ is the weighted average of the effect estimate. The Q test is assumed to follow a χ^2 distribution with k-1 degrees of freedom and rejection of null hypothesis is interpreted as presence of heterogeneity [245]. On the contrary, the I^2 statistic measures the extent of heterogeneity by estimating the proportion of total variation across the included studies that is not attributable to chance. It is based on the Q statistic and is calculated by the formula:

$$I^2 = 100\% \frac{Q - df}{Q}$$

Where Q is the Cohran Q test and df are the degrees of freedom [246]. Using the estimate of I^2, a rough categorization of heterogeneity is possible with I^2 upper limits of 25%, 50%, 75% and 100% reflecting low, moderate, substantial and considerable heterogeneity respectively [247]. In the presence of substantial heterogeneity, random effects models should be used to synthesize results as this method will account for the heterogeneity and result in wider and more conservative confidence intervals [248]. High heterogeneity may also lead the reviewer to consider a number of options prior interpreting the meta-analysis results. It is possible that meta-analysis should not have been performed along with the systematic review as the individual studies may be so heterogeneous that meta-analysis may not be a suitable option to address the scientific question in hand especially in the presence of conflicting results. Another option for the reviewer is to explore this heterogeneity and understand its sources through the performance of subgroup analyses. Subgroup analyses usually regard to separate

analyses for a subset of studies and can be useful in understanding the source of the heterogeneity in the original estimate, in answering specific questions about particular subgroups (i.e. particular patient groups, particular intervention settings) [202].

Statistical methods in meta-analysis of proportions

For the meta-analysis of studies with data that follow a binomial distribution (YES/NO responses, success/failure) a number of different statistical methods have been proposed. The most common method to model binomial data such as probabilities (proportions), is to use the normal approximation to the binomial distribution which applies when n is large enough and probability is not close to either margins (either 0 or 1). The *metan* command in STATA software allows for the implementation of many meta-analytic procedures and includes procedures for meta-analysis of binomial data using either fixed or random effects models. However it is designed and mostly used to pool estimates from studies that compare a dichotomous outcome such as the OR, the RR or the difference of two proportions (risk difference, RD) across two groups using the formulas presented here [249].

$$OR_i = \frac{\frac{p_{Ti}}{1-p_{Ti}}}{\frac{p_{Ci}}{1-p_{Ci}}}$$

$$RR_i = \frac{p_{Ti}}{p_{Ci}}$$

$$RD_i = p_{Ti} - p_{Ci}$$

Where p_{Ti} is the proportion of "success" in the treatment group and p_{Ci} is the proportion of success in the control group of study i. In the case of pooling proportion estimates (such as

prevalence estimates for a specific disease) across one group without a comparator group the following formula provides the pooled estimate [250].

$$p_i = \frac{r_i}{n_i}$$

Where r_i is the number of successes and n_i is the total number of observations.

In this case, the *metan* command is not applicable as it is limited by the fact that it does not have the ability to pool studies at the two extremes, in other words with proportions at either 0% (e.g. no-one has the disease) or 100% (everyone has the disease). Consequently, the *metan* command excludes these studies and results in a biased pooled estimate [250]. A recently developed *metaprop* command builds upon the *metan* command and provides the framework to pool proportions along with 95% confidence interval for the pooled estimate. Meta-analysis of proportions performed with *metaprop* allows the calculation of a pooled estimate (weighted average) of the prevalence of a disease across studies containing binomial data and estimates 95% confidence intervals. The weighted summary proportion (in this case prevalence) is calculated as a fraction with the numerator defined as the number of patients identified with the condition and the denominator as the total number of patients with or without the disease. A continuity correction allows for studies with a proportion at the extremes (0% or a 100%) not to be excluded thus providing a more robust estimate of the pooled proportion. Two types of confidence intervals (exact and transformed) for the weighted summary proportion can be calculated. Exact confidence intervals provide more conservative estimates [250] but result in confidence intervals that are problematic when p_i is close to the extremes [251]. The lowest exact confidence interval (CI) is calculated as the $\frac{a}{2}$ quintile of Beta distribution (x_i, n_i - x_i + 1) and the upper exact CI as the 1- $\frac{a}{2}$ quantile of the Beta distribution (x_i + 1, n_i - x_i) [252]. To make the distribution of data even more similar to

the Normal distribution the Freeman-Tukey double arcsine transformation is frequently used using the formula below [253] and calculation of the weighted summary proportion under and the calculation of confidence intervals under the fixed and random effects model [242].

$$p = sin^{-1} * \sqrt{\frac{r_i}{n_i + 1}} + sin^{-1} \sqrt{\frac{r_i + 1}{n_i + 1}}$$

Where r_i is the number of successes and n_i is the total number of observations. To transform the values back to proportions the inverse of Freeman-Tukey double arcsine transformation is used [254].

$$p = \frac{1}{2} * \left[1 - \sin(cos_t) * \sqrt{\left[1 - \left(sin_t + \frac{sin_t - \frac{1}{sin_t}}{n} \right)^2 \right]} \right]$$

Where *t* refers to the transformed value and *n* refers to sample size.

Systematic review and meta-analysis of diagnostic accuracy studies

Diagnostic accuracy studies are performed in order to understand how good a diagnostic test is, in distinguishing subjects with a specific disease and subjects without it. This type of studies is usually performed when a new diagnostic test has been developed and the accuracy of the test has not yet been examined in the clinical setting. The main motivation behind the development of new diagnostic tests in the era of constrained healthcare resources is the

requirement for cheaper, better and faster to perform tests [255]. In addition, it some cases current gold standard tests may be not feasible or unethical to perform such as biopsies in some cases of brain tumors and in the case of Alzheimer disease [256] and thus new tests in the form of better imaging (e.g. Magnetic Resonance Imaging, High Resolution Computed Tomography or Positron Emission Tomography) or novel biomarkers are sought.

Diagnostic accuracy studies primarily focus on the two statistical measures of diagnostic accuracy: (a) the sensitivity of the test and (b) the specificity of the test. Sensitivity refers to the proportion of subjects with the disease that have an abnormal (positive) test result while specificity refers to the proportion of subjects without the disease that have a normal (negative) test result [257]. Another set of test characteristics that are also usually reported in diagnostic accuracy studies is Positive Predictive Value (PPV) and Negative Predictive Value (NPV). PPV refers to the likelihood that a subject has the disease given that the test result was positive and NPV refers to the likelihood that a subject does not have the disease given a negative test. For each test, a 2x2 table can be constructed to demonstrate the cross tabulation of diagnostic test result and disease status. Table 3.3 demonstrates one such cross-tabulation for a specific index test versus disease status as obtained by a "gold standard" test. Gold standard test is the term used for a test that serves as the unqualifiedly the most accurate diagnostic procedure which reveals the absolute truth about disease status although in reality it such level of accuracy is difficult to achieve with for any biological test. Thus, a gold standard test just represents the best currently available tool to classify disease status [258, 259].

Table 3.3: Cross-tabulation of test result and disease status

Index Test result	**Diseased (D+)**	**Non-diseased (D-)**	**Total**
Test Positive (T+)	True Positives (TP)	False Positives (FP)	Test Positives (TP+FP)
Test Negative (T-)	False Negatives (FN)	True Negatives (TN)	Test Negatives (FN+TN)
Total	Disease Positives (TP+FN)	Disease Negatives (FP+TN)	N (TP+FN+FP+TN)

For the calculation of Sensitivity, Specificity, PPV and NPV, the number of True Positives (TP), False Positives (FP), True Negatives (TN) and False Negatives (FN) needs to be estimated. TP refer to patients with the disease and an abnormal (positive) test result, FP refer to patients without the disease and an abnormal (positive) test result, TN refer to patients without the disease and a normal (negative) test result and FN refer to patients with the disease and a normal (negative) test result. The following formulas demonstrate how TP, FP, TN, FN give rise to Sensitivity, Specificity, PPV and NPV [260].

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$PPV = \frac{TP}{TP + FP}$$

$$NPV = \frac{TN}{TN + FN}$$

The measures of sensitivity and specificity can be combined in the terms positive Likelihood Ratio (LR) and negative LR, which provide another estimate of the test accuracy Positive LR describes how much more likely it is for a patient who tests positive to have the disease

compared with a patient who tests negative. Negative LR describes how much less likely it is for a patient who tests negative to have the disease compared with a patient who tests positive. Positive LR is a number >1 and negative LRs is a number <1 and are calculated as follows [260]:

$$Positive\ LR = \frac{Sensitivity}{1 - Specificity}$$

$$Negative\ LR = \frac{1 - Sensitivity}{Specificity}$$

Nevertheless all measures of the above mentioned diagnostic accuracy measures correspond to a specific but many times arbitrarily selected threshold (cut-off point) for each diagnostic test. The threshold refers to the cut-off point above or below which the test is considered abnormal (positive) and its selection affects the numbers of TP, FP, TN and FN and consequently the sensitivity and specificity of the test. For example, as in the case of a diagnostic test results where results below a specific threshold are abnormal, increasing and lowering the threshold would differentially affect sensitivity and specificity. When the threshold is increased, fewer FP and more FN are expected and thus the test is considered highly specific but less sensitive. On the contrary, when the threshold is lowered fewer FP and more FN are expected and thus the test becomes highly sensitive but less specific [261]. Sometimes there is no biologically relevant threshold to be considered as the primary diagnostic cut-off and many diagnostic accuracy studies use different values to define normal and abnormal test results and present their results in the form of Receiver Operator Characteristic (ROC) curves. ROC curves are plots of sensitivity on the y-axis and 1-specificity on the x-axis for all possible thresholds. Figure 3.8 demonstrates an array of ROC curves. The use of ROC curves to compare diagnostic tests allows comparisons taking into account the test accuracy across a range of thresholds. The area under each ROC curve

(AUC) is a combined measure of sensitivity and specificity and represents the overall accuracy of each diagnostic test, and values close to an AUC equal to 1.0 (closer to the upper left-hand corner in an ROC graph) indicate high sensitivity and specificity. The point across an ROC curve that is closer to the upper left-hand corner represents the best combination of Sensitivity and Specificity [262]

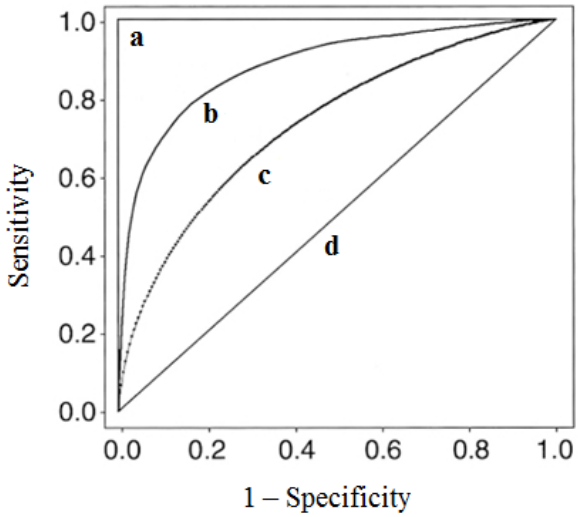


Figure 3.8: *ROC curves from four different tests. A perfect test (a) has an AUC equal to 1. The diagonal ROC (d) has an AUC equal to 0.5 and corresponds to a test that has no discriminant ability. The remaining ROC curves (b) and (c) refer to tests with some discriminant ability (0.5<AUC<1). ROC curve (b) has a higher AUC compared to (c) demonstrating that the corresponding test b has an overall better diagnostic performance compared to test (c). Adopted from Park SH et al* [263].

Systematic review and meta-analysis of diagnostic accuracy studies primarily aim to describe how well the test classifies subjects (TP, FP, TN, FN), provide summary estimates of sensitivity and specificity based on the results of the individual included studies, describe the uncertainty of around these estimates and describe how test accuracy varies depending on the test threshold. Similarly with meta-analysis of intervention effects, the increased sample size resulting from pooling individual studies together allows results in increased statistical power and for more precise estimates of sensitivity and specificity [264]. However, systematic reviews and meta-analyses of studies of diagnostic accuracy are different compared to other kinds of systematic reviews and meta-analyses both in the method of addressing study quality

as well as in the statistical approaches required to combine results from individual studies [238].

Quality assessment in systematic reviews of diagnostic accuracy studies

Diagnostic accuracy studies frequently suffer from methodological weaknesses. Such weaknesses may include the use of an inappropriate "gold standard" test, selection bias (diagnostic test evaluated in a subset of subjects that may not be characteristic of the population the test is supposed to apply), lack of diagnostic test user blinding (the user is familiar with the disease status of the subjects prior the performance of the diagnostic test), and insufficient definition of what is a positive or negative test result (cut-off value not defined, or definition may be disputed) [265]. Furthermore, diagnostic accuracy studies may not describe the "gold standard" and index test in detail, may not have performed or may not have reported information for all evaluated patients may have "gold standard" and index test performed after the initiation of treatment and may or may not have reported "inconclusive" findings of the test procedures [34].

To address all the above weaknesses that may lead to different types of biases and allow the reviewer to describe the quality of the included studies in terms of reporting or risk of different biases, the Standards for Reporting of Diagnostic Accuracy (STARD) and the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tools have been developed [218, 223]. Since 2011 the QUADAS-2 quality assessment tool has been developed from the original QUADAS and is currently the most widely applied quality assessment tool in systematic reviews of diagnostic accuracy[218]. It is made up of four key domains that cover patient selection (domain 1), index test (domain 2), "gold standard" test (domain 3) and flow of patients through the study and timing of the "gold standard" and index test (domain 4).

Each domain, with the help of signaling questions, is rated in terms of risk of bias and regarding the applicability to the research question at hand. Table 3.4 summarizes the components of QUADAS-2 assessment tool [219].

Table 3.4: Components of QUADAS-2 assessment tool

Domain	Category	Signaling Questions	Quality Assessment
Patient Selection	Risk of Bias	Was a consecutive/random sample of patients enrolled? Was a case-control design avoided? Did the study avoid inappropriate exclusions?	Could the selection of patients have introduced bias? RISK: Low/High/Unclear
	Applicability	Prior testing, Presentation, intended use of index test and setting	Is there concern that the included patients do not match the review question? CONCERN: Low/High/Unclear
Index Test(s)	Risk of Bias	Were the index test results interpreted without knowledge of the results of the reference standard? If a threshold was used, was it pre-specified?	Could the conduct or interpretation of the index test have introduced bias? RISK: Low/High/Unclear
	Applicability	n/a	Is there concern that the index test, its conduct, or interpretation differ from the review question? CONCERN: Low/High/Unclear
Reference Standard ("Gold Standard")	Risk of Bias	Is "gold standard" likely to correctly classify the target condition? Were the reference standard results interpreted without knowledge of the results of the index test?	Could the reference standard, its conduct, or its interpretation have introduced bias? RISK: Low/High/Unclear
	Applicability	n/a	Is there concern that the target condition as defined by the reference standard does not match the review question? CONCERN: Low/High/Unclear
Flow and Timing	Risk of Bias	Was there an appropriate interval between index test(s) and reference standard? Did all patients receive a reference standard? Did all patients receive the same reference standard? Were all patients included in the analysis?	Could the patient flow have introduced bias? RISK: Low/High/Unclear

Source: [219]

Statistical Methods in meta-analysis of Diagnostic accuracy studies

Diagnostic accuracy meta-analyses primarily aim to calculate pooled estimates of sensitivity and sensitivity, taking into account the variability across studies. Sensitivity and specificity can be calculated for a specific common threshold (referred to as the average operating point) or an ROC curve can be computed across m thresholds. However, in the case that included studies are characterized by mix thresholds, estimating only a summary sensitivity and specificity threshold will relate to an unspecified average of the thresholds used in the included studies which is unhelpful and must be avoided [264].

A number of statistical methods have been suggested perform diagnostic accuracy meta-analysis. Among them are simple pooling (fixed-effects) of sensitivity and specificity, random-effects meta-analysis separately for Sensitivity and separately for Specificity, separate meta-analysis of positive and negative likelihood ratios, the Littenberg-Moses summary ROC curve, bivariate random effects meta-analysis and calculation of Hierarchical Summary ROC curves (HSROC) [266]. Simple pooling requires the reviewer to construct a 2x2 table as the one presented in Table 3.3 by summing up the numbers of TP,FP,TN,FN from the individual studies and subsequently calculate Sensitivity and Specificity as though all the data originated from an individual study. This method basically ignores any between study heterogeneity but also ignores any correlation between the two measures of diagnostic accuracy. Sensitivity is inversely correlated with Specificity due to the trade-off between these measures as the test threshold varies. More importantly, when Sensitivity and Specificity are reported at different thresholds in the included studies, ignoring this correlation may bias the pooled estimates significantly. The same limitation applies for separately pooling Sensitivity and Specificity using a random effects model which although accounts for the between study heterogeneity, it does not account for the correlation between the two measures [238]. The Moses and Littenberg summary ROC curve (Moses-Littenberg

SROC) is based on logit trasformations of TP and FP [267], makes use of simple linear regression to generate a SROC curve. This method assumes that observed differences across studies result from the different thresholds used but it is considered an approximation as the assumptions of simple linear regression are not met [268, 269].

Bivariate random effects meta-analysis for diagnostic accuracy studies, provides average sensitivity and specificity estimates across studies and has been first described by Reitsma et al in 2005 [270]. It allows for pairs of sensitivity and specificity from each included study to be jointly analyzed using a random effects approach and as a result it incorporates both the between study heterogeneity and the correlation between the two measures. Furthermore, it provides a confidence and prediction region for Sensitivity and Specificity. The confidence region provides a measure of the uncertainty in the pooled estimates and the prediction region informs about the region within which it is expected for the Sensitivity and Specificity of a future study to lie [271]. The Bivariate model involves statistical considerations in two levels. At the first level, the cell counts obtained from the 2x2 tables of each included study are extracted using binomial distributions and logistic transformations of proportions while in the second level, the random effects model accounts for the heterogeneity between the studies beyond the sampling variability accounted at the first level [264]. The Bivariate model is specified as follows [272]:

$$\begin{pmatrix} \mu A_i \\ \mu B_i \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_A \\ \mu_B \end{pmatrix}, \Sigma_{\alpha\beta} \right)$$

$$\Sigma_{\alpha\beta} = \begin{pmatrix} {\sigma^2}_A & \sigma_{AB} \\ \sigma_{AB} & {\sigma^2}_B \end{pmatrix}$$

Where μA_i refers to log-transformed sensitivity and μB_i refers to log-transformed specificity for an included study i. Variables ${\sigma^2}_A$ and ${\sigma^2}_B$ refer to the between study variability in log-

transformed sensitivity and specificity while σ_{AB} refers to the covariance between log-transformed sensitivity and specificity.

The fitting of Hierarchical Summary ROC curves (HSROC) as a mean to summarize results from individual studies of diagnostic accuracy has been first described by Rutner and Gatsonis in 2001 [268]. This method extends the Moses-Littenberg SROC but is more flexible and performs better in incorporating both within and between study variability. The resulting HSROC describes the relationship between sensitivity and specificity derived from the individual ROC of each included study. The HSROC model is specified as a two-level mixed logistic regression model conditional on the sensitivity and the specificity of each study (within-study model) and a bivariate normal model for the sensitivity and specificity between studies (between study model). The within-study model is specified as follows.

$$logit\left(\pi_{ij}\right) = \left(\theta_i + a_i X_{ij}\right)\exp(-\beta X_{ij})$$

Where π_{ij} is the probability that a patient with disease status j in study $_i$ will test positive. Disease status is categorised as diseased (j=1) or non-diseased (j=0) thus for study I π_{i0} equals the FP rate and π_{i1} is the TP rate and accounts for within study variability at first level. Variables θ_i and a_i are the cut-off points and accuracy parameters that are allowed to vary between studies (accounting for between study variability). Variable X_{ij} serves as the true disease status of a patient in study i with disease status j and variable β is a scale parameter that allows the modelling of the asymmetry in the ROC curve. Similarly to the bivariate model, HSROC analysis can be used to derive summary estimates of Sensitivity and Specificity as well as confidence and prediction regions along with summary ROC curves [272].

Given that both of bivariate model and HSROC analysis are statistically rigorous procedures, specialised statistical software is required to run them. The *metandi* command (along with the *metandiplot* command that produces accompanying graphs), runs the hierarchical logistic

models and presents the results in both bivariate and HSROC parameterization and has been developed for STATA software by Harbord and Whiting in 2009 [273] and is now used widely in meta-analysis of diagnostic accuracy meta-analyses. The output of the two commands includes summary estimates and confidence intervals for Sensitivity and Specificity as well as estimates of LR+ and LR- at the summary point. The *metandi* command also provides the diagnostic Odd Ratio (DOR) which is ratio of the odds of a positive test given that the patient has a disease relative to the odds of a positive test given that the patient does not have the disease and is considered a single measure of the effectiveness of the test. The *metandiplot* command produces the graph of the model fit by *metandi*. Figure 3.9 demonstrates the output of *metandiplot* command.

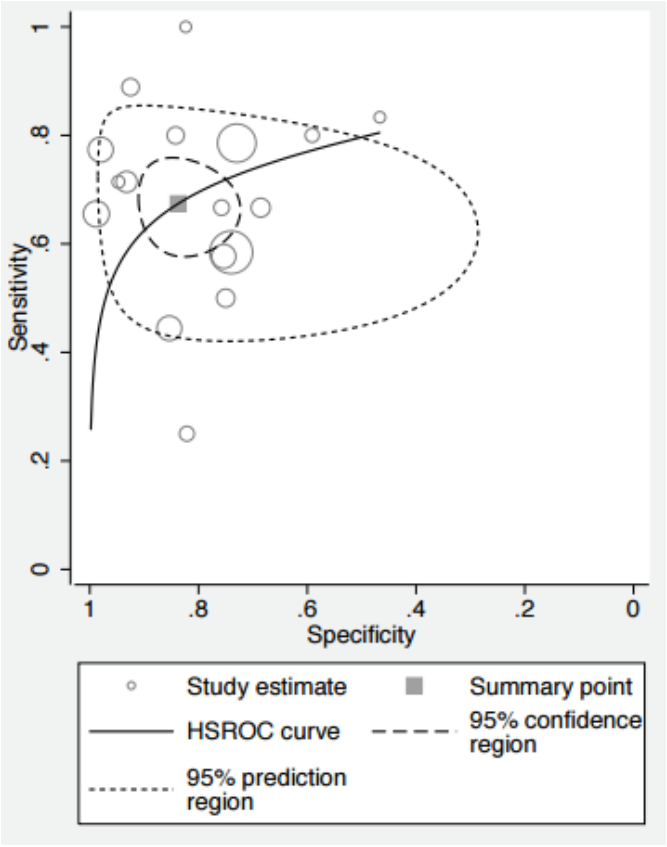


Figure 3.9: *Command metandiplot output in STATA. Circles correspond to Sensitivity and (1-Specificity) results from the individual studies. The size of the circle corresponds to the weight that each study is assigned in the model. Output includes the summary estimate of Sensitivity and Specificity along with 95% Confidence and 95% prediction regions and the HSROC. Adapted from Harbord and Whiting* [273].

Chapter 4: Methodology - Cost-Effectiveness Analysis in Healthcare

Decision making in health and medicine

In the past, decision making in medicine was largely driven by the experience (and sometimes the paternalism) of individual medical practitioners who were aware of a narrow range of possible diseases, were equipped with a limited number of diagnostic tests and had access to few possible treatments to choose from. However, during the last decades, following the advancements of medical knowledge and technology, medical practitioners have access to an ever expanding spectrum of diagnostic tests and treatments for all kind of diseases, complex or monogenic, common or rare. This combination of wide range of diseases and availability of different but frequently not perfect tools to diagnose and treat them results in more complex and difficult decisions for the medical practitioner, the hospital manager or the healthcare policy decision maker to make.

Decision making in healthcare is difficult not only due to the wide range of the possible options that are now available but because of the important consequences that these decisions have and the significant uncertainties and trade-offs they involve. Uncertainties may arise from limitations in the accuracy of diagnostic tests, from the ambiguity in data collection or from the uncertainty in the effectiveness of different treatments. The treatment effectiveness can be described as the number of patients achieving a clinical improvement target (such as 10 mmHg fall in blood pressure), the number of patients experiencing exacerbations of their disease, the number of deaths averted or the number life years saved. Trade-offs in medical decisions may refer to the contrast between the health benefits and side-effects of a specific treatments (e.g. chemotherapy) or may refer to the trade-offs between Sensitivity and Specificity of a test that need to be considered during the introduction of new diagnostic guidelines for a specific disease. Furthermore, the trade-off may refer to the choice between

treatments that may have different health benefits but also may be accompanied by significantly different economic cost for the patient or the healthcare system [255]. The physician, in order to make the best possible decision needs to have access to the best evidence available that inform about the many parameters of the problem in question and illuminate all relevant uncertainties and trade-offs. The best sources of evidence are systematic reviews of individual studies that addressed a specific issue or parts of it. If accompanied by a meta-analysis, a quantitative summary along with an estimate of its uncertainty will be available and can be used in informed decision making. In the absence of meta-analytic estimates, single studies of good quality are appropriate sources of evidence and finally, in the case of absence of good single studies, subjective estimates about specific parameters of the problem based on personal experiences can be used [255]. However, even if evidence of good quality is available, a logical and structured method to combine these data and reach to conclusions is still required. Formal decision analysis, uses this method as a quantitative technique to systematically integrate all available evidence that relate to particular decision and aims to facilitate decision making by identifying the course of action that is expected to maximize the desired outcome [274]. Decision analysis can be comprised by the following four distinct steps [260]:

a. Development of a decision tree (includes formulating the decision problem, assigning probabilities and measuring outcomes).
b. Calculation of the expected value of each decision alternative.
c. Select the decision alternative with the highest expected value.
d. Perform sensitivity analysis to test the decision analysis conclusions.

In more detail, a decision tree is a graphic representation of the decision problem which displays all outcomes of each decision. It consists of a set of decision (box symbol) and chance nodes (circle symbol) along with connecting branches. Decision nodes represent the

points in the tree at which several alternative decisions can be made and chance nodes represent points in the tree at which chance (probabilities) determine which outcome will occur. Some examples of different decision alternatives could be whether to start a new screening program for breast cancer or not, whether to fund an alternative clinical intervention for cardiovascular disease management at the expense of another intervention or whether purchase a new diagnostic modality to replace an already acquired equipment. Terminal nodes are represented by triangles and signify the end of a specific decision branch where the expected value of the decision outcome is calculated. Different examples of outcomes could be the number of cases detected, the number of lives saved, the number of life-years saved or the number of Quality Adjusted Life Years (QALYs) saved. Decision trees are usually written from left to right starting from the initial decision node at the far left and terminal nodes at the far right. An illustrative example of a decision tree is presented in Figure 4.1.

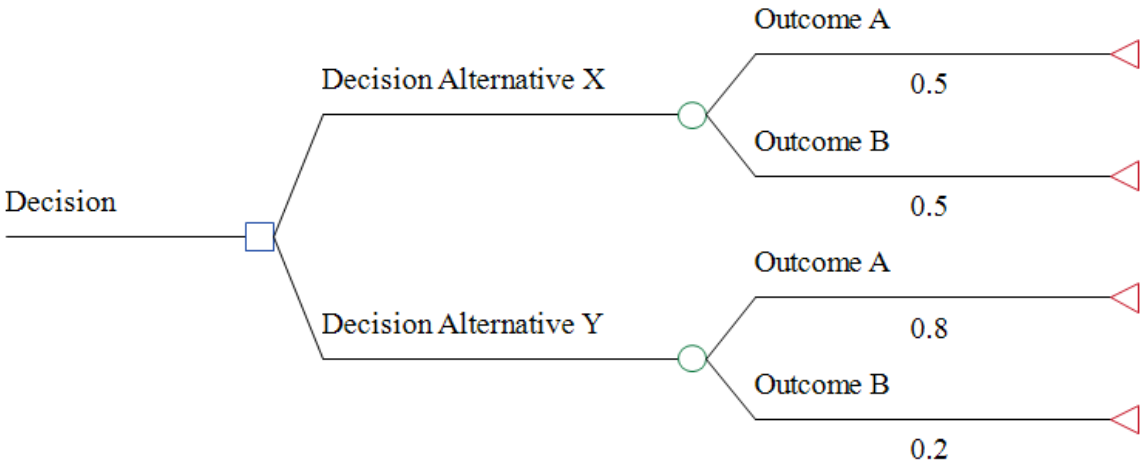


Figure 4.1: *An example of a decision tree (developed by TreeAge PRO software). Numbers below outcomes represent the probability that the outcome will occur.*

The basic decision tree framework described here can describe well the sequence of events and outcomes but only if there are no recursions. In case of recurring events (including recurring decisions and outcomes) such as in the case of cancer patients which may face

cancer recurrence after treatment, decision trees become increasingly complex and hard to calculate [255]. Relatively simple methods such as Markov models [275] as well as more advanced methods (discrete event simulation, dynamic transmission model) have been developed to account for decision problems with recurring events [276, 277].

The term expected value is defined as the sum of the values of all the outcomes resulting from a specific decision alternative with each value weighted by the probability that the consequence will occur. It is a term that closely relates to the term weighted average and is calculated by the following formula:

$$E(X) = P(B) * U(A) + P(B) * U(B)$$

Where E(X) is the expected value of decision alternative X with possible outcomes A and B. P(A) and P(B) refer to the probability that outcome A or B will occur respectively and U(A) and U(B) refer to the value of the outcome occurring. The decision alternative with the highest expected value represents the best choice [260]. Finally sensitivity analysis can be employed if there is considerable uncertainty regarding the assumptions that were made and there are doubts whether the conclusions made by the analysis are valid and generalizable. Towards evaluating whether the same conclusions would apply under different assumptions, sensitivity analysis is performed, that is the analysis is repeated after substituting a range of parameter values and assess whether the final conclusion of the analysis is altered. An unaltered conclusion provides reassurance that the original conclusions are valid while conclusions that are sensitive to small alterations in some of the parameters values may prompt for additional refinements in the original analysis [255].

Economic evaluation in healthcare

Given that healthcare costs take up a significant part of national expenditures and the healthcare industry is growing in a planet with an increasingly aging population, it has been increasingly recognized that the effectiveness of an intervention should not be the only component of health care decision making. As a result, instead of just comparing health benefits and harms of the different interventions and selecting the decision that leads to the greatest health benefit, the economic cost of the intervention has to be taken into account and decision could be based in other criteria apart from the health benefit. The economic cost of the intervention refers to any economic resource (such as medical or non-medical equipment, consumables, human labor and the use of buildings/energy) that is consumed for the implementation of the intervention [255]. A number of different economic evaluation approaches have been developed and are used in health care decision making. The four main approaches are (a) cost minimization analysis (CMA), (b) cost benefit analysis (CBA), (c) cost-effectiveness analysis (CEA) [278].

CMA focuses on comparing the recourse cost of alternative interventions and usually applies when the two (or more) alternative interventions are assumed to have the same effectiveness and safety or tolerability in regards to adverse effects. Although CMA is easier from all other types of economic evaluations as it does not require the quantification of health benefits which frequently is difficult to perform, it is now rarely used mainly due to concerns about the applicability of assuming that alternative interventions result in equivalent health outcomes [279]. CBA does not rely on any prior assumption regarding the health benefit of each alternative intervention and can be used to compare interventions with different effectiveness and safety outcomes. In order to do so, CBA requires for monetary values to be placed on health benefits by taking into account how much society is willing to pay for the particular benefits. Interventions are then compared based on their net monetary benefit

(monetary value of the cost of intervention plus the monetary value of the health effect of the intervention). An intervention with negative net monetary benefit (costs outweigh the benefits) should not be implemented and an intervention with positive net monetary benefit (benefits outweigh costs) should be implemented. Similarly an intervention with a higher positive net benefit should be preferred instead of an intervention with a lower positive net benefit [280]. CBA is a useful tool for decision makers as it is the most comprehensive form of economical evaluation compared to CMA or CEA and by using a single measurement, it provides a clear answer to the question "Is this intervention worth doing?" [280, 281]. In addition, it allows the comparison of health care interventions that may also span out of the healthcare sector into the education, transport or any other non-healthcare economic sector [281]. However, the use of CBA is limited by several ethical and practical considerations. The primary factor that deters decision makers to make use of CBA is the requirement to assign monetary values to health effects such as human lives or the quality of human life. Methods such as the human capital approach as well as the observed and stated preferences approach have been used to facilitate monetary valuation of health outcomes but not without significant criticism. The human capital approach considers human beings as capital equipment (similar to machinery in manufacturing industry) that are expected to produce a flow of productive activity in the future (mirrored in the individual's annual compensation) and the health benefit can be valued in terms of future income that would have been lost if the healthcare intervention was not applied [282]. The criticism about the human capital approach, focuses on the facts that rates of compensation are used as a measure of productivity in the human capital approach while rates of compensation reflect productivity only when certain market conditions apply and that valuing health effects in terms of rate of compensation downplays the value of health effects in individuals that are not employed or are retired [280]. The observe preferences approach relies on observing how much money are

individuals paid to undertake a job that entrails a significant amount of risk (e.g. deep sea divers) or how much money do individuals pay different safety measures (e.g. safety features in cars) and the resulting relationships between amount of money and risk can provide a value to value health benefits of interventions. Finally the stated preferences approach requires individuals to choose among specified choices in monetary terms, a method that is also described with the term "willingness-to-pay approach". This approach associates monetary values to health outcomes by asking individuals how much they would be willing to pay to avoid a negative health outcome or obtain a positive health outcome [280]. The main criticism in the observed preferences approach is that it can be evaluated in only a very limited number of situations and thus it is difficult to generalize the results obtained while for the willingness to pay approach criticism focuses to its reduced sensitivity to the magnitude of the benefit, the differences in the responses when individuals are asked about a health outcome in isolation compared when individuals are asked about multiple health outcomes [283] and to the fact that (absolute) willingness to pay is affected by the individual's ability to pay thus resulting in equity issues [282]. In addition to the above practical reasons, the widespread reluctance of healthcare policy makers to rely on monetary valuation of health outcomes resulted in a limited use of CBA as opposed to CEA analysis which is the most frequently used method of economic evaluation in the field of healthcare [284].

Cost Effectiveness Analysis

In CEA, both costs and health outcomes are assessed separately and interventions are compared based on their differences in net cost and net effectiveness and the result is presented as a ratio (cost per unit of health outcome) [285]. When the intervention is compared against a "do nothing" the ratio is usually called Average Cost Effectiveness Ratio

(ACER) or plainly Cost Effectiveness Ratio. However it is more common to compare a new intervention compared to an intervention that is currently being in used and very frequently more than two interventions are compared. Given that the comparisons are also mutually exclusive the incremental differences are of interest and thus the terminology typically used includes the terms incremental costs, incremental health effects and incremental cost-effectiveness ratio (ICER). The ICER can be calculated as the difference in costs divided by the difference in effects of two interventions (intervention (a) and intervention (b)):

$$ICER = \frac{Cost_a - Cost_b}{Effect_a - Effect_b} = \frac{\Delta Cost}{\Delta Effect}$$

When results of cost effectiveness analysis are presented graphically the series of calculated ICERs that connect the mutually exclusive interventions, the characteristic cost-effectiveness frontier (CEF) is formed. All interventions that are placed along the CEF are cost-effective options at different ceiling ratios. Ceiling ratios are limits on the cost that a decision maker sets per unit of outcome, frequently also called the maximum acceptable value of an ICER. Interventions that lie to the left of the CE frontier are considered dominated. Dominance exists when an intervention is both more costly and less expensive compared to an alternative intervention. It is possible that an intervention will be more costly and less effective than a combination of two other interventions. This case is called extended dominance and practically means that any combination of the other two interventions (that are along the CEF) is always more cost-effective compared to the dominated intervention. Figure 4.2 presents a characteristic scatter plot of costs vs total Life Years saved for four different interventions and the CEF. Based on the amount that the decision maker wants to spend (ceiling ratio), both interventions B and D are cost-effective options and the ICER for intervention D is the slope of the line that connects interventions B and D. Intervention A is dominated and intervention C is characterized by extended dominance.

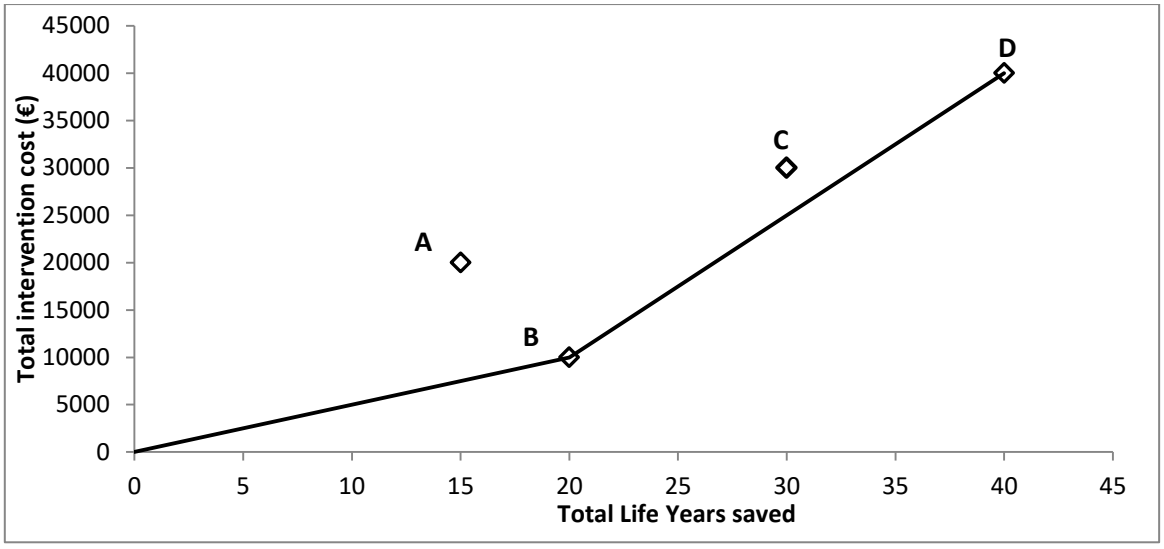


Figure 4.2: *The continuous black line that connects the origin, point B and point D is called the CEF. All interventions that are placed along the CEF are cost-effective at different ceiling ratios. Intervention A is dominated by intervention B and intervention C is dominated by the combination of interventions B and D.*

Estimation of Costs in CEA

The calculation of intervention costs that make up the numerator of the ICER ratio usually includes health related but also non-health related costs of care and is influenced by the decision-making perspective that is adopted by the CEA. The different perspectives that are usually used are the payer (patient) perspective, the provider (hospital or health-system) and the societal perspective and depending on the perspective used, different set of costs are included in the analysis. The payer and the provider perspectives are considered narrower in scope as only costs incurred upon the patient (e.g. per-visit charge or co-payment) or only costs incurred upon the health care provider (e.g. cost of patient visit or cost of surgery) are included respectively. On the contrary the societal perspective in considered wider in scope and includes all health (e.g. cost of surgery) and non-health related costs (e.g. the time of unpaid caregivers) and effects resulting from the intervention and are borne by any individuals in the society [286]. Direct health related direct costs usually include the costs of resources used to implement a health care intervention such as the costs of medical personnel,

the cost of diagnostic tests, the cost of equipment and consumables, the cost of medication or the cost of inpatient admission while direct non-health related costs (sometimes called indirect costs) may refer to the child-care cost borne by a parent undergoing treatment or cost of time a family member spend to care for a diseased relative [287]. Many non-health care costs may also refer to productivity costs (or opportunity costs) that can be translated as costs that occur because time is not used productively such as in the case of the time "lost" by the caregiver to care for a diseased relative. While impaired productivity is monetized in CBA, CEA treats impaired productivity as effect and is measured in terms of improvement in patient's quality of life and life extension [255].

Practically, the cost calculation in CEA is a three step process which includes the identification of the different categories of resources that are required for the intervention, the measurement of the extent of recourse use and the valuation of the resource use through the estimation of the resource unit cost [288]. Which resources will be identified for the cost calculation is influenced by whether a gross-costing or a micro-costing approach is chosen. The gross-costing approach views the different resources as bundles (e.g total surgery cost or total diagnostic test cost) while the micro-costing approach requires the recognition of all the underlying activities that make up a specific intervention or specific procedure. Data availability usually dictates which of the two approaches will be used and it is not unlikely that both approaches are combined in the same CEA [289, 290]. Following the identification of the resource use (e.g hospital admission), the calculation of intervention costs requires the estimation of the unit cost (cost per inpatient day) and the estimation of recourse use (e.g. number of days in hospital). The product of resource use and unit cost provides the total cost estimate for the particular resource [290]. Information about resource use could be routinely collected by the hospital administration, could be extracted from medical notes or could be gathered directly from patients through the use of specific questionnaires such as the SMILE

resource use questionnaire. Resource use questionnaires are made up of a series of questions that may regard to the frequency and duration of hospital and community care visits, medication use but also the loss of earnings and the occurrence of other out of pocket expenses for the individual. [291]. Information about unit cost can be obtained from current market prices in the form of drug prices, salaries for staff or list prices for in-hospital procedures.

Nevertheless, two significant concerns about the calculation of costs resulting from the implementation of a specific health intervention need to be addressed. The first regards to decisions that involve the purchase of equipment such as a Computed Tomography (CT) scanner or Magnetic Resonance Imaging equipment that are considered to have a life-span of many years. In this case the cost of the equipment is considered to spread over the lifetime of the equipment and reflects the often real-life scenario of buying equipment through obtaining a loan and repaying the loan through annual mortgage payments. The amount that has to be paid every year during the lifespan of the equipment in order to repay the equipment and the interest on a loan is given by the amortized annual cost formula [255]:

$$M = P * i * \frac{(1+i)^{N-1}}{(1+i)^{N} - 1}$$

Where *M* is the amortized annual cost of payment incurred at the beginning of the year, *P* is the purchase price, *N* is the equipment lifespan and *i* is the annual interest rate. The second concern regards to the different manipulation that is required for costs that are incurred in the present and costs that are expected to be incurred in the future as it is known that there is a preference to incur costs later in the future instead of now [292]. The underlying reason behind this preference is the opportunity cost of money which can demonstrated by the fact that an amount of money that is not spent in the present can be invested to produce a larger

amount of money in the future. In order to account for this preference, future costs are required to be converted to their present values. The process of discounting allows the calculation of the present value of an amount of money that is spend in the future by taking into account a discount rate and the how far into the future the costs are incurred. The discount rate corresponds to the rate at which an amount of money is discounted per year and usually is equal to the annual interest rate that the same amount of money would yield if invested. Although discount rates used in economic evaluations has varied considerably, as low as 1% and as high as 10%, the most frequently used rates was 5% in the past 3% in more recent economic evaluations [255]. The present value of an amount of money spent in the future is calculated as follows [260]:

$$PV = \frac{S}{(1+r)^N}$$

Where S is the amount of a future expense, PV is the present value of the future expense, r is the discount rate (per year) and N is the number of years until the expense is incurred.

Estimation of Health Effects in CEA

The health effectiveness measure can be any kind of effectiveness measure depending on the interventions that are being tested such as cases detected when comparing different diagnostic procedures, number of cases prevented when comparing vaccination programs or years of life saved when comparing clinical interventions. However since decision makers are usually interested in allocation of constrained resources across different healthcare areas, a common effectiveness measure based on expected utility theory, the Quality Adjusted Life Year (QALY), has gained widespread use in CEA analysis [293]. When the primary health effect

outcome in CEA is measured in QALYS, these analyses are often called Cost Utility Analysis (CUA).

The main benefit of using QALYs as the main effectiveness measure is its ability to capture not only the intervention's impact on length of life (as it happens with measure "years of life saved") but also the intervention's impact on quality of life In principle, QALYs are a measure of health effectiveness which assigns to particular lifetime periods a weight (health utility), ranging from 0 to 1. The assigned health utility ideally corresponds to the health related quality of life (HRQoL) during that particular lifetime period. A health utility equal to 1 corresponds to full health while a health utility of 0 corresponds to a health state equivalent to death. For health states that may be considered worse than death a health utility with a value less than 0 can be used. Additionally, any change along the interval scale between 0 and 1 should be considered equivalent to any other change of the same magnitude regardless of the point along the scale that relates to (e.g a change between 0.3 to 0.4 is equivalent to a change between 0.7 to 0.8) [294]. QALYs are calculated by summing the products of the time spend in a particular time period with the specific health utility of that particular time period across a lifetime. Similarly the Quality Adjusted Life Expectancy (QALE) for an individual at age *x* can be calculated as follows [295]:

$$QALE = \sum_{t=x}^{x+L} Q_t$$

Where *L* is the life expectancy of the individual at age *x* and *t* corresponds to single years within that life expectancy range and *Q* is the health utility for each year *t*. A graphical representation of how QALYs are calculated is presented in Figure 4.3.

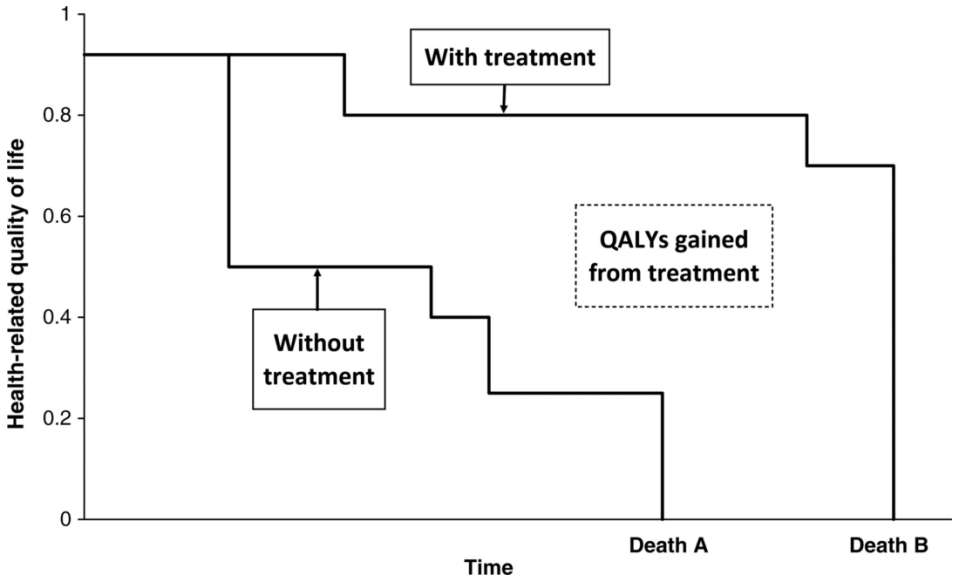


Figure 4.3: *A diagram that represents how QALYs are calculated. The area under each curve (AuC) calculated as health utility x Time represent the total QALYs lived by the individual. The difference between the two AuC corresponds to the total QALYs gained by the individual due to the intervention. Adapted from Whitehead et al* [294]

The health utilities that are part of the QALY are calculated in differently compared to the more widely used descriptive HRQoL measures that are derived from more widespread tools such as the Short Form 36-item questionnaire (SF-36) [296] or disease specific HRQoL questionnaires [194, 195]. Descriptive HRQoL tools usually provide a summary of scores for several quality of life domains (such as self-functioning, physical activity and social activity). On the contrary health utilities do not just describe the characteristics of the health state as such or how it affects the ability of the individual to function but instead reflect how an individual values (or feels about) a specific health state, in other words describe what is the individual's health preference about that specific health state [255]. Health utilities can be estimated with a number of methods. These can be distinguished in methods that direct elicit utilities such as the Time-Trade-Off (TTO) approach, the Visual Analog Scale (VAS) and the Standard Gamble (SG) approach [297] and multi-attribute utility scales such as the Euro Qol-5D (EQ-5D) that can indirectly obtain utility estimates.

The simplest of the direct elicitation methods is the VAS which consists of a rating scale as the one presented in figure 4.4 and the question "On a scale where 0 represents the worst imaginable health state and 100 represents the best imaginable health state, what number best describes your current health status?".

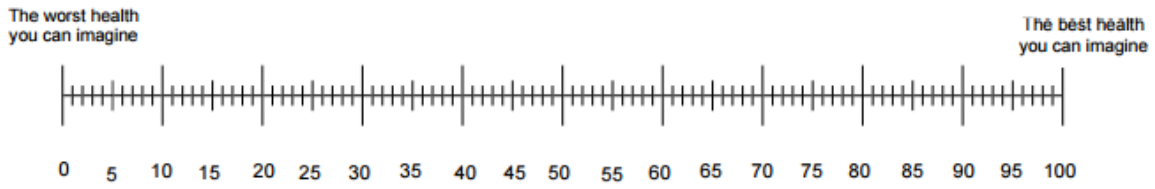


Figure 4.4: *An example of a Visual Analog Scale. Adapted from the EuroQol group EQ-5D questionnaire* [298].

Responders are asked to place an X on the VAS for the value of death and an X for the value which best describes the health state. Utility Weights for the specific disease health state are them given by the formula [299]:

$$HU = \frac{x - d}{100 - d}$$

Where HU refers to Health Utility, x refers to the placement on the VAS for the health state and d refers to the placement on the VAS for death.

The TTO approach includes assessment of the health state utility by asking how much time (usually years of life) would the responder trade to improve his current health state to a health state of perfect health. . The usefulness of this approach relies on the assumptions that the impact of health problems/symptoms can be quantified and that the more life years someone is willing to sacrifice (trade off), the worse his/her health status is [300]. The point where the responder is indifferent towards living a full life with the disease or a shorter life with perfect health is the point of indifference and indicates the preference of the health state. As opposed to VAS and TTO approach, the SG relies directly on utility theory and is the original method

of measuring utilities [301]. Utility theory considers a utility scale (or utility function) as an assignment of numerical values to a set of outcomes, given that if the expected value of the utilities assigned to the outcomes in one branch of the decision tree in greater than the expected value of the utilities assigned to outcomes in another branch of the decision tree, then the first branch (alternative decision) is preferred over the second one [255]. The principle that underlies SG application in health preferences elicitation is that it assesses the utility by asking a respondent to choose between living in a specific health state and a gamble between perfect health and death. Subsequently the worse the health state assessed is, the higher risk of immediate death one will be willing to accept in order to avoid it. The decision tree presented in Figure 4.5 demonstrates the alternative choices.

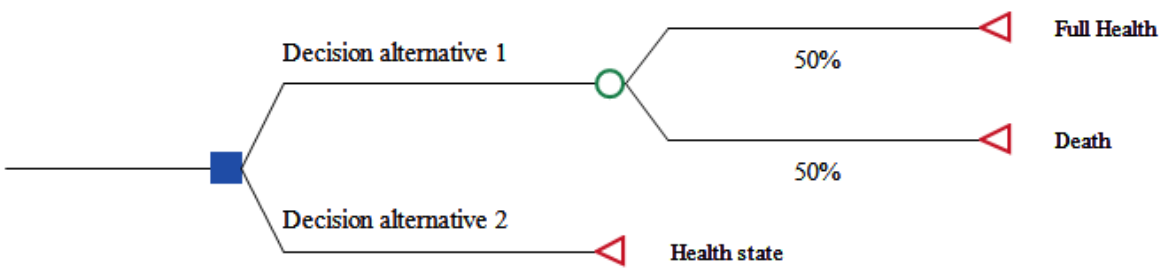


Figure 4.5: *A decision tree diagram demonstrating the alternative choices under Standard Gamble with a 50% of immediate death. Developed with TreeAge PRO software.*

The point of indifference, which is the probability value of immediate death at which the responder cannot choose between the two alternative decisions, is considered the responder's utility for the particular health state [301].

Indirect elicitation of health utilities is possible using multi-attribute utility scales such as the Euro Qol-5D-5L (EQ-5D-5L). Multi-attribute scales are frequently also called Health Indexes (HI) and constitute a classification instrument that resembles hybrid between traditional measurement of HRQoL with questionnaires such as the SF-36 and direct elicitation of health

utilities using the TTO approach. HI categorizes health status in different categories and the patient is asked to respond to each specific category. For example the EQ-5D-5L categorizes health status into the following five domains: (a) Mobility, (b) Self-care, (c) Usual activities, (d) Pain/discomfort, (e) Anxiety depression and for each domain, the responder chooses among five levels: (a) No problem, (b) Some or limited problems, (c) Moderate problems, (d) Severe problems, (e) Unable or extreme problems [298]. The patients respond to the HI questionnaire but the different health states described by the HI are assigned preferences by polling the reference population (e.g. the general public) using a direct elicitation method such as the TTO. All direct elicitation approaches and the HI method involve relevant but not identical concepts so their results may vary and caution is required during interpretation. The main distinction that needs to be considered is that direct elicitation of health utilities using the VAS, TTO or SG assessed the preference of the affected individual for the specific health state while HI assesses the societal preference for the specific health state. Between the direct elicitation methods, VAS is the easiest to perform while SG and TTO may often confuse responders so interviewers usually rely on visual props or use questioning techniques to train the responder and avoid the introduction of bias [297].

Uncertainty in Cost-Effectiveness Analysis

All information that has been collected is combined and with the use of mathematical modeling the final ICER is reported. However, given that the model structure itself, as well as the model parameters, is the result of various choices, assumptions and simplifications made by the analyst towards completing the CEA, it is expected that the final result of the CEA will be characterized by uncertainty. The main two approaches that are commonly used to address the issue of uncertainty are sensitivity analysis in deterministic modeling and uncertainty propagation using probabilistic sampling. In deterministic modeling, all parameters are

defined by discrete values while in probabilistic modeling; parameters are defined by a probability distribution of values rather than discrete values.

In principle, sensitivity analysis involves the systematic evaluation of how the uncertainty about the deterministic estimates of several parameters and assumptions influences the final model results. In its simplest form, called one –way sensitivity analysis, the analyst varies the estimate of one parameter across a range of values while in the meantime all other variables remain at their "best estimate" value. The range of values over which the parameter is varied could be from minimum to maximum, across the 95% CI or across any other range the analyst considers meaningful [302]. Although one-way sensitivity analysis is useful tool that allows an improved insight into the factors that influence the results and provides a validity check to address the effect of parameters taking on their extreme values, it faces important limitations as it is known to grossly underestimate overall uncertainty and provide a somewhat misleading picture by not taking into account the interaction between multiple parameters in the model [290]. An example of one way sensitivity analysis and the resulting Tornado graph that is used to present its results is presented in Figure 4.6.

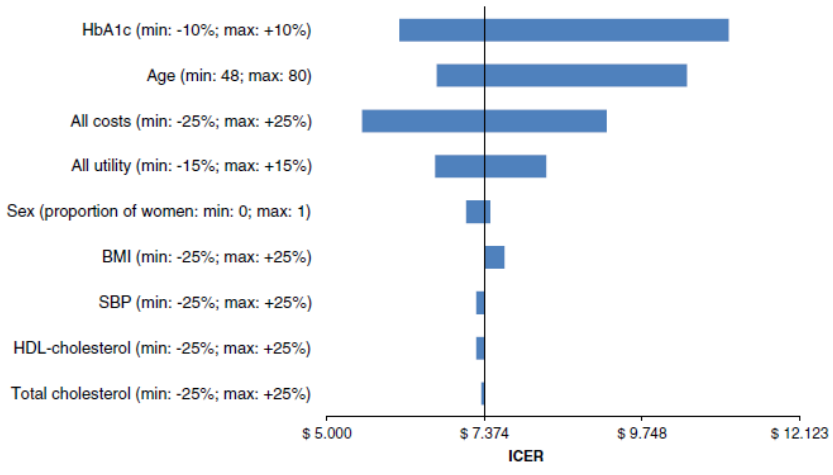


Figure 4.6: *A characteristic example of a Tornado diagram resulting from one way Sensitivity Analysis performed in regards to a cost effectiveness analysis for the treatment of type II diabetes with saxagliptin in Argentina. The Tornado graph presents the examined parameters and how these were varied on the left hand side of the graph while on the other side it presents the variation of the ICER. The parameters are ranked relative to the magnitude of their impact on the ICER. Adapted from Elgart JF et al* [303].

Other forms of sensitivity analysis such as the two-way and multi-way sensitivity analysis can also be used. Two-way sensitivity analysis is a useful alternative when the analyst already knows of specific two key parameters that are correlated and wants to avoid the misleading view of one-way sensitivity analysis. Examples of such correlated parameters in the case of CEA for cancer treatment interventions could be the hazard ratio for survival without disease progression and overall survival, or health utility estimates for moderate and severe disease states [304]. Multi-way sensitivity analysis (also called scenario analysis or extreme scenario analysis) involves simultaneously setting each parameter to take the most extreme values as if model parameters were perfectly correlated. This approach allows the analyst to both assess the best case and worst case scenarios. However, in real life, parameters neither vary in isolation nor are perfectly correlated and as one-way sensitivity analysis underestimates the overall uncertainty, multi-way sensitivity analysis overestimates it [302]. In general, all deterministic sensitivity analysis approaches cannot provide the analyst with an indication of the likelihood of the result but rather provide an array of different results associated with varying one or more estimates (a process which is strongly dependent on additional arbitrary choices by the analyst) [290]. Furthermore, it relies to the analyst to decide what a noteworthy difference in the sensitivity analysis results is, while the potentially extensive array of results poses additional difficulties in terms of presentation and interpretation [255].

On the contrary, uncertainty propagation using probabilistic sampling (also called Probabilistic Sensitivity Analysis (PSA)) allows the simultaneous assessment of all uncertainties in the model and provides the analyst with an estimate of the uncertainty in the model output which resulted from the uncertainties in the model inputs. In order to do so, the analyst needs to provide model parameter inputs in the form of probability distributions instead of discrete values and these probability distributions should reflect the uncertainty

around the best estimate (i.e. mean and standard error). PSA usually relies on a robust sampling technique such as Monte Carlo simulations to propagate uncertainty through the model. The principle underlying the Monte Carlo technique is that uncertainty is propagated by randomly selecting values from the probability distributions for each uncertain parameter and this is repeated for sufficiently large number of iteration (>1000) and the results are represented as distribution of incremental effects, incremental costs and eventually as a distribution of ICERs [290]. Figure xx presents a representation of the principle that underlies Monte Carlo simulation.

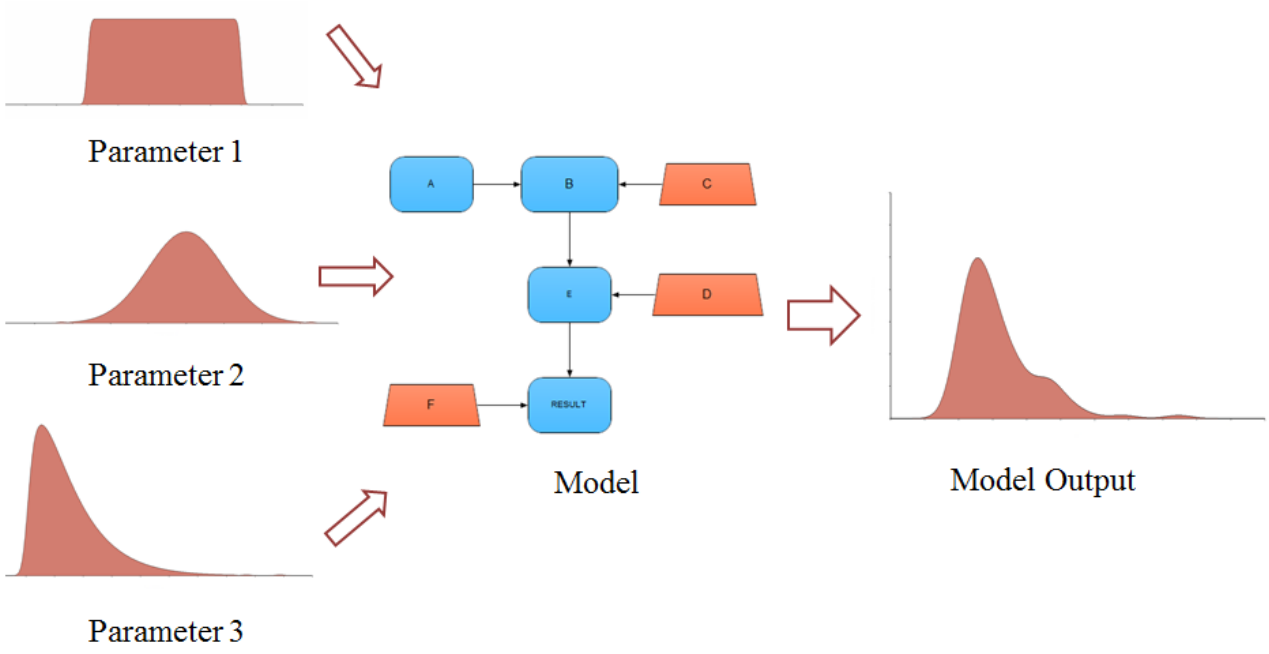


Figure 4.7: *Representation of PSA in CEA using Monte Carlo. Model Output could be incremental effect or incremental costs. Adapted from Hunnick et al 2014* [255]

Following PSA, the results are usually presented as ICERs with a corresponding 95% CI and/or as cost-effectiveness acceptability curve (CEAC). CEACs provide a graphical representation of the probability that the intervention is cost effective (y-axis) across a range of ceiling ratios. Following PSA the incremental costs and effects distributions are combined into a series of net benefit distributions and the probability that the specific intervention is

cost-effective is plotted across a series of ceiling ratios [305]. A characteristic example of a CEAC is presented in Figure 4.8 [306]. It regards to a study about the cost-effectiveness of the use of adjunctive cognitive therapy for relapse prevention in chronic depression (versus clinical management and medication alone). The curve indicates the probability that cognitive therapy is more cost effective compared to the alternative for a range of ceiling ratios (maximum acceptable value for ICER). It can be deducted from the CEAC that if the decision maker wants to spend £6000 per additional relapse avoided the adjunctive cognitive therapy has a probability of being more cost efficient of 60% while if the decision maker is willing to spend £8500, the probability rises to 80% [37].

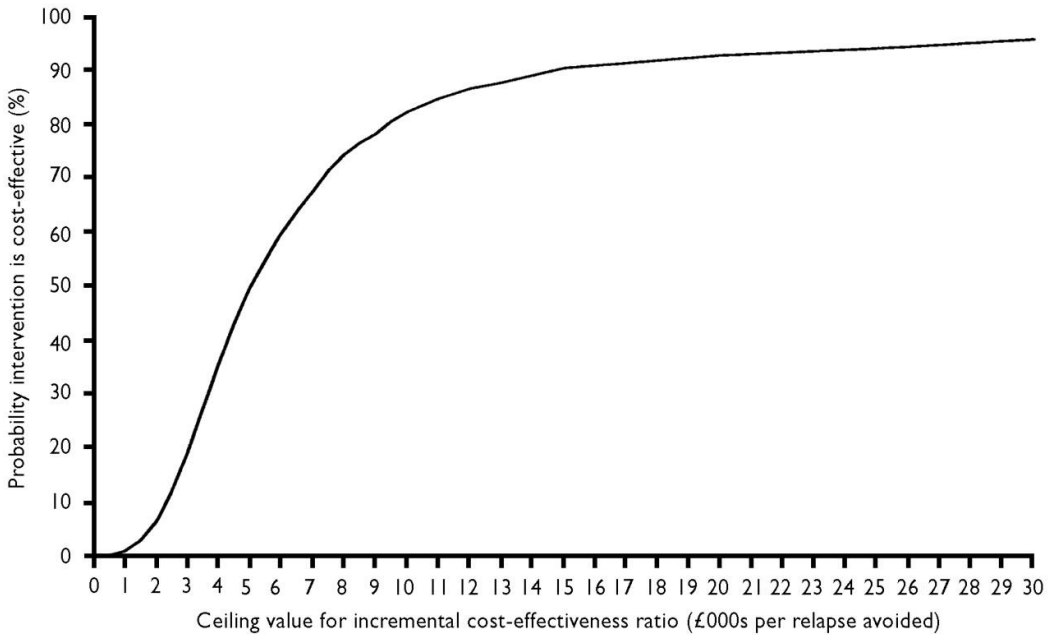


Figure 4.8: *A characteristic example CEAC which regards a study about the cost-effectiveness of the use of adjunctive cognitive therapy for relapse prevention in chronic depression (versus clinical management and medication alone). The X axis displays the ICER ceiling values and the Y axis the probability of the intervention being cost-effective. The curve indicates the probability that cognitive therapy is more cost effective compared to the alternative for a range of ceiling ratios (maximum acceptable value for ICER). Adapted from Scott J et al 2013* [306].

Chapter 5: Diagnostic accuracy of nasal nitric oxide for establishing diagnosis of primary ciliary dyskinesia: A meta-analysis.

Abstract

Background: To date, diagnosis of Primary Ciliary Dyskinesia (PCD) remains difficult and challenging. We systematically evaluated the diagnostic performance of nasal Nitric Oxide (nNO) measurement for the detection of PCD, using either velum-closure (VC) or non-velum-closure (non-VC) techniques.

Methods: All major electronic databases were searched from inception until March 2015 using appropriate terms. The sensitivity and specificity of nNO measurement was calculated in PCD patients diagnosed by transmission electron microscopy, high speed video-microscopy or genetic testing. Summary receiver operating characteristic (HSROC) curves were drawn using the parameters of the fitted models.

Results: Twelve studies provided data for thirteen different populations, including nine case-control (n=793) and four prospective cohorts (n=392). The overall sensitivity of nNO measured by VC techniques was 0.95 (95% CI 0.91-0.97), while specificity was 0.94 (95% CI 0.88-0.97). The positive likelihood ratio (LR+) of the test was 15.8 (95% CI 8.1-30.6), whereas the negative likelihood ratio (LR-) was 0.06 (95% CI 0.04-0.09). For non-VC techniques, the overall sensitivity of nNO measurement was 0.93 (95% CI 0.89-0.96) whereas specificity was 0.95 (95% CI 0.82-0.99). The LR+ of the test was 18.5 (95% CI 4.6-73.8) whereas the LR- was 0.07 (95% CI 0.04-0.12).

Conclusions: Diagnostic accuracy of nNO measurement both with VC and non-VC maneuvers is high and can be effectively employed in the clinical setting to detect PCD even in young children, thus potentiating early diagnosis. Measurement of nNO merits to be part of a revised diagnostic algorithm with the most efficacious combination of tests to achieve PCD diagnosis.

Introduction

Primary ciliary dyskinesia (PCD) is a rare, hereditary disorder characterized by impaired mucociliary clearance [307]. Apart from situs inversus in ~50% of the cases, the main manifestations of the disease are not specific. Nevertheless, the associated recurrent sinopulmonary infections eventually lead to severe chronic lung disease and development of bronchiectasis [308, 309].

While some centers began using targeted genetic testing,[310] the diagnosis of PCD in the majority of centers currently relies on an array of different sophisticated tests namely the High Speed Video Microscopy (HSVM) for ciliary motility assessment,[311] Transmission Electron Microscopy (TEM) for the examination of cilia ultrastructure [312] and nasal nitric oxide (nNO) measurement [313]. The diversity of the employed diagnostic tests reflects the lack of a golden diagnostic standard and the weaknesses and inaccuracies that characterize each of these tests. In particular, TEM examination of ciliary axonemes exhibits normal ultrastructure in confirmed patients with biallelic mutations in certain disease-causing genes such as DNAH11, [314] while the motility patterns observed by HSVM vary widely depending on the implicated genetic variant [315, 316].

Nasal nitric oxide (nNO) is abnormally low in PCD patients [317] and it has been part of the diagnostic work-up in many PCD centers.[318] Current American Thoracic Society/European Respiratory Society (ATS/ERS) guidelines for nNO measurements recommend air aspiration via a nasal probe while the subject exhales through the mouth against resistance in order to maintain velum closure. Alternative techniques to maintain velum closure such as breath hold or pursed- lip breathing via the mouth are also acceptable [319]. However, velum closure requires cooperation and this precludes the performance of these techniques in young children. Few reports have investigated the discriminative ability of nNO measurements with

the velum open as in the case of tidal breathing [309, 320] with encouraging findings for the usefulness of this technique in screening for PCD in younger children and adults unable to perform velum closure.

In view of the above specific restrictions and weaknesses, for the clinicians and the patients it remains of key importance to appraise the potential diagnostic value of each of the available diagnostic tests for PCD, in order to find its place in the armamentarium for elicitation of the diagnosis of the disease. A recent systematic review and meta-analysis summarized the published evidence on the measurement of nNO in PCD and reported on the mean difference of nNO production values obtained during velum closure techniques in PCD patients versus healthy controls (231 nL/min, 95% CI: 193.3-268.9) and cystic fibrosis patients (114.1 nL/min, 95% CI: 101.5-126.8) [321]. However, that report did not perform a meta-analysis on the diagnostic accuracy of nNO measurements in order to provide synthesized data on the potential diagnostic value this test may have in future algorithms for PCD diagnosis, which would be particularly informative in clinical decision making. The aim of this study was to systematically evaluate the diagnostic performance of nNO measurement as obtained either with a velum-closure or a non-velum-closure technique in screening for PCD so as to provide appropriate summary estimates of diagnostic accuracy with each breathing technique and demonstrate the summary trade-off between sensitivity and specificity across the included studies.

Methods

Search strategy and selection criteria

The electronic databases PubMed, SCOPUS, Cochrane Database of Systematic Reviews and Google Scholar were searched from inception until March 2015 using the keywords: ‘nasal nitric oxide, “nNO”, “nasal NO”, “Primary Ciliary Dyskinesia”, “PCD”, “lung”,

"pulmonary", "pulm*", "cilia" either in the title or the abstract or using MeSH terms. The references of eligible studies were further examined for possible missing articles. We included studies which were identified after two reviewers (PK, SIP) independently screened the title and abstract of the obtained search results. Final selection was based on full text evaluation. Any disagreements were resolved by discussion and in case of discrepancy, by a third researcher (PKY). As this study is based on a systematic review of the previously published literature, an ethical approval was not obtained, since there is no potential of participant identification and ethical approval and consent was already obtained at the individual study level. The guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) were followed.

The validity of each primary study was assessed using the Quality Assessment of Diagnostic Accuracy Studies -2 (QUADAS-2) tool [322], that evaluates the risk of bias and applicability of diagnostic accuracy studies. It consists of four key domains: patient selection, index test, reference standard, flow and timing. Each is assessed in terms of risk of bias and the first three in terms of issues regarding applicability.

Studies were considered eligible if they provided data on the sensitivity and specificity of nNO for the diagnosis of PCD in order to construct a 2 x 2 table for each study calculating true positives (TP), false negatives (FN), false positives (FP) and true negatives (FN) for the presence or not of PCD according to nNO values set as a cut-off in each study. In some studies, the numbers were not provided per se but it was possible to extract them from other manuscript data sources. In case of incomplete information, we contacted the authors of the primary studies. Studies that reported only mean values of nNO were not included in our analyses as they did not provide data for computing summary diagnostic accuracy estimates (sensitivity, specificity, positive and negative likelihood ratio). Disease status in each selected study was required to have been confirmed by TEM and/or HSVM or genetic testing.

Additional information on NO analyzer type, flow rate and breathing maneuver was also collected and used in data synthesis. Studies that did not report the equipment and flow rate used were not considered eligible as well as studies that may have used flow rate outside of the ATS/ERS recommended range (0.25-3L/min)[319]. Cut-off values for the nNO test, were usually reported in parts per billion (ppb) and were transformed to NO production rate units (nl/min), using the conversion formula concentration (ppb) x sampling rate (L/min) as used previously [313], in order to account for the used different flow rates. Breathing maneuvers such as breath hold (BH) and exhalation against resistance (ER) were categorized as velum closure (VC) techniques and in case of both maneuvers performed by the study subjects; only results for the ER maneuver were included as the most validated technique according to ATS/ERS guidelines [319]. For studies employing the non-velum closure (non-VC) technique, only results of nNO measurements that were performed during tidal breathing (TB) with mouth open were included in the meta-analysis.

Data extraction

The name of author, study design, publication year, country of origin, study population sample size, age distribution of study population subgroups, nNO cut-off levels, information on the measurement method and the test(s) used for the diagnosis of PCD were recorded for each study. Data on the values of TP, TN, FP, FN were extracted independently by two reviewers (PK, SIP). A third investigator (PKY) settled any discrepancies and consensus was reached for all data.

Analysis

A bivariate model was used to calculate estimates of overall sensitivity and overall specificity. We fitted a two-level mixed logistic regression model conditional on the sensitivity and the specificity of each study and a bivariate normal model for the sensitivity and specificity between studies [323]. This method combines information from multiple thresholds and the output is expressed as a hierarchical summary receiver operator curve (HSROC). The HSROC describes the relationship between sensitivity and specificity derived from the individual receiver operator curves (ROC) of each study. Following this method, it describes the 'average' relationship between a continuous cut-off value and discriminatory ability in the 'average' population. This maximizes the amount of information used in the evidence synthesis and better represents the available data. The advantage of this method is that it allows clinicians to estimate how changing thresholds will alter the diagnostic utility of the test under study. All calculations are performed using STATA (Version 12, StataCorp, College Station, Texas) with the commands metandi and metandi plots for analyses of four studies and above [324].

We also reported the summary likelihood ratios across all studies. These measures also combine in their calculation both sensitivity and specificity. Positive likelihood ratio (LR+) is the ratio of sensitivity / (1-specificity), whereas negative likelihood ratio (LR-) is defined as the ratio of (1-sensitivity)/specificity. When there is absolutely no discriminating ability for a diagnostic test, both ratios are equal to 1. The discriminating ability is better with higher LR+ and lower LR-. A good diagnostic test has typically LR+ greater than 5.0 and LR- less than 0.2 [325].

VC and non-VC measurements were analyzed separately and this allowed us to arrive at estimates on overall sensitivity, specificity and likelihood ratios for nNO depending on VC

status. We also performed a sensitivity analysis including only studies in which PCD status was defined by TEM and at least one more diagnostic test, with the rationale to examine whether the diagnostic accuracy of nNO measurement differs with the inclusion of a more representative spectrum of PCD population. Measurements of nNO were compared to PCD diagnosis obtained through a combination of tests which included TEM and HSVM or DNA testing.

Results

Eligible Studies

Of the 1940 items retrieved through online search, 1866 were excluded based on the title and abstract and the remaining 74 were downloaded for detailed, full text assessment. Two additional studies were identified through references screening and were also evaluated. Studies with overlapping populations were cross-checked and final selection was based on the largest number of participating PCD patients. In summary, 26 studies did not provide data on sensitivity and specificity, 13 items involved overlapping populations, 15 items were review papers while the remaining items that were excluded were case reports (2), editorials (3) and guidelines papers (2) (Figure 1). Of the total 76 studies assessed in detail, 15 provided enough data for the construction of a 2x2 table. Among these, two studies did not report type of NO analyzer and flow rate and despite our effort to obtain this information after contacting the authors, this was not feasible and they were excluded from the analysis [326, 327]. Finally, quantitative synthesis included data on thirteen different populations from twelve studies (Marthin et al included data on more than one population) and two separate analyses were carried out, based on the employed breathing maneuver.

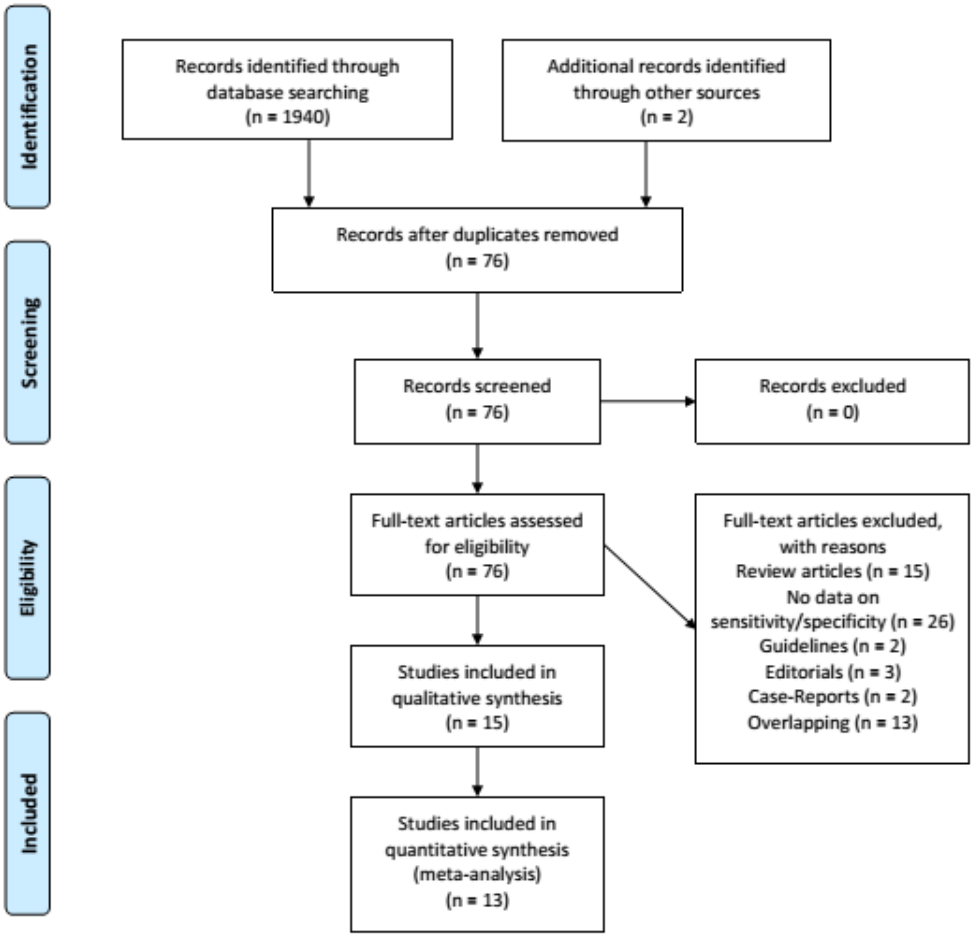


Figure 5.1: *PRISMA diagram for the search strategy and selected studies*

Study characteristics

Descriptive characteristics of the included studies are presented in Table 1. From twelve studies, 325 PCD patients and 711 non-PCD subjects were included in the meta-analysis for the diagnostic performance of VC nNO testing. In the case of non-VC nNO testing, 210 PCD patients and 471 non-PCD subjects from seven studies were included. The majority of the studies were performed in Western Europe and only two in North America. Four studies evaluated the diagnostic efficacy of nNO in cohorts of referred suspect patients for PCD testing [313, 328-330] whereas the rest of the studies had a case-control design. Controls were non-PCD subjects, either healthy subjects only [328, 331-334] or healthy subjects and patients with other respiratory diseases [320, 335-337].

Table 5.1: *Characteristics of the included studies*

#	Author/Study	Country	Study Design	Study Population [a]	Age range (yrs) (Mean, range/SD)	Analyzer	Flow rate (L/min)	Measurement Method [b]	Cut – off (nL/min)	Inclusion criteria/ Diagnosis
1	Narang I [26] (2002)	United Kingdom	Case - Control	31 PCD 53 HC	PCD: 11.0 (5.5-17.3) HC: 10.7 (5.5-19.0)	LR 2000	0.25	BH	62.5	HSVM and TEM
2	Corbeli R [23] (2004)	Switzerland	Prospective Cohort	17 PCD 17 non PCD (BE,B)	All: 11.4 (1.2)	CLD88sp	1.20	BH	126	TEM
3	Piacentini G [25] (2008)	Italy	Case - Control	10 PCD 27 HC	PCD: 17 (-) HC: 7 (-)	NIOX Flex	0.30	BH	21.3	TEM
4	Mateos Coral D [31] (2011)	Canada	Case – Control (with longitudinal follow-up in a subsample)	20 PCD 65 non PCD (CF,BE,HC)	PCD: 11.4 (3.5) HC: 11.0 (3.7) CF: 11.0 (3.4) BE: 10.9(3.3)	CLD88sp	0.33	ER & TB	ER: 58.5 TB: 37.1	TEM
5	Marthin JK [22] (2011) Substudy 3	Denmark	Prospective Cohort	20 PCD 97 non PCD	All: 6.9 (0.0-62.4)[c]	NIOX Flex	0.30	BH & TB	BH: 52.5 TB: 47.4	HSVM and TEM
6	Marthin JK [22] (2011) Substudy 2	Denmark	Case - Control	59 PCD 57 HC	PCD: 17.4 (3.6-65.8)[c] Non PCD:29.5 (3.1-63.6)[c]	NIOX Flex	0.30	BH & TB	BH: 52.5 TB: 47.4	HSVM and/or TEM
7	Leigh M [7] (2013)	United States	Prospective Cohort	71 PCD 84 non-PCD	PCD: 23.3(18) Non PCD: 31.8 (22.3)	Sievers 280i CLD88sp NIOX Flex	0.50 0.33 0.30	ER	76.9	TEM and DNA
8	Boon M [30] (2014)	Belgium	Case - Control	38 PCD 188 non PCD (HC, CF, Asthma, HID)	PCD: 14.3 (8.8-18.1)[c] HC: 14.9 (10.8-20.4)[c] CF: 14.0 (9.2-17.9)[c] Asthma: 12.1 (9.8-16.5)[c] HID: 10.7 (8.2-15.6)[c]	CLD88sp	0.30	ER & TB	ER: 90 TB: 60	HSVM and TEM (and culture)
9	Harris A [14] (2014)	United Kingdom	Case - Control	13 PCD 37 non PCD (HC,CF, CSLD)	PCD: 23 (5-71) HC: 31 (8-65) CF: 15(6-29) CSLD: 36 (8-79)	NIOX Flex NIOX MINO	0.30	BH & TB	BH: 38 TB: 30	HSVM and TEM (and culture for some)
10	Montella S [27] (2012)	Italy	Case - Control	23 PCD 23 HC	PCD: 15.8 (4.6-32.8)[c] HC: 15.7 (4.3-32.1)[c]	NIOX MINO	0.30	TB	17.4	HSVM and TEM
11	Santamaria F [28] (2008)	Italy	Case - Control	14 PCD 14 HC	PCD: 15 (7-27) HC:16 (7-27)	NIOX Flex	0.28	BH	7.2	TEM
12	Moreno Caldo A [29] (2010)	Spain	Case Control	9 PCD 112 non PCD (HC, CF, Asthma ,BE)	PCD: - (7-14) HC: - (-) CF: - (6-14) Asthma: (6-17) BE: - (6-14)	LR2000	0.25	BH	28	TEM
13	Beydon M [24] 2015	France	Prospective Cohort	49 PCD 37 non-PCD	PCD: 11.4 (7,13.9) [d] Non PCD: 7.9 (4.9,11.6) [d]	NIOX Flex Endono 8000	0.30	BH/ER TB	BH/ER: 82.2 TB: 39.9	HSVM, TEM and/or DNA

PCD: Primary Ciliary Dyskinesia, HC: Healthy Controls, B: Bronchitis, CF: Cystic Fibrosis, BE: non CF non PCD Bronchiectasis, CSLD: Chronic Suppurative Lung Disease, HID: Humoral Immunodeficiency Disorders, TEM: Transmission Electron Microscopy, HSVM: High Speed Video Microscopy, DNA: Genetic testing

[a] Study population refers to subgroups that comparisons (sensitivity, specificity, PPV, NPV) were reported for in the original articles.

[b] Measurements methods taken into account for the meta-analysis, BH: Breath Hold, ER: Exhalation against Resistance, TB: Tidal Breathing

[c] Median (range)

[d] Median (IQR)

The number of PCD patients (range: 9-59) and controls (range: 14-188) per case-control study varied widely. All case-control studies confirmed PCD status by TEM findings while in 55% of them HSVM was also performed. Of the four prospective studies, Beydon et al used a combination of TEM, HSVM and genetic testing to confirm PCD diagnosis [330] while Marthin et al used TEM and HSVM in their cohort of consecutive referrals [328]. Leigh et al confirmed PCD via a combination of ultrastructure assessment and genetic testing [313] while the smallest cohort study confirmed PCD only via ultrastructural assessment [329]. The sensitivity and specificity of each included study with VC and non-VC technique are shown in Figure 2.

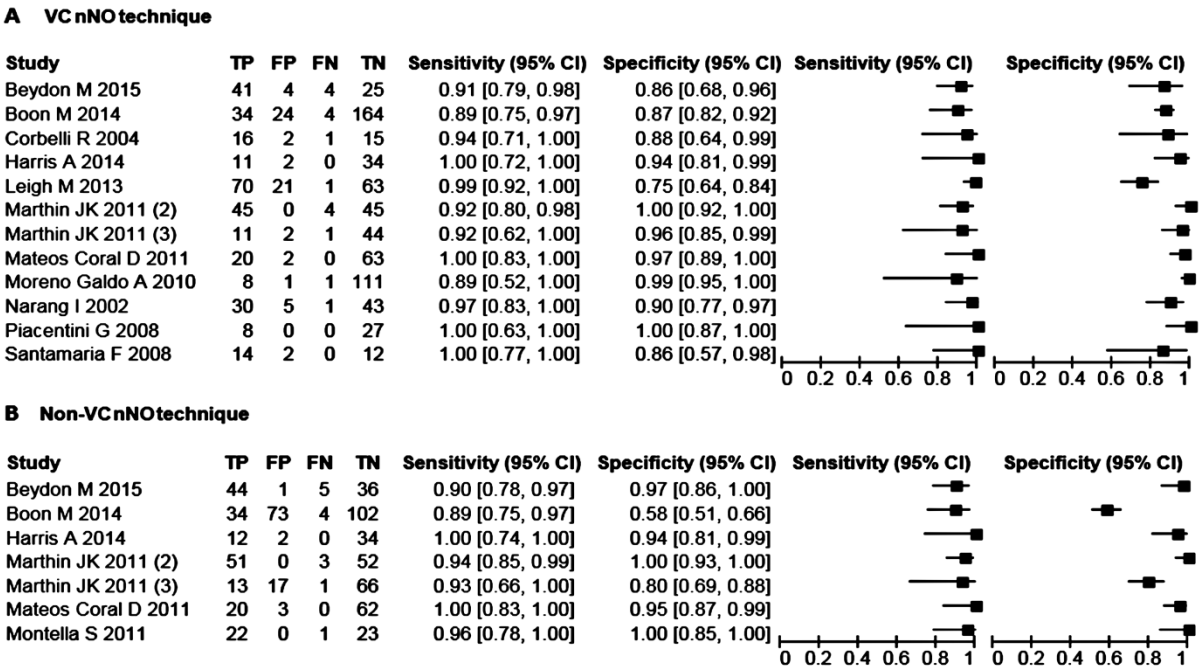


Figure 5.2: *Forest plots for sensitivity and specificity. Forest plot of sensitivity and specificity of nNO for detecting PCD with the 95% CI for each population of the included studies. A) Forest plot for studies employing a VC breathing technique and B) Forest plot for studies employing a non-VC breathing technique*

Quality assessment

Reporting of the meta-analysis is based on PRISMA guidelines.[338] Based on the QUADAS-2 tool, the quality assessment of the primary studies is shown in Table 2. In

general, the analyzed studies had overall reasonably good methodology and this offers relative reassurance that results have not been substantially influenced from bias.

Table 5. 2: *QUADAS-2 Quality Assessment results*

Study	RISK OF BIAS				APPLICABILITY CONCERNS		
	PATIENT SELECTION	INDEX TEST	EFERENCE TANDARD	FLOW AND TIMING	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD
Boon 2014	L	U	L	L	L	L	L
Mateos Coral 2011	U	L	L	L	U	L	L
Piacentini 2008	U	L	L	L	U	L	L
Santamaria 2008	L	U	L	L	L	L	L
Montela 2012	U	L	U	L	L	L	L
Corbelli 2004	L	U	L	L	L	L	L
Narang 2002	L	L	L	L	L	L	L
Harris 2014	L	U	L	L	L	L	L
Leigh 2013	L	L	L	L	L	L	L
Marthin 2011	L	U	L	L	L	L	L
Moreno Galdo 2010	U	U	L	L	U	L	L
Beydon M 2015	L	L	L	U	L	L	L

QUADAS 2 consists of four key domains covering patient selection, index test, reference standard and flow of patients through the study and timing of the index test and reference standard ("flow and timing"). Each domain is assessed in terms of the risk of bias and the first three are also assessed in terms of concerns regarding applicability
U: Unknown, L: Low, H:High

Data Synthesis

The overall sensitivity of abnormal (low) nNO measured by VC techniques for all the included studies was 0.95 (95% CI 0.91-0.97), while the specificity was 0.94 (95% CI 0.88-0.97). The LR+ of the test was 15.8 (95% CI 8.1-30.6), whereas the LR- was 0.06 (95% CI 0.04-0.09). The HSROC curve is shown in Figure 3.

For the non-VC techniques the overall sensitivity of nNO to detect PCD was 0.93 (95% CI 0.89-0.96) whereas the specificity was 0.95 (95% CI 0.82-0.99). The LR+ of the test was 18.5 (95% CI 4.6-73.8) whereas the LR- was 0.07 (95% CI 0.04-0.12). The HSROC curve is shown in Figure 3.

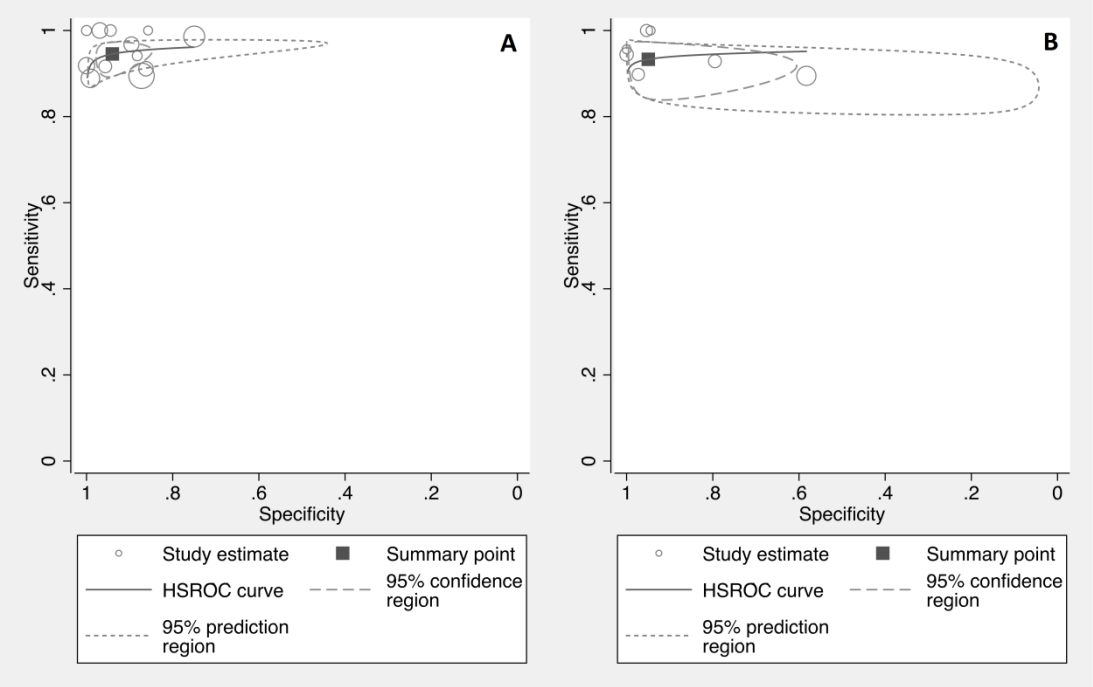


Figure 5.3: *VC and non-VC HSROC curves. HSROC curves for the included studies. A) HSROC curve for studies employing a VC breathing technique and B) HSROC curve for studies employing a non-VC breathing technique*

When we performed a sensitivity analysis, to calculate the nNO diagnostic accuracy in studies that only included PCD populations diagnosed by more than one test (combination of TEM and HSVM or genetic testing) [313, 320, 328, 330, 332, 333, 336], the results did not change significantly. Similarly, after performing a post hoc sensitivity analysis, with the exclusion of studies that used the electrochemical devise NIOX MINO [320, 333], the resulting estimates of overall sensitivity and overall specificity for the non-VC maneuver do not significantly differ from the estimates of the main analysis.

Discussion

In this meta-analysis we demonstrated that nNO measurement with VC techniques has overall sensitivity of 95% and a specificity of 94% whereas nNO measurement with the non-

VC technique has comparable and very similar sensitivity (93%) and specificity (95%). We applied a different approach to the one employed in a previous report [321] and evaluated nNO diagnostic performance metrics by using all the available evidence in the literature. These summary estimates allow us to make comparisons between the various proposed and established diagnostic tests for PCD, which are essential for clinical decision making. We also provide a graphical representation of our results using the hierarchical summary receiver operating characteristic (HSROC) curve incorporating the different cut-offs between primary studies. The clinical utility of nNO measurement is underlined by the high LR+ (VC: 15.8, non-VC: 18.5) and low LR- (VC: 0.06, non-VC: 0.07), meaning that an abnormal (low) nNO leads to a steep increase in the post-test probability of PCD, compared to the pretest probability, while in the case of a normal nNO measurement the opposite is also true [325]. However, since the sensitivity and specificity of the test are not 100%, in the presence of strong clinical suspicion for PCD [339], even in the case of a negative nNO test, a more detailed diagnostic work-up (HSVM, TEM, genetics) is indicated.

Current ATS/ERS recommendations for nNO include only VC maneuvers although recent evidence [320, 336, 337], that is supported by the results of this meta-analysis, highlights the discriminative ability of nNO during TB. TB is the only method available to obtain nNO measurements in young children (<5 years), which is particularly important as disease manifestations appear very early in life. Of course, the validity of nNO measurements in infants (<6 months) has been questioned, as nNO output in infancy is reduced due to the partial development of paranasal sinuses[331] where the majority of NO is produced [317] whereas the number of patients under 5 years which were evaluated in these studies [320, 328, 330, 333, 336, 337] is very small . The usefulness of nNO during TB has been demonstrated in the Danish cohort of 117 consecutive referrals with median age 6.9 years, where 83% were able to perform TB versus 50% for BH and 31% for ER [328]. There is

evidence that patients that have earlier diagnosis of this disease might have better clinical and functional outcomes [340, 341] and the application of this promising screening method in preschool children could not only lead to diagnosis at an earlier age but could also contribute towards the reduction of unnecessary cilia biopsies. Nevertheless, our results for the non-VC techniques should be interpreted with caution. Their low 95% CI limit for specificity is at 82% which suggests that a significant number of suspect PCD referrals is possible to give a falsely low nNO and prompt further diagnostic testing, thus increasing costs both to the healthcare system and the patient. Only seven studies were eligible for inclusion in the meta-analysis of non-VC maneuver, as opposed to twelve studies for the VC maneuver, and the low 95% CI limit of the former may be due to the limited sample size. Furthermore, two of the non-VC studies used the NIOX MINO portable device which uses electrochemical analysis of NO as opposed to the better validated chemiluminescence method of the stationary devices (NIOX FLEX, Ecomedics CLD88). Nevertheless, we performed a post hoc analysis with the exclusion of these studies which did not influence the diagnostic accuracy of the non-VC maneuver and the quantitative synthesis includes data from all seven non-VC studies.

NIOX MINO is a simpler and cheaper tool for measuring nNO, and validation studies have already been published [320, 342, 343]. However, as NIOX MINO was designed for exhaled NO measurement in asthmatic individuals, issues relating to its accuracy [343] and repeatability [320] have been reported when used for nNO measurement in subjects referred for PCD evaluation. In addition, while the manufacturer recommends measurements of nNO with BH for at least 45 seconds, this is usually not possible by many patients and instead NIOX MINO is frequently used with the alternative TB maneuver regardless of patient's age [320, 343]. These limitations question the suitability of NIOX MINO as a stand-alone diagnostic test and additional studies on the diagnostic accuracy of NIOX MINO

measurements during TB are needed to confirm the validity of this method. However, the low cost and simple use potentiate the consideration of NIOX MINO as a promising first line screening test in a future diagnostic algorithm for PCD.

Currently, there is no universally accepted cutoff for abnormally low nNO. The included studies in this meta-analysis proposed a variety of cutoffs for nNO production by VC (7.2-126 nl/min) and non-VC (17.4-60 nl/min) techniques. This variability demonstrates the need for standardization of nNO measurements and agreement on cutoffs for the different breathing maneuvers. A recent, large, multicenter study has proposed a cutoff equal to 77 nl/min for VC [313], whereas the meta-analysis by Collins et al reported that a cutoff of 75.2 nl/min would include 99.85% of PCD patients performing VC maneuvers [321]. Regarding the non-VC maneuver however, no cutoff value has been proposed by a large enough study, thus additional studies are needed for the establishment of such cutoffs and further standardization of the technique.

Our study has some limitations. The main limitation is the heterogeneity and the weaknesses of the diagnostic standard that the published studies are employing for the definition of PCD status. As a result, the captured spectrum of the disease might not be totally representative of the true PCD population. TEM, which is the most commonly used test for PCD status definition in published studies, misses approximately 30% of patients with PCD [344] and this should be taken into account in the assessment of the diagnostic efficacy of nNO. However, in the sensitivity analysis, when we included the studies that had employed more than one (in addition to TEM) diagnostic test to establish PCD diagnosis, our results did not change substantially thus providing relative certainty to the accuracy of nNO as a diagnostic test. There is considerable variation between individual studies in the number of cases, total sample size and cut-off values. However, the bivariate meta-analysis and HSROC curve analyses take explicitly this diversity into account and can accommodate studies with

populations of different risks and different definition thresholds. Another issue for the synthesis of the data is that the majority of the included studies were diagnostic case-control studies. Empirical evidence has shown that case-control studies, as opposed to cohort studies, may overestimate the diagnostics Odds Ratio (DOR) [345]. Nevertheless, we think that the possibility of overestimation is limited, as the case-control studies included here were diagnostic studies designed to assess the test accuracy and not to provide evidence on associations between a risk factor and the disease.[346] Additionally, given the rarity of PCD, it is expected that the majority of studies will have a case-control design. Due to the same reason, both case-control and prospective cohort studies included relatively small numbers of subjects. However the synthesis of the included studies led to the inclusion of data for several hundreds of PCD and non-PCD subjects and allowed the use of the appropriate statistical models and provided the estimates we report. It should be underlined of course that these estimates apply provided that the ATS/ERS guidelines are followed for the performance of the test and the obtained values are compared to the normal values obtained from samples of healthy subjects in the respective populations.

Conclusions

In summary, measurement of nNO, both with VC and non-VC maneuvers, has high overall diagnostic accuracy and provides a clinically significant diagnostic tool for large uninvestigated populations of suspect cases worldwide where access to TEM and HSVM is not easy. Furthermore, the high overall diagnostic accuracy of nNO calls for re-evaluation of the diagnostic accuracy of each of the available diagnostic tests for PCD (nNO, TEM and HSVM) with the aim to develop an algorithm with the most efficacious combination of tests to achieve PCD diagnosis.

Chapter 6: Prevalence of Primary Ciliary Dyskinesia in consecutive referrals of suspect cases and the Transmission Electron Microscopy detection rate: A systematic review and meta-analysis

Abstract

Background

Diagnostic testing for Primary Ciliary Dyskinesia (PCD) usually includes Transmission Electron Microscopy (TEM), nasal Nitric Oxide, High Speed Video Microscopy and genetics. Diagnostic performance of each test should be assessed towards the development of PCD diagnostic algorithms. We systematically reviewed the literature and quantified PCD prevalence among referrals and TEM detection rate in confirmed PCD patients.

Methods

Major electronic databases were searched until December 2015 using appropriate terms. Included studies described cohorts of consecutive PCD referrals in which PCD was confirmed by at least TEM and one additional test, in order to compare the index test performance with other test(s). Meta-analyses of pooled PCD prevalence and TEM detection rate across studies were performed.

Results

PCD prevalence among referrals was 32% (95%CI:25%-39%, I^2=92%). TEM detection rate among PCD patients was 83% (95%CI:75%-90%, I^2=90%). Exclusion of studies reporting isolated inner dynein arm defects as PCD, reduced TEM detection rate and explained an important fraction of observed heterogeneity (74%, 95%CI:66%-83%, I^2=66%).

Conclusion:

Approximately, one third of referrals, are diagnosed with PCD. Among PCD patients, a significant percentage, at least as high as 26%, is missed by TEM, a limitation that should be accounted towards the development of an efficacious PCD diagnostic algorithm.

Introduction

Primary Ciliary Dyskinesia (PCD) is caused by dysfunctional motile cilia and it is characterized by impaired mucociliary clearance which predisposes patients to recurrent respiratory infections. Patients usually suffer from lifelong rhinorrhea, chronic wet cough, progressive loss of lung function and eventually develop structural damage of the airways and bronchiectasis [347]. The main clinical manifestations that lead to consideration of PCD diagnostic testing are situs abnormalities, a history of neonatal respiratory distress syndrome, a family history of PCD, male infertility and chronic productive cough in the absence of more common causes of chronic lung disease [348].

Confirmation of a positive PCD diagnosis remains challenging as no single diagnostic test has been shown to have 100% sensitivity and specificity, thus a combination of diagnostic tests is usually needed for a final decision [349]. Specialized diagnostic testing is currently available only in few specialized centers and includes the measurement of nasal Nitric Oxide (nNO) [350], assessment of ciliary motility [351] and ciliary ultrastructure [352], while a few centers have also introduced genetic testing in clinical practice [353]. Overall, PCD diagnostic testing is expensive and time consuming [354, 355], which underlines the need to estimate the prevalence of PCD among referrals. This estimate is useful to know for cost-benefit analyses as higher prevalence of PCD among referrals corresponds to a lower proportion of non-PCD patients that are referred for PCD diagnostic testing and lower economic burden for the healthcare system and/or the patient family and vice versa. Furthermore, different centers follow various diagnostic algorithms for PCD diagnosis [356], indicating the need for the development of an evidence-based diagnostic decision tree for PCD. Such an approach requires the prior assessment of summary estimates of the diagnostic performance of each individual diagnostic test. Application of the Bayes Theorem on estimates of diagnostic performance along with information about the prior probability of

disease (prevalence of PCD among referrals) will allow the calculation of positive predictive values and negative predictive values for different diagnostic tests and algorithms [357].

In the past, assessment of ciliary ultrastructure abnormalities with Transmission Electron Microscopy (TEM) was considered to be the "gold standard" for the diagnosis of PCD [352]. However, for several years now, guidelines highlight that TEM cannot be considered as a gold-standard test [348] as a substantial subset of PCD patients display normal axonemal ultrastructure and cannot be identified through TEM [358]. These patients usually carry biallelic mutations in the *DNAH11* gene and their ciliary motility is characterized by a flickering movement [359]. Furthermore, another subset of PCD patients may also remain unidentified by TEM, as specific ultrastructural defects such as nexin link defects, are not be easily discernible by standard TEM [360]. Several studies reporting TEM findings in different cohorts of PCD patients demonstrate wide variation in the percentage of PCD patients missed by TEM ranging from below 10% [361, 362] to over 30% [363, 364]. Although some reviews and editorial papers reported that this percentage is equal to approximately 30% [349, 365], this estimate is not based on a systematic review of the entire published evidence.

This study systematically reviewed the published evidence aiming to quantify the prevalence of PCD diagnosis in cohorts of suspect cases referred for PCD diagnostic testing and to estimate the diagnostic detection rate of TEM in PCD patients in whom diagnosis was confirmed with a combination of tests.

Methods

Search strategy and selection criteria

The electronic databases PubMed, SCOPUS and Google Scholar were searched from inception until December 2015 using combinations of the keywords ‘Electron Microscopy’ and ‘Ciliary Motility Disorders’ as Medical Subject Headings (MeSH) or individual terms and ‘Primary Ciliary Dyskinesia’ OR “PCD” and combinations either in the title or in the abstract. The reference lists of the retrieved studies and reviews were further searched for additional reports. The included studies were identified after two reviewers (PK, SIP) independently screened the title and abstract of the obtained electronic search results and final selection was based on full text evaluation. A third researcher (PKY) resolved any discrepancies. The guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) were followed [366].

For the prevalence of PCD among referrals estimate, studies were selected according to the following inclusion criteria: Cohorts of consecutive referrals for PCD testing and, in addition to TEM, at least one more test such as High Speed Video Microscopy (HSVM), nNO or genetic testing for confirmation of PCD diagnosis.

For the detection rate of TEM among confirmed PCD patients estimate, studies were selected according to the following criteria: Cohorts of consecutive PCD patients, reporting of TEM findings and confirmation of PCD diagnosis with at least one more test such as High Speed Video Microscopy (HSVM), nNO or genetic testing.

Confirmation of PCD diagnosis was set to rely on TEM and at least one additional test, thus potentiating the comparison of the index test against other diagnostic test(s). Studies that confirmed PCD with TEM only were excluded. Studies with overlapping patient populations were cross-checked and only the study with the largest and most recent population was selected.

Data Extraction

The year of publication, name of author, study design, country of origin, study population, number of patients referred for PCD testing, number of patients confirmed as PCD and number of patients with a reported ultrastructural defect identified by TEM were recorded. The distribution of ultrastructural defects among the TEM positive patients and the age range of participating patients were recorded additionally where available. The data were extracted independently by two reviewers (PK, SIP) and consensus was reached for all data. Reporting of the included studies underwent quality assessment based on the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) criteria [367]. Not all of the MOOSE criteria were examined, as the assessment of reporting quality involved only the specific criteria that were relevant to the purposes of this study. Quality characteristics were assessed descriptively in order to detect any low quality evidence that could influence the results.

Analysis

In order to combine the data on PCD prevalence among patients referred for PCD diagnostic testing, we performed a meta-analysis of proportions using a random effects model. Meta-analysis of proportions allows the calculation of the pooled prevalence of PCD across studies containing binomial data with the numerator defined as the number of PCD patients identified by a combination of diagnostic tests (TEM and nNO or HSVM or genetic testing) and the denominator as the total number of consecutive referrals for PCD testing [368]. The random effects allow for each study to be assigned a weight which includes the within study variance and the between studies variance [369]. Furthermore, when only limited numbers of studies were available, and heterogeneity was $I^2 > 0\%$, we applied the Hartung, Knapp, Sidik

and Jonkman (HKSJ) approach which uses a Student T distribution, instead of a Normal distribution for the effects' estimates. This method applies an ad-hoc correction and yields more conservative results. [370]

The same method was used for the estimation of the pooled percentage of patients that are identified by TEM with the numerator defined as the number of PCD patients identified by abnormal TEM and the denominator defined as the number of PCD patients identified by a combination of diagnostic tests (TEM and nNO or HSVM or genetic testing). Lastly, the fraction of PCD patients with different ultrastructural defects (isolated ODA, combined ODA+IDA, MTD) was calculated.

Heterogeneity was assessed with the I^2 which describes the proportion of total variation in the effect estimate that results from the between-studies heterogeneity and ranges from 0-100% [371]. Subgroup analysis was planned a priori based on factors that a) could affect referral or diagnostic patterns, b) could lead to a different cohort of referrals or c) could detect a different spectrum of the examined disease. Such factors were the region specific referral patterns (in series of tests results from a specific country were excluded one at a time), the number of tests used to confirm PCD (exclusion of studies that performed only two tests), the sample size (exclusion of studies with number of referrals below the median number of referrals of all included studies) and whether an isolated IDA defect was considered diagnostic for PCD (exclusion of studies that reported isolated IDA defects as diagnostic). Isolated IDA defects as a diagnostic feature for PCD remains to date controversial because IDA are usually characterized by low contrast [352, 372]. In addition, none of the reported 32 genes, which harbor pathogenic mutations for PCD, have been found to affect only IDA [373]. All calculations were performed using STATA (Version 12, StataCorp, College Station, Texas) with the command *metaprop* for binomial data [368].

Results

Eligible Studies

A total of 2253 studies were retrieved through online search and 6 additional studies were identified through references' screening. Among the retrieved studies, 2097 were excluded based on title or abstract. Of the 161 studies that were assessed in full detail, 12 were review studies, 3 were guidelines or editorials, 56 were case-control or case-series studies, 33 did not include data on TEM results and 5 did not include data on PCD patients. A total of 36 studies were included in the qualitative synthesis. Of these, 25 studies were excluded at the last step and 11 were eventually included in the quantitative synthesis (Figure 1).

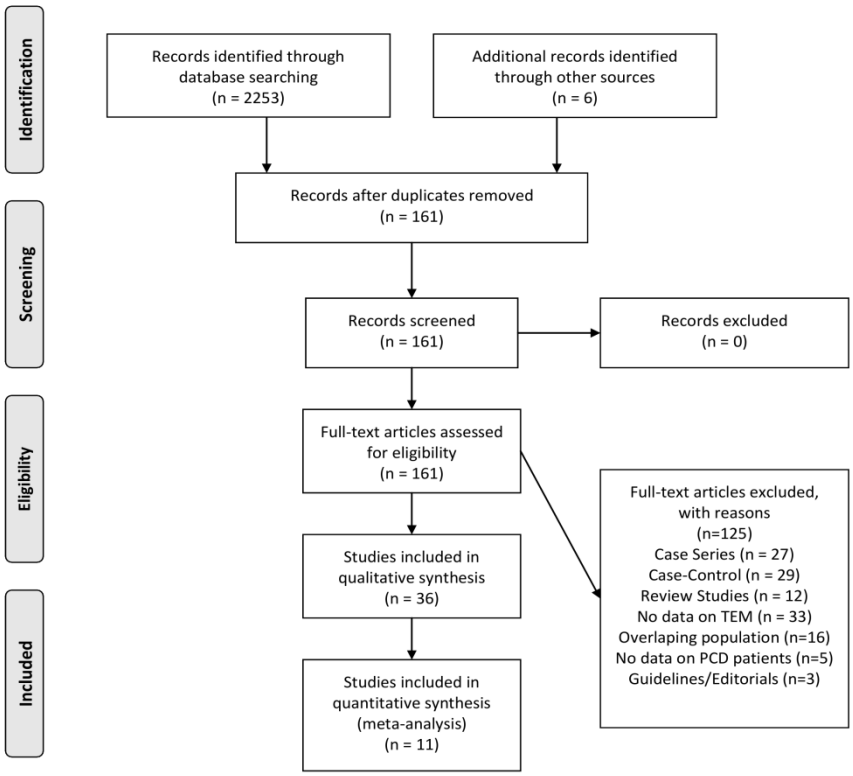


***Figure 6.1**: PRISMA diagram for the search strategy and selected studies.*

The 25 studies that were excluded at the last step prior to quantitative synthesis, as well as the main reason for their exclusion, are presented separately in Supplemental Table S1 (online). The majority of the studies (n=19) presented in Supplemental Table S1 (online), were excluded from further analysis because PCD diagnosis was established only based on the TEM results, although in some of these studies (n=8), additional tests were partly also performed.

Study Characteristics

Included studies and their descriptive characteristics are presented in Table 1. All the studies confirmed PCD diagnosis with at least one more test in addition to TEM. The vast majority of studies used HSVM as an additional test with the exception of Leigh et al which used genetic testing [362]. Furthermore, approximately 36% of the studies also included nNO in the diagnostic work-up. Three studies were performed in the United Kingdom (UK) [361, 374, 375] and three studies were performed in the Americas [362, 376, 377] while the remaining were performed in other European countries. The majority of the studies provided information about the number of consecutive referrals that underwent PCD diagnostic testing and were included in the meta-analysis regarding PCD prevalence among cohorts of respiratory referrals (n=2475) [361, 364, 374-378]. The study by Stannard et al. although focused on PCD patients with a positive TEM diagnosis, it also performed HSVM and reported patients diagnosed with abnormal beating and normal ultrastructure and as a result, it was also included in the PCD prevalence meta-analysis [361]. All the included studies, with the exception of Shappiro et al. [376] provided data on ultrastructural assessment for all the PCD diagnosed patients and were also included in the meta-analysis of the TEM detection rate (n= 728) [350, 361, 362, 364, 374, 375]. The most common ultrastructural findings in the assessed studies were isolated Outer Dynein Arm (ODA) defects and combined ODA and IDA (ODA+IDA) defects as well as tubular defects and normal ultrastructure (NU). A subset

of studies also reported isolated IDA defects [361, 374, 379]. Lastly, two of the included studies also reported a phenotype with lack of multiple cilia (acilia) in some patients [361, 363].

Table 6.1: Characteristics of included studies

#	Author, Year	Region	Study Design	Age at diagnosis (mean±sd)	Suspect patients	PCD patients	TEM confirmed	TEM result within PCD	Diagnostic Tests
1	Boon M. 2014 [17]	Belgium	All PCD cohort	9.9 (3.7- 23.4) [a]	-	206	138	ODA:82 ODA&IDA:8 CP:41 Acilia:6 MTD:1 NU:68	TEM, HSVM
2	Shappiro AJ. 2010 [22]	USA	Retrospective Referrals Cohort	12 (0.1-79) [b]	444	174	nr	nr	TEM, DNA, nNO
3	Djakow J. 2012 [25]	Czech Republic	All PCD cohort	9.1 (6-17) [c]	-	30	29	ODA:7 BOTH/DA:15 IDA:2 CP:2 RS:1 OTHER:2 NU:1	TEM, HSVM
4	Jackson C.L. 2015 [21]	United Kingdom	Retrospective Referrals Cohort	nr	368	72	57	ODA:19 ODA&IDA?:7 ODA&IDA:22 CP:3 MTD:6 NU:15	TEM, HSVM, nNO
5	Leigh M. 2013 [16]	North America	Prospective Referrals Cohort	nr	155	71	65	nr	TEM, DNA (n=6) [d]
6	Nauta F. 2011 [26]	Netherlands	All PCD Cohort	3.8 (0.1-18) [a]	-	63	45	nr	TEM, HSVM

7	Shoemark A. 2011 [20]	United Kingdom	Retrospective Referrals Cohort	PCD: 10 (0.1-77) [b]	1182	275	242	ODA:105 ODA&IDA:57 IDA:30 MTD:28 RS:22 NU:33	TEM, HSVM, nNO
8	Stannard W.A. 2010 [15]	United Kingdom	Retrospective Referrals Cohort	8.0	371	74	72	ODA=18 ODA&IDA=23 IDA=14 CP=3 RS=4 MTD=6 DO=2 Acilia=2 NU=2	TEM, HSVM,
9	Olm M.A. 2011 [23]	Brazil	Retrospective Referrals Cohort	1-19 [c]	24	12	11	ODA=3 RS=5 MTD=3 NU=1	TEM,HSVM
10	Pifferi M. 2011 [24]	Italy	Retrospective Referrals Cohort	Children 10.7±2.9 Adults 32.0±9.2	86	41	30	nr	TEM, HSVM
11	Yiallouros P.K. 2015 [18]	Cyprus	Retrospective Referrals Cohort	13.9 (0.1- 58.4) [b]	76	30	19	ODA=3 ODA&IDA=13 CP=3 NU:11	TEM, HSVM, nNO

[a] Median (IQR)
[b] Median (range)
[c] Mean (range)
[d] DNA testing was applied to patients with normal ultrastructure (n=6)
DO: Disorientation NU: Normal ultrastructure

Assessment of Reporting Quality

Quality assessment results for the included studies are presented in Figure 2. Overall, the analyzed studies were characterized by good methodology and all of them appropriately described the diagnostic tests performed. On the other hand, there were some studies that did not report the recruitment period or did not describe in detail the patients' characteristics. Only a few of the studies discussed potential study limitations although this was expected, as most of them were of a descriptive nature. Finally, a small number of studies [362, 374, 375, 377, 378] reported some efforts to reduce bias during the evaluation of diagnostic tests such as blinded assessment or assessment by more than one evaluator. The results of reporting quality assessments for each study are presented in detail in Supplemental table S2 (online).

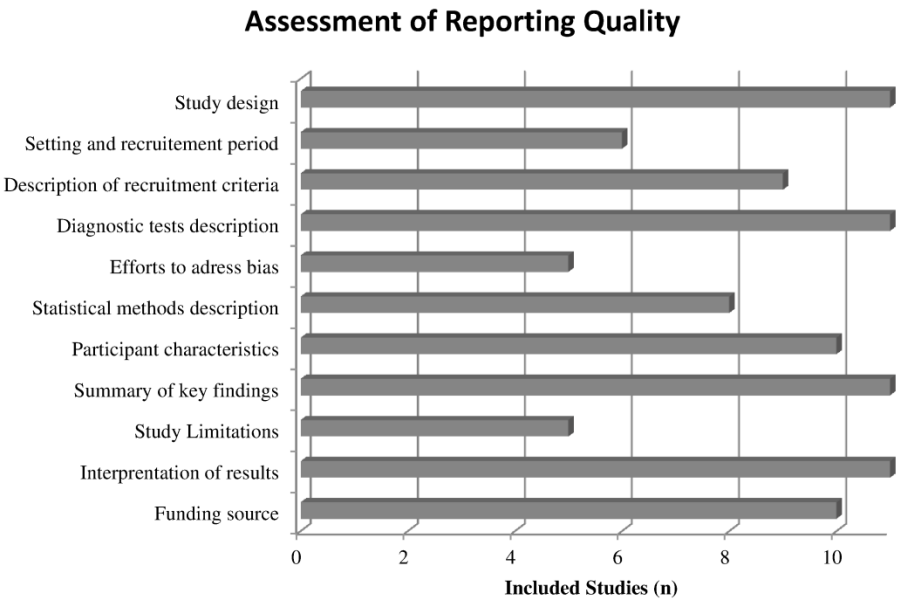


Figure 6.2: Quality assessment results for the included studies

Data Synthesis

The pooled prevalence of newly diagnosed PCD patients in cohorts of consecutive referrals of suspect cases was 32% (95%CI: 25% - 39%, I^2 = 92%) (Figure 3). The ad-hoc correction using Hartung, Knapp, Sidik and Jonkman (HKSJ) resulted in the same pooled prevalence

estimate but wider confidence intervals (32%, 95%CI: 20% - 44%, I^2 = 92%). A series of subgroup analyses were performed by excluding each time studies from each individual country, studies that performed only two tests for PCD confirmation, and studies with low sample size. Overall, subgroup analyses did not yield significant differences from the original analysis (Supplemental Figures S1 and S2 (online)) but the heterogeneity in the effect estimate was explained by the exclusion of UK studies (prevalence: 41%, 95%CI: 37% - 45%, I^2 = 0%) (Supplemental Figure S3 (online)).

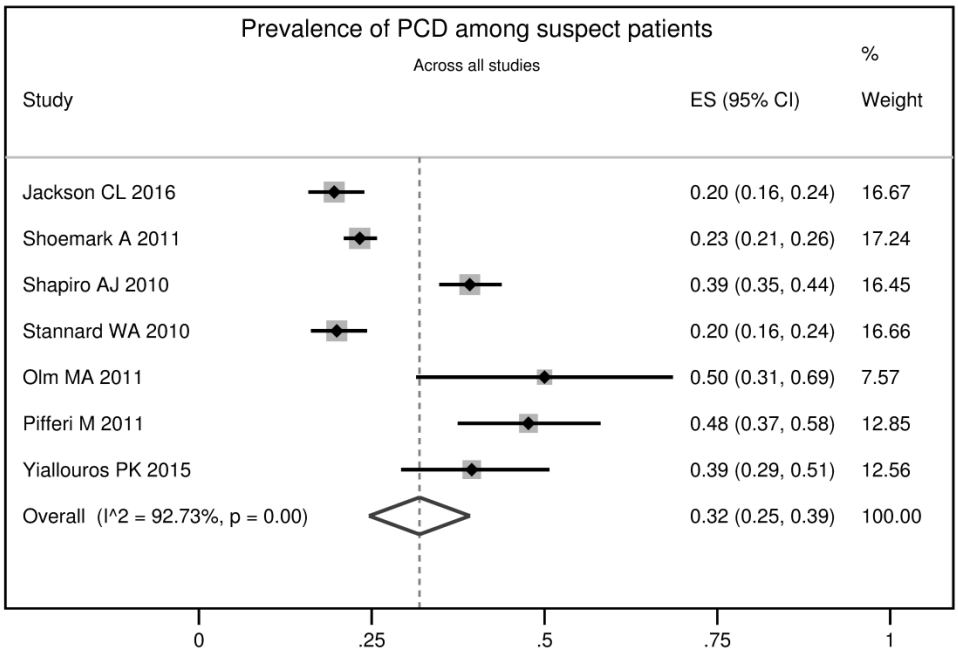


Figure 6.3: *Forest plot of the prevalence of PCD among cohorts of suspect patients. Forest plot of the proportion of referred patients for PCD testing that have eventually PCD confirmed.*

The detection rate of TEM in PCD diagnosed patients was 83% (95% CI: 75% - 90%, I^2= 90%), (Figure 4). The detection rate of TEM in PCD after the ad-hoc HKSJ correction was also 83% (95% CI: 74% - 92%, I^2= 90%).

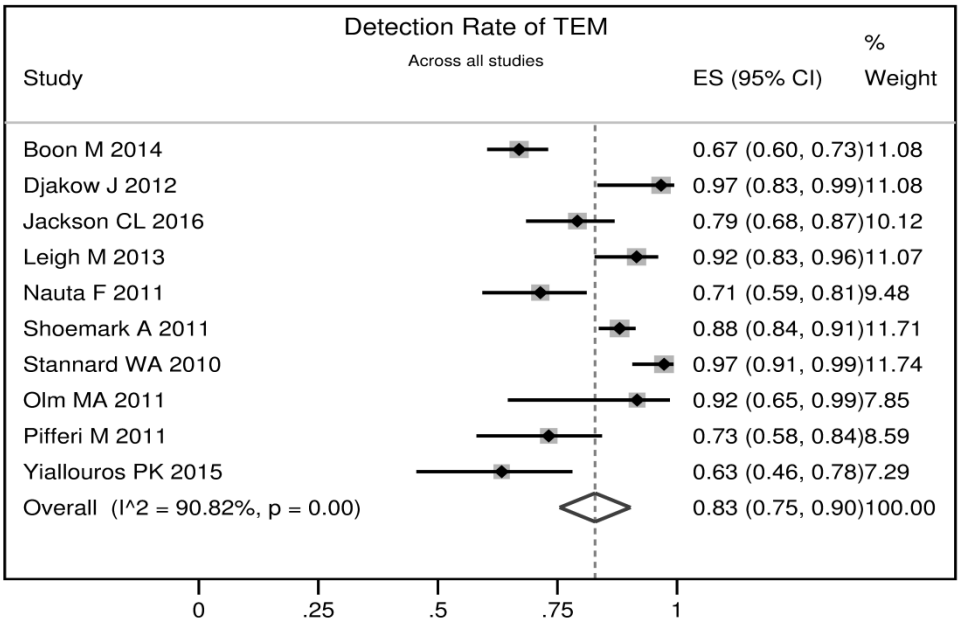


Figure 6.4: Forest plot of the detection rate of TEM across all studies. Forest plot of the detection rate of TEM in cohorts of patients that have PCD confirmed with a combination of diagnostic tests (all included studies)

Subgroup analyses were also performed for this estimate by excluding each time, studies that performed only two tests for PCD confirmation, studies with low sample size and studies which reported isolated IDA defects. The subgroup analyses, with the exception of one, did not demonstrate significant differences compared to the original analysis (Supplemental Figures S4 and S5 (online)). The subgroup analysis that included studies that did not report isolated IDA defects resulted in a marked reduction in the detection rate of TEM and explained an important fraction of the observed heterogeneity (detection rate: 74%, 95% CI: 66% - 83%, I^2 = 66%), (Figure 5). Similarly as before, the HKSJ ad-hoc correction resulted in wider confidence intervals (detection rate: 74%, 95% CI: 61% - 87%, I^2 = 66%).

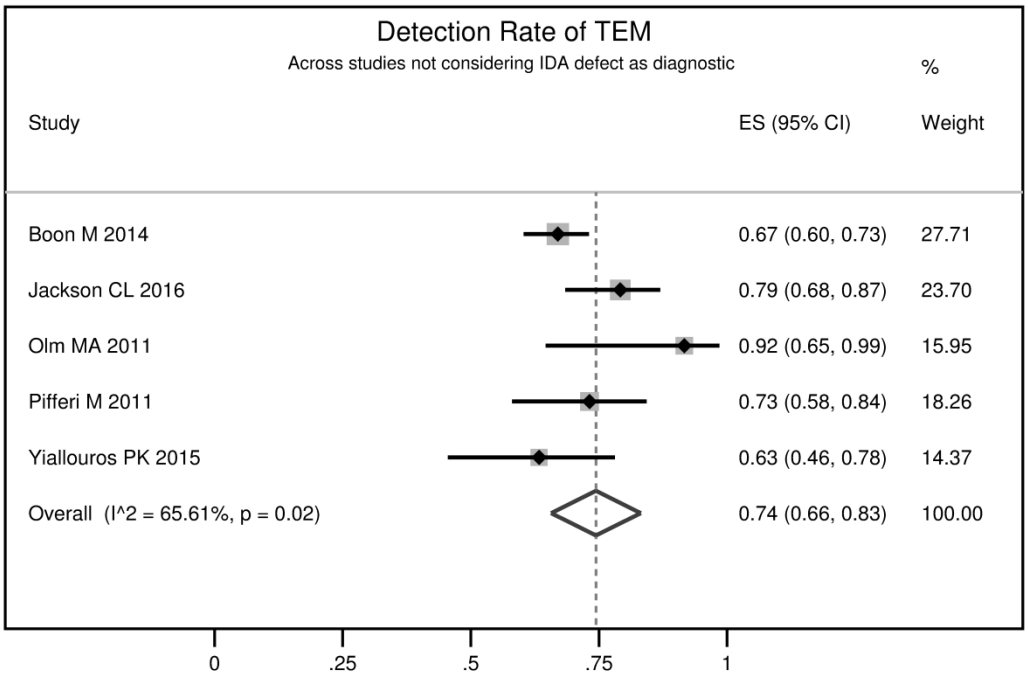


Figure 6.5: Forest plot of the detection rate of TEM across studies excluding IDA. Forest Plot of the detection rate of TEM across the included studies that did not report an isolated IDA defect

Among PCD patients, an isolated ODA defect was identified in 28% (95% CI: 19%-36%, I^2 = 80%) (Supplemental Figure S6 (online)) while a combined ODA+IDA defect was identified in 26% (95%: 14%-39%, I^2=95%) (Supplemental Figure S7 (online)). HKSJ ad-hoc correction resulted in an estimate of 28% (95%CI: 18%-38%) for the fraction of patients with isolated ODA and in an estimate of 26% (95%CI: 9%-43%) for the patients with a combined ODA+IDA defect. The fraction of patients identified with a tubular defect was 10% (95%CI: 3% - 18%, I^2 = 93%, HKSJ ad-hoc correction: 10%, 95%CI 0% -25%) (Supplemental Figure S8 (online)).

Discussion

In this systematic review and meta-analysis, we found the pooled prevalence of PCD diagnosis in cohorts of consecutive referrals of suspect cases for PCD diagnostic testing to be

32%. This finding suggests that one third of the patients suspected for PCD, are indeed affected by the disease. The high heterogeneity observed in the estimate can be at least partly explained by the inhomogeneity in diagnostic protocols and referral patterns between the different studies. Some studies used different number or combination of diagnostic tests for eliciting PCD diagnosis. The study by Shappiro et al. used TEM along with genetics and nNO for confirmation of PCD diagnosis [376] while other studies used TEM and HSVM [361, 363, 377, 379, 380] or TEM, HSVM and nNO [364, 374, 375, 378]. Different referral patterns between the included studies may have also influenced the heterogeneity in the final estimate. Although the referral of patients for PCD diagnostic testing should be based on combinations of classical PCD features [349], to date the decision to refer patients does not result from a suspect manifestations scoring system but rather from the clinicians' awareness of PCD and personal experience with the disease. The lack of such a scoring system is reflected in differences in the referral patterns across different countries or different centers [364, 381, 382]. The recent publication by Leigh et al. regarding the association of specific clinical features with the likelihood of PCD [383] and the development of PICADAR clinical scoring tool [384] constitute the first steps towards the introduction of a universal clinical scoring system and referral algorithm for PCD in the primary care clinical setting. The performance of future referral algorithms can be compared with the current estimate of the prevalence of PCD among referrals reported here, which is essentially compromised by the variability in the referral patterns of the different centers.

Furthermore, among consecutive diagnosed PCD patients that underwent ultrastructural assessment, we calculated the detection rate of TEM to be 83%. In all the included studies, positive PCD diagnosis was based on a combination of at least two or three diagnostic tests and ciliary ultrastructural assessment using TEM was part of the diagnostic work-up. The estimated detection rate means that approximately 17% of PCD patients do not exhibit ciliary

abnormalities on TEM analysis. This analysis also displayed significant heterogeneity across studies. In subgroup analysis, after exclusion of studies which reported isolated IDA defects as an abnormal diagnostic TEM finding, we found the detection rate estimate to be 74% and the heterogeneity to be markedly reduced. The resulting pooled estimate suggests that 26% of PCD patients do not exhibit abnormal ultrastructure and this estimate is much closer to the empirically quoted estimate of 30% [365]. Among the confirmed PCD patients, 28% were characterized by an isolated ODA defect, 26% were characterized by a combined ODA+IDA defect and 10% by tubular defects. However, among the remaining two categories of ultrastructural defects reported in the individual studies, Central Pair (CP) and Radial Spoke (RS) defects, it is possible that some of the patients also have microtubular disorganization (MTD) as some genetic mutations result in both or either CP and MTD defects through the disruption of radial spokes [385, 386]. As a result, the estimated 10% fraction for tubular defects could be an underestimation, probably affected by the underlying categorisation of certain TEM defects such as CP, RS and MTD within the included studies.

The results of this analysis suggest that PCD diagnosis cannot rely only on TEM examination. As a test of characterization of morphologic features, TEM holds substantial subjectivity and may be influenced by the overall quality of the obtained sample [387]. The other routinely used PCD diagnostic tests, the measurement of nNO and HSVM, have both been reported to perform better than TEM. More specifically, a recent meta-analysis of nNO measurements in PCD patients has demonstrated a pooled sensitivity of 93% to 95% depending on vellum closure status during testing [388] while a number of studies have reported high sensitivity values for HSVM ranging from 89% [389] to 100% [375]. Although our findings suggest that PCD diagnosis should not rely only on TEM, performance of this test may still be beneficial since determination of an ultrastructural defect or confirmation of its absence may guide genetic testing wherever this is available. This is particularly important

for centers that lack access to whole exome or whole genome sequencing and rely on genotyping of specific PCD genes. In this setting, prior identification of ultrastructural defects allows prioritization of which genes to be sequenced based on known genotype-TEM findings correlations [390]. In addition, recent studies have suggested associations between the ultrastructural phenotype and severity of clinical features and disease progression [363, 391] thus highlighting the potential of TEM analysis to facilitate the identification of clinically significant phenotypic subgroups among PCD patients. Furthermore, a recent study by Knowles et al. has highlighted the milder clinical phenotype in PCD individuals with bi-allelic mutations in *RSPH1* and mainly CP defects in a small subset of cross-sections [392].

IDA imaging with TEM is more difficult compared to ODA due to the low contrast of IDA coupled with the frequent presence of non-specific biological or technical artifacts [352, 372]. Furthermore, as ODA and IDA are multiprotein complexes of different axonemal dynein polypeptides which include heavy, intermediate and light chain polypeptides, the composition and variability of which, may affect the visualization of these structures under TEM. In more detail, it has been shown that IDA composition is more diverse compared to ODA [393] as well as that the periodicity of IDA is higher compared to the periodicity of ODA along the axoneme [394]. As a result specific tools have been proposed in order to allow for the clearer visualization of IDA (and ODA) in electron micrographs such as averaged TEM pictures [395] and Markham rotation [396] although none has received widespread application. Overall, the presence of an isolated IDA defect in PCD remains controversial as, up to now, none of the reported 32 genes which harbor pathogenic mutations for PCD, has been found to affect only IDA [373]. Several genes affect both ODA and IDA [373], while CCDC39 and CCDC40, which have been recently shown to cause loss of IDA, cause as well disruption of the axonemal organization [397] and more severe disease [391]. As a result, the analysis which

excluded studies reporting isolated IDA defects provides a more reliable estimate of the TEM detection rate.

This is to our knowledge, the first study that summarizes the evidence from cohorts of consecutive referrals and informs about the prevalence of PCD among these cohorts and the detection rate of TEM among these patients. This meta-analysis benefited from including data from a large number of suspect cases referrals and PCD cases from many centers. However, as most of the included studies were retrospective, the possibility of selection or misclassification bias cannot be ruled out. Furthermore, it is acknowledged that, although the number of studies reporting TEM findings in PCD patients is quite extensive, the studies that were finally included in this meta-analysis are only eleven. However, the goal of this systematic review was to estimate the detection rate of TEM in PCD confirmed cases, thus the included studies should have had both criteria, consecutive referrals and the PCD diagnosis confirmed by a combination of diagnostic tests and not only by TEM. This design may have led to a smaller number of included studies but it enabled the calculation of a more reliable estimate for the detection rate of TEM in PCD. In 8 of the 10 included cohorts it was clearly stated that all PCD patients were diagnosed by at least one more diagnostic test in addition to TEM. In the remainder 2 cohorts (Leigh et al. 2013 and Shoemark et al. 2011), a fraction of their patients were not diagnosed by another diagnostic test (in addition to TEM). This represents a limitation of the analysis probably leading to a slight overestimation of the diagnostic performance of TEM. The TEM detection rate in the sensitivity analysis which excluded cohorts reporting isolated IDA defects as a diagnostic finding is not affected, as these two studies [362, 374] were not included in the analysis.

In this analysis, for the vast majority of PCD patients who had normal ultrastructure and were missed by TEM, the genetic defect is not specified. This represents an important limiting factor as we do not know if all the responsible genetic defects known to date to cause PCD

and normal cilia structure were represented in this subgroup. In a recent review of different PCD populations, the frequency of genetic defects, which are known to cause PCD but not detectable ultrastructure changes by TEM, was found to be approximately 30% [398]. This percentage is slightly higher but close to the 26% reported here. Additional studies in large cohorts of PCD patients, reporting diagnostic tests results and responsible genetic defects are needed to inform about the precise diagnostic accuracy of TEM.

In summary, among cohorts of consecutive referrals of suspect cases for PCD testing, approximately one third are eventually confirmed as PCD patients. Among PCD cases that underwent TEM studies, a significant percentage, at least as high as 26%, were not identified by TEM. This limitation of TEM should be taken into account during the development of a universal and efficacious diagnostic algorithm for PCD.

Chapter 6: Supplemental Material

Supplementary table 6.1: Characteristics of studies excluded at the last step prior the quantitative analysis

#	Author, Year	Country	Study Design	Age range (mean±sd)	Suspect patients	PCD patient	TEM confirmed	Diagnostic Tests	Exclusion Reason
1	Escudier E. 2002 [399]	France	Retrospective Referrals cohort	nr	40	26	26	TEM,LM	Diagnosis based only on TEM
2	Olin J.T. 2011 [400]	USA, Canada	Prospective Referrals Cohort	nr	448	155	155	TEM	Diagnosis based only on TEM
3	Kawakami M. 1996 [401]	Japan	Retrospective PCD cohort	38 (17-72)	-	45	45	TEM, LM in few	Diagnosis based only on TEM
4	Papon J.F. 2012 [402]	France	Retrospective Referrals Cohort	32.5 ± 15.8	34	10	10	TEM, HSVM, nNO	Diagnosis based only on TEM
5	Coste A. 1997 [403]	France	Retrospective Referrals Cohort	nr	106	6	6	TEM, LM	Diagnosis based only on TEM, poor TEM data
6	Hosie P.H. 2014 [404]	Australia	Retrospective Referrals Cohort	PCD: 6.4 (0.1-18.2)	1037	84	81	TEM, LM	Ciliary motility assessed via photometer only
7	Sirvanci S. 2008 [405]	Turkey	Retrospective Referrals cohort	nr	34	10	10	TEM	Diagnosis based only on TEM
8	Noone P.G. 2004 [406]	North America	Retrospective Referrals Cohort	Adults PCD:36(19-73)[b] Children PCD: 8 (1-17)[b]	94	78	78	TEM, DNA (no DNAH11), nNO for some	Diagnosis based on TEM and DNA (no DNAH11)

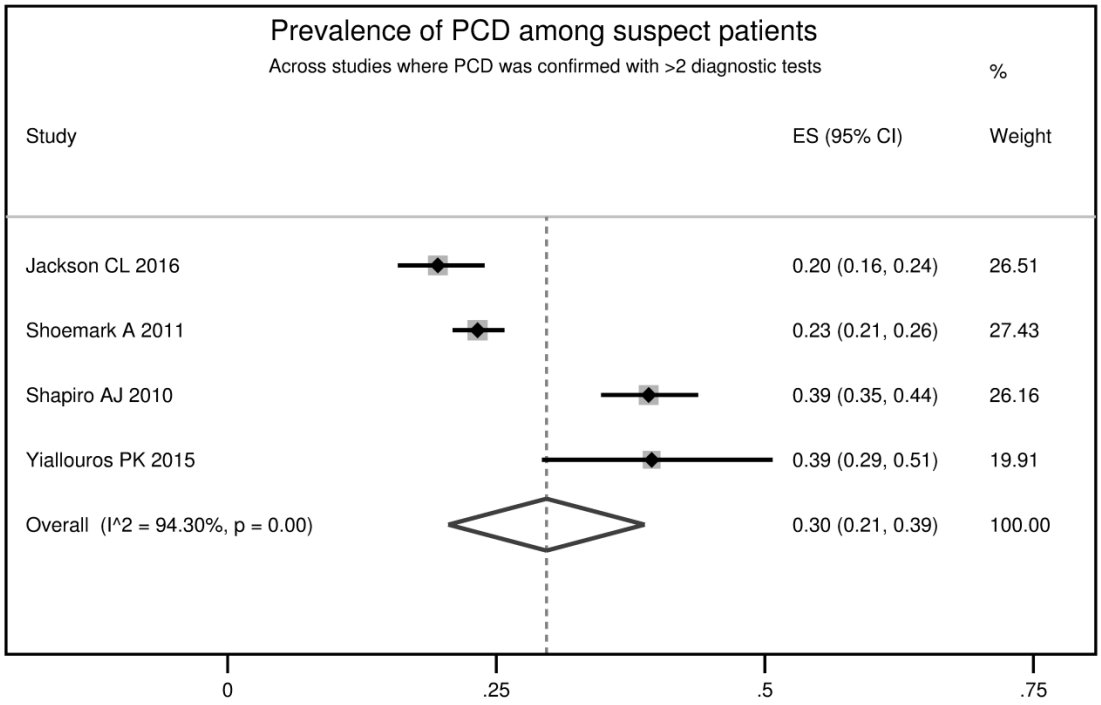
9	Chin G.Y. 2002 [407]	North America	Retrospective Referrals Cohort	Range: 2 - 48 years	118	73	73	TEM	Diagnosis based only on TEM
10	Beydon N. 2015 [408]	France	Prospective Referrals Cohort	11.4 (7,13.9) [b]	142	49	44	TEM, HSVM, DNA, nNO	Study population included suspect and already diagnosed PCD patients
11	Carda C. 2004 [409]	Spain	Retrospective Referrals Cohort	nr	200	14	10	TEM, 99 m TC	Diagnosis based on Kartagener Syndrome
12	Papon J.F. 2010 [410]	France	Retrospective Referrals Cohort	nr	820	245	245	TEM	Diagnosis based only on TEM
13	Rubio M.T.R. 2011 [411]	Spain	Prospective Referrals Cohort	nr	79	4	1	TEM,HSVM	Low number of events
14	Simoneau T. 2013 [412]	USA	Retrospective Referrals Cohort	6.3 (0.5-29)[a]	187	4	4	TEM	Diagnosis based only on TEM, Low number of events
15	Corbelli R. 2004 [413]	Switzerland	Retrospective Referrals Cohort	PCD: 12±2 Non-PCD: 10.5±1.8	34	17	17	TEM,LM	Diagnosis based only on TEM
16	Davis S.D. 2015 [414]	North America	Retrospective PCD cohort	nr	-	118	118	TEM, DNA	Diagnosis based only on TEM
17	Chilvers M.A. 2003 [415]	United Kingdom	Retrospective PCD cohort	4.7 (0.1-14)[a]	-	56	56	TEM,HSVM	Diagnosis based only on TEM
18		Korea	Retrospective	nr	17	4	4	TEM	Diagnosis based only on

	Shin S.A. 2006 [416]		Referrals Cohort						TEM
19	Theegarten D. 2011 [417]	Germany	Retrospective Referrals Cohort	PCD: 7.7 (0.1-50)[b]	742	134	134	TEM	Diagnosis based only on TEM
20	Chi J. G. 1993 [418]	Korea	Retrospective Referrals Cohort	PCD: 11 (5–15)[a]	80	17	17	TEM	Diagnosis based only on TEM
21	Bent J.P. 1997 [419]	USA	Retrospective Referrals Cohort	nr	20	3	3	TEM,LM	Ciliary motility assessed via LM, Low number of events
22	Pizzi S. 2003 [420]	Italy	Retrospective Referrals Cohort	7.1	34	2	2	TEM	Diagnosis based only on TEM, Low number of events
23	Busquets R.M. 2013 [421]	Spain	Retrospective Referrals Cohort	3.6 (0.1-19)	63	35	35	TEM	Diagnosis based only on TEM
24	Plesec T. 2008 [422]	USA	Retrospective Referrals Cohort	19.6 (1-54)[a]	150	21	21	TEM, LM	Diagnosis based only on TEM
25	Daniels M.L.A. 2011 [423]	USA	Retrospective Referrals Cohort	nr	551	206	206	TEM	Diagnosis based only on TEM

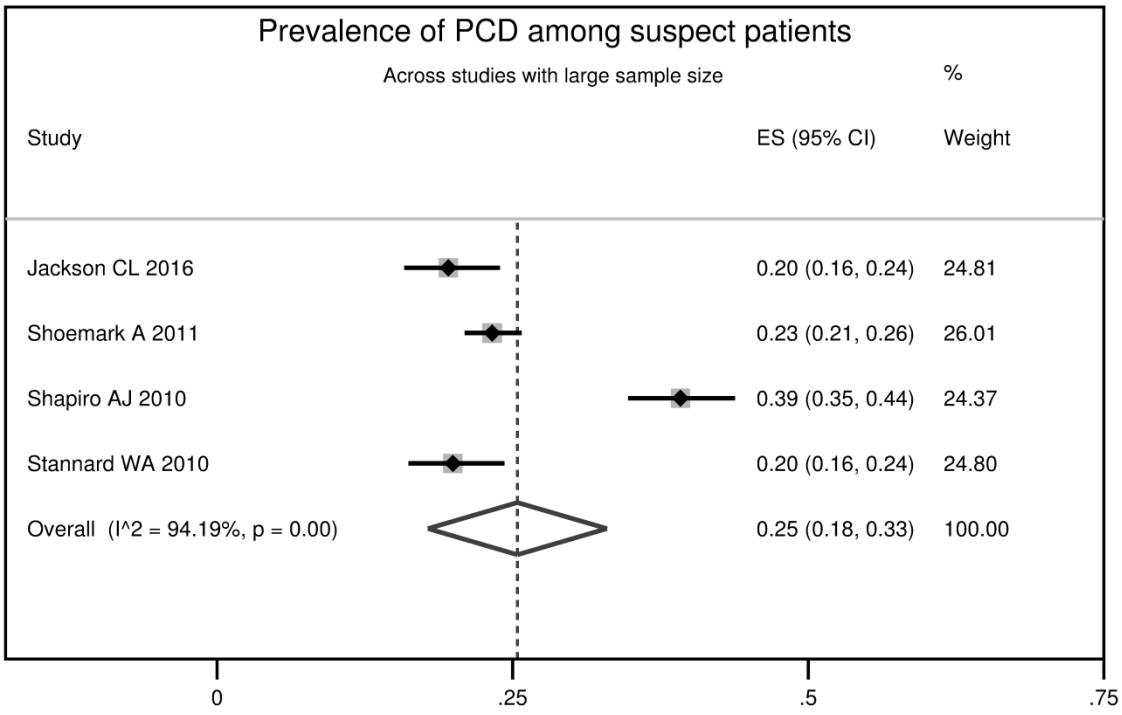
[a]Mean (range)
[b]Median (range)
nr: non reported
ODA: Outer Dynein Arm Defect, IDA: Inner Dynein Arm Defect, ODA&IDA: Outer and Inner Dynein Arm Defect, CP: Central Pair Defect, RS: Radia Spoke defect,TD: Transposition defect, DO: Disorientation, NU: Normal ultrastructure, TEM: Transmission Electron Microscopy, HSVM: High Speed Video Microscopy, LM: Light Microscopy, DNA: Genetic Testing, 99 m TC: 99m TC-labelled serum albumin

Supplementary table 6.2: Assessment of Reporting Quality

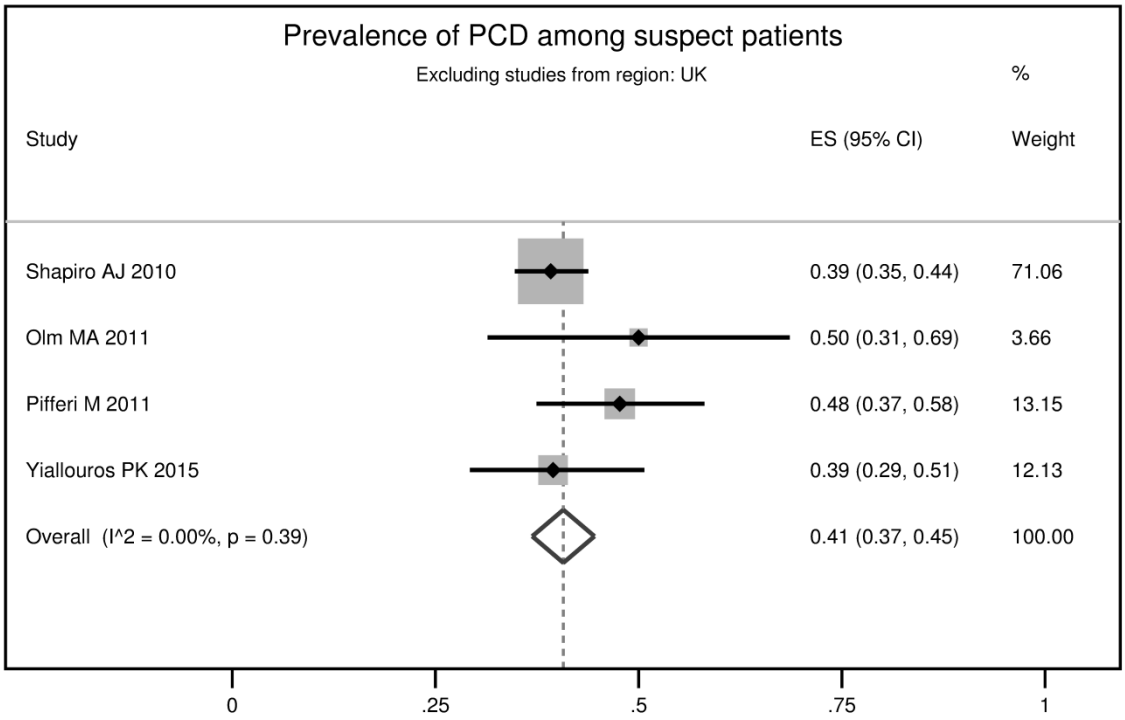
Variable	Boon M. 2014 [17]	Shapiro AJ. 2010 [22]	Djakow J. 2012 [25]	Jackson C.L. 2015 [21]	Leigh M. 2013 [16]	Nauta F. 2011 [26]	Shoemark A. 2011 [20]	Stannard W.A. 2010 [15]	Olm M.A. 2011 [23]	Pifferi M. 2011 [24]	Yiallouros P.K. 2015 [18]
Study Design	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Setting and recruitment period		✓		✓	✓		✓	✓			✓
Diagnostic tests description	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Efforts to address bias				✓	✓		✓		✓	✓	
Statistical methods description	✓		✓	✓	✓			✓	✓	✓	✓
Participants characteristics	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓
Summary of key findings	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Study limitations	✓		✓	✓					✓		✓
Interpretation of results	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Funding source	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓



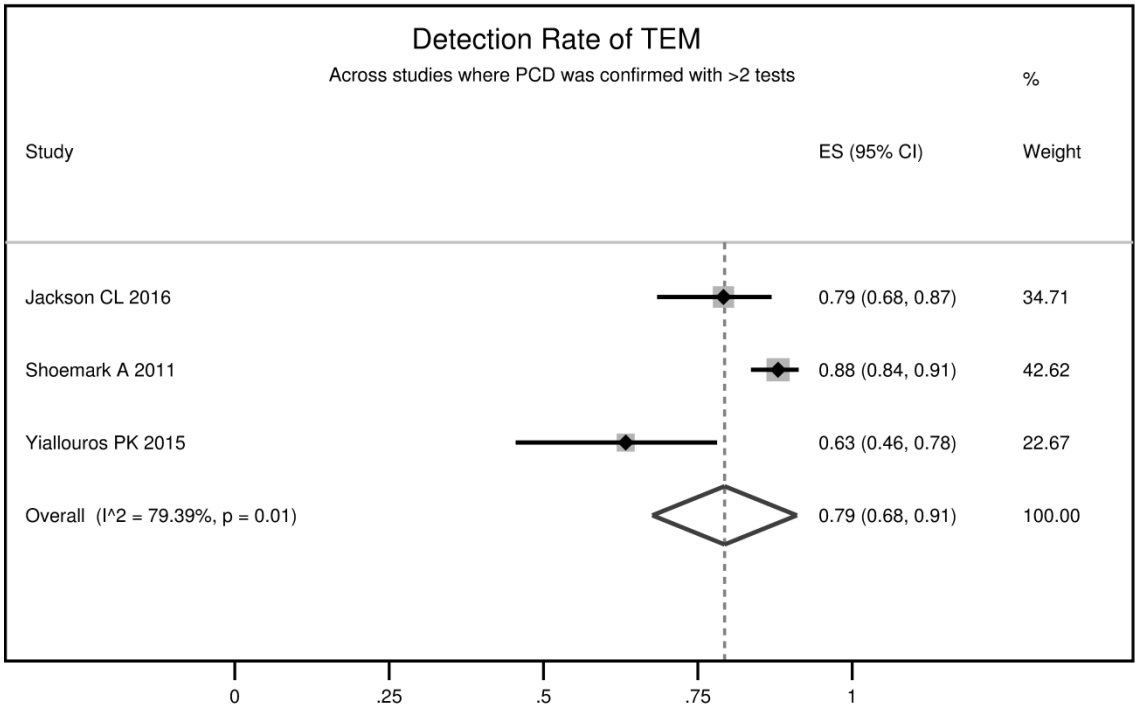
Supplemental Figure 6.1: Forest plot of the prevalence of PCD among cohorts of suspect patients (across studies where PCD was confirmed with >2 diagnostic tests). Forest plot of the proportion of referred patients for PCD testing that have eventually PCD confirmed.



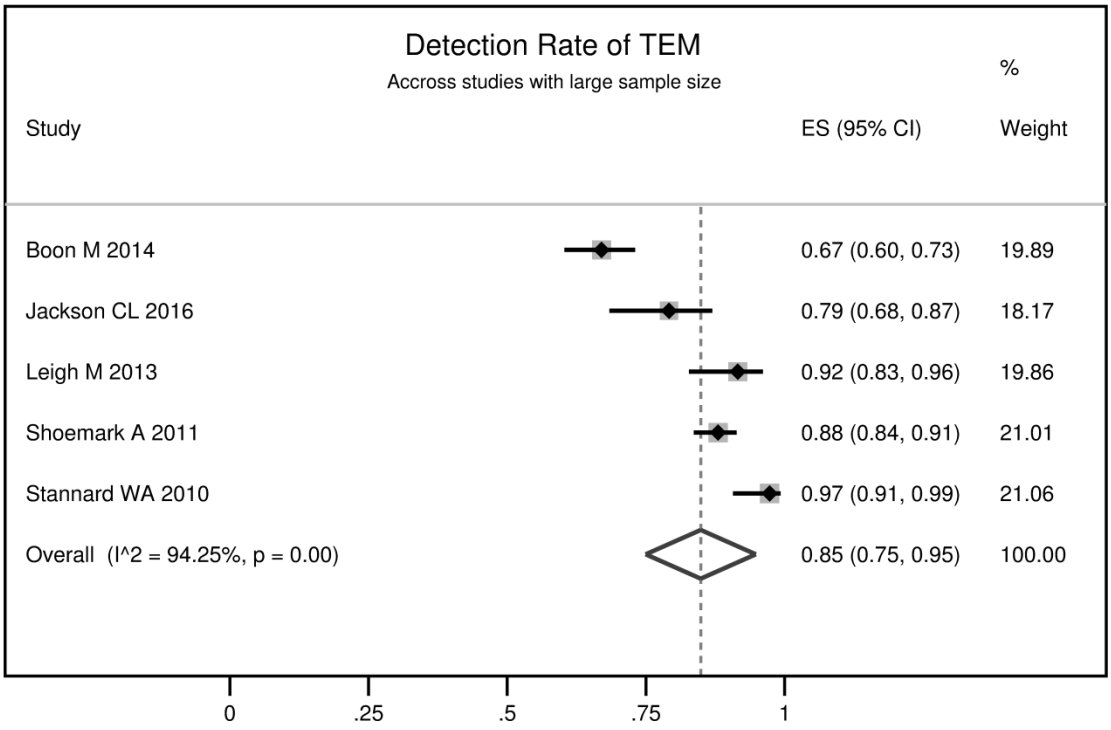
Supplemental Figure 6.2: Forest plot of the prevalence of PCD among cohorts of suspect patients (across studies with large sample size). Forest plot of the proportion of referred patients for PCD testing that have eventually PCD confirmed.



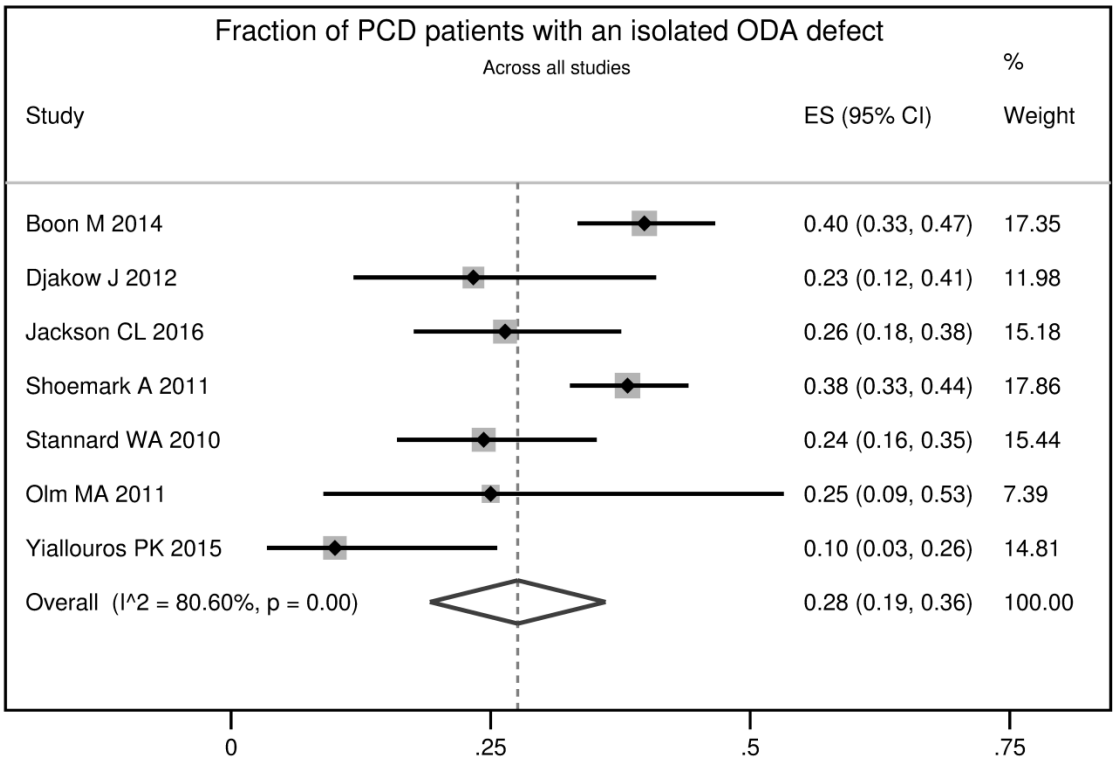
Supplemental Figure 6.3: Forest plot of the prevalence of PCD among cohorts of suspect patients (excluding UK region studies). Forest plot of the proportion of referred patients for PCD testing that have eventually PCD confirmed.



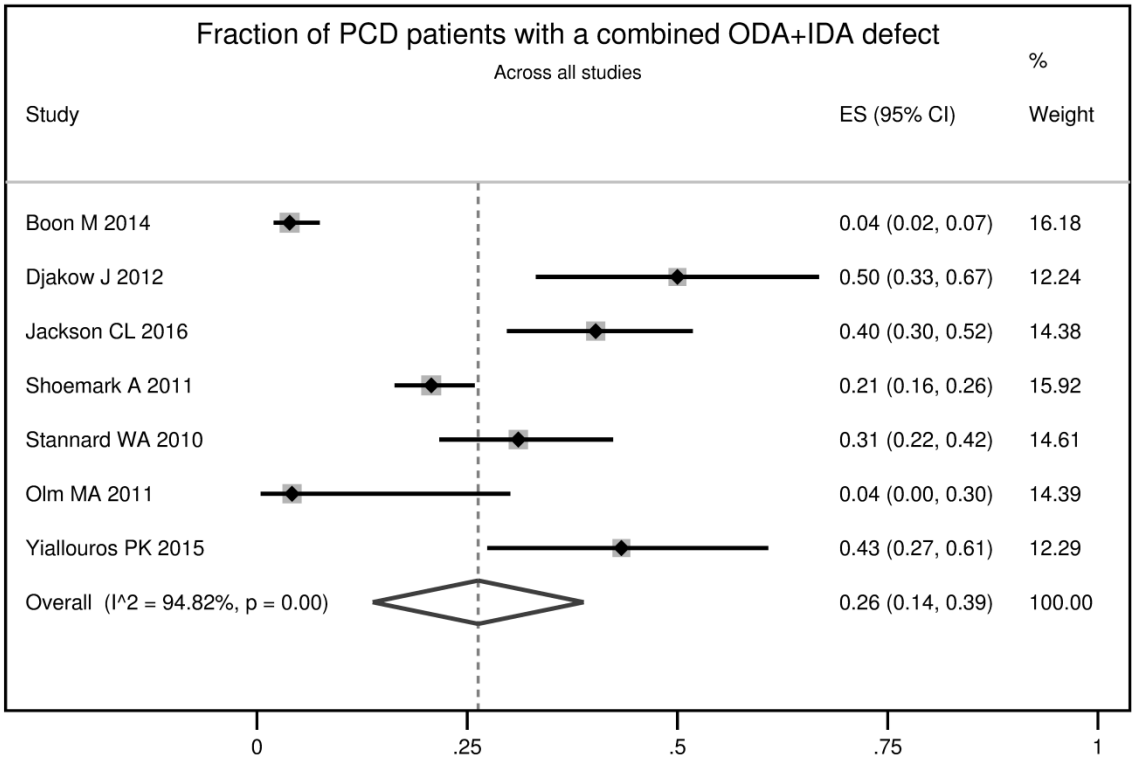
Supplemental Figure 6.4: Forest plot of the detection rate of TEM (across studies where PCD was confirmed with >2 diagnostic tests). Forest plot of the detection rate of TEM in cohorts of patients that have PCD confirmed with a combination of diagnostic tests.



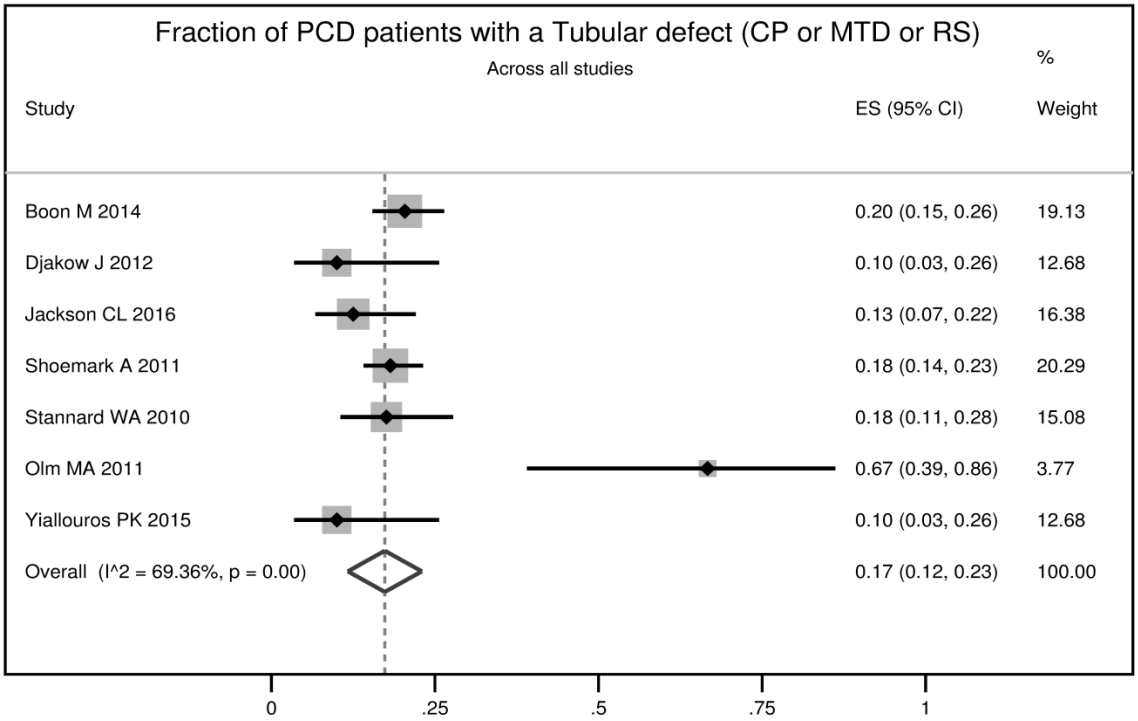
Supplemental Figure 6.5: Forest plot of the detection rate of TEM (across studies with large sample size). Forest plot of the detection rate of TEM in cohorts of patients that have PCD confirmed with a combination of diagnostic tests.



Supplemental Figure 6.6: Forest plot of the fraction of PCD patients with an isolated ODA defect (across all studies). Forest plot of the fraction of PCD patients with an isolated ODA defect in cohorts of patients that have PCD confirmed with a combination of diagnostic tests.



Supplemental Figure 6.7: Forest plot of the fraction of PCD patients with a combined ODA+IDA defect (across all studies). Forest plot of the fraction of PCD patients with a combined ODA+IDA defect in cohorts of patients that have PCD confirmed with a combination of diagnostic tests.



Supplemental Figure 6.8: Forest plot of the fraction of PCD patients with a tubular defect (across all studies). Forest plot of the fraction of PCD patients with a tubular defect in cohorts of patients that have PCD confirmed with a combination of diagnostic tests.

Chapter 7: Cost-effectiveness Analysis of Three Algorithms for Diagnosing Primary Ciliary Dyskinesia

Abstract

Background

Primary Ciliary Dyskinesia (PCD) diagnosis relies on a combination of tests which may include (a) nasal Nitric Oxide (nNO), (b) High Speed Video Microscopy (HSVM) and (c) Transmission Electron Microscopy (TEM). There is variability in the availability of these tests and lack of universal agreement whether diagnostic tests should be performed in sequence or in parallel. We assessed three different combinations of tests for PCD diagnosis and estimated net sensitivity and specificity as well as cost-effectiveness (CE) and incremental cost-effectiveness (ICE) ratios.

Methods

A hypothetical initial population of 1000 referrals (expected 320 PCD patients) was followed through a probabilistic decision analysis model which was created to assess the CE of three diagnostic algorithms (a) nNO+TEM in sequence, (b) nNO+HSVM in sequence and (c) nNO/HSVM in parallel followed, in cases with conflicting results, by confirmatory TEM (nNO/HSVM+TEM). Model inputs were obtained from published meta-analyses and large studies. Number of PCD patients identified, CE and ICE ratios were calculated using Monte Carlo analysis in ANALYTICA.

Results

Out of 320 PCD patients, 311 were identified by nNO/HSVM+TEM, 274 with nNO+HSVM and 198 with nNO+TEM. The nNO/HSVM+TEM had the higher mean cost (€97K) followed by nNO+TEM (€56K) and nNO+HSVM (€39K). The nNO+HSVM algorithm dominated the nNO+TEM algorithm (less costly and more effective). The ICE ratio for nNO/HSVM+EM was €1.6K per additional PCD patient identified.

Conclusion:

The diagnostic algorithm (nNO/HSVM+TEM) with parallel testing outperforms algorithms with tests in sequence. Decision analysis methods can facilitate the discussion towards the development of the most efficacious diagnostic algorithm for PCD.

Introduction

Primary Ciliary Dyskinesia (PCD) is a rare, genetically heterogeneous disorder that affects one in approximately 15 000 live births [424]. It results from the dysfunction of respiratory motile cilia, and the consequent compromise of mucociliary clearance which is a critical innate respiratory defense mechanism [425]. PCD is characterized by chronic sinopulmonary symptoms and development of bronchiectasis, recurrent otitis, male infertility and situs inversus [426]. Defective components of the ciliary axoneme (e.g. dynein arms) as well as dysfunctional regulatory or transport proteins have been implicated in the etiology of PCD and to date 32 genes have been found to be causative for PCD [427]. This genetic heterogeneity translates into a wide spectrum of ciliary structural and beating abnormalities and a diverse phenotype.

Diagnostic testing for PCD relies on a combination of tests which primarily include nasal Nitric Oxide (nNO) [428], High Speed Video Microscopy (HSVM) [429, 430] and Transmission Electron Microscopy (TEM) [431, 432] as no single test has been shown to have 100% sensitivity and specificity [433]. “Because of this and because many centers lack either the needed equipment or the expertise to perform all required tests, some of which are quite laborious and time consuming, different diagnostic algorithms for diagnosis of PCD have been adopted by diagnostic centers across the world. [434]. Recently, nNO has been proposed as the screening test of choice in cohorts of patients with PCD-suspect manifestations due to its high ability to discriminate between PCD and non-PCD subjects [428, 435]. Although the cost of a (validated) chemiluminescence NO analyser is quite high (approximately €40000 per piece), the recent development of handheld and cheaper electrochemical NO analysers [436] and publication of relevant technical guidelines by the American Thoracic Society (ATS) and the European Respiratory Society (ERS) [437] may further enhance the potential of nNO measurement to be used as a screening test in the

clinical setting and especially in countries with limited resources or in areas that lack or are distant from PCD-specialist centers [438]. However, the use of a non-perfect screening test such as nNO in isolation may allow for some PCD patients with false negative results to be missed entirely or some non-PCD patients with false-positive results to undergo further diagnostic tests. For this reason, the diagnostic algorithm described as part of Standardized Operating Procedures for PCD diagnosis developed by the EU funded FP7 project BESTCILIA, in 2016, proposed standardized operating procedures for PCD diagnosis and a diagnostic algorithm which recommended that nNO should be performed in parallel with HSVM and confirmatory TEM assessment should follow in case of conflicting results [439]. Similarly, the recent ERS guidelines for the diagnosis of Primary Ciliary Dyskinesia also recommend a diagnostic algorithm which includes as a first step the parallel performance of both NO and HSVM and confirmation with TEM in a second step [440]. Nevertheless, such algorithms require the performance of a significantly higher number of nasal brushings for HSVM and would result in higher costs compared to algorithms that require the performance of a confirmatory test (HSVM or TEM) only following a positive result of the screening test.

To better illuminate the decision making process, the overall costs involved with the performance of each algorithm and the resulting health benefits for PCD patients need to be addressed and compared. This study aimed to evaluate the cost effectiveness and incremental cost-effectiveness of three distinct diagnostic algorithms for patients referred for PCD diagnosting testing across the European Union through a probabilistic decision analysis framework.

Materials and Methods

Decision tree model

Using a probabilistic decision tree model, three diagnostic algorithms were evaluated versus each other and against a baseline of not performing any diagnostic testing for PCD. The three diagnostic algorithms evaluated were a) Sequential testing with nNO screening followed by HSVM only when NO was positive (nNO+HSVM), b) Sequential testing with nNO screening followed by TEM only when NO was positive (nNO+TEM), c) nNO performed in parallel with HSVM and followed, in cases with conflicting results, by confirmatory TEM (nNO/HSVM+TEM). The decision tree displaying the evaluated three diagnostic algorithms in this study is presented in Figure 1.

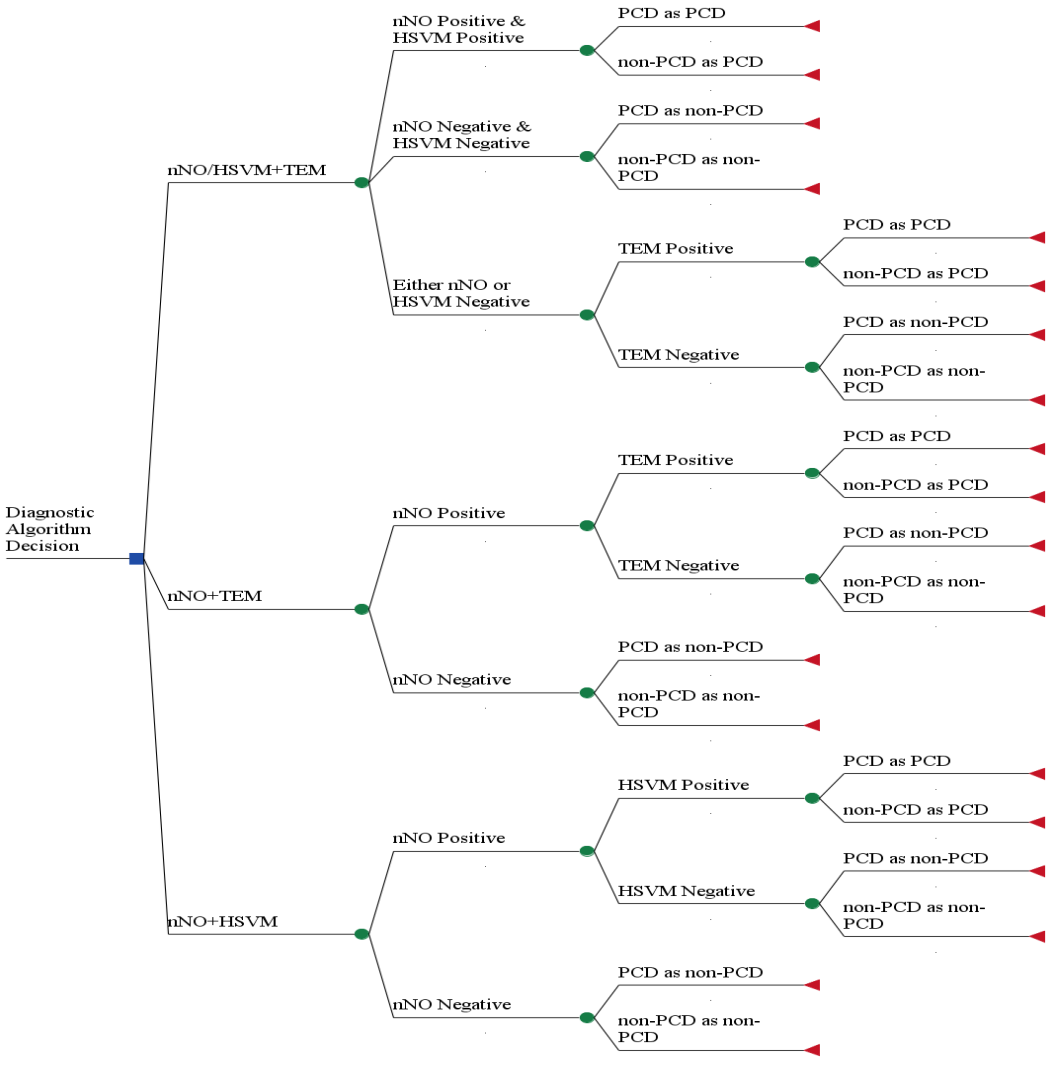


Figure7. 1: Decision Tree diagram for the three different diagnostic algorithms for PCD. *The decision tree begins from the left side and the decision whether to perform nNO+TEM, nNO+HSVM or nNO/HSVM+TEM. Squares represent decision nodes, circles represent chance nodes and triangles represent outcome nodes.*

The starting population of referrals for PCD diagnosting testing that enters the model was defined as one thousand per year for the whole of the European Union (EU). To estimate the classification of patients under each diagnostic algorithm, Bayes' Theorem was used. Bayes' Theorem allows the calculation of probability of suffering from PCD or not given the pre-test probability of disease and given a positive or negative diagnostic test [441]. The formula for estimating the probability of disease given positive diagnostic test is:

$$P(PCD|Test+) = \frac{P(Test+|PCD) * P(PCD)}{P(Test+|PCD) * P(PCD) + P(Test+|nonPCD) * P(nonPCD)}$$

Where P(Test+|PCD) is the probability of positive test given PCD is present (test sensitivity), P(PCD) is the prevalence of PCD in the tested population, P(Test+|nonPCD) is the probability of positive test given disease is not present (1-specificity of the test) and P(non-PCD) is the probability of not having PCD in the tested population. The formula can be rearranged accordingly to calculate probability of PCD given positive diagnostic test, probability of PCD given negative diagnostic test and probability of non-PCD given negative diagnostic test as well as probability of non-PCD given positive diagnostic test.n To model the sequence of diagnostic tests in each diagnostic algorithm the resulting probability of PCD given a positive first test as calculated using Bayes' Theorem was used as the pre-test probability of PCD for the second test. The final modeled health outputs regarding the effectiveness of each diagnostic algorithm included the number of PCD patients confirmed as PCD (True Positive - TP), PCD patients missed (False Negative - FN), non-PCD patients wrongly diagnosed as PCD (False Positive - FP), and non-PCD patients that had a diagnosis of PCD excluded (True Negative - TN). In addition, the annual total cost outcome (in Euros) was calculated for each diagnostic algorithm using a micro-costing approach. This approach involves the recognition of all underlying activities that make up a specific healthcare procedure and the product of resource cost and resource use provides the total cost estimate

for the procedure [442]. A detailed description of the diagnostic cost analysis is presented in Technical Appendix I.

Amortized annual cost estimates for the various pieces of diagnostic equipment (nNO analyzers, HSVM systems and TEM) were calculated by the formula [443]:

$$M = P * i * \frac{(1+i)^{N-1}}{(1+i)^{N} - 1}$$

Where P is the capital cost, i is the annual discount rate (treated as 0.05 per year in this analysis) and N is the useful life (lifespan) of the equipment. Amortized annual cost was added to the annual operating and maintainance costs (includes labor costs, consumables and maintenance) to estimate the total annual diagnostic cost for each diagnostic algorithm.

The incremental Cost Effectiveness Ratios (ICER) were calculated as the ratio of incremental costs to incremental effectiveness, i.e. [444]:

$$ICER = \frac{Cost_A - Cost_B}{Effect_A - Effect_B}$$

Here, $Cost_A$ and $Cost_B$ are the total annual per-patient costs of performing test algorithms A and B, respectively, and $Effect_A$ and $Effect_B$ are the number of PCD patients correctly diagnosed with PCD for the same diagnostic algorithms.

We conducted a secondary analysis that aimed to (1) broaden our characterization of costs to include all healthcare expenditures, and (2) quantify effectiveness in terms of quality adjusted life years (QALYs), a metric used broadly in the health economics literature [445]. We estimated QALYs for PCD patients as [446]:

$$QALYs = LE_{PCD} * HU_{PCD}$$

Where LE_{PCD} is the Life Expectancy for PCD patients and HU_{PCD} is the Health Utility for PCD. The total QALYs saved for each algorithm were calculated using the relationship:

$$Total\ QALYs = TP * PCD_{treated}QALYs + FN * PCD_{not\ treated}QALYs$$

Where TP and FN are the number of tested patients characterized respectively as True Positive and False Negative per year, while $PCD_{treated}$ QALYs are the number of QALYs saved when a PCD patient receives PCD-specific care and $PCD_{not\ treated}$ QALYs are the number of QALYs saved when a PCD patients does not receive PCD-specific care. The treatment procedures followed in PCD management were derived from recent reviews and consensus papers from Europe and North America and calculation of annual treatment cost for an average diagnosed PCD patient was possible. [433, 447-449] However there are no data regarding the treatment received by undiagnosed PCD patients and calculation of treatment cost for this subset of patients is not feasible.. To overcome this lack of evidence we performed a sensitivity analysis that provided a range of model results (CER and ICERs) while allowing the treatment cost for missed PCD patients to range from 3 times lower to 3 times higher as compared to the treatment cost for correctly diagnosed PCD patients. This sensitivity analysis reflects both the possible scenario of a PCD patient who has not been diagnosed and does not adhere to PCD specific treatment protocols, causing reduced healthcare costs as a result of, among other, less clinical appointments and physiotherapy consultations, reduced prescription of prophylactic antibiotics and less CT scans and the possible scenario of the same non-diagnosed patient causing higher healthcare costs due to higher frequency of exacerbations that may result to a higher frequency of hospitalizations and intravenous antibiotic administration, unnecessary tests and surgical procedures (e.g. grommets insertion, lung resection or lung transplantation) that could have otherwise been avoided [450].

Model Parameter Inputs

The prevalence of PCD in the general population was assumed to be 1/15000 births and the prevalence of PCD among patients referred for diagnostic testing was allocated a probability of 0.32 (95% CI: 0.26-0.39) as reported before [451]. Data regarding the diagnostic accuracy of each test were derived from systematic reviews and meta-analyses, when possible, and from alternative data sources such as large studies and multiple sources when meta-analytic estimates were not available. The parameter inputs for sensitivity and specificity of nNO during Velum Closure (VC) were 0.95 (95% CI: 0.91-0.97) and 0.94 (0.88-0.97) respectively, based on published meta-analytic estimates [452]. For HSVM, the parameter inputs for sensitivity and specificity were 1.0 (95% CI: 0.89-1.00) and 0.92 (95% CI: 0.86-0.96) based on published evidence provided by Boon et al. 2013 and Jackson et al. 2016 [453, 454]. For assessment of ciliary ultrastructure with TEM, the parameter inputs for sensitivity and specificity were 0.74 (95% CI: 0.68-0.80) and 0.91 (95% CI: 0.86-0.96) respectively based on a recent meta-analysis of 11 studies [451]. Sensitivity and specificity values for HSVM and TEM following a positive nNO result were obtained from the study by Jackson et al. 2016 [453]. Table 1 summarizes all parameter values that were part of the basic model. In the extended model, estimates about the QALYs saved in PCD and the treatment cost involved in PCD were required. Estimates about QALYs saved due to a specific treatment are usually available following randomized clinical trials (RCT) and life-long follow up of patients to inform about the life expectancy of the disease. In the case of PCD, RCT data are still scarce and information about life expectancy in PCD is non- existent. However, since life expectancy in many PCD patients is assumed to be normal or near-normal [455, 456] and using health utility (HU) estimates from other diseases with similar symptoms with PCD such as Cystic Fibrosis (CF) (HU ~ 0.75) [457], chronic bronchitis and chronic sinusitis (HU ~ 0.80) [458] and chronic obstructive pulmonary disease (HU ~ 0.85) [459], a calculation

regarding the number of QALYs saved in a PCD patient was possible. The calculation relies on the assumption that out of a full life expectancy of 80 years, 5 years (95%CI: 1.5-8.5) are lost due to PCD and that HU_{PCD} equals approximately 0.20 (95% CI: 0.16-0.24). The cost of the different resources (i.e. physiotherapy equipment and consultations, specialist consultations, antibiotic administration, hearing aids) that are part of PCD specific treatment were obtained through national and international databases (NHS Reference Costs [460], Germany Federal Health Monitoring [461], World Health Organization CHOICE database [462]). For all resources, we aimed to calculate European average resource prices by averaging prices from 3 countries (UK, Germany and Italy or Greece) and resource use was based on the experience of the authors. Mean estimates and 95% CI for the health utility for PCD patients, YLL in PCD and treatment cost parameters that were included in the extended model are also presented in Table 1.

Table 7.1: Model parameter inputs

Parameter description	Best Estimate (95% CI)	Probability distribution	Source
PCD prevalence among suspect patients	0.32 (0.25-0.39)	Normal (μ: 0.32,SD:0.028)	[28]
Diagnostic Accuracy			
nNO (VC) sensitivity	0.95 (0.91-0.97)	Normal (μ: 0.95, SD: 0.015)	[29]
nNO (VC) specificity	0.94 (0.88-0.97)	Normal (μ: 0.94, SD: 0.021)	[29]
TEM sensitivity	0.74 (0.66-0.83)	Normal (μ: 0.74,SD: 0.03)	[28]
TEM specificity	0.91 (0.86-0.96)	Normal (μ: 0.91,SD: 0.02)	[28]
HSVM sensitivity	1.00 (0.89- 1.00) 0.89 (-)	Truncate: Normal (μ: 0.98,SD: 0.035) Min: 0.89, Max: 1.00)	[30,31]
HSVM specificity	0.92 (0.86-0.96)	Normal (μ: 0.92,SD: 0.017)	[30]
Effectiveness Outcomes			

Years of Life Lost in PCD	5 (2-8)	Normal (μ: 5,SD: 1.5)	Assumption
Health Utility in PCD	0.15 (0.10-0.20)	Normal (μ:0.15, SD: 0.2)	[34, 35, 36]
Diagnostic Costs			
nNO related cost parameters			
nNO Ecomedics CLD88sp (VC) capital cost (€)	40000 (36000-44000)	Normal (μ: 40000,SD: 2000)	Market Value
nNO Ecomedics CLD88sp (VC) consumables per patient (€)	15 (9-21)	Normal (μ: 15,SD: 3)	Market Value
nNO operators rate (€/hour)	20 (10-30)	Normal (μ: 20,SD: 5)	Eurostat
nNO Ecomedics CLD88sp (VC) test duration (hours)	0.5 (0.3-0.7)	Normal (μ: 0.5,SD: 0.1)	Based on ATS/ERS [14]
nNO Ecomedics CLD88sp equipment lifespan (years)	15 (13-17)	Normal (μ: 15, SD: 1)	Market Value
nNO Ecomedics CLD88sp (VC) annual maintenance (€)	1300 (1100-1500)	Normal (μ: 1300, SD: 100)	Market Value
HSVM related cost parameters			
Capital cost HSVM – SAVA system (€)	5000 (3000-7000)	Normal (μ: 5000, SD: 1000)	Market Value (incl. camera and software)
HSVM consumables (€)	20 (16-24)	Normal (μ: 20, SD: 5)	Market Value
HSVM operators rate (€/hour)	20 (10-30)	Normal (μ: 20, SD: 5)	Eurostat
HSVM equipment lifespan (years)	15 (10-20)	Normal (μ: 15, SD: 2)	Assumption
HSVM test duration (hours)	2 (1.6-2.4)	Normal (μ: 2, SD: 0.2)	Based on Sisson J 2003 [7]
TEM related cost parameters			
TEM capital cost (€)	100000 (90000-110000)	Normal (μ:100000, SD:5000)	Market Value
TEM consumables (€)	110 (90-130)	Normal (μ:110, SD:10)	Market Value
Brushing Time (hours)	0.2 (-)	Constant: (Brushing Time: 0.2)	Assumption
TEM operators rate (€/hour)	20 (10-30)	Normal (μ: 20, SD: 5)	Eurostat
TEM test duration (hours)	10	LogNormal	[8, 30]

	(6-18)	(Median: 10, gsd: 1.3)	
TEM equipment lifespan (years)	30 (20-40)	Normal (μ: 30, SD: 5)	Assumption
Physician's rate (€/hour)	50 (30-70)	Normal (μ: 50, SD: 10)	Eurostat
TEM annual maintenance (€)	2000 (1300-2600)	Normal (μ: 2000, SD: 300)	Assumption

Characterization of Uncertainty

Reported uncertainty around pooled estimates of the meta-analyses of diagnostic effectiveness and uncertainties about the true value of costs and other parameters are reflected by the probability distributions around the parameter means which are used in this model. A Cost-Effectiveness Acceptability Curve was used to demonstrate the uncertainty in the estimation of the ICER [463]. All parameters and equations constitute the final model which was developed with ANALYTICA 101 edition (Lumina decision systems, CA, United States). The model was executed with 1000 iterations per "model run" using Latin Hypercube sampling to generate samples from the underlying parameter probability distributions. The model can be assessed online and a model overview is presented in supplementary figure 1.

Results

The model output for TP, FN, TN and FP and estimates of net sensitivity, net specificity, net positive predictive value and net negative predictive value for the application of each diagnostic algorithm in a hypothetical cohort of 1000 patients suspected of PCD is presented in Table 2. Table 3 compares mean diagnostic costs with the number of PCD cases identified and reports relevant CERs and ICERs. Deterministic comparison for mean costs and effects demonstrated that the nNO/HSVM+TEM was the most effective algorithm but also the most costly (311 PCD cases identified/year, 97,400 €/year). nNO+HSVM was the second most

effective (274 PCD cases identified/year, 38,600 €/year) while nNO+TEM was the least effective (198 PCD cases identified/year, 55,800 €/year). The most cost-effective algorithm was nNO+HSVM with a CER of 140 €/PCD case identified, followed by nNO+TEM (280 €/PCD case identified) and nNO/HSVM+TEM (315 €/PCD case identified).

Table 7.2: Diagnostic accuracy of nNO+TEM, nNO+HSVM and nNO/HSVM+TEM algorithms

	Diagnostic Algorithm		
Classification	**NO+TEM**	**NO+HSVM**	**NO/HSVM+TEM**
PCD as PCD (% of PCD)	198 (62%)	274 (86%)	311 (97%)
PCD as non-PCD (% of PCD)	122 (38%)	46 (14%)	9 (3%)
Non-PCD as non-PCD (% of non-PCD)	680 (100%)	680 (100%)	677 (99.5%)
Non-PCD as PCD (% of non-PCD)	0 (0%)	0 (0%)	3 (0.5%)
Net Sensitivity	65%	86%	97%
Net Specificity	100%	100%	99.5%
Net PPV	100%	100%	99%
Net NPV	85%	94%	99%

PPV: Positive Predictive Value
NPV: Negative Predictive Value

Table 7.3: Diagnostic costs per year, identified PCD cases per year

			ICER (€/PCD case identified	
Diagnostic Algorithm	**Diagnostic Cost per annum in € (SD)**	**PCD cases identified per annum (SD)**	**Compared to No screening**	**Compared to next most effective algorithm***
Do nothing	0	0	-	-
NO+HSVM	38,600 (6150)	274 (24)	140 (22)	140 (22)
NO+TEM	55,800 (6242)	198 (18)	285 (32)	Dominated
NO/HSVM+TEM	97,400 (11920)	311 (27)	315 (45)	1620 (337)

*ICER compared to next less expensive algorithm omits from consideration those algorithms that are "dominated" (make health worse and cost more). Hence, we compare NO/HSVM+TEM (last row) to NO+HSVM (2nd row) and not to NO+TEM (3rd row) because NO+TEM is dominated (it costs more than NO+HSVM but identifies fewer cases).

The cost effectiveness frontier in presented in Figure 2 and the resulting ICER for nNO/HSVM+TEM compared to nNO+HSVM, the second most effective algorithm, is 1620 € per additional PCD case identified. The nNO+TEM algorithm is dominated (simple dominance) by nNO+HSVM as it is more expensive but less effective compared to nNO+HSVM.

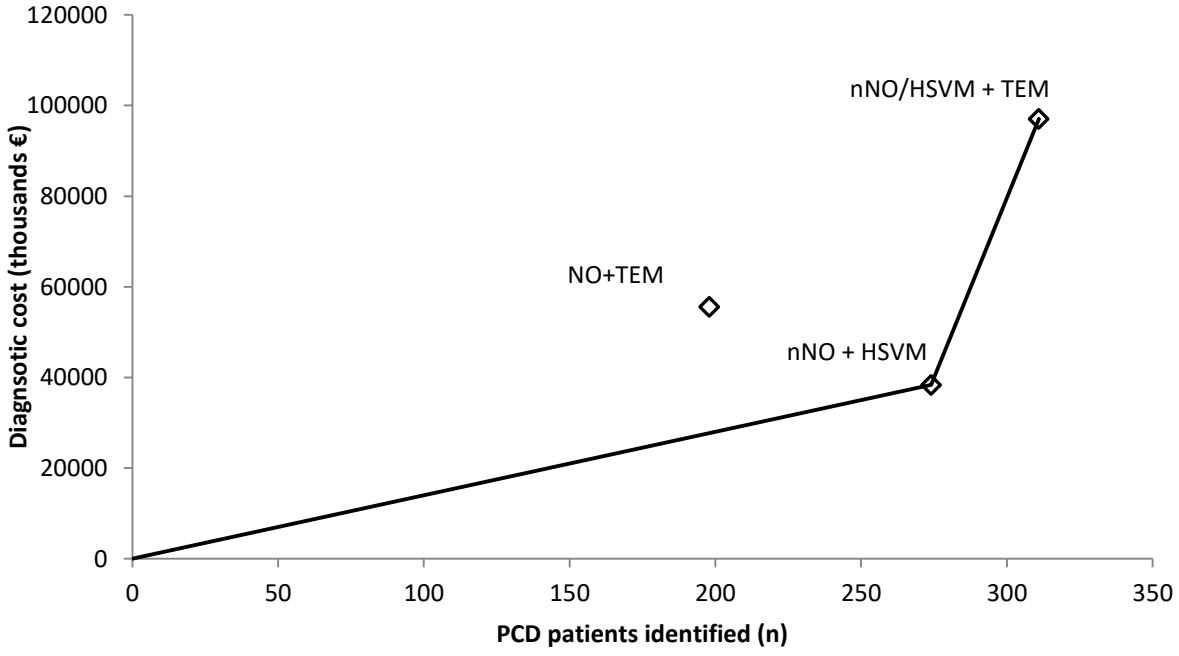


Figure 7.2: Cost-effectiveness frontier for the three different diagnostic algorithms for PCD. Diagnostic algorithms nNO+HSVM snd nNO/HSVM+TEM are cost-effective alternatives at different WTP thresholds. Diagnostic algorithm nNO+TEM is dominated by nNO+HSVM.

Figure 3 presents the cost-effectiveness acceptability curve (CEAC) for nNO/HSVM+TEM. The CEAC demonstrates the uncertainty in the estimation of ICER and provides information about the probability of nNO/HSVM+TEM being more cost effective compared to nNO+HSVM for a range of potential monetary amounts (termed willingness to pay (WTP) thresholds) that a decision maker might be willing to pay to correctly diagnose an additional PCD case. For a WTP threshold equal to €2000 the probability of nNO/HSVM+TEM being

cost effective is over 80% and for a WTP threshold equal to €2500 the probability is over 97%.

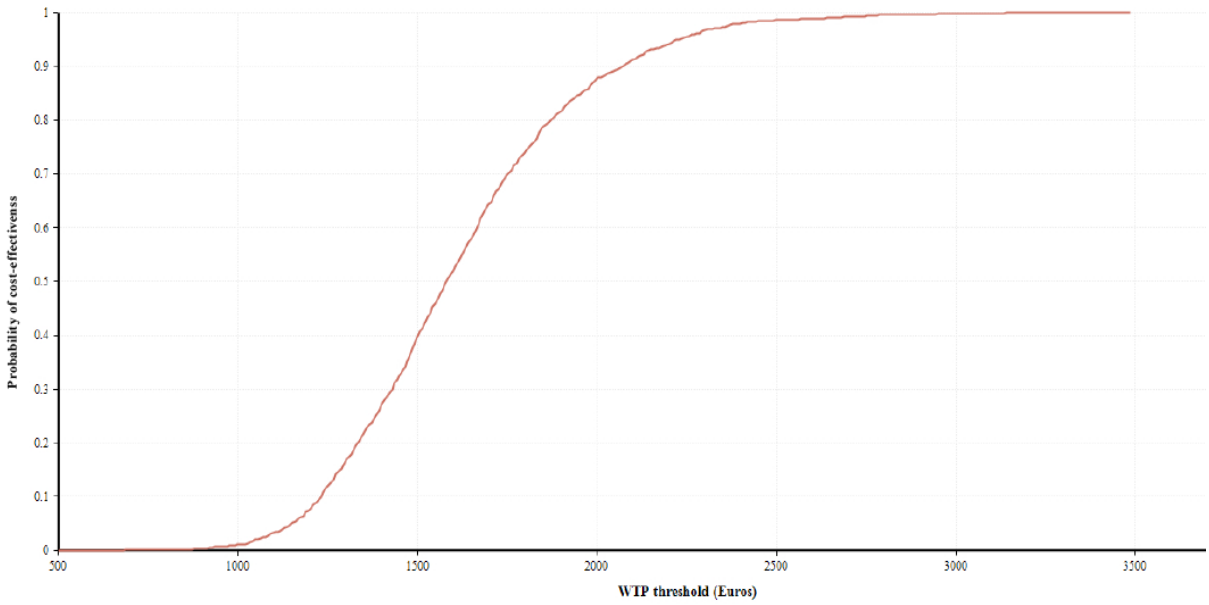


Figure 7.3: Cost Effectiveness Acceptability Curve for nNO/HSVM+TEM. The probability that diagnostic algorithm nNO/HSVM+TEM is cost-effective for a range of WTP thresholds.

The development of the extended model, allowed the calculation of ICERs which included the effectiveness and cost of PCD specific treatment as well. Table 4 presents the lifetime effect (in QALYs saved) and the present value of lifetime costs (in €) for the cohort of 1000 referrals and resulting CER.

Table 7.4: Total QALYs saved for a cohort of 1000 referrals (baseline analysis)

Diagnostic Algorithm	QALYs saved (SD)	CER (€/PCD case identified) (SD)*
Do nothing	0	-
NO+TEM	4075 (586)	10,170 (2621)
NO+HSVM	5519 (805)	6863 (1305)
NO/HSVM+TEM	6224 (909)	5708 (959)

*Compared to not performing any diagnostic procedure, CER: Average Cost-effectiveness Ratio

With the basic model, the most effective algorithm was nNO/HSVM+TEM with a mean of 6224 QALYs saved for the whole cohort followed by nNO+HSVM with 5519 QALYs and nNO+TEM with 4075 QALYs. The most cost-effective algorithm was nNO/HSVM+TEM followed by nNO+HSVM and nNO+TEM with mean CERs equal to 5700 €/QALY saved, 6860 €/QALY saved and 10170 €/QALY saved respectively. The sensitivity analysis that allowed the ratio of treatment cost between undiagnosed PCD patient and diagnosed PCD patient to vary between 0.3 and 3 demonstrated that nNO/HSVM+TEM was always the most cost-effective algorithm but the differences in resulting CERs were greater when treatment cost for an undiagnosed PCD patient far exceeded the treatment cost for a diagnosed PCD patient (Table 5).

Table 7.5: Total lifetime costs, total QALYs saved for a cohort of 1000 referrals (sensitivity analysis)

Diagnostic Algorithm	Total Lifetime Cost in million € (SD)		QALYs saved (SD)	ACER (€/PCD case identified) (SD)*	
Ratio of undiagnosed PCD TC to diagnosed PCD TC	0.3	3	-	0.3	3
Do nothing	0	0	0	-	-
NO+TEM	25.07 (3.63)	60.36 (8.73)	4075 (586)	6232 (1016)	15010 (2481)
NO+HSVM	30.74 (4.45)	44.11 (6.43)	5519 (805)	5644 (932)	8104 (1373)
NO/HSVM+TEM	33.57 (4.84)	36.25 (5.27)	6224 (909)	5467 (906)	5905 (995)

* Compared to not performing any diagnostic procedure, CER: Average Cost-effectiveness Ratio, TC: lifetime treatment cost

Discussion

The high genetic heterogeneity that characterizes PCD and the resulting inability to rely on a single test to confirm or exclude diagnosis of the disease has led to increased research interest in specialized diagnostic testing for PCD in recent years. This study compares three diagnostic strategies currently in use for diagnosing PCD and reports on their effectiveness

and cost-effectiveness. Data were drawn primarily from meta-analyses of diagnostic effectiveness or published estimates from large studies and were synthesized in a probabilistic cost effectiveness model.

The results presented here demonstrate that when the effectiveness outcome is defined as the number of PCD patients identified, nNO/HSVM+TEM is the most effective diagnostic algorithm followed closely by nNO+HSVM. Both nNO/HSVM+TEM and nNO+HSVM are significantly more effective compared to the third diagnostic strategy evaluated, nNO+TEM. Mean estimates of CERs demonstrate that nNO+HSVM was the most cost-effective option and a decision maker should expect to pay on average an amount equal to 1620 € per additional case identified if nNO/HSVM+TEM is implemented. Whether the effectiveness outcome is defined as number of PCD patients identified or as the number of QALYs saved nNO/HSVM+TEM was still the most effective algorithm followed by nNO+HSVM and nNO+TEM. Nevertheless, the results of the extended model, which are expressed in Euros per QALY saved, demonstrate that all three diagnostic algorithms appear to be very cost-effective. Compared to no screening, the cost per QALY gained for the three diagnostic algorithms examined here ranged from 5700 to 10170, an estimate which is much lower than WTP thresholds commonly used by regulatory authorities around the world. Such WTP thresholds range between 20,000 and 30,000 pounds per QALY saved in the UK [464] or the more conventional WTP threshold of 50000 dollars per QALY saved, commonly used in the US [465].

Diagnostic algorithms including nNO measurement during VC as an initial screening could be cost-effective. However, our results demonstrate that nNO screening is more effective when the confirmatory test is HSVM and not TEM. Although in the past TEM was considered the gold standard [37] TEM analysis it is now known to miss an important fraction of PCD patients [451], mainly those with biallelic mutations in *DNAH11* gene [466]

and those with specific ultrastructural abnormalities (nexin link defects) that are not easily detectable by standard TEM [467]. Furthermore, TEM analysis requires access to a specialized lab with experienced personnel in staining and interpretation of TEM micrographs and consequently involves considerable resource allocation [468]. Furthermore, TEM studies are usually time consuming and results are usually obtained and communicated to patients considerably later than results of other tests thus contributing to patient distress [469]. HSVM is easier, considerably faster and cheaper than TEM as it is usually performed on the same day following nasal brushing and the equipment required consists of standard microscope, a high speed video camera and a standard computer loaded with specialized software. It has also been reported to be a highly sensitive and specific test [453] thus it significantly outperforms TEM as a confirmatory test both in terms of overall effectiveness and cost. However, extra caution is required with HSVM as it may be affected by observer subjectivity and non-PCD specific findings which may interfere with the motility interpretation [434]. Overall, the parallel performance of two highly specific and sensitive tests such as nNO and HSVM during the first step of the diagnostic algorithm, followed by confirmatory TEM in only the few cases of conflicting findings, results in the identification of most PCD patients and does not require the performance of the more expensive and time consuming TEM analysis for the largest part of the cohort of suspect patients.

In this study we did not include diagnostic algorithms that included immunofluorescence (IF) and/or genetic testing for PCD. Although a recent study has reported the first diagnostic accuracy and cost estimates for immunofluorescence testing in PCD [470], the use of this test is still very limited (as it is performed only in a small number of few highly specialized centers around the world). Genetic testing, on the other hand, is available in many centers around the world. However, as yet, there is little standardization of procedures for the conduct and interpretation of results. Different centers may use different technologies and

may not test for the same number of genes [471, 472]. Thus estimation of the effectiveness or the cost of genetic testing as diagnostic for PCD was not possible at this stage and it was not included in the diagnostic algorithms considered in our analysis. This approach is in line with the recent guidelines published by the ERS where genetic testing was recommended as a last step following abnormal TEM primarily for further characterization of the underlying defect or as a final diagnostic test if all other tests were inconclusive. For immunofluorescence there was no ERS recommendation towards its use as a diagnostic test given the scarcity of evidence [440].

The main strength of this study is that it makes use of evidence-based estimates and individual good quality studies on the diagnostic accuracy of nNO, TEM and HSVM and the prevalence of PCD among cohorts of referred suspect patients. With the use of Bayes' Theorem, it was possible to estimate the diagnostic effectiveness of sequential tests and to compare the effectiveness of diagnostic algorithms instead of simply comparing the effectiveness of isolated tests, as had been done in the past. In addition, our analysis of the costs involved in diagnostic testing followed standard approaches for economic analysis of healthcare procedures [442] and made use of the extensive literature on the effort, equipment and consumables involved in the performance of nNO [473, 474], HSVM [453, 475] and TEM [432]. Based on this evidence, we were able to calculate effectiveness and economic outcomes (number of PCD patients identified, total diagnostic costs) as well as robust CERs, ICERs and identify the cost-effectiveness frontier.

Nonetheless, this study also has several limitations. Most of these relate to the considerable uncertainty of the parameters that make up the extended model and for this reason the results of basic and extended models are presented separately. The model parameter regarding life expectancy in PCD is based on our own assumption as data about the life expectancy of this disease are currently not available. This lack of information could be attributed to the fact that

PCD is a disease that has been studied primarily in the pediatric setting and long-term studies of adult PCD patients are rare. The recently developed prospective PCD registry [476] provides a useful tool to follow pediatric and adult patients in time in order to monitor disease progression and life-expectancy in PCD. Furthermore, although empirical evidence about various approaches for the treatment of PCD is beginning to accumulate, at the moment there are no widely recognized PCD-specific treatment protocols. The efficacy of a few treatment approaches are now under investigation through RCTs, for example, those now underway on the effect of Azithromycin for antibiotic prophylaxis [477]. Furthermore, there are no published estimates of the annual (or lifetime) cost of various options for treatment of PCD. Although we used credible sources [460-462] to estimate the cost of each procedure (resource cost) we had to rely on our own experience with the disease to characterize the typical frequency of treatment (resource use). To address this limitation, the underlying uncertainty in each parameter was characterized and included in the model. Through Latin Hypercube sampling and Monte Carlo analysis, these uncertainties in individual parameters were propagated through the model and are reflected in the uncertainty in final model outputs.

Evidence about treatment costs is especially weak. We found no evidence of the cost of treatment of PCD patients who manage to remain undiagnosed; and only limited evidence about the cost of treatment of PCD patients who are properly diagnosed. A sensitivity analysis was conducted to determine whether differences in the overall costs of treatment of diagnosed and undiagnosed PCD patients affected the estimates of cost-effectiveness from the extended model. The overall order of diagnostic algorithms was not affected and nNO/HSVM+TEM was the most cost efficient algorithm in all scenarios. However, the magnitude of the difference in the cost effectiveness of the three algorithms was significantly affected, with nNO/HSVM+TEM becoming relatively more cost-effective when it was assumed that the cost of treating undiagnosed PCD patients was at least 3 times greater than

the treatment cost for properly-diagnosed PCD patients. This highlights the importance of future studies which address the economic cost of treatment in PCD patients before and after diagnosis.

Conclusions

Many centers for the diagnosis and treatment of PCD in the developed world follow a variety of algorithms for diagnosing PCD. In some low income countries, most likely, there is a complete lack of specialized diagnostic testing. The results of this study suggest that a diagnostic algorithm which includes nNO during VC as a screening test followed by confirmatory HSVM identifies approximately 86% of PCD patients with a mean CER of 140€ per PCD case identified. The algorithm which maximizes the number of PCD patients identified involves parallel performance of nNO and HSVM as the first step, followed by TEM as a confirmatory test for the few cases where nNO and HSVM yield conflicting results, with a corresponding ICER of 1620€ per additional PCD patient identified. These findings can inform the dialogue about the development of evidence-based guidelines for PCD diagnostic testing and can illuminate discussions about how these guidelines can best be implemented across various healthcare systems.

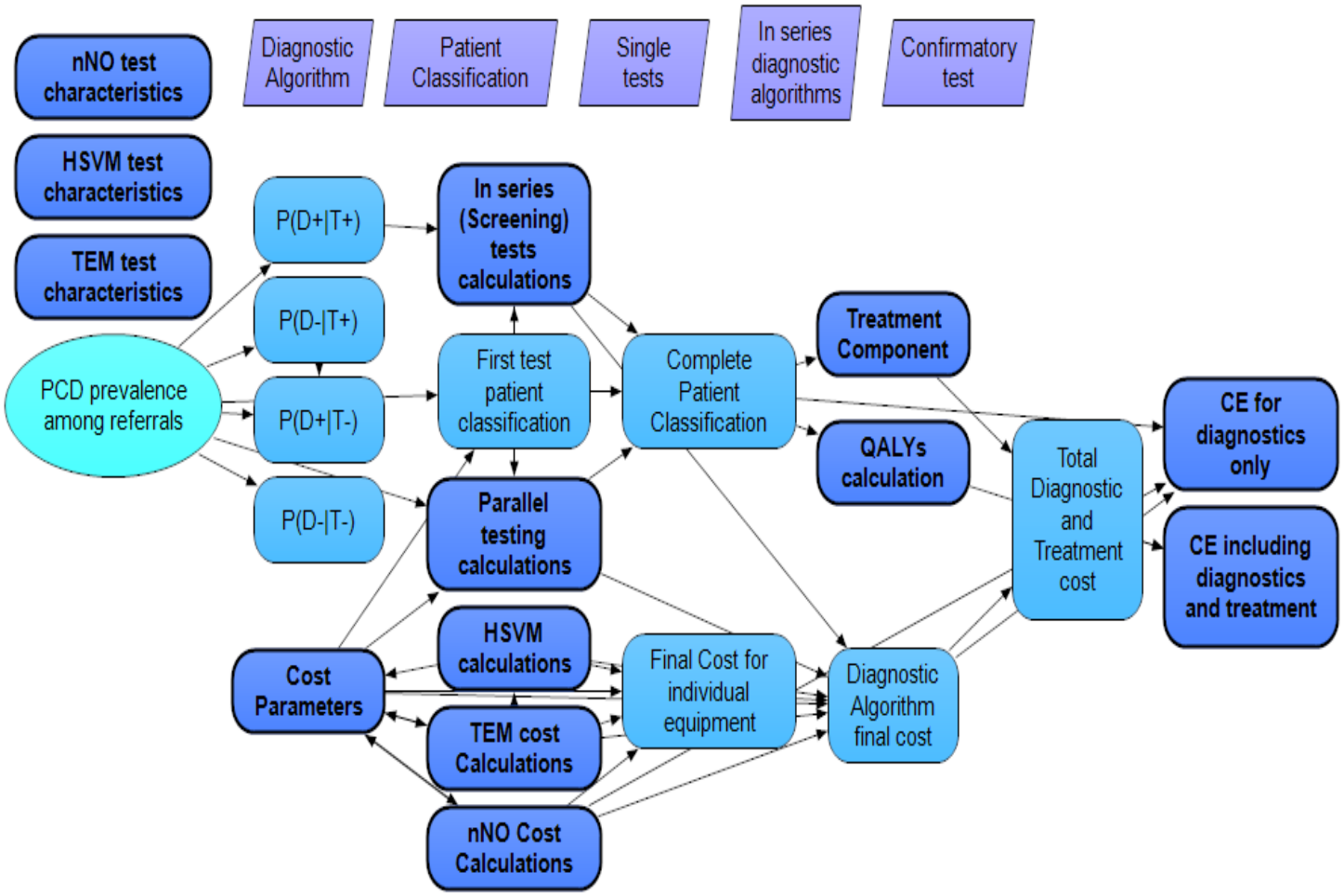
Acknowledgments

P.K. was supported by the European Union's Seventh Framework Program EC-GA No. 305404 BESTCILIA. Joshua Cohen provided advice on cost-effectiveness analysis methodology and contributed towards the final version of the manuscript. John S. Evans supervised the model development and interpretation of the findings and contributed

intellectually towards the final version of the manuscript. The sponsors had no role or involvement in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. All authors wish to declare that they have no competing interests.

Contributions

P.K. performed the developed the decision tree model under the supervision of J.S.E, extracted relevant information from the published literature and prepared the first draft of the manuscript. S.I.P. contributed to the model development, extraction of relevant information form the published literature contributed intellectually towards the final version of the manuscript. N.M., A.H. and K.K. contributed to the interpretation of the findings and critically revised the final version of the manuscript. G.Y. assisted in the economic analysis and contributed towards the final version of the manuscript. P.K.Y. conceived the hypothesis of the manuscript, contributed to the model development and interpretation of the findings and contributed intellectually towards the final version of the manuscript



Supplementary Figure 7.1: Model Overview. Schematic Overview of ANALYTICA model

Technical Appendix

For

Cost-effectiveness Analysis of Three Algorithms for Diagnosing Primary Ciliary Dyskinesia

Diagnostic Cost Analysis:

The analysis was conducted from the perspective of an average European Union (EU) Healthcare System and thus it is confined to direct medical costs based on EU rates. The annual total diagnostic cost for a hypothetical cohort of 1000 suspect patients referred for specialised PCD testing was calculated using the micro-costing approach which involved the following steps:

(a) Identification of all cost generating procedures (recourses) involved in diagnosting testing for PCD

(b) Estimate unit resource cost based on average EU prices for different resources.

(c) Estimate the resource usage based on model output.

(d) Combine the information on unit recourse costs with resource usage and sum across all identified resources as described by the general formula:

$$\sum_{i=1}^{k} RC_i + RU_i$$

Where *k* refers to the total number of resources involved in PCD diagnosting testing and RC_i and RU_i refer to the resource cost and resource use estimated for resource *i*.

Direct medical costs included in PCD diagnostic testing

Direct medical costs include the cost of nasal Nitric Oxide (nNO) measurement, the physician office visit cost for the performance of nasal brushing and the laboratory cost involved in the performance of the High Speed Video Microscopy (HSVM) and Transmission Electron

Microscopy (TEM) analyses on the acquired sample. The cost of nNO, HSVM and TEM equipment purchase, disposables, labour and maintenance constitute the direct medical cost and the detailed breakdown of these diagnostic procedures into separate resources and their pricing is described in Table 1.

Calculation of diagnostic costs:

The electron microscope used for TEM analysis and the analyser used for NO measurement are not only used for PCD diagnosis but could be used for other applications as well. On the contrary HSVM is expected to solemnly be utilised for PCD diagnosis. As a result, two slightly different approaches are followed for the calculation of diagnostic costs for HSVM and nNO and TEM. More specifically, in the case of nNO and TEM the final cost of the diagnostic procedure calculated across all resources (e.g. capital cost, labour, consumables etc) has to be corrected using the ratio of equipment use to the total equipment efficiency which is calculated as follows:

$$\frac{Number\ of\ patients\ tested * T_{Test}}{Efficiency_{Equipment}}$$

Where T_{Test} refers to the duration of the test in hours and $Efficiency_{Equipment}$ refers to the total time (hours) that the equipment could theoretically be used over a year.

Diagnostic cost (HSVM):

The operational cost per patient for each diagnostic test was calculated using the formula:

$$OC_{per\ patient} = \left[T_{Test}\ (hours) * CR\ (\frac{€}{hour})\right] + Consumables_{per\ patient}$$

Where $OC_{per\ patient}$ refers to Operational Cost per patient, T_{Test} refers to the duration of the test in hours and *CR* refers to Compensation Rate for the diagnostician in €/hour. The total annual operational cost (TOC) was calculated as follows:

$$TOC = OC_{per\ patient} * Number\ of\ patients\ tested$$

Total annual Operational and Maintenance Cost (TOCM) which includes the operational cost as well as annual payments towards the capital investment and annual maintenance is calculated as:

$$TOCM = TOC + AC + AM$$

Where *TOC* refers to Total Operational Cost, *AC* refers to amortized cost and *AM* refers to expected annual maintenance of the equipment.

Diagnostic cost (nNO and TEM):

The operational cost per patient for each diagnostic test was calculated using the formula:

$$OC_{per\ patient} = \left[T_{Test}\ (hours) * CR\ (\frac{€}{hour})\right] + Consumables_{per\ patient}$$

Where $OC_{per\ patient}$ refers to Operational Cost per patient, T_{Test} refers to the duration of the test in hours and *CR* refers to Compensation Rate for the diagnostician in €/hour. The total annual operational cost (TOC) was calculated as follows:

$$TOC = OC_{per\ patient} * Efficiency_{Equipment}$$

Total annual Operational and Maintenance Cost (TOCM) which includes the operational cost as well as annual payments towards the capital investment and annual maintenance is calculated as:

$$TOCM = TOC + AC + AM$$

For nNO measurement:

$$\left[\frac{Number\ of\ patients\ tested * T_{Test}}{Efficiency_{Equipment}}\right] * TOCM$$

For TEM analysis (also includes sample processing cost (PC) prior TEM analysis):

$$(Number\ of\ patients\ tested * PC) + \left[\frac{Number\ of\ patients\ tested * T_{Test}}{Efficiency_{Equipment}}\right] * TOCM$$

Chapter 8: Concluding Remarks, Limitations and Future Research

Concluding Remarks

Although many of the manifestations of PCD present early in life, diagnosis is often delayed or missed completely, primarily due to the low specificity of some symptoms (e.g. cough, rhinorrhea) and lack of awareness for PCD among clinicians [5]. In addition, difficulties in establishing PCD diagnosis, both due to lack of equipment and or lack of expertise, might further contribute towards misdiagnosis [11]. Up to date, several tests have been developed for PCD diagnosis such as TEM, HSVM, nNO measurement, immunofluorescence analysis and genetic testing but none of which has been found to be characterized by 100% sensitivity and 100% specificity. As a result, a combination of tests is usually required for the establishment of PCD diagnosis [18]. Nevertheless, due to the lack of either the needed equipment or the expertise to perform all tests as well as due to the lack of universal agreement whether diagnostic tests should be performed in sequence or in parallel, different diagnostic algorithms for PCD have been adopted by diagnostic centers across the world [16].

This study focuses on the description of the diagnostic properties of the three most widely used tests for PCD (TEM, HSVM, nNO measurement) through the performance of systematic reviews and meta-analyses and the comparison of three different diagnostic algorithms (nNO+TEM, NO+HSVM and NO/HSVM +TEM) through a probabilistic decision tree model and the estimation of cost-effectiveness and incremental cost-effectiveness of the evaluated algorithms.

In chapter 5, a systematic review and a diagnostic accuracy meta-analysis of NO measurement during VC or non-VC for the diagnosis of PCD was carried out. Twelve individual studies were synthesized and provided data for thirteen different populations, including nine case-control (n=793) and four prospective cohorts (n=392). The overall

sensitivity of nNO measured during VC was 0.95 (95% CI 0.91-0.97), while specificity was 0.94 (95% CI 0.88-0.97). For non-VC techniques, the overall sensitivity of nNO measurement was 0.93 (95% CI 0.89-0.96) whereas specificity was 0.95 (95% CI 0.82-0.99). The results indicate that that measurement of nNO, both with VC and non-VC maneuvers, has high overall diagnostic accuracy and provides a clinically significant diagnostic tool for large uninvestigated populations of suspect cases worldwide where access to TEM and HSVM is not easy.

In chapter 6, the published literature was systematically reviewed and pooled estimates of the PCD prevalence among referrals and TEM detection rate in confirmed PCD patients were derived through meta-analysis of proportions. PCD prevalence among referrals was found to be 32% (95%CI:25%-39%, I^2=92%) and the TEM detection rate among PCD patients was 83% (95%CI:75%-90%, I^2=90%). Exclusion of studies reporting an isolated inner dynein arm defect in TEM which is considered unreliable for PCD diagnosis [478], reduced the pooled TEM detection rate and explained an important fraction of the observed heterogeneity (74%, 95%CI:66%-83%, I^2=66%). This analysis demonstrated that among cohorts of consecutive referrals of suspect cases for PCD testing, approximately one third are eventually confirmed as PCD patients and among PCD cases that underwent TEM studies, a significant percentage, at least as high as 26%, were not identified by TEM (ref).

The evidence regarding the diagnostic properties of nNO and TEM as well as the evidence regarding the prevalence of PCD among suspect patients were combined along with diagnostic accuracy estimates for HSVM from individual studies to develop a probabilistic decision model that allowed the calculation of net sensitivity and specificity as well as the cost-effectiveness (CE) for three diagnostic algorithms that were characterized by different combinations of nNO, TEM and HSVM. The three diagnostic algorithms evaluated were (a) nNO+TEM in sequence, (b) nNO+HSVM in sequence and (c) nNO/HSVM in parallel

followed, in cases with conflicting results, by confirmatory TEM (nNO/HSVM+TEM). The results presented in chapter 7 indicate that out of a hypothetical starting population of 1000 referrals (of whom 320 are PCD patients), 311 were correctly identified as PCD by nNO/HSVM+TEM, 274 with nNO+HSVM and 198 with nNO+TEM. The nNO/HSVM+TEM had the higher mean cost (€97K) followed by nNO+TEM (€56K) and nNO+HSVM (€39K). The nNO+HSVM algorithm dominated the nNO+TEM algorithm (less costly and more effective). The ICE ratio for nNO/HSVM+EM was €1.6K per additional PCD patient identified. Furthermore, in an extended analysis that also took into account the treatment effectiveness and accompanying treatment cost and expressed the cost effectiveness ratios in units of Euros per QALY saved, not only identified nNO/HSVM+TEM as the most cost-effective diagnostic algorithm but also indicated that all three evaluated diagnostic algorithms appear to be very cost–effective when compared frequently used WTP thresholds. Compared to no screening, the mean cost per QALY gained for all three diagnostic algorithms evaluated ranged from 5700 to 10170 Euros while widely used WTP thresholds range between 20,000 and 30,000 pounds per QALY saved in the UK [464] or 50000 dollars per QALY saved, in the US [465].

Limitations

Primary Ciliary Dyskinesia is a rare disease and usually the number of patients included in the individual studies that were synthesized in Chapter 5 and 6 were small. In addition, as was the case in Chapter 5 many of the studies followed a case-control design which may lead to overestimation of the diagnostics Odds Ratio [479]. However, due to the fact the case-control studies included in the meta-analysis for nNO measurements were studies designed to assess the test accuracy and not to provide evidence on associations between a risk factor and

the disease, the possibility of overestimation is limited [480]. In Chapter 5, there was also considerable heterogeneity between included studies regarding the diagnostic standard used to classify patients as PCD as well as in the number of cases, total sample size and cut-off values. In order to address these limitations a series of sensitivity analyses were performed and demonstrated that meta-analysis results were not sensitive to variations in the diagnostic standard used in the individual studies and HSROC were employed to account for the different cut-offs reported in individual studies.

In Chapter 6, the majority of the studies included in the meta-analysis for PCD prevalence among suspect patients and the meta-analysis for the detection rate of TEM were retrospective cohort studies and this type of studies suffer from selection and misclassification bias. Another important limitation of the analysis presented in Chapter 6 is that for the vast majority of PCD patients who had normal ultrastructure and were missed by TEM, the genetic defect is not specified. This represents an important limiting factor as we do not know if all the responsible genetic defects known to date to cause PCD and normal cilia structure were represented in this subgroup.

The cost-effectiveness analysis, presented in Chapter 7 also has some limitations. While the primary analysis that estimated the diagnostic cost per PCD case identified was characterized by good data quality, the secondary analysis (extended model) which estimated the total cost per QALY saved was limited by the considerable uncertainty of the model parameters. Such parameters that are not based on good quality data were the life expectancy in PCD, the Health utility in PCD the efficacy of PCD specific treatment and the typical frequency of each treatment procedure. To address this limitation, the underlying uncertainty in each parameter was characterized and included in the model. Through Latin Hypercube sampling and Monte Carlo analysis, these uncertainties in individual parameters were propagated through the model and are reflected in the uncertainty in final model outputs. Furthermore, no

data were available about the treatment cost in PCD patients that remain undiagnosed. A sensitivity analysis was conducted to determine whether differences in the overall costs of treatment of diagnosed and undiagnosed PCD patients affected the estimates of cost-effectiveness from the extended model. The overall order of diagnostic algorithms was not affected and nNO/HSVM+TEM was the most cost efficient algorithm in all scenarios. However, the magnitude of the difference in the cost effectiveness of the three algorithms was significantly affected, with nNO/HSVM+TEM becoming increasingly more cost-effective when the cost of treating undiagnosed PCD patients was relatively higher compared to the cost of treating diagnosed PCD patients.

Future Research

During the last decade the great advancements in genetic sequencing has driven the progress in the field of Primary Ciliary Dyskinesia. It has allowed improved understanding of disease etiology and along with the improvements in diagnostic techniques (primarily in HSVM [150] and the development of newer applications such as cryo-electron tomography of respiratory cilia [481], it has driven the improved understanding of PCD pathophysiology. Up to this date, genetic testing is not considered as a stand-alone test for PCD and current recommendations consider genetic testing as mostly confirmatory test [18]. In the future however, with ever reducing costs for DNA sequencing, genetics are expected to have an even more prominent role during diagnostic testing. Not surprisingly, in recent years, different studies have already evaluated whether targeted NGS panels or whole exome sequencing could be applied in routine PCD diagnostics [482-484] with promising results. Furthermore, even post-diagnosis genetic investigations will continue to be of critical value given that with advancements in the field of gene therapy, it could be possible for

personalized interventions at the molecular levels to be developed and provided to the patients.

Progress in the understanding of PCD pathophysiology and genetic etiology is also a consequence of research community improved understanding of cilia biology. Basic scientists working with different animal models with a focus on cilia development and cilia motility will continue to provide insights about novel PCD candidate genes and will further contribute towards the continuing efforts for the developments of treatments for PCD. Currently basic scientists with an interest in cilia motility are utilizing an array of animal models such as *Chlamydomonas reinhardtii* [122], *Xenopus laevis* [124], *Zebrafish* [126] and mice [127] while other animal models that can be useful for cilia motility research have been described recently such as the planarian flatworm *Schmidtea mediterranea* [485]. Apart from animal models basic scientists involved in cilia biology research are taking advantage of in vitro cultures of ciliated cells either in the form of Air Liquid Interface cell cultures (ALI cultures) or suspension cultures. All animal models and cell culture systems can be used as pre-clinical models that can be utilized for the deeper characterization of the genotype–phenotype relationships and the development of therapies for PCD [486]. Pre-clinical models, especially species with lungs such as mice and rats can provide important insights not only on cilia motility parameters but also on other physiological parameters (oxidative stress, inflammation development of bronchiectasis) that may affect respiratory health in animals and humans alike. A major drawback of using mice or rats to model PCD is the high mortality in the animals due to the extremely high prevalence of hydrocephalus, a phenomenon that is attributed to the concurrent presence of ciliary dyskinesia and the small diameter of brain ventricles (compared to human) [141]. As a result newer studies have focused on developing (postnatal knockdowns) mouse models that show no evidence of hydrocephalus such as the Dnaic1[iv] mouse [487], the Dnahc11[iv] [488] and the Dnahc6[iv] [84].

This trend is expected to continue since animal models that make it to adulthood allow researchers to model disease progression and thus more murine postnatal knockdowns for more PCD related genes are expected to be described. Furthermore the recent identification of the genetic defects that are responsible for the occurrence of acilia phenotypes (lack of cilia phenotypes), have demonstrated the importance of basic science research regarding multiciliated cells differentiation and cilia development. Pathogenic mutations in MCIDAS and CCNO have been reported to cause reduced number of multiple motile cilia along the cell surface due to disruption of Basal Bodies amplification and epithelial cell differentiation to multiciliated cells [81, 82]. These observations, highlight the importance collaboration between clinicians that encounter often peculiar and hard to explain (frequently syndromic) phenotypes and basic scientists that have the ability to model and explain these phenotypes in preclinical models.

A separate aspect of basic science that radically influences the clinical condition of the patient is the microbiome of the diseased lung. Given that PCD is characterized by failure of an innate defense mechanism (mucociliary clearance) and establishment of pathogens in the lung is uncomplicated, the microbiome of the lung is usually one of the key determinants of disease progression [489]. In PCD, only few studies have examined the lung microbiome of PCD patients [490] but undergoing collaborative projects in the field of PCD have characterized the study of the lung microbiome as top priority [491]. Lastly, although PCD is a rare disease and research in PCD can be considered as a narrow field of research, implications of research findings regarding ciliary motility may have widespread applications in other respiratory diseases. Ciliary functionality has been found to be affected in asthma [492] and COPD [493] in a condition that is often described as secondary dyskinesia, diseases with important prevalence worldwide and very high morbidity and mortality burden [494, 495]. The identification of novel compounds or methods that can improve ciliary

functionality not only in PCD patients but also in patients with secondary ciliary dyskinesia such as ciliostimulatory compounds could have significant consequences for millions of patients around the world.

For the purposes of this thesis, chapters 5 and 6 summarized the evidence regarding the diagnostic properties of TEM and nNO. A similar approach was not possible as very few studies have examined the diagnostic accuracy of HSVM although this technique has been used in several centers for many years [496]. HSVM is widely used in European centers but in the US, consideration about the subjectivity of HSVM interpretation limits its application in PCD diagnostic testing [16]. As additional evidence becomes available for HSVM similar meta-analytic approaches as those used in Chapter 5 and 6 can be utilized towards estimating pooled estimates of sensitivity and specificity for HSVM. Similar approaches could also be used in the case of the diagnostic tests that have only been developed recently but are increasingly gaining ground in PCD diagnosis such as genetic testing and immunofluorescence analysis. Important research needs were identified through the efforts to estimate the cost-effectiveness of different diagnostic algorithms as described in Chapter 7. The life expectancy in PCD is considered normal or near normal [455, 456] but there are no evidence to support this statement. The recently developed prospective PCD registry [497] provides a useful tool to follow pediatric and adult patients in time in order to monitor disease progression and life-expectancy in PCD and future studies will certainly focus on these outcomes and the factors that may influence them. Health utilities in PCD have never been studied and future studies can make use of the different direct methods hor health utility elicitation that have been developed such as the SG, VAS and TTO [297] or make use of standardized questionnaires that have been developed for indirect HU elicitation such as the EQ-5D [298]. Gaps in the knowledge regarding the efficacy of different treatment approaches are expected to slowly be bridged as the results of the first RCTs in PCD will become

available and evidence-based treatment recommendations will be made available for PCD [498, 499]. Lastly the financial costs that are bared by the individual or the healthcare system have never been studied in PCD patients. PCD is a chronic disease characterized by high morbidity but low mortality and it is likely that the costs involved in PCD management will be significant. Current treatment protocols, although mostly empirical and derived from CF, among other, advocate a multidisciplinary approach to PCD treatment (involving pediatric or adult pulmonologist, cardiologist, ENT specialist, physiotherapist and fertility specialist), performance of frequent microbiological assessment and less frequent assessment of bronchiectasis development through HRCT imaging [146]. A cost of illness analysis for PCD with a focus on the costs involved in running a PCD clinic similar to studies performed for CF [500, 501] will provide valuable information for future cost-effectiveness analyses in the field of PCD.

References

1. Knowles MR, Daniels LA, Davis SD, Zariwala MA, Leigh MW: **Primary ciliary dyskinesia. Recent advances in diagnostics, genetics, and characterization of clinical disease.** American journal of respiratory and critical care medicine 2013, **188**(8):913-922.

2. Fliegauf M, Benzing T, Omran H: **When cilia go bad: cilia defects and ciliopathies.** Nature reviews Molecular cell biology 2007, **8**(11):880-893.

3. Boon M, Jorissen M, Proesmans M, De Boeck K: **Primary ciliary dyskinesia, an orphan disease.** Eur J Pediatr 2013, **172**(2):151-162.

4. Pennekamp P, Menchen T, Dworniczak B, Hamada H: **Situs inversus and ciliary abnormalities: 20 years later, what is the connection?** Cilia 2015, **4**(1):1.

5. Kuehni CE, Frischer T, Strippoli MP, Maurer E, Bush A, Nielsen KG, Escribano A, Lucas JS, Yiallouros P, Omran H, Eber E, O'Callaghan C, Snijders D, Barbato A, ERS Task Force on Primary Ciliary Dyskinesia in Children: **Factors influencing age at diagnosis of primary ciliary dyskinesia in European children.** Eur Respir J 2010, **36**(6):1248-1258.

6. Kennedy MP, Omran H, Leigh MW, Dell S, Morgan L, Molina PL, Robinson BV, Minnix SL, Olbrich H, Severin T, Ahrens P, Lange L, Morillas HN, Noone PG, Zariwala MA, Knowles MR: **Congenital heart disease and other heterotaxic defects in a large cohort of patients with primary ciliary dyskinesia.** Circulation 2007, **115**(22):2814-2821.

7. Leigh MW, Ferkol TW, Davis SD, Lee H, Rosenfeld M, Dell SD, Sagel SD, Milla C, Olivier KN, Sullivan KM: **Clinical features and associated likelihood of primary ciliary dyskinesia in children and adolescents.** Annals of the American Thoracic Society 2016, **13**(8):1305-1313.

8. Goutaki M, Meier AB, Halbeisen FS, Lucas JS, Dell SD, Maurer E, Casaulta C, Jurca M, Spycher BD, Kuehni CE: **Clinical manifestations in primary ciliary dyskinesia: systematic review and meta-analysis.** Eur Respir J 2016, **48**(4):1081-1095.

9. Munro NC, Currie DC, Lindsay KS, Ryder TA, Rutman A, Dewar A, Greenstone MA, Hendry WF, Cole PJ: **Fertility in men with primary ciliary dyskinesia presenting with respiratory infection.** Thorax 1994, **49**(7):684-687.

10. Conraads V, Galdermans DI, Kockx MM, Jacob WA, Van Schaardenburg C, Coolen D: **Ultrastructurally normal and motile spermatozoa in a fertile man with Kartagener's syndrome.** CHEST Journal 1992, **102**(5):1616-1618.

11. Leigh MW, O'callaghan C, Knowles MR: **The challenges of diagnosing primary ciliary dyskinesia.** Proceedings of the American Thoracic Society 2011, **8**(5):434-437.

12. O'Callaghan C, Chilvers M, Hogg C, Bush A, Lucas J: **Diagnosing primary ciliary dyskinesia.** Thorax 2007, **62**(8):656-657.

13. Boaretto F, Snijders D, Salvoro C, Spalletta A, Mostacciuolo ML, Collura M, Cazzato S, Girosi D, Silvestri M, Rossi GA: **Diagnosis of Primary Ciliary Dyskinesia by a Targeted Next-Generation Sequencing Panel: Molecular and Clinical Findings in Italian Patients.** The Journal of Molecular Diagnostics 2016, **18**(6):912-922.

14. Shoemark A, Frost E, Dixon M, Ollosson S, Kilpin K, Patel M, Scully J, Rogers AV, Mitchison HM, Bush A: **Accuracy of Immunofluorescence in the Diagnosis of Primary Ciliary Dyskinesia.** American Journal of Respiratory And Critical Care Medicine 2017, (ja).

15. Strippoli MP, Frischer T, Barbato A, Snijders D, Maurer E, Lucas JS, Eber E, Karadag B, Pohunek P, Zivkovic Z, Escribano A, O'Callaghan C, Bush A, Kuehni CE, ERS Task Force onPrimary Ciliary Dyskinesia in Children: **Management of primary ciliary dyskinesia in European children: recommendations and clinical practice.** Eur Respir J 2012, **39**(6):1482-1491.

16. Lucas JS, Leigh MW: **Diagnosis of primary ciliary dyskinesia: searching for a gold standard.** Eur Respir J 2014, **44**(6):1418-1422.

17. Kyriacou K, Yiallouros P, Kouis P, Papatheodorou S, Omran H: **Better experimental screening and treatment for primary ciliary dyskinesia: The FP7 BESTCILIA project.** Ultrastruct Pathol 2017, **41**(1):110-112.

18. Lucas JS, Barbato A, Collins SA, Goutaki M, Behan L, Caudri D, Dell S, Eber E, Escudier E, Hirst RA, Hogg C, Jorissen M, Latzin P, Legendre M, Leigh MW, Midulla F, Nielsen KG, Omran H, Papon JF, Pohunek P, Redfern B, Rigau D, Rindlisbacher B, Santamaria F, Shoemark A, Snijders D, Tonia T, Titieni A, Walker WT, Werner C, Bush A, Kuehni CE: **European Respiratory Society guidelines for the diagnosis of primary ciliary dyskinesia.** Eur Respir J 2017, **49**(1):10.1183/13993003.01090-2016. Print 2017 Jan.

19. Kouis P, Papatheodorou SI, Yiallouros PK: **Diagnostic accuracy of nasal nitric oxide for establishing diagnosis of primary ciliary dyskinesia: a meta-analysis.** BMC pulmonary medicine 2015, **15**(1):153.

20. Kouis P, Yiallouros PK, Middleton N, Evans JS, Kyriacou K, Papatheodorou SI: **Prevalence of primary ciliary dyskinesia in consecutive referrals of suspect cases and the transmission electron microscopy detection rate: a systematic review and meta-analysis.** Pediatr Res 2016, .

21. Ibanez-Tallon I, Heintz N, Omran H: **To beat or not to beat: roles of cilia in development and disease.** Hum Mol Genet 2003, **12 Spec No 1**:R27-35.

22. Luck DJ: **Genetic and biochemical dissection of the eucaryotic flagellum.** J Cell Biol 1984, **98**(3):789-794.

23. Ostrowski LE, Blackburn K, Radde KM, Moyer MB, Schlatzer DM, Moseley A, Boucher RC: **A proteomic analysis of human cilia: identification of novel components.** Mol Cell Proteomics 2002, **1**(6):451-465.

24. Leigh MW, Pittman JE, Carson JL, Ferkol TW, Dell SD, Davis SD, Knowles MR, Zariwala MA: **Clinical and genetic aspects of primary ciliary dyskinesia/Kartagener syndrome.** Genetics in Medicine 2009, **11**(7):473-487.

25. Singla V, Reiter JF: **The primary cilium as the cell's antenna: signaling at a sensory organelle.** Science 2006, **313**(5787):629-633.

26. Gerdes JM, Davis EE, Katsanis N: **The vertebrate primary cilium in development, homeostasis, and disease.** Cell 2009, **137**(1):32-45.

27. Veland IR, Awan A, Pedersen LB, Yoder BK, Christensen ST: **Primary cilia and signaling pathways in mammalian development, health and disease.** Nephron Physiol 2009, **111**(3):p39-53.

28. Nauli SM, Alenghat FJ, Luo Y, Williams E, Vassilev P, Li X, Elia AE, Lu W, Brown EM, Quinn SJ: **Polycystins 1 and 2 mediate mechanosensation in the primary cilium of kidney cells.** Nat Genet 2003, **33**(2):129-137.

29. Pan J, Snell W: **The primary cilium: keeper of the key to cell division.** Cell 2007, **129**(7):1255-1257.

30. Quarmby LM, Parker JD: **Cilia and the cell cycle?** J Cell Biol 2005, **169**(5):707-710.

31. Mans DA, Voest EE, Giles RH: **All along the watchtower: is the cilium a tumor suppressor organelle?** Biochimica et Biophysica Acta (BBA)-Reviews on Cancer 2008, **1786**(2):114-125.

32. Christensen ST, Pedersen LB, Schneider L, Satir P: **Sensory cilia and integration of signal transduction in human health and disease.** Traffic 2007, **8**(2):97-109.

33. Hirokawa N, Tanaka Y, Okada Y, Takeda S: **Nodal flow and the generation of left-right asymmetry.** Cell 2006, **125**(1):33-45.

34. Cohen MM: **The hedgehog signaling network.** American Journal of Medical Genetics Part A 2003, **123**(1):5-28.

35. Hildebrandt F, Benzing T, Katsanis N: **Ciliopathies.** N Engl J Med 2011, **364**(16):1533-1543.

36. Waters AM, Beales PL: **Ciliopathies: an expanding disease spectrum.** Pediatric Nephrology 2011, **26**(7):1039-1056.

37. Quinlan RJ, Tobin JL, Beales PL: **Modeling ciliopathies: Primary cilia in development and disease.** Curr Top Dev Biol 2008, **84**:249-310.

38. Braiman A, Priel Z: **Efficient mucociliary transport relies on efficient regulation of ciliary beating.** Respiratory physiology & neurobiology 2008, **163**(1):202-207.

39. Teilmann SC, Byskov AG, Pedersen PA, Wheatley DN, Pazour GJ, Christensen ST: **Localization of transient receptor potential ion channels in primary and motile cilia of the female murine reproductive organs.** Mol Reprod Dev 2005, **71**(4):444-452.

40. Ross AJ, May-Simera H, Eichers ER, Kai M, Hill J, Jagger DJ, Leitch CC, Chapple JP, Munro PM, Fisher S: **Disruption of Bardet-Biedl syndrome ciliary proteins perturbs planar cell polarity in vertebrates.** Nat Genet 2005, **37**(10):1135-1140.

41. Evans MJ, Fanucchi MV, Van Winkle LS, Baker GL, Murphy AE, Nishio SJ, Sannes PL, Plopper CG: **Fibroblast growth factor-2 during postnatal development of the tracheal basement membrane zone.** Am J Physiol Lung Cell Mol Physiol 2002, **283**(6):L1263-70.

42. Afzelius BA: **Situs inversus and ciliary abnormalities. What is the connection?** Int J Dev Biol 1995, **39**(5):839-844.

43. Brown NA, Wolpert L: **The development of handedness in left/right asymmetry.** Development 1990, **109**(1):1-9.

44. Wanner A, Salathe M, O'Riordan TG: **Mucociliary clearance in the airways.** Am J Respir Crit Care Med 1996, **154**(6 Pt 1):1868-1902.

45. Oberdörster G, Oberdörster E, Oberdörster J: **Nanotoxicology: an emerging discipline evolving from studies of ultrafine particles.** Environ Health Perspect 2005, :823-839.

46. Martin TR, Frevert CW: **Innate immunity in the lungs.** Proceedings of the American Thoracic Society 2005, **2**(5):403-411.

47. Wright JR: **Immunomodulatory functions of surfactant.** Physiol Rev 1997, **77**(4):931-962.

48. Houtmeyers E, Gosselink R, Gayan-Ramirez G, Decramer M: **Regulation of mucociliary clearance in health and disease.** European Respiratory Journal 1999, **13**(5):1177-1188.

49. Livraghi A, Randell SH: **Cystic fibrosis and other respiratory diseases of impaired mucus clearance.** Toxicol Pathol 2007, **35**(1):116-129.

50. Chilvers MA, Rutman A, O'callaghan C: **Ciliary beat pattern is associated with specific ultrastructural defects in primary ciliary dyskinesia.** J Allergy Clin Immunol 2003, **112**(3):518-524.

51. Guo H, Nawroth J, Ding Y, Kanso E: **Cilia beating patterns are not hydrodynamically optimal.** Phys Fluids 2014, **26**(9):091901.

52. Guirao B, Joanny J: **Spontaneous creation of macroscopic flow and metachronal waves in an array of cilia.** Biophys J 2007, **92**(6):1900-1917.

53. Munkholm M, Mortensen J: **Mucociliary clearance: pathophysiological aspects.** Clinical physiology and functional imaging 2014, **34**(3):171-177.

54. Stoltz DA, Meyerholz DK, Welsh MJ: **Origins of cystic fibrosis lung disease.** N Engl J Med 2015, **372**(4):351-362.

55. Siewert A: **Über einen Fall von Bronchiectasie bei einem Patienten mit situs inversus viscerum.** Berliner klinische Wochenschrift 1904, **41**:139-141.

56. Kartagener M: **Zur Pathogenese der Bronchiektasien. I. Mitteilung: Bronchiektasien bei Situs viscerum inversus.** Beiträge zur Klinik und Erforschung der Tuberkulose und der Lungenkrankenheiten 1933, **83**:489-501.

57. AFZELIUS B, ELIASSON R: **LACK OF DYNEIN ARMS IN IMMOTILE.** J Cell Biol 1975, **66**(9):225-232.

58. Afzelius BA: **A human syndrome caused by immotile cilia.** Science 1976, **193**(4250):317-319.

59. Pedersen M: **Specific types of abnormal ciliary motility in Kartagener's syndrome and analogous respiratory disorders. A quantified microphoto-oscillographic investigation of 27 patients.** Eur J Respir Dis Suppl 1983, **127**:78-90.

60. Kuehni CE, Frischer T, Strippoli MP, Maurer E, Bush A, Nielsen KG, Escribano A, Lucas JS, Yiallouros P, Omran H, Eber E, O'Callaghan C, Snijders D, Barbato A, ERS Task Force on Primary Ciliary Dyskinesia in Children: **Factors influencing age at diagnosis of primary ciliary dyskinesia in European children.** Eur Respir J 2010, **36**(6):1248-1258.

61. Marthin JK, Petersen N, Skovgaard LT, Nielsen KG: **Lung function in patients with primary ciliary dyskinesia: a cross-sectional and 3-decade longitudinal study.** American journal of respiratory and critical care medicine 2010, **181**(11):1262-1268.

62. Frija-Masson J, Bassinet L, Honore I, Dufeu N, Housset B, Coste A, Papon JF, Escudier E, Burgel PR, Maitre B: **Clinical characteristics, functional respiratory decline and follow-up in adult patients with primary ciliary dyskinesia.** Thorax 2017, **72**(2):154-160.

63. Yiallouros PK, Kouis P, Middleton N, Nearchou M, Adamidi T, Georgiou A, Eleftheriou A, Ioannou P, Hadjisavvas A, Kyriacou K: **Clinical features of primary ciliary dyskinesia in Cyprus with emphasis on lobectomized patients.** Respir Med 2015, **109**(3):347-356.

64. Ellerman A, Bisgaard H: **Longitudinal study of lung function in a cohort of primary ciliary dyskinesia.** Eur Respir J 1997, **10**(10):2376-2379.

65. Harness-Brumley CL, Elliott AC, Rosenbluth DB, Raghavan D, Jain R: **Gender differences in outcomes of patients with cystic fibrosis.** Journal of women's health 2014, **23**(12):1012-1020.

66. Chotirmall SH, Smith SG, Gunaratnam C, Cosgrove S, Dimitrov BD, O'neill SJ, Harvey BJ, Greene CM, McElvaney NG: **Effect of estrogen on pseudomonas mucoidy and exacerbations in cystic fibrosis.** N Engl J Med 2012, **366**(21):1978-1986.

67. Jain R, Ray JM, Pan J, Brody SL: **Sex hormone–dependent regulation of cilia beat frequency in airway epithelium.** American journal of respiratory cell and molecular biology 2012, **46**(4):446-453.

68. Pennarun G, Escudier E, Chapelin C, Bridoux A, Cacheux V, Roger G, Clément A, Goossens M, Amselem S, Duriez B: **Loss-of-function mutations in a human gene related to Chlamydomonas reinhardtii dynein IC78 result in primary ciliary dyskinesia.** The American Journal of Human Genetics 1999, **65**(6):1508-1519.

69. Olbrich H, Häffner K, Kispert A, Völkel A, Volz A, Sasmaz G, Reinhardt R, Hennig S, Lehrach H, Konietzko N: **Mutations in DNAH5 cause primary ciliary dyskinesia and randomization of left–right asymmetry.** Nat Genet 2002, **30**(2):143-144.

70. Hornef N, Olbrich H, Horvath J, Zariwala MA, Fliegauf M, Loges NT, Wildhaber J, Noone PG, Kennedy M, Antonarakis SE: **DNAH5 mutations are a common cause of primary ciliary dyskinesia with outer dynein arm defects.** American journal of respiratory and critical care medicine 2006, **174**(2):120-126.

71. Dougherty GW, Loges NT, Klinkenbusch JA, Olbrich H, Pennekamp P, Menchen T, Raidt J, Wallmeier J, Werner C, Westermann C: **DNAH11 localization in the proximal region of respiratory cilia defines distinct outer dynein arm complexes.** American journal of respiratory cell and molecular biology 2016, **55**(2):213-224.

72. Pifferi M, Michelucci A, Conidi ME, Cangiotti AM, Simi P, Macchia P, Boner AL: **New DNAH11 mutations in primary ciliary dyskinesia with normal axonemal ultrastructure.** Eur Respir J 2010, **35**(6):1413-1416.

73. Oda T, Yanagisawa H, Yagi T, Kikkawa M: **Mechanosignaling between central apparatus and radial spokes controls axonemal dynein activity.** J Cell Biol 2014, **204**(5):807-819.

74. Yang P, Diener DR, Yang C, Kohno T, Pazour GJ, Dienes JM, Agrin NS, King SM, Sale WS, Kamiya R, Rosenbaum JL, Witman GB: **Radial spoke proteins of Chlamydomonas flagella.** J Cell Sci 2006, **119**(Pt 6):1165-1174.

75. Castleman VH, Romio L, Chodhari R, Hirst RA, de Castro SC, Parker KA, Ybot-Gonzalez P, Emes RD, Wilson SW, Wallis C: **Mutations in radial spoke head protein genes RSPH9 and RSPH4A cause primary ciliary dyskinesia with central-microtubular-pair abnormalities.** The American Journal of Human Genetics 2009, **84**(2):197-209.

76. Olbrich H, Cremers C, Loges NT, Werner C, Nielsen KG, Marthin JK, Philipsen M, Wallmeier J, Pennekamp P, Menchen T: **Loss-of-function GAS8 mutations cause primary ciliary dyskinesia and disrupt the nexin-dynein regulatory complex.** The American Journal of Human Genetics 2015, **97**(4):546-554.

77. Heuser T, Raytchev M, Krell J, Porter ME, Nicastro D: **The dynein regulatory complex is the nexin link and a major regulatory node in cilia and flagella.** J Cell Biol 2009, **187**(6):921-933.

78. Olbrich H, Schmidts M, Werner C, Onoufriadis A, Loges NT, Raidt J, Banki NF, Shoemark A, Burgoyne T, Al Turki S: **Recessive HYDIN mutations cause primary ciliary dyskinesia without randomization of left-right body asymmetry.** The American Journal of Human Genetics 2012, **91**(4):672-684.

79. Kott E, Duquesnoy P, Copin B, Legendre M, Dastot-Le Moal F, Montantin G, Jeanson L, Tamalet A, Papon J, Siffroi J: **Loss-of-function mutations in LRRC6, a gene essential for proper axonemal assembly of inner and outer dynein arms, cause primary ciliary dyskinesia.** The American Journal of Human Genetics 2012, **91**(5):958-964.

80. Horani A, Ferkol TW, Shoseyov D, Wasserman MG, Oren YS, Kerem B, Amirav I, Cohen-Cymberknoh M, Dutcher SK, Brody SL: **LRRC6 mutation causes primary ciliary dyskinesia with dynein arm defects.** PLoS One 2013, **8**(3):e59436.

81. Wallmeier J, Al-Mutairi DA, Chen C, Loges NT, Pennekamp P, Menchen T, Ma L, Shamseldin HE, Olbrich H, Dougherty GW: **Mutations in CCNO result in congenital mucociliary clearance disorder with reduced generation of multiple motile cilia.** Nat Genet 2014, **46**(6):646-651.

82. Boon M, Wallmeier J, Ma L, Loges NT, Jaspers M, Olbrich H, Dougherty GW, Raidt J, Werner C, Amirav I: **MCIDAS mutations result in a mucociliary clearance disorder with reduced generation of multiple motile cilia.** Nature communications 2014, **5**.

83. Paff T, Loges NT, Aprea I, Wu K, Bakey Z, Haarman EG, Daniels JM, Sistermans EA, Bogunovic N, Dougherty GW: **Mutations in PIH1D3 Cause X-Linked Primary Ciliary Dyskinesia with Outer and Inner Dynein Arm Defects.** The American Journal of Human Genetics 2017, **100**(1):160-168.

84. Li Y, Yagi H, Onuoha EO, Damerla RR, Francis R, Furutani Y, Tariq M, King SM, Hendricks G, Cui C: **DNAH6 and its interactions with PCD genes in heterotaxy and primary ciliary dyskinesia.** PLoS genetics 2016, **12**(2):e1005821.

85. Kurkowiak M, Zietkiewicz E, Witt M: **Recent advances in primary ciliary dyskinesia genetics.** J Med Genet 2015, **52**(1):1-9.

86. Lai M, Pifferi M, Bush A, Piras M, Michelucci A, Di Cicco M, del Grosso A, Quaranta P, Cursi C, Tantillo E, Franceschi S, Mazzanti MC, Simi P, Saggese G, Boner A, Pistello M: **Gene editing of**

DNAH11 restores normal cilia motility in primary ciliary dyskinesia. J Med Genet 2016, **53**(4):242-249.

87. Ostrowski LE, Yin W, Patel M, Sechelski J, Rogers T, Burns K, Grubb BR, Olsen JC: **Restoring ciliary function to differentiated primary ciliary dyskinesia cells with a lentiviral vector.** Gene Ther 2014, **21**(3):253-261.

88. Chhin B, Negre D, Merrot O, Pham J, Tourneur Y, Ressnikoff D, Jaspers M, Jorissen M, Cosset F, Bouvagnet P: **Ciliary beating recovery in deficient human airway epithelial cells after lentivirus ex vivo gene therapy.** PLoS genetics 2009, **5**(3):e1000422.

89. Phillips MI, Burns AB: **The emergence of gene therapy for rare diseases.** Expert Opinion on Orphan Drugs 2014, **2**(11):1197-1209.

90. Cox DBT, Platt RJ, Zhang F: **Therapeutic genome editing: prospects and challenges.** Nat Med 2015, **21**(2):121-131.

91. Weinberg MS, Morris KV: **Transcriptional gene silencing in humans.** Nucleic Acids Res 2016, **44**(14):6505-6517.

92. Marson FA, Bertuzzo CS, Ribeiro JD: **Personalized or precision medicine? The example of cystic fibrosis.** Frontiers in pharmacology 2017, **8**.

93. Kispert A, Petry M, Olbrich H, Volz A, Ketelsen UP, Horvath J, Melkaoui R, Omran H, Zariwala M, Noone PG, Knowles M: **Genotype-phenotype correlations in PCD patients carrying DNAH5 mutations.** Thorax 2003, **58**(6):552-554.

94. Davis SD, Ferkol TW, Rosenfeld M, Lee H, Dell SD, Sagel SD, Milla C, Zariwala MA, Pittman JE, Shapiro AJ: **Clinical features of childhood primary ciliary dyskinesia by genotype and ultrastructural phenotype.** American journal of respiratory and critical care medicine 2015, **191**(3):316-324.

95. Loges NT, Olbrich H, Fenske L, Mussaffi H, Horvath J, Fliegauf M, Kuhl H, Baktai G, Peterffy E, Chodhari R: **DNAI2 mutations cause primary ciliary dyskinesia with defects in the outer dynein arm.** The American Journal of Human Genetics 2008, **83**(5):547-558.

96. Mazor M, Alkrinawi S, Chalifa-Caspi V, Manor E, Sheffield VC, Aviram M, Parvari R: **Primary ciliary dyskinesia caused by homozygous mutation in DNAL1, encoding dynein light chain 1.** The American Journal of Human Genetics 2011, **88**(5):599-607.

97. Duriez B, Duquesnoy P, Escudier E, Bridoux AM, Escalier D, Rayet I, Marcos E, Vojtek AM, Bercher JF, Amselem S: **A common variant in combination with a nonsense mutation in a member of the thioredoxin family causes primary ciliary dyskinesia.** Proc Natl Acad Sci U S A 2007, **104**(9):3336-3341.

98. Hjeij R, Lindstrand A, Francis R, Zariwala MA, Liu X, Li Y, Damerla R, Dougherty GW, Abouhamed M, Olbrich H: **ARMC4 mutations cause primary ciliary dyskinesia with randomization of left/right body asymmetry.** The American Journal of Human Genetics 2013, **93**(2):357-367.

99. Loges NT, Olbrich H, Becker-Heck A, Häffner K, Heer A, Reinhard C, Schmidts M, Kispert A, Zariwala MA, Leigh MW: **Deletions and point mutations of LRRC50 cause primary ciliary**

dyskinesia due to dynein arm defects. The American Journal of Human Genetics 2009, **85**(6):883-889.

100. Omran H, Kobayashi D, Olbrich H, Tsukahara T, Loges NT, Hagiwara H, Zhang Q, Leblond G, O'toole E, Hara C: **Ktu/PF13 is required for cytoplasmic pre-assembly of axonemal dyneins.** Nature 2008, **456**(7222):611-616.

101. Mitchison HM, Schmidts M, Loges NT, Freshour J, Dritsoula A, Hirst RA, O'Callaghan C, Blau H, Al Dabbagh M, Olbrich H: **Mutations in axonemal dynein assembly factor DNAAF3 cause primary ciliary dyskinesia.** Nat Genet 2012, **44**(4):381-389.

102. Horani A, Druley TE, Zariwala MA, Patel AC, Levinson BT, Van Arendonk LG, Thornton KC, Giacalone JC, Albee AJ, Wilson KS: **Whole-exome capture and sequencing identifies HEATR2 mutation as a cause of primary ciliary dyskinesia.** The American Journal of Human Genetics 2012, **91**(4):685-693.

103. Tarkar A, Loges NT, Slagle CE, Francis R, Dougherty GW, Tamayo JV, Shook B, Cantino M, Schwartz D, Jahnke C: **DYX1C1 is required for axonemal dynein assembly and ciliary motility.** Nat Genet 2013, **45**(9):995-1003.

104. Zariwala MA, Gee HY, Kurkowiak M, Al-Mutairi DA, Leigh MW, Hurd TW, Hjeij R, Dell SD, Chaki M, Dougherty GW: **ZMYND10 is mutated in primary ciliary dyskinesia and interacts with LRRC6.** The American Journal of Human Genetics 2013, **93**(2):336-345.

105. Knowles MR, Ostrowski LE, Loges NT, Hurd T, Leigh MW, Huang L, Wolf WE, Carson JL, Hazucha MJ, Yin W: **Mutations in SPAG1 cause primary ciliary dyskinesia associated with defective outer and inner dynein arms.** The American Journal of Human Genetics 2013, **93**(4):711-720.

106. Austin-Tse C, Halbritter J, Zariwala MA, Gilberti RM, Gee HY, Hellman N, Pathak N, Liu Y, Panizzi JR, Patel-King RS: **Zebrafish ciliopathy screen plus human mutational analysis identifies C21orf59 and CCDC65 defects as causing primary ciliary dyskinesia.** The American Journal of Human Genetics 2013, **93**(4):672-686.

107. Merveille A, Davis EE, Becker-Heck A, Legendre M, Amirav I, Bataille G, Belmont J, Beydon N, Billen F, Clément A: **CCDC39 is required for assembly of inner dynein arms and the dynein regulatory complex and for normal ciliary motility in humans and dogs.** Nat Genet 2011, **43**(1):72-78.

108. Becker-Heck A, Zohn IE, Okabe N, Pollock A, Lenhart KB, Sullivan-Brown J, McSheene J, Loges NT, Olbrich H, Haeffner K: **The coiled-coil domain containing protein CCDC40 is essential for motile cilia function and left-right axis formation.** Nat Genet 2011, **43**(1):79-84.

109. Knowles MR, Leigh MW, Ostrowski LE, Huang L, Carson JL, Hazucha MJ, Yin W, Berg JS, Davis SD, Dell SD: **Exome sequencing identifies mutations in CCDC114 as a cause of primary ciliary dyskinesia.** The American Journal of Human Genetics 2013, **92**(1):99-106.

110. Wirschell M, Olbrich H, Werner C, Tritschler D, Bower R, Sale WS, Loges NT, Pennekamp P, Lindberg S, Stenram U: **The nexin-dynein regulatory complex subunit DRC1 is essential for motile cilia function in algae and humans.** Nat Genet 2013, **45**(3):262-268.

111. Hjeij R, Onoufriadis A, Watson CM, Slagle CE, Klena NT, Dougherty GW, Kurkowiak M, Loges NT, Diggle CP, Morante NF: **CCDC151 mutations cause primary ciliary dyskinesia by**

disruption of the outer dynein arm docking complex formation. The American Journal of Human Genetics 2014, **95**(3):257-274.

112. Panizzi JR, Becker-Heck A, Castleman VH, Al-Mutairi DA, Liu Y, Loges NT, Pathak N, Austin-Tse C, Sheridan E, Schmidts M: **CCDC103 mutations cause primary ciliary dyskinesia by disrupting assembly of ciliary dynein arms.** Nat Genet 2012, **44**(6):714-719.

113. Kott E, Legendre M, Copin B, Papon J, Dastot-Le Moal F, Montantin G, Duquesnoy P, Piterboth W, Amram D, Bassinet L: **Loss-of-function mutations in RSPH1 cause primary ciliary dyskinesia with central-complex and radial-spoke defects.** The American Journal of Human Genetics 2013, **93**(3):561-570.

114. Castleman VH, Romio L, Chodhari R, Hirst RA, de Castro SC, Parker KA, Ybot-Gonzalez P, Emes RD, Wilson SW, Wallis C: **Mutations in radial spoke head protein genes RSPH9 and RSPH4A cause primary ciliary dyskinesia with central-microtubular-pair abnormalities.** The American Journal of Human Genetics 2009, **84**(2):197-209.

115. Moore A, Escudier E, Roger G, Tamalet A, Pelosse B, Marlin S, Clement A, Geremek M, Delaisi B, Bridoux AM, Coste A, Witt M, Duriez B, Amselem S: **RPGR is mutated in patients with a complex X linked phenotype combining primary ciliary dyskinesia and retinitis pigmentosa.** J Med Genet 2006, **43**(4):326-333.

116. Budny B, Chen W, Omran H, Fliegauf M, Tzschach A, Wisniewska M, Jensen LR, Raynaud M, Shoichet SA, Badura M: **A novel X-linked recessive mental retardation syndrome comprising macrocephaly and ciliary dysfunction is allelic to oral–facial–digital type I syndrome.** Hum Genet 2006, **120**(2):171-178.

117. El Khouri E, Thomas L, Jeanson L, Bequignon E, Vallette B, Duquesnoy P, Montantin G, Copin B, Dastot-Le Moal F, Blanchon S: **Mutations in DNAJB13, encoding an HSP40 family member, cause primary ciliary dyskinesia and male infertility.** The American Journal of Human Genetics 2016, **99**(2):489-500.

118. Jeanson L, Copin B, Papon J, Dastot-Le Moal F, Duquesnoy P, Montantin G, Cadranel J, Corvol H, Coste A, Désir J: **RSPH3 mutations cause primary ciliary dyskinesia with central-complex defects and a near absence of radial spokes.** The American Journal of Human Genetics 2015, **97**(1):153-162.

119. Wallmeier J, Shiratori H, Dougherty GW, Edelbusch C, Hjeij R, Loges NT, Menchen T, Olbrich H, Pennekamp P, Raidt J: **TTC25 deficiency results in defects of the outer dynein arm docking machinery and primary ciliary dyskinesia with left-right body asymmetry randomization.** The American Journal of Human Genetics 2016, **99**(2):460-469.

120. Emiralioglu N, Wallmeier J, Olbrich H, Omran H, Ozcelik U: **DYNC2H1 mutation causes Jeune syndrome and recurrent lung infections associated with ciliopathy.** The Clinical Respiratory Journal 2017, .

121. Ostrowski LE, Dutcher SK, Lo CW: **Cilia and models for studying structure and function.** Proceedings of the American Thoracic Society 2011, **8**(5):423-429.

122. Huang B, Piperno G, Luck DJ: **Paralyzed flagella mutants of Chlamydomonas reinhardtii. Defective for axonemal doublet microtubule arms.** J Biol Chem 1979, **254**(8):3091-3099.

123. Rajagopalan V, Corpuz EO, Hubenschmidt MJ, Townsend CR, Asai DJ, Wilkes DE: **Analysis of properties of cilia using Tetrahymena thermophila.** Cytoskeleton Methods and Protocols 2010, :283-299.

124. Werner M, Mitchell B: **Understanding ciliated epithelia: the power of Xenopus.** Genesis 2012, **50**(3):176-185.

125. Kobayashi D, Takeda H: **Primary Ciliary Dyskinesia in Fish: The Analysis of a Novel Medaka Mutant Kintoun.** In *Medaka.* Edited by Anonymous Springer; 2011:131-156.

126. Kramer-Zucker AG, Olale F, Haycraft CJ, Yoder BK, Schier AF, Drummond IA: **Cilia-driven fluid flow in the zebrafish pronephros, brain and Kupffer's vesicle is required for normal organogenesis.** Development 2005, **132**(8):1907-1921.

127. Norris DP, Grimes DT: **Mouse models of ciliopathies: the state of the art.** Dis Model Mech 2012, **5**(3):299-312.

128. Torikata C, Kijimoto C, Koto M: **Ultrastructure of respiratory cilia of WIC-Hyd male rats. An animal model for human immotile cilia syndrome.** Am J Pathol 1991, **138**(2):341-347.

129. Pazour GJ, Agrin N, Leszyk J, Witman GB: **Proteomic analysis of a eukaryotic cilium.** J Cell Biol 2005, **170**(1):103-113.

130. Bayly P, Lewis B, Kemp P, Pless R, Dutcher S: **Efficient spatiotemporal analysis of the flagellar waveform of Chlamydomonas reinhardtii.** Cytoskeleton 2010, **67**(1):56-69.

131. Huang B, Ramanis Z, Luck DJ: **Suppressor mutations in Chlamydomonas reveal a regulatory mechanism for flagellar function.** Cell 1982, **28**(1):115-124.

132. Howard DR, Habermacher G, Glass DB, Smith EF, Sale WS: **Regulation of Chlamydomonas flagellar dynein by an axonemal protein kinase.** J Cell Biol 1994, **127**(6 Pt 1):1683-1692.

133. Wyatt TA, Spurzem JR, May K, Sisson JH: **Regulation of ciliary beat frequency by both PKA and PKG in bovine airway epithelial cells.** Am J Physiol 1998, **275**(4 Pt 1):L827-35.

134. Li D, Shirakami G, Zhan X, Johns RA: **Regulation of ciliary beat frequency by the nitric oxide–cyclic guanosine monophosphate signaling pathway in rat airway epithelial cells.** American Journal of Respiratory Cell and Molecular Biology 2000, **23**(2):175-181.

135. Hayashi T, Kawakami M, Sasaki S, Katsumata T, Mori H, Yoshida H, Nakahari T: **ATP regulation of ciliary beat frequency in rat tracheal and distal airway epithelium.** Exp Physiol 2005, **90**(4):535-544.

136. Wallingford JB: **Planar cell polarity signaling, cilia and polarized ciliary beating.** Curr Opin Cell Biol 2010, **22**(5):597-604.

137. Antic D, Stubbs JL, Suyama K, Kintner C, Scott MP, Axelrod JD: **Planar cell polarity enables posterior localization of nodal cilia and left-right axis determination during mouse and Xenopus embryogenesis.** PloS one 2010, **5**(2):e8999.

138. McIntyre JC, Davis EE, Joiner A, Williams CL, Tsai I, Jenkins PM, McEwen DP, Zhang L, Escobado J, Thomas S: **Gene therapy rescues cilia defects and restores olfactory function in a mammalian ciliopathy model.** Nat Med 2012, **18**(9):1423-1428.

139. Tobin JL, Beales PL: **Restoration of renal function in zebrafish models of ciliopathies.** Pediatric nephrology 2008, **23**(11):2095-2099.

140. Rachel RA, May-Simera HL, Veleri S, Gotoh N, Choi BY, Murga-Zamalloa C, McIntyre JC, Marek J, Lopez I, Hackett AN, Zhang J, Brooks M, den Hollander AI, Beales PL, Li T, Jacobson SG, Sood R, Martens JR, Liu P, Friedman TB, Khanna H, Koenekoop RK, Kelley MW, Swaroop A: **Combining Cep290 and Mkks ciliopathy alleles in mice rescues sensory defects and restores ciliogenesis.** J Clin Invest 2012, **122**(4):1233-1245.

141. Ibañez-Tallon I, Gorokhova S, Heintz N: **Loss of function of axonemal dynein Mdnah5 causes primary ciliary dyskinesia and hydrocephalus.** Hum Mol Genet 2002, **11**(6):715-721.

142. Tan SY, Rosenthal J, Zhao XQ, Francis RJ, Chatterjee B, Sabol SL, Linask KL, Bracero L, Connelly PS, Daniels MP, Yu Q, Omran H, Leatherbury L, Lo CW: **Heterotaxy and complex structural heart defects in a mutant mouse model of primary ciliary dyskinesia.** J Clin Invest 2007, **117**(12):3742-3752.

143. Lucas JS, Burgess A, Mitchison HM, Moya E, Williamson M, Hogg C, National PCD Service U: **Diagnosis and management of primary ciliary dyskinesia.** Arch Dis Child 2014, **99**(9):850-856.

144. Canciani M, Barlocco E, Mastella G, De Santi M, Gardi C, Lungarella G: **The saccharin method for testing mucociliary function in patients suspected of having primary ciliary dyskinesia.** Pediatr Pulmonol 1988, **5**(4):210-214.

145. Bush A, Cole P, Hariri M, Mackay I, Phillips G, O'Callaghan C, Wilson R, Warner JO: **Primary ciliary dyskinesia: diagnosis and standards of care.** Eur Respir J 1998, **12**(4):982-988.

146. Barbato A, Frischer T, Kuehni CE, Snijders D, Azevedo I, Baktai G, Bartoloni L, Eber E, Escribano A, Haarman E, Hesselmar B, Hogg C, Jorissen M, Lucas J, Nielsen KG, O'Callaghan C, Omran H, Pohunek P, Strippoli MP, Bush A: **Primary ciliary dyskinesia: a consensus statement on diagnostic and treatment approaches in children.** Eur Respir J 2009, **34**(6):1264-1276.

147. Fliegauf M, Olbrich H, Horvath J, Wildhaber JH, Zariwala MA, Kennedy M, Knowles MR, Omran H: **Mislocalization of DNAH5 and DNAH9 in respiratory cells from patients with primary ciliary dyskinesia.** American journal of respiratory and critical care medicine 2005, **171**(12):1343-1349.

148. Horani A, Ferkol TW, Dutcher SK, Brody SL: **Genetics and biology of primary ciliary dyskinesia.** Paediatric respiratory reviews 2016, **18**:18-24.

149. Stannard WA, Chilvers MA, Rutman AR, Williams CD, O'callaghan C: **Diagnostic testing of patients suspected of primary ciliary dyskinesia.** American journal of respiratory and critical care medicine 2010, **181**(4):307-314.

150. Sisson JH, Stoner J, Ammons B, Wyatt TA: **All-digital image capture and whole-field analysis of ciliary beat frequency.** J Microsc 2003, **211**(2):103-111.

151. Olm MA, Kogler JE,Jr, Macchione M, Shoemark A, Saldiva PH, Rodrigues JC: **Primary ciliary dyskinesia: evaluation using cilia beat frequency assessment via spectral analysis of digital microscopy images.** J Appl Physiol (1985) 2011, **111**(1):295-302.

152. Dimova S, Maes F, Brewster ME, Jorissen M, Noppe M, Augustijns P: **High-speed digital imaging method for ciliary beat frequency measurement.** J Pharm Pharmacol 2005, **57**(4):521-526.

153. Raidt J, Wallmeier J, Hjeij R, Onnebrink JG, Pennekamp P, Loges NT, Olbrich H, Haffner K, Dougherty GW, Omran H, Werner C: **Ciliary beat pattern and frequency in genetic variants of primary ciliary dyskinesia.** Eur Respir J 2014, **44**(6):1579-1588.

154. Papon JF, Coste A, Roudot-Thoraval F, Boucherat M, Roger G, Tamalet A, Vojtek AM, Amselem S, Escudier E: **A 20-year experience of electron microscopy in the diagnosis of primary ciliary dyskinesia.** Eur Respir J 2010, **35**(5):1057-1063.

155. Lundberg J, Farkas-Szallasi T, Weitzberg E, Rinder J, Lidholm J, Änggård A, Hökfelt T, Lundberg J, Alving K: **High nitric oxide production in human paranasal sinuses.** Nat Med 1995, **1**(4):370-373.

156. Collins SA, Gove K, Walker W, Lucas JS: **Nasal nitric oxide screening for primary ciliary dyskinesia: systematic review and meta-analysis.** Eur Respir J 2014, **44**(6):1589-1599.

157. Ricciardolo FL: **Multiple roles of nitric oxide in the airways.** Thorax 2003, **58**(2):175-182.

158. Zihlif N, Paraskakis E, Tripoli C, Lex C, Bush A: **Markers of airway inflammation in primary ciliary dyskinesia studied using exhaled breath condensate.** Pediatr Pulmonol 2006, **41**(6):509-514.

159. Walker WT, Jackson CL, Lackie PM, Hogg C, Lucas JS: **Nitric oxide in primary ciliary dyskinesia.** Eur Respir J 2012, **40**(4):1024-1032.

160. Csoma Z, Bush A, Wilson NM, Donnelly L, Balint B, Barnes PJ, Kharitonov SA: **Nitric oxide metabolites are not reduced in exhaled breath condensate of patients with primary ciliary dyskinesia.** CHEST Journal 2003, **124**(2):633-638.

161. Pifferi M, Bush A, Caramella D, Di Cicco M, Zangani M, Chinellato I, Macchia P, Boner AL: **Agenesis of paranasal sinuses and nasal nitric oxide in primary ciliary dyskinesia.** Eur Respir J 2011, **37**(3):566-571.

162. Santamaria F, De Stefano S, Montella S, Barbarano F, Iacotucci P, Ciccarelli R, Sofia M, Maniscalco M: **Nasal nitric oxide assessment in primary ciliary dyskinesia using aspiration, exhalation, and humming.** Med Sci Monit 2008, **14**(2):CR80-85.

163. Pifferi M, Bush A, Maggi F, Michelucci A, Ricci V, Conidi ME, Cangiotti AM, Bodini A, Simi P, Macchia P, Boner AL: **Nasal nitric oxide and nitric oxide synthase expression in primary ciliary dyskinesia.** Eur Respir J 2011, **37**(3):572-577.

164. Gücüyener K, Ergenekon E, Erbas D, Pinarli G, Serdaroğlu A: **The serum nitric oxide levels in patients with Duchenne muscular dystrophy.** Brain and Development 2000, **22**(3):181-183.

165. Barnes PJ, Loukides S, Kharitonov S, Wodehouse T, Cole PJ: **Effect of arginine on mucociliary function in primary ciliary dyskinesia.** The Lancet 1998, **352**(9125):371-372.

166. Grasemann H, Gärtig S, Wiesemann H, Teschler H, Konietzko N, Ratjen F: **Effect of L-arginine infusion on airway NO in cystic fibrosis and primary ciliary dyskinesia syndrome.** European Respiratory Journal 1999, **13**(1):114-118.

167. Torgersen J: **Transposition of Viscera–Bronchiectasis and Nasal Polyps: A Genetical Analysis and a Contribution to the Problem of Constitution.** Acta Radiol 1947, (1):17-24.

168. Afzelius BA, Stenram U: **Prevalence and genetics of immotile-cilia syndrome and left-handedness.** Int J Dev Biol 2006, **50**(6):571-573.

169. Katsuhara K, Kawamoto S, Wakabayashi T, Belsky JL: **Situs inversus totalis and Kartagener's syndrome in a Japanese population.** Chest 1972, **61**(1):56-61.

170. Polineni D, Davis SD, Dell SD: **Treatment recommendations in primary ciliary dyskinesia.** Paediatric respiratory reviews 2016, **18**:39-45.

171. Anonymous *Proceedings of the Seminars in Respiratory and Critical Care Medicine:* Copyright© 2003 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, NY 10001, USA. Tel.: 1 (212) 584-4662; 2003.

172. Phillips GE, Thomas S, Heather S, Bush A: **Airway response of children with primary ciliary dyskinesia to exercise and beta2-agonist challenge.** Eur Respir J 1998, **11**(6):1389-1391.

173. Paff T, Daniels JM, Weersink EJ, Lutter R, Vonk Noordegraaf A, Haarman EG: **A randomised controlled trial on the effect of inhaled hypertonic saline on quality of life in primary ciliary dyskinesia.** Eur Respir J 2017, **49**(2):10.1183/13993003.01770-2016. Print 2017 Feb.

174. Smit HJ, Schreurs AJ, Van den Bosch, Jules MM, Westermann CJ: **Is resection of bronchiectasis beneficial in patients with primary ciliary dyskinesia?** Chest 1996, **109**(6):1541-1544.

175. Anonymous *Proceedings of the BMC Proceedings:* BioMed Central; 2016.

176. Sidney Filho LA, Machuca TN, Camargo JdJ, Felicetti JC, Camargo SM, Perin FA, Sanchez LB, Schio SM: **Lung transplantation without the use of cardiopulmonary bypass in a patient with Kartagener syndrome.** Jornal Brasileiro de Pneumologia 2012, **38**(6):806-809.

177. Hayes D,Jr, Reynolds SD, Tumin D: **Outcomes of lung transplantation for primary ciliary dyskinesia and Kartagener syndrome.** J Heart Lung Transplant 2016, **35**(11):1377-1378.

178. Date H, Yamashita M, Nagahiro I, Aoe M, Andou A, Shimizu N: **Living-donor lobar lung transplantation for primary ciliary dyskinesia.** Ann Thorac Surg 2001, **71**(6):2008-2009.

179. Hughes D: **Primary ciliary dyskinesia.** Paediatr Child Health 2008, **13**(8):672-674.

180. Dell SD: **Primary ciliary dyskinesia: Myths and realities.** Paediatr Child Health 2008, **13**(8):668-670.

181. Maglione M, Bush A, Nielsen KG, Hogg C, Montella S, Marthin JK, Di Giorgio A, Santamaria F: **Multicenter analysis of body mass index, lung function, and sputum microbiology in primary ciliary dyskinesia.** Pediatr Pulmonol 2014, **49**(12):1243-1250.

182. Schunemann HJ, Dorn J, Grant BJ, Winkelstein W, Trevisan M: **Pulmonary function is a long-term predictor of mortality in the general population: 29-year follow-up of the Buffalo Health Study.** CHEST Journal 2000, **118**(3):656-664.

183. Mannino DM, Buist AS, Petty TL, Enright PL, Redd SC: **Lung function and mortality in the United States: data from the First National Health and Nutrition Examination Survey follow up study.** Thorax 2003, **58**(5):388-393.

184. Riekert KA, Bartlett SJ, Boyle MP, Krishnan JA, Rand CS: **The association between depression, lung function, and health-related quality of life among adults with cystic fibrosis.** Chest Journal 2007, **132**(1):231-237.

185. Abbott J, Hurley MA, Morton AM, Conway SP: **Longitudinal association between lung function and health-related quality of life in cystic fibrosis.** Thorax 2013, **68**(2):149-154.

186. Ketelaars CA, Schlosser MA, Mostert R, Huyer Abu-Saad H, Halfens RJ, Wouters EF: **Determinants of health-related quality of life in patients with chronic obstructive pulmonary disease.** Thorax 1996, **51**(1):39-43.

187. Pereira EDB, Pinto R, Alcantara M, Medeiros M, Mota RMS: **Influence of respiratory function parameters on the quality of life of COPD patients.** Jornal Brasileiro de Pneumologia 2009, **35**(8):730-736.

188. Birch DS, Saleh HA, Wodehouse T, Simpson IN, Mackay IS: **Assessing the quality of life for patients with chronic rhinosinusitis using the "Rhinosinusitis Disability Index".** Rhinology 2001, **39**(4):191-196.

189. French CL, Irwin RS, Curley FJ, Krikorian CJ: **Impact of chronic cough on quality of life.** Arch Intern Med 1998, **158**(15):1657-1661.

190. Majithia A, Fong J, Hariri M, Harcourt J: **Hearing outcomes in children with primary ciliary dyskinesia—a longitudinal study.** Int J Pediatr Otorhinolaryngol 2005, **69**(8):1061-1064.

191. Behan L, Rubbo B, Lucas JS, Galvin AD: **The patient's experience of primary ciliary dyskinesia: a systematic review.** Quality of Life Research 2017, :1-21.

192. Pifferi M, Bush A, Di Cicco M, Pradal U, Ragazzo V, Macchia P, Boner AL: **Health-related quality of life and unmet needs in patients with primary ciliary dyskinesia.** Eur Respir J 2010, **35**(4):787-794.

193. McManus IC, Mitchison HM, Chung EM, Stubbings GF, Martin N: **Primary ciliary dyskinesia (Siewert's/Kartagener's syndrome): respiratory symptoms and psycho-social impact.** BMC Pulmonary Medicine 2003, **3**(1):4.

194. Behan L, Leigh MW, Dell SD, Dunn Galvin A, Quittner AL, Lucas JS: **Validation of a health-related quality of life instrument for primary ciliary dyskinesia (QOL-PCD).** Thorax 2017, .

195. Dell SD, Leigh MW, Lucas JS, Ferkol TW, Knowles MR, Alpern A, Behan L, Morris AM, Hogg C, DunnGalvin A: **Primary ciliary dyskinesia: first health-related quality-of-life measures for pediatric patients.** Annals of the American Thoracic Society 2016, **13**(10):1726-1735.

196. Bakker C, van der Linden S: **Health related utility measurement: an introduction.** J Rheumatol 1995, **22**(6):1197-1199.

197. Rabin R, Charro Fd: **EQ-SD: a measure of health status from the EuroQol Group.** Ann Med 2001, **33**(5):337-343.

198. Bradley JM, Blume SW, Balp MM, Honeybourne D, Elborn JS: **Quality of life and healthcare utilisation in cystic fibrosis: a multicentre study.** Eur Respir J 2013, **41**(3):571-577.

199. Khan K, Kunz R, Kleijnen J, Antes G: *Systematic reviews to support evidence-based medicine:* Crc Press; 2011.

200. Grant MJ, Booth A: **A typology of reviews: an analysis of 14 review types and associated methodologies.** Health Information & Libraries Journal 2009, **26**(2):91-108.

201. Oxman AD, Guyatt GH: **The science of reviewing research.** Ann N Y Acad Sci 1993, **703**(1):125-134.

202. Higgins JP, Green S: *Cochrane handbook for systematic reviews of interventions:* John Wiley & Sons; 2011.

203. Haidich AB: **Meta-analysis in medical research.** Hippokratia 2010, **14**(Suppl 1):29-37.

204. Meline T: **Selecting studies for systematic review: Inclusion and exclusion criteria.** Contemporary issues in communication science and disorders 2006, **33**(21-27).

205. Chang SM, Matchar DB, Smetana GW, Umscheid CA: **Methods guide for medical test reviews.** Rockville: Agency for Healthcare Research and Quality 2012, .

206. Lowe HJ, Barnett GO: **Understanding and using the medical subject headings (MeSH) vocabulary to perform literature searches.** JAMA 1994, **271**(14):1103-1108.

207. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne JA, Cochrane Bias Methods Group, Cochrane Statistical Methods Group: **The Cochrane Collaboration's tool for assessing risk of bias in randomised trials.** BMJ 2011, **343**:d5928.

208. Egger M, Dickersin K, Smith GD: **Problems and limitations in conducting systematic reviews.** Systematic Reviews in Health Care: Meta-Analysis in Context, Second Edition 2001, :43-68.

209. Gluud LL: **Bias in clinical intervention research.** Am J Epidemiol 2006, **163**(6):493-501.

210. Hammer GP, du Prel J, Blettner M: **Avoiding bias in observational studies.** Dtsch Arzteblatt Int 2009, **106**:664-668.

211. Brueton V, Tierney J, Stenning S, Nazareth I, Meredith S, Harding S, Rait G: **Strategies to reduce attrition in randomised trials.** Trials 2011, **12**(1):A128.

212. Song F, Hooper L, Loke Y: **Publication bias: what is it? How do we measure it? How do we avoid it?** Open Access Journal of Clinical Trials 2013, **2013**(5):71-81.

213. Skelly AC, Dettori JR, Brodt ED: **Assessing bias: the importance of considering confounding.** Evidence-based spine-care journal 2012, **3**(01):9-12.

214. Zeng X, Zhang Y, Kwong JS, Zhang C, Li S, Sun F, Niu Y, Du L: **The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: a systematic review.** Journal of Evidence-Based Medicine 2015, **8**(1):2-10.

215. Sanderson S, Tatt ID, Higgins JP: **Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography.** Int J Epidemiol 2007, **36**(3):666-676.

216. Slim K, Nini E, Forestier D, Kwiatkowski F, Panis Y, Chipponi J: **Methodological index for non-randomized studies (MINORS): development and validation of a new instrument.** ANZ J Surg 2003, **73**(9):712-716.

217. Stang A: **Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses.** Eur J Epidemiol 2010, **25**(9):603-605.

218. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J: **The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews.** BMC medical research methodology 2003, **3**(1):25.

219. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM: **QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies.** Ann Intern Med 2011, **155**(8):529-536.

220. Hooijmans CR, Rovers MM, de Vries RB, Leenaars M, Ritskes-Hoitinga M, Langendam MW: **SYRCLE's risk of bias tool for animal studies.** BMC medical research methodology 2014, **14**(1):43.

221. Vandenbroucke JP, Von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, Poole C, Schlesselman JJ, Egger M, Strobe Initiative: **Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration.** PLoS Med 2007, **4**(10):e297.

222. Schulz KF, Altman DG, Moher D: **CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials.** BMC medicine 2010, **8**(1):18.

223. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Moher D, Rennie D, de Vet HC, Lijmer JG, Standards for Reporting of Diagnostic Accuracy: **The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration.** Clin Chem 2003, **49**(1):7-18.

224. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF: **Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement.** Revista Española de Salud Pública 2000, **74**(2):00-00.

225. Moher D, Liberati A, Tetzlaff J, Altman DG, Prisma Group: **Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.** PLoS med 2009, **6**(7):e1000097.

226. Berman NG, Parker RA: **Meta-analysis: neither quick nor easy.** BMC Medical Research Methodology 2002, **2**(1):10.

227. Gandini S, Boniol M, Haukka J, Byrnes G, Cox B, Sneyd MJ, Mullie P, Autier P: **Meta-analysis of observational studies of serum 25-hydroxyvitamin D levels and colorectal, breast and prostate cancer and colorectal adenoma.** International Journal of Cancer 2011, **128**(6):1414-1424.

228. Cipriani A, Hawton K, Stockton S, Geddes JR: **Lithium in the prevention of suicide in mood disorders: updated systematic review and meta-analysis.** BMJ 2013, **346**:f3646.

229. Tomlinson CL, Patel S, Meek C, Herd CP, Clarke CE, Stowe R, Shah L, Sackley CM, Deane KH, Wheatley K: **Physiotherapy versus placebo or no intervention in Parkinson's disease.** The Cochrane Library 2013, .

230. Zhang MW, Ho RC, Cheung MW, Fu E, Mak A: **Prevalence of depressive symptoms in patients with chronic obstructive pulmonary disease: a systematic review, meta-analysis and meta-regression.** Gen Hosp Psychiatry 2011, **33**(3):217-223.

231. Verhagen AP, Ferreira ML: **Forest plots.** Journal of physiotherapy 2014, **60**(3):170-173.

232. Spector TD, Thompson SG: **The potential and limitations of meta-analysis.** J Epidemiol Community Health 1991, **45**(2):89-92.

233. Jin Z, Zhou X, He J: **Statistical methods for dealing with publication bias in meta-analysis.** Stat Med 2015, **34**(2):343-360.

234. Harrison F: **Getting started with meta-analysis.** Methods in Ecology and Evolution 2011, **2**(1):1-10.

235. Greco T, Zangrillo A, Biondi-Zoccai G, Landoni G: **Meta-analysis: pitfalls and hints.** Heart Lung Vessel 2013, **5**(4):219-225.

236. Schwarzer G, Carpenter JR, Rücker G: **Small-Study Effects in Meta-Analysis.** In *Meta-Analysis with R.* Edited by Anonymous Springer; 2015:107-141.

237. Sterne JA, Egger M, Smith GD: **Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis.** BMJ 2001, **323**(7304):101-105.

238. Deeks JJ: **Systematic reviews of evaluations of diagnostic and screening tests.** BMJ: British Medical Journal 2001, **323**(7305):157.

239. Borenstein M, Hedges L, Rothstein H: **Meta-analysis: Fixed effect vs. random effects.** Meta-analysis.com 2007, .

240. Deeks JJ, Altman DG, Bradburn MJ: **Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis.** Systematic Reviews in Health Care: Meta-Analysis in Context, Second Edition 2008, :285-312.

241. Mantel N, Haenszel W: **Statistical aspects of the analysis of data from retrospective studies of disease.** 1959, .

242. DerSimonian R, Laird N: **Meta-analysis in clinical trials.** Control Clin Trials 1986, **7**(3):177-188.

243. Melsen W, Bootsma M, Rovers M, Bonten M: **The effects of clinical and statistical heterogeneity on the predictive values of results from meta-analyses.** Clinical Microbiology and Infection 2014, **20**(2):123-129.

244. Cochran WG: **The combination of estimates from different experiments.** Biometrics 1954, **10**(1):101-129.

245. Hoaglin DC: **Misunderstandings about Q and 'Cochran's Q test'in meta-analysis.** Stat Med 2016, **35**(4):485-495.

246. Higgins J, Thompson SG: **Quantifying heterogeneity in a meta-analysis.** Stat Med 2002, **21**(11):1539-1558.

247. Huedo-Medina TB, Sánchez-Meca J, Marín-Martínez F, Botella J: **Assessing heterogeneity in meta-analysis: Q statistic or I² index?** Psychol Methods 2006, **11**(2):193.

248. Song F, Sheldon TA, Sutton AJ, Abrams KR, Jones DR: **Methods for exploring heterogeneity in meta-analysis.** Eval Health Prof 2001, **24**(2):126-151.

249. Harris R, Bradburn M, Deeks J, Harbord R, Altman D, Sterne J: **Metan: fixed-and random-effects meta-analysis.** Stata journal 2008, **8**(1):3.

250. Nyaga VN, Arbyn M, Aerts M: **Metaprop: a Stata command to perform meta-analysis of binomial data.** Archives of Public Health 2014, **72**(1):39.

251. Newcombe RG: **Two-sided confidence intervals for the single proportion: comparison of seven methods.** Stat Med 1998, **17**(8):857-872.

252. Brown LD, Cai TT, DasGupta A: **Interval estimation for a binomial proportion.** Statistical science 2001, :101-117.

253. Freeman MF, Tukey JW: **Transformations related to the angular and the square root.** The Annals of Mathematical Statistics 1950, :607-611.

254. Miller JJ: **The inverse of the Freeman–Tukey double arcsine transformation.** The American Statistician 1978, **32**(4):138-138.

255. Hunink MM, Weinstein MC, Wittenberg E, Drummond MF, Pliskin JS, Wong JB, Glasziou PP: *Decision making in health and medicine: integrating evidence and values:* Cambridge University Press; 2014.

256. Kerr WT, Lau EP, Owens GE, Trefler A: **The future of medical diagnostics: large digitized databases.** Yale J Biol Med 2012, **85**(3):363-377.

257. Altman DG, Bland JM: **Diagnostic tests. 1: Sensitivity and specificity.** BMJ 1994, **308**(6943):1552.

258. Cardoso JR, Pereira LM, Iversen MD, Ramos AL: **What is gold standard and what is ground truth?** Dental press journal of orthodontics 2014, **19**(5):27-30.

259. Claassen JA: **The gold standard: not a golden standard.** BMJ: British Medical Journal 2005, **330**(7500):1121.

260. Sox HC: *Medical decision making:* ACP Press; 1988.

261. Parikh R, Mathai A, Parikh S, Chandra Sekhar G, Thomas R: **Understanding and using sensitivity, specificity and predictive values.** Indian J Ophthalmol 2008, **56**(1):45-50.

262. Hajian-Tilaki K: **Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation.** Caspian J Intern Med 2013, **4**(2):627-635.

263. Park SH, Goo JM, Jo CH: **Receiver operating characteristic (ROC) curve: practical review for radiologists.** Korean J Radiol 2004, **5**(1):11-18.

264. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y: **Cochrane handbook for systematic reviews of diagnostic test accuracy.** Version 0.9.0.London: The Cochrane Collaboration 2010, .

265. Cook C, Cleland J, Huijbregts P: **Creation and critique of studies of diagnostic accuracy: use of the STARD and QUADAS methodological quality assessment tools.** Journal of Manual & Manipulative Therapy 2007, **15**(2):93-102.

266. Harbord RM, Whiting P, Sterne JA, Egger M, Deeks JJ, Shang A, Bachmann LM: **An empirical comparison of methods for meta-analysis of diagnostic accuracy showed hierarchical models are necessary.** J Clin Epidemiol 2008, **61**(11):1095-1103.

267. Moses LE, Shapiro D, Littenberg B: **Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations.** Stat Med 1993, **12**(14):1293-1316.

268. Rutter CM, Gatsonis CA: **A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations.** Stat Med 2001, **20**(19):2865-2884.

269. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA: **A unification of models for meta-analysis of diagnostic accuracy studies.** Biostatistics 2007, **8**(2):239-251.

270. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH: **Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews.** J Clin Epidemiol 2005, **58**(10):982-990.

271. Chu H, Cole SR: **Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach.** J Clin Epidemiol 2006, **59**(12):1331-1332.

272. Guideline E: **Meta-analysis of diagnostic test accuracy studies.** European Network for Health Technology Assessment 2014, .

273. Harbord RM, Whiting P: **metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression.** Stata Journal 2009, **9**(2):211.

274. Anonymous *Proceedings of the Seminars in Spine Surgery:* Elsevier; 2009.

275. Sonnenberg FA, Beck JR: **Markov models in medical decision making: a practical guide.** Med Decis Making 1993, **13**(4):322-338.

276. Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Möller J: **Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–4.** Medical decision making 2012, **32**(5):701-711.

277. Pitman R, Fisman D, Zaric GS, Postma M, Kretzschmar M, Edmunds J, Brisson M, ISPOR-SMDM Modeling Good Research Practices Task Force: **Dynamic transmission modeling: a report of the ISPOR-SMDM modeling good research practices task force-5.** Value in health 2012, **15**(6):828-834.

278. Robinson R: **Economic evaluation and health care. What does it mean?** BMJ 1993, **307**(6905):670-673.

279. Briggs AH, O'Brien BJ: **The death of cost-minimization analysis?** Health Econ 2001, **10**(2):179-184.

280. Robinson R: **Cost-benefit analysis.** BMJ 1993, **307**(6909):924-926.

281. Hutton J: **Cost-benefit analysis in health care expenditure decision-making.** Health Econ 1992, **1**(4):213-216.

282. McIntosh E, Donaldson C, Ryan M: **Recent advances in the methods of cost-benefit analysis in healthcare.** Pharmacoeconomics 1999, **15**(4):357-367.

283. Cookson R: **Willingness to pay methods in health care: a sceptical view.** Health Econ 2003, **12**(11):891-894.

284. Elixhauser A: *Health Care Cost Benefit and Cost Effectiveness Analysis:(CBA/CEA); from 1979 to 1990: a Bibliography:* Lippincott; 1993.

285. Robinson R: **Cost-effectiveness analysis.** BMJ 1993, **307**(6907):793-795.

286. Polimeni JM, Vichansavakul K, Iorgulescu RI, Chandrasekara R: **Why perspective matters in health outcomes research analyses.** The International Business & Economics Research Journal (Online) 2013, **12**(11):1503.

287. Neumann PJ: **Costing and perspective in published cost-effectiveness analysis.** Med Care 2009, **47**(7 Suppl 1):S28-32.

288. Drummond M, Sculpher M, Torrance G, O'Brien B, Stoddart G: **Methods for the economic evaluation of health care programmes: Oxford University Press.** New York 2005, .

289. Adam T, Evans DB, Koopmanschap MA: **Cost-effectiveness analysis: can we reduce variability in costing methods?** Int J Technol Assess Health Care 2003, **19**(02):407-420.

290. Gray AM, Clarke PM, Wolstenholme JL, Wordsworth S: *Applied methods of cost-effectiveness analysis in healthcare:* OUP Oxford; 2010.

291. Crawley E, Mills N, Hollingworth W, Deans Z, Sterne JA, Donovan JL, Beasant L, Montgomery A: **Comparing specialist medical care with specialist medical care plus the Lightning Process® for chronic fatigue syndrome or myalgic encephalomyelitis (CFS/ME): study protocol for a randomised controlled trial (SMILE Trial).** Trials 2013, **14**(1):444.

292. Cairns J: *Discounting in economic evaluation:* Oxford University Press: New York; 2001.

293. Weinstein MC, Torrance G, McGuire A: **QALYs: the basics.** Value in health 2009, **12**(s1):S5-S9.

294. Whitehead SJ, Ali S: **Health outcomes in economic evaluation: the QALY and utilities.** Br Med Bull 2010, **96**:5-21.

295. Sassi F: **Calculating QALYs, comparing QALY and DALY calculations.** Health Policy Plan 2006, **21**(5):402-408.

296. Ware Jr JE, Sherbourne CD: **The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection.** Med Care 1992, :473-483.

297. Torrance GW: **Measurement of health state utilities for economic appraisal: a review.** J Health Econ 1986, **5**(1):1-30.

298. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X: **Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L).** Quality of life research 2011, **20**(10):1727-1736.

299. Cho S, Kim H, Kim S, Ock M, Oh Y, Jo M: **Utility estimation of hypothetical chronic obstructive pulmonary disease health states by the general population and health professionals.** Health and quality of life outcomes 2015, **13**(1):34.

300. Buitinga L, Braakman-Jansen LM, Taal E, van de Laar, Mart AFJ: **Construct validity of the interview Time Trade-Off and computer Time Trade-Off in patients with rheumatoid arthritis: A cross-sectional observational pilot study.** BMC musculoskeletal disorders 2012, **13**(1):112.

301. Torrance GW: **Utility approach to measuring health-related quality of life.** J Chronic Dis 1987, **40**(6):593-600.

302. Briggs AH, Gray AM: **Handling uncertainty in economic evaluations of healthcare interventions.** BMJ: British Medical Journal 1999, **319**(7210):635.

303. Elgart JF, Caporale JE, Gonzalez L, Aiello E, Waschbusch M, Gagliardino JJ: **Treatment of type 2 diabetes with saxagliptin: a pharmacoeconomic evaluation in Argentina.** Health economics review 2013, **3**(1):11.

304. York Health Economics Consortium, York: **Two Way Sensitivity Analysis [online].** http://www.yhec.co.uk/glossary/two-way-sensitivity-analysis/ 2016, .

305. Fenwick E, Claxton K, Sculpher M: **Representing uncertainty: the role of cost-effectiveness acceptability curves.** Health Econ 2001, **10**(8):779-787.

306. Scott J, Palmer S, Paykel E, Teasdale J, Hayhurst H: **Use of cognitive therapy for relapse prevention in chronic depression. Cost-effectiveness study.** Br J Psychiatry 2003, **182**:221-227.

307. Leigh MW, Pittman JE, Carson JL, Ferkol TW, Dell SD, Davis SD, Knowles MR, Zariwala MA: **Clinical and genetic aspects of primary ciliary dyskinesia/Kartagener syndrome.** Genetics in Medicine 2009, **11**(7):473-487.

308. Noone PG, Leigh MW, Sannuti A, Minnix SL, Carson JL, Hazucha M, Zariwala MA, Knowles MR: **Primary ciliary dyskinesia: diagnostic and phenotypic features.** American journal of respiratory and critical care medicine 2004, **169**(4):459-467.

309. Boon M, Jorissen M, Proesmans M, De Boeck K: **Primary ciliary dyskinesia, an orphan disease.** Eur J Pediatr 2013, **172**(2):151-162.

310. Kim RH, A. Hall D, Cutz E, Knowles MR, Nelligan KA, Nykamp K, Zariwala MA, Dell SD: **The Role of Molecular Genetic Analysis in the Diagnosis of Primary Ciliary Dyskinesia.** Annals of the American Thoracic Society 2014, **11**(3):351-359.

311. Stannard WA, Chilvers MA, Rutman AR, Williams CD, O'Callaghan C: **Diagnostic testing of patients suspected of primary ciliary dyskinesia.** American journal of respiratory and critical care medicine 2010, **181**(4):307-314.

312. Papon JF, Coste A, Roudot-Thoraval F, Boucherat M, Roger G, Tamalet A, Vojtek AM, Amselem S, Escudier E: **A 20-year experience of electron microscopy in the diagnosis of primary ciliary dyskinesia.** Eur Respir J 2010, **35**(5):1057-1063.

313. Leigh MW, Hazucha MJ, Chawla KK, Baker BR, Shapiro AJ, Brown DE, LaVange LM, Horton BJ, Qaqish B, Carson JL: **Standardizing nasal nitric oxide measurement as a test for primary ciliary dyskinesia.** Annals of the American Thoracic Society 2013, **10**(6):574-581.

314. Schwabe GC, Hoffmann K, Loges NT, Birker D, Rossier C, De Santi MM, Olbrich H, Fliegauf M, Failly M, Liebers U: **Primary ciliary dyskinesia associated with normal axoneme ultrastructure is caused by DNAH11 mutations.** Hum Mutat 2008, **29**(2):289-298.

315. Lucas JS, Adam EC, Goggin PM, Jackson CL, Powles-Glover N, Patel SH, Humphreys J, Fray MD, Falconnet E, Blouin J: **Static respiratory cilia associated with mutations in Dnahc11/DNAH11: a mouse model of PCD.** Hum Mutat 2012, **33**(3):495-503.

316. Raidt J, Wallmeier J, Hjeij R, Onnebrink JG, Pennekamp P, Loges NT, Olbrich H, Haffner K, Dougherty GW, Omran H, Werner C: **Ciliary beat pattern and frequency in genetic variants of primary ciliary dyskinesia.** Eur Respir J 2014, .

317. Lundberg J, Farkas-Szallasi T, Weitzberg E, Rinder J, Lidholm J, Änggård A, Hökfelt T, Lundberg J, Alving K: **High nitric oxide production in human paranasal sinuses.** Nat Med 1995, **1**(4):370-373.

318. Strippoli MP, Frischer T, Barbato A, Snijders D, Maurer E, Lucas JS, Eber E, Karadag B, Pohunek P, Zivkovic Z, Escribano A, O'Callaghan C, Bush A, Kuehni CE, ERS Task Force onPrimary Ciliary Dyskinesia in Children: **Management of primary ciliary dyskinesia in European children: recommendations and clinical practice.** Eur Respir J 2012, **39**(6):1482-1491.

319. American Thoracic Society, European Respiratory Society: **ATS/ERS recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005.** Am J Respir Crit Care Med 2005, **171**(8):912-930.

320. Harris A, Bhullar E, Gove K, Joslin R, Pelling J, Evans HJ, Walker WT, Lucas JS: **Validation of a portable nitric oxide analyzer for screening in primary ciliary dyskinesias.** BMC pulmonary medicine 2014, **14**(1):18.

321. Collins SA, Gove K, Walker W, Lucas JS: **Nasal nitric oxide screening for primary ciliary dyskinesia: systematic review and meta-analysis.** Eur Respir J 2014, .

322. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM: **QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies.** Ann Intern Med 2011, **155**(8):529-536.

323. Chu H, Cole SR: **Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach.** J Clin Epidemiol 2006, **59**(12):1331-1332.

324. Harbord R: **METANDI: Stata module to perform meta-analysis of diagnostic accuracy.** Statistical Software Components 2008, .

325. Deeks JJ, Altman DG: **Diagnostic tests 4: likelihood ratios.** BMJ 2004, **329**(7458):168-169.

326. Pifferi M, Bush A, Maggi F, Michelucci A, Ricci V, Conidi ME, Cangiotti AM, Bodini A, Simi P, Macchia P, Boner AL: **Nasal nitric oxide and nitric oxide synthase expression in primary ciliary dyskinesia.** Eur Respir J 2011, **37**(3):572-577.

327. Morgan L, Buddle L, Rogers P, Seccombe L, MacKenney K, Hughes L: **Nasal nitric oxide is useful in an Australian Primary Ciliary Dyskinesia clinic.** Cilia 2012, **1**(Suppl 1):P7.

328. Marthin JK, Nielsen KG: **Choice of nasal nitric oxide technique as first-line test for primary ciliary dyskinesia.** Eur Respir J 2011, **37**(3):559-565.

329. Corbelli R, Bringolf-Isler B, Amacher A, Sasse B, Spycher M, Hammer J: **Nasal nitric oxide measurements to screen children for primary ciliary dyskinesia.** CHEST Journal 2004, **126**(4):1054-1059.

330. Beydon N, Chambellan A, Alberti C, de Blic J, Clément A, Escudier E, Le Bourgeois M: **Technical and practical issues for tidal breathing measurements of nasal nitric oxide in children.** Pediatr Pulmonol 2015, :n/a-n/a.

331. Piacentini GL, Bodini A, Peroni D, Rigotti E, Pigozzi R, Pradal U, Boner AL: **Nasal nitric oxide for early diagnosis of primary ciliary dyskinesia: practical issues in children.** Respir Med 2008, **102**(4):541-547.

332. Narang I, Ersu R, Wilson NM, Bush A: **Nitric oxide in chronic airway inflammation in children: diagnostic use and pathophysiological significance.** Thorax 2002, **57**(7):586-589.

333. Montella S, Alving K, De Stefano S, Di Micco LL, Di Giorgio A, Santamaria F: **Nasal nitric oxide measurement using continuous aspiration by hand-held device discriminates patients with primary ciliary dyskinesia from healthy subjects.** European Respiratory Journal 2012, **40**(Suppl 56):P3338.

334. Santamaria F, De Stefano S, Montella S, Barbarano F, Iacotucci P, Ciccarelli R, Sofia M, Maniscalco M: **Nasal nitric oxide assessment in primary ciliary dyskinesia using aspiration, exhalation, and humming.** Med Sci Monit 2008, **14**(2):CR80-85.

335. Moreno Galdo A, Vizmanos Lamotte G, Reverte Bover C, Gartner S, Cobos Barroso N, Rovira Amigo S, Linan Cortes S, Lloreta Trull J, Busquets Monge R: **Value of nasal nitric oxide in the diagnosis of primary ciliary dyskinesia.** An Pediatr (Barc) 2010, **73**(2):88-93.

336. Boon M, Meyts I, Proesmans M, Vermeulen FL, Jorissen M, De Boeck K: **Diagnostic accuracy of nitric oxide measurements to detect primary ciliary dyskinesia.** Eur J Clin Invest 2014, **44**(5):477-485.

337. Mateos-Corral D, Coombs R, Grasemann H, Ratjen F, Dell SD: **Diagnostic value of nasal nitric oxide measured with non-velum closure techniques for children with primary ciliary dyskinesia.** J Pediatr 2011, **159**(3):420-424.

338. Moher D, Liberati A, Tetzlaff J, Altman DG: **Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.** Ann Intern Med 2009, **151**(4):264-269.

339. Barbato A, Frischer T, Kuehni CE, Snijders D, Azevedo I, Baktai G, Bartoloni L, Eber E, Escribano A, Haarman E, Hesselmar B, Hogg C, Jorissen M, Lucas J, Nielsen KG, O'Callaghan C, Omran H, Pohunek P, Strippoli MP, Bush A: **Primary ciliary dyskinesia: a consensus statement on diagnostic and treatment approaches in children.** Eur Respir J 2009, **34**(6):1264-1276.

340. Yiallouros PK, Kouis P, Middleton N, Nearchou M, Adamidi T, Georgiou A, Eleftheriou A, Ioannou P, Hadjisavvas A, Kyriacou K: **Clinical features of primary ciliary dyskinesia in Cyprus with emphasis on lobectomized patients.** Respir Med 2015, **109**(3):347-356.

341. Coren M, Meeks M, Morrison I, Buchdahl R, Bush A: **Primary ciliary dyskinesia: age at diagnosis and symptom history.** Acta Paediatrica 2002, **91**(6):667-669.

342. Maniscalco M, Laurentiis Gd, Weitzberg E, Lundberg J, Sofia M: **Validation study of nasal nitric oxide measurements using a hand-held electrochemical analyser.** Eur J Clin Invest 2008, **38**(3):197-200.

343. Marthin JK, Nielsen KG: **Hand-held tidal breathing nasal nitric oxide measurement–a promising targeted case-finding tool for the diagnosis of primary ciliary dyskinesia.** PloS one 2013, **8**(2):e57262.

344. Knowles MR, Leigh MW, Carson JL, Davis SD, Dell SD, Ferkol TW, Olivier KN, Sagel SD, Rosenfeld M, Burns KA, Minnix SL, Armstrong MC, Lori A, Hazucha MJ, Loges NT, Olbrich H, Becker-Heck A, Schmidts M, Werner C, Omran H, Zariwala MA, Genetic Disorders of Mucociliary Clearance Consortium: **Mutations of DNAH11 in patients with primary ciliary dyskinesia with normal ciliary ultrastructure.** Thorax 2012, **67**(5):433-441.

345. Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen, Jan HP, Bossuyt PM: **Empirical evidence of design-related bias in studies of diagnostic tests.** JAMA 1999, **282**(11):1061-1066.

346. Rutjes AW, Reitsma JB, Vandenbroucke JP, Glas AS, Bossuyt PM: **Case-control and two-gate designs in diagnostic accuracy studies.** Clin Chem 2005, **51**(8):1335-1341.

347. Boon M, Jorissen M, Proesmans M, De Boeck K: **Primary ciliary dyskinesia, an orphan disease.** Eur J Pediatr 2013, **172**(2):151-162.

348. Barbato A, Frischer T, Kuehni CE, Snijders D, Azevedo I, Baktai G, Bartoloni L, Eber E, Escribano A, Haarman E, Hesselmar B, Hogg C, Jorissen M, Lucas J, Nielsen KG, O'Callaghan C, Omran H, Pohunek P, Strippoli MP, Bush A: **Primary ciliary dyskinesia: a consensus statement on diagnostic and treatment approaches in children.** Eur Respir J 2009, **34**(6):1264-1276.

349. Lucas JS, Burgess A, Mitchison HM, Moya E, Williamson M, Hogg C, National PCD Service U: **Diagnosis and management of primary ciliary dyskinesia.** Arch Dis Child 2014, **99**(9):850-856.

350. Boon M, Meyts I, Proesmans M, Vermeulen FL, Jorissen M, De Boeck K: **Diagnostic accuracy of nitric oxide measurements to detect primary ciliary dyskinesia.** Eur J Clin Invest 2014, **44**(5):477-485.

351. Santamaria F, Santi Md, Grillo G, Sarnelli P, Caterino M, Greco L: **Ciliary motility at light microscopy: a screening technique for ciliary defects?** Acta paediatrica 1999, **88**(8):853-857.

352. Leigh MW, O'Callaghan C, Knowles MR: **The challenges of diagnosing primary ciliary dyskinesia.** Proceedings of the American Thoracic Society 2011, **8**(5):434-437.

353. Kim RH, A. Hall D, Cutz E, Knowles MR, Nelligan KA, Nykamp K, Zariwala MA, Dell SD: **The Role of Molecular Genetic Analysis in the Diagnosis of Primary Ciliary Dyskinesia.** Annals of the American Thoracic Society 2014, **11**(3):351-359.

354. Lucas JS, Paff T, Goggin P, Haarman E: **Diagnostic methods in primary ciliary dyskinesia.** Paediatric respiratory reviews 2015, .

355. Hogg C: **Primary ciliary dyskinesia: when to suspect the diagnosis and how to confirm it.** Paediatric respiratory reviews 2009, **10**(2):44-50.

356. Lucas JS, Leigh MW: **Diagnosis of primary ciliary dyskinesia: searching for a gold standard.** Eur Respir J 2014, **44**(6):1418-1422.

357. Sox HC: *Medical decision making:* ACP Press; 1988.

358. Horani A, Brody SL, Ferkol TW: **Picking up speed: advances in the genetics of primary ciliary dyskinesia.** Pediatr Res 2013, **75**(1-2):158-164.

359. Schwabe GC, Hoffmann K, Loges NT, Birker D, Rossier C, De Santi MM, Olbrich H, Fliegauf M, Failly M, Liebers U: **Primary ciliary dyskinesia associated with normal axoneme ultrastructure is caused by DNAH11 mutations.** Hum Mutat 2008, **29**(2):289-298.

360. Olbrich H, Cremers C, Loges NT, Werner C, Nielsen KG, Marthin JK, Philipsen M, Wallmeier J, Pennekamp P, Menchen T: **Loss-of-Function GAS8 Mutations Cause Primary Ciliary Dyskinesia and Disrupt the Nexin-Dynein Regulatory Complex.** The American Journal of Human Genetics 2015, **97**(4):546-554.

361. Stannard WA, Chilvers MA, Rutman AR, Williams CD, O'Callaghan C: **Diagnostic testing of patients suspected of primary ciliary dyskinesia.** American journal of respiratory and critical care medicine 2010, **181**(4):307-314.

362. Leigh MW, Hazucha MJ, Chawla KK, Baker BR, Shapiro AJ, Brown DE, LaVange LM, Horton BJ, Qaqish B, Carson JL: **Standardizing nasal nitric oxide measurement as a test for primary ciliary dyskinesia.** Annals of the American Thoracic Society 2013, **10**(6):574-581.

363. Boon M, Smits A, Cuppens H, Jaspers M, Proesmans M, Dupont LJ, Vermeulen FL, Van Daele S, Malfroot A, Godding V: **Primary ciliary dyskinesia: critical evaluation of clinical symptoms and diagnosis in patients with normal and abnormal ultrastructure.** Orphanet J Rare Dis 2014, **9**(11).

364. Yiallouros PK, Kouis P, Middleton N, Nearchou M, Adamidi T, Georgiou A, Eleftheriou A, Ioannou P, Hadjisavvas A, Kyriacou K: **Clinical features of primary ciliary dyskinesia in Cyprus with emphasis on lobectomized patients.** Respir Med 2015, **109**(3):347-356.

365. Werner C, Onnebrink JG, Omran H: **Diagnosis and management of primary ciliary dyskinesia.** Cilia 2015, **4**(1):2-014-0011-8. eCollection 2015.

366. Moher D, Liberati A, Tetzlaff J, Altman DG: **Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.** Ann Intern Med 2009, **151**(4):264-269.

367. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB: **Meta-analysis of observational studies in epidemiology: a proposal for reporting.** JAMA 2000, **283**(15):2008-2012.

368. Nyaga VN, Arbyn M, Aerts M: **Metaprop: a Stata command to perform meta-analysis of binomial data.** Arch Public Health 2014, **72**(1):39-3258-72-39. eCollection 2014.

369. Borenstein M, Hedges L, Rothstein H: **Meta-analysis: fixed effect vs. random effects.** Meta-Analysis.com 2007, .

370. IntHout J, Ioannidis JP, Borm GF: **The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method.** BMC medical research methodology 2014, **14**(1):1.

371. Higgins JP, Thompson SG: **Quantifying heterogeneity in a meta-analysis.** Stat Med 2002, **21**(11):1539-1558.

372. O'Callaghan C, Rutman A, Williams G, Hirst R: **Inner dynein arm defects causing primary ciliary dyskinesia: repeat testing required.** European Respiratory Journal 2011, **38**(3):603-607.

373. Kurkowiak M, Zietkiewicz E, Witt M: **Recent advances in primary ciliary dyskinesia genetics.** J Med Genet 2015, **52**(1):1-9.

374. Shoemark A, Dixon M, Corrin B, Dewar A: **Twenty-year review of quantitative transmission electron microscopy for the diagnosis of primary ciliary dyskinesia.** J Clin Pathol 2012, **65**(3):267-271.

375. Jackson CL, Behan L, Collins SA, Goggin PM, Adam EC, Coles JL, Evans HJ, Harris A, Lackie P, Packham S, Page A, Thompson J, Walker WT, Kuehni C, Lucas JS: **Accuracy of diagnostic testing in primary ciliary dyskinesia.** Eur Respir J 2016, **47**(3):837-848.

376. Davis S, Olivier K, Ferkol T, Dell S, Sagel S, Rosenfeld M, Milla C, Atkinson J, Shapiro M: **Clinical Symptoms Associated With Primary Ciliary Dyskinesia–Results Of A Multi-Centered Study.** Am J Respir Crit Care Med 2010, **181**:A6728.

377. Olm MA, Kogler JE,Jr, Macchione M, Shoemark A, Saldiva PH, Rodrigues JC: **Primary ciliary dyskinesia: evaluation using cilia beat frequency assessment via spectral analysis of digital microscopy images.** J Appl Physiol (1985) 2011, **111**(1):295-302.

378. Pifferi M, Bush A, Caramella D, Di Cicco M, Zangani M, Chinellato I, Macchia P, Boner AL: **Agenesis of paranasal sinuses and nasal nitric oxide in primary ciliary dyskinesia.** Eur Respir J 2011, **37**(3):566-571.

379. Djakow J, Svobodová T, Hrach K, Uhlík J, Cinek O, Pohunek P: **Effectiveness of sequencing selected exons of DNAH5 and DNAI1 in diagnosis of primary ciliary dyskinesia.** Pediatr Pulmonol 2012, **47**(9):864-875.

380. Nauta F, Pals G, Daniels H, Nagelkerke A, Haarman E: **Diagnosis of primary ciliary dyskinesia in a Dutch cohort of 63 pediatric patients: An overview.** European Respiratory Journal 2011, **38**(Suppl 55):p511.

381. O'Callaghan C, Chilvers M, Hogg C, Bush A, Lucas J: **Diagnosing primary ciliary dyskinesia.** Thorax 2007, **62**(8):656-657.

382. Kuehni CE, Frischer T, Strippoli MP, Maurer E, Bush A, Nielsen KG, Escribano A, Lucas JS, Yiallouros P, Omran H, Eber E, O'Callaghan C, Snijders D, Barbato A, ERS Task Force on Primary Ciliary Dyskinesia in Children: **Factors influencing age at diagnosis of primary ciliary dyskinesia in European children.** Eur Respir J 2010, **36**(6):1248-1258.

383. Leigh MW, Ferkol TW, Davis SD, Lee H, Rosenfeld M, Dell SD, Sagel SD, Milla C, Olivier KN, Sullivan KM: **Clinical Features and Associated Likelihood of Primary Ciliary Dyskinesia in Children and Adolescents.** Annals of the American Thoracic Society 2016, (ja).

384. Behan L, Dimitrov BD, Kuehni CE, Hogg C, Carroll M, Evans HJ, Goutaki M, Harris A, Packham S, Walker WT, Lucas JS: **PICADAR: a diagnostic predictive tool for primary ciliary dyskinesia.** Eur Respir J 2016, **47**(4):1103-1112.

385. Onoufriadis A, Shoemark A, Schmidts M, Patel M, Jimenez G, Liu H, Thomas B, Dixon M, Hirst RA, Rutman A, Burgoyne T, Williams C, Scully J, Bolard F, Lafitte JJ, Beales PL, Hogg C, Yang P, Chung EM, Emes RD, O'Callaghan C, UK10K, Bouvagnet P, Mitchison HM: **Targeted NGS gene panel identifies mutations in RSPH1 causing primary ciliary dyskinesia and a common mechanism for ciliary central pair agenesis due to radial spoke defects.** Hum Mol Genet 2014, **23**(13):3362-3374.

386. Castleman VH, Romio L, Chodhari R, Hirst RA, de Castro SC, Parker KA, Ybot-Gonzalez P, Emes RD, Wilson SW, Wallis C: **Mutations in radial spoke head protein genes RSPH9 and RSPH4A cause primary ciliary dyskinesia with central-microtubular-pair abnormalities.** The American Journal of Human Genetics 2009, **84**(2):197-209.

387. Olin JT, Burns K, Carson JL, Metjian H, Atkinson JJ, Davis SD, Dell SD, Ferkol TW, Milla CE, Olivier KN: **Diagnostic yield of nasal scrape biopsies in primary ciliary dyskinesia: a multicenter experience.** Pediatr Pulmonol 2011, **46**(5):483-488.

388. Kouis P, Papatheodorou SI, Yiallouros PK: **Diagnostic accuracy of nasal nitric oxide for establishing diagnosis of primary ciliary dyskinesia: a meta-analysis.** BMC Pulmonary Medicine 2015, **15**(1):153.

389. Boon M, Jorissen M, Jaspers M, Cuppens H, De Boeck K: **Is the sensitivity of primary ciliary dyskinesia detection by ciliary function analysis 100%?** Eur Respir J 2013, **42**(4):1159-1161.

390. Horani A, Ferkol TW, Dutcher SK, Brody SL: **Genetics and biology of primary ciliary dyskinesia.** Paediatric respiratory reviews 2016, **18**:18-24.

391. Davis SD, Ferkol TW, Rosenfeld M, Lee H, Dell SD, Sagel SD, Milla C, Zariwala MA, Pittman JE, Shapiro AJ: **Clinical features of childhood primary ciliary dyskinesia by genotype and ultrastructural phenotype.** American journal of respiratory and critical care medicine 2015, **191**(3):316-324.

392. Knowles MR, Ostrowski LE, Leigh MW, Sears PR, Davis SD, Wolf WE, Hazucha MJ, Carson JL, Olivier KN, Sagel SD: **Mutations in RSPH1 cause primary ciliary dyskinesia with a unique clinical and ciliary phenotype.** American journal of respiratory and critical care medicine 2014, **189**(6):707-717.

393. Ibanez-Tallon I, Heintz N, Omran H: **To beat or not to beat: roles of cilia in development and disease.** Hum Mol Genet 2003, **12 Spec No 1**:R27-35.

394. Bui KH, Sakakibara H, Movassagh T, Oiwa K, Ishikawa T: **Asymmetry of inner dynein arms and inter-doublet links in Chlamydomonas flagella.** J Cell Biol 2009, **186**(3):437-446.

395. Keith Funkhouser III W, Niethammer M, Carson JL, Burns KA, Knowles MR, Leigh MW, Zariwala MA, Funkhouser Jr WK: **A new tool improves diagnostic test performance for transmission EM evaluation of axonemal dynein arms.** Ultrastruct Pathol 2014, **38**(4):248-255.

396. Escudier E, Couprie M, Duriez B, Roudot-Thoraval F, Millepied M, Prulière-Escabasse V, Labatte L, Coste A: **Computer-assisted Analysis Helps Detect Inner Dynein Arm abnormalities.** American Journal of Respiratory and Critical Care Medicine 2002, **166**(9):1257-1262.

397. Antony D, Becker-Heck A, Zariwala MA, Schmidts M, Onoufriadis A, Forouhan M, Wilson R, Taylor-Cox T, Dewar A, Jackson C: **Mutations in CCDC39 and CCDC40 are the major cause of primary ciliary dyskinesia with axonemal disorganization and absent inner dynein arms.** Hum Mutat 2013, **34**(3):462-472.

398. Zariwala MA, Knowles MR, Leigh MW: **Primary Ciliary Dyskinesia.** In *GeneReviews(R).* Edited by Pagon RA, Adam MP, Ardinger HH, et al. Seattle (WA): University of Washington, Seattle; 1993:.

399. Escudier E, Couprie M, Duriez B, Roudot-Thoraval F, Millepied M, Prulière-Escabasse V, Labatte L, Coste A: **Computer-assisted analysis helps detect inner dynein arm abnormalities.** American journal of respiratory and critical care medicine 2002, **166**(9):1257-1262.

400. Olin JT, Burns K, Carson JL, Metjian H, Atkinson JJ, Davis SD, Dell SD, Ferkol TW, Milla CE, Olivier KN: **Diagnostic yield of nasal scrape biopsies in primary ciliary dyskinesia: a multicenter experience.** Pediatr Pulmonol 2011, **46(**5):483-488.

401. Kawakami M, Hattori Y, Nakamura S: **Reflection of structural abnormality in the axoneme of respiratory cilia in the clinical features of immotile cilia syndrome.** Internal Medicine 1996, **35**(8):617-623.

402. Papon J, Bassinet L, Cariou-Patron G, Zerah-Lancner F, Vojtek A, Blanchon S, Crestani B, Amselem S, Coste A, Housset B: **Quantitative analysis of ciliary beating in primary ciliary dyskinesia: a pilot study.** Orphanet journal of rare diseases 2012, **7**(1):1.

403. Coste A, Millepied M, Chapelin C, Reinert P, Poron F, Boucherat M, Escudier E: **Incidence of primary ciliary dyskinesia in children with recurrent respiratory diseases.** Annals of Otology, Rhinology & Laryngology 1997, **106**(10):854-858.

404. Hosie PH, Fitzgerald DA, Jaffe A, Birman CS, Rutland J, Morgan LC: **Presentation of primary ciliary dyskinesia in children: 30 years' experience.** J Paediatr Child Health 2015, **51**(7):722-726.

405. Sirvanci S, Uyan ZS, Ercan F, Karadag B, Ersu R, Karakoc F, Dagli E, San T: **Quantitative analysis of ciliary ultrastructure in patients with primary ciliary dyskinesia.** Acta Histochem 2008, **110**(1):34-41.

406. Noone PG, Leigh MW, Sannuti A, Minnix SL, Carson JL, Hazucha M, Zariwala MA, Knowles MR: **Primary ciliary dyskinesia: diagnostic and phenotypic features.** American journal of respiratory and critical care medicine 2004, **169(**4):459-467.

407. Chin GY, Karas DE, Kashgarian M: **Correlation of presentation and pathologic condition in primary ciliary dyskinesia.** Archives of Otolaryngology–Head & Neck Surgery 2002, **128**(11):1292-1294.

408. Beydon N, Chambellan A, Alberti C, de Blic J, Clément A, Escudier E, Le Bourgeois M: **Technical and practical issues for tidal breathing measurements of nasal nitric oxide in children.** Pediatr Pulmonol 2015, **50**(12):1374-1382.

409. Carda C, Armengot M, Escribano A, Peydró A: **Ultrastructural patterns of primary ciliar dyskinesia syndrome.** Ultrastruct Pathol 2005, **29**(1):3-8.

410. Papon JF, Coste A, Roudot-Thoraval F, Boucherat M, Roger G, Tamalet A, Vojtek AM, Amselem S, Escudier E: **A 20-year experience of electron microscopy in the diagnosis of primary ciliary dyskinesia.** Eur Respir J 2010, **35**(5):1057-1063.

411. Rubio MTR, Carceller MA, Batalla CC, Montaner AE, Paya JM: **What is the gold standard in the diagnosis of primary ciliary dyskinesia syndrome?** European Respiratory Journal 2011, **38**(Suppl 55):p512.

412. Simoneau T, Zandieh SO, Rao DR, Vo P, Palm KE, McCown M, Kopel LS, Dias A, Casey A, Perez-Atayde AR: **Impact of cilia ultrastructural examination on the diagnosis of primary ciliary dyskinesia.** Pediatric and Developmental Pathology 2013, **16**(5):321-326.

413. Corbelli R, Bringolf-Isler B, Amacher A, Sasse B, Spycher M, Hammer J: **Nasal nitric oxide measurements to screen children for primary ciliary dyskinesia.** CHEST Journal 2004, **126**(4):1054-1059.

414. Davis SD, Ferkol TW, Rosenfeld M, Lee H, Dell SD, Sagel SD, Milla C, Zariwala MA, Pittman JE, Shapiro AJ: **Clinical features of childhood primary ciliary dyskinesia by genotype and ultrastructural phenotype.** American journal of respiratory and critical care medicine 2015, **191**(3):316-324.

415. Chilvers MA, Rutman A, O'Callaghan C: **Ciliary beat pattern is associated with specific ultrastructural defects in primary ciliary dyskinesia.** J Allergy Clin Immunol 2003, **112**(3):518-524.

416. Shin S, Yang S, Oh J, Lee H, Park C, Kwan J: **Ultrastructure of nasal cilia in children with recurrent or persistent respiratory diseases.** Korean Journal of Pediatrics 2006, **49**(4):410-416.

417. Theegarten D, Ebsen M: **Ultrastructural pathology of primary ciliary dyskinesia: report about 125 cases in Germany.** Diagn Pathol 2011, **6**:115.

418. Chi JG, Yoon CJ: **Primary Ciliary Dyskinesia (Immotile Cilia Syndrome): Clinical and electron microscopic analysis of 17 cases.** Korean Journal of Pathology 1993, **27**(2):99-107.

419. Bent JP,3rd, Smith RJ: **Intraoperative diagnosis of primary ciliary dyskinesia.** Otolaryngol Head Neck Surg 1997, **116**(1):64-67.

420. Pizzi S, Cazzato S, Bernardi F, Mantovani W, Cenacchi G: **Clinico-pathological evaluation of ciliary dyskinesia: diagnostic role of electron microscopy.** Ultrastruct Pathol 2003, **27**(4):243-252.

421. Busquets RM, Caballero-Rabasco MA, Velasco M, Lloreta J, García-Algar Ó: **Primary Ciliary Dyskinesia: clinical criteria indicating ultrastructural studies.** Archivos de Bronconeumología (English Edition) 2013, **49**(3):99-104.

422. Plesec TP, Ruiz A, McMahon JT, Prayson RA: **Ultrastructural abnormalities of respiratory cilia: a 25-year experience.** Arch Pathol Lab Med 2008, **132**(11):1786-1791.

423. Baker B, Minnix S, Dell S, Ferkol T, Milla C, Olivier K, Rosenfeld M, Sagel S, Carson J, Daniels S: **The diagnostic dilemma of primary ciliary dyskinesia: Findings and experience of the**

Genetic Disorders of Mucociliary Clearance Consortium. Am J Respir Crit Care Med 2011, **183**:A1217.

424. Knowles MR, Daniels LA, Davis SD, Zariwala MA, Leigh MW: **Primary ciliary dyskinesia. Recent advances in diagnostics, genetics, and characterization of clinical disease.** American journal of respiratory and critical care medicine 2013, **188**(8):913-922.

425. Knowles MR, Boucher RC: **Mucus clearance as a primary innate defense mechanism for mammalian airways.** J Clin Invest 2002, **109**(5):571-577.

426. Noone PG, Leigh MW, Sannuti A, Minnix SL, Carson JL, Hazucha M, Zariwala MA, Knowles MR: **Primary ciliary dyskinesia: diagnostic and phenotypic features.** American journal of respiratory and critical care medicine 2004, **169**(4):459-467.

427. Kurkowiak M, Zietkiewicz E, Witt M: **Recent advances in primary ciliary dyskinesia genetics.** J Med Genet 2015, **52**(1):1-9.

428. Collins SA, Gove K, Walker W, Lucas JS: **Nasal nitric oxide screening for primary ciliary dyskinesia: systematic review and meta-analysis.** Eur Respir J 2014, **44**(6):1589-1599.

429. Chilvers MA, Rutman A, O'callaghan C: **Ciliary beat pattern is associated with specific ultrastructural defects in primary ciliary dyskinesia.** J Allergy Clin Immunol 2003, **112**(3):518-524.

430. Sisson JH, Stoner J, Ammons B, Wyatt T: **All-digital image capture and whole-field analysis of ciliary beat frequency.** J Microsc 2003, **211**(2):103-111.

431. Shoemark A, Dixon M, Corrin B, Dewar A: **Twenty-year review of quantitative transmission electron microscopy for the diagnosis of primary ciliary dyskinesia.** J Clin Pathol 2012, **65**(3):267-271.

432. Papon JF, Coste A, Roudot-Thoraval F, Boucherat M, Roger G, Tamalet A, Vojtek AM, Amselem S, Escudier E: **A 20-year experience of electron microscopy in the diagnosis of primary ciliary dyskinesia.** Eur Respir J 2010, **35**(5):1057-1063.

433. Lucas JS, Burgess A, Mitchison HM, Moya E, Williamson M, Hogg C, National PCD Service U: **Diagnosis and management of primary ciliary dyskinesia.** Arch Dis Child 2014, **99**(9):850-856.

434. Lucas JS, Leigh MW: **Diagnosis of primary ciliary dyskinesia: searching for a gold standard.** Eur Respir J 2014, **44**(6):1418-1422.

435. Marthin JK, Nielsen KG: **Choice of nasal nitric oxide technique as first-line test for primary ciliary dyskinesia.** Eur Respir J 2011, **37**(3):559-565.

436. Marthin JK, Nielsen KG: **Hand-held tidal breathing nasal nitric oxide measurement–a promising targeted case-finding tool for the diagnosis of primary ciliary dyskinesia.** PLoS One 2013, **8**(2):e57262.

437. American Thoracic Society: **European Respiratory Society. ATS/ERS recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005.** Am J Respir Crit Care Med 2005, **171**:912-930.

438. Rumman N, Jackson C, Collins S, Goggin P, Coles J, Lucas JS: **Diagnosis of primary ciliary dyskinesia: potential options for resource-limited countries.** Eur Respir Rev 2017, **26**(143):10.1183/16000617.0058-2016. Print 2017 Jan.

439. Kyriacou K, Yiallouros P, Kouis P, Papatheodorou S, Omran H: **Better experimental screening and treatment for primary ciliary dyskinesia: The FP7 BESTCILIA project.** Ultrastruct Pathol 2017, **41**(1):110-112.

440. Lucas JS, Barbato A, Collins SA, Goutaki M, Behan L, Caudri D, Dell S, Eber E, Escudier E, Hirst RA, Hogg C, Jorissen M, Latzin P, Legendre M, Leigh MW, Midulla F, Nielsen KG, Omran H, Papon JF, Pohunek P, Redfern B, Rigau D, Rindlisbacher B, Santamaria F, Shoemark A, Snijders D, Tonia T, Titieni A, Walker WT, Werner C, Bush A, Kuehni CE: **European Respiratory Society guidelines for the diagnosis of primary ciliary dyskinesia.** Eur Respir J 2017, **49**(1):10.1183/13993003.01090-2016. Print 2017 Jan.

441. Sox HC: *Medical decision making:* ACP Press; 1988.

442. Xu X, Nardini HKG, Ruger JP: **Micro-costing studies in the health and medical literature: protocol for a systematic review.** Systematic reviews 2014, **3**(1):47.

443. Hunink MM, Weinstein MC, Wittenberg E, Drummond MF, Pliskin JS, Wong JB, Glasziou PP: *Decision making in health and medicine: integrating evidence and values:* Cambridge University Press; 2014.

444. Owens DK: **Interpretation of cost-effectiveness analyses.** Journal of general internal medicine 1998, **13**(10):716-717.

445. Neumann PJ, Thorat T, Shi J, Saret CJ, Cohen JT: **The changing face of the cost-utility literature, 1990–2012.** Value in health 2015, **18**(2):271-277.

446. Prieto L, Sacristán JA: **Problems and solutions in calculating quality-adjusted life years (QALYs).** Health and quality of life outcomes 2003, **1**(1):80.

447. Strippoli MP, Frischer T, Barbato A, Snijders D, Maurer E, Lucas JS, Eber E, Karadag B, Pohunek P, Zivkovic Z, Escribano A, O'Callaghan C, Bush A, Kuehni CE, ERS Task Force onPrimary Ciliary Dyskinesia in Children: **Management of primary ciliary dyskinesia in European children: recommendations and clinical practice.** Eur Respir J 2012, **39**(6):1482-1491.

448. Shapiro AJ, Zariwala MA, Ferkol T, Davis SD, Sagel SD, Dell SD, Rosenfeld M, Olivier KN, Milla C, Daniel SJ: **Diagnosis, monitoring, and treatment of primary ciliary dyskinesia: PCD foundation consensus recommendations based on state of the art review.** Pediatr Pulmonol 2016, **51**(2):115-132.

449. Barbato A, Frischer T, Kuehni CE, Snijders D, Azevedo I, Baktai G, Bartoloni L, Eber E, Escribano A, Haarman E, Hesselmar B, Hogg C, Jorissen M, Lucas J, Nielsen KG, O'Callaghan C, Omran H, Pohunek P, Strippoli MP, Bush A: **Primary ciliary dyskinesia: a consensus statement on diagnostic and treatment approaches in children.** Eur Respir J 2009, **34**(6):1264-1276.

450. Yiallouros PK, Kouis P, Middleton N, Nearchou M, Adamidi T, Georgiou A, Eleftheriou A, Ioannou P, Hadjisavvas A, Kyriacou K: **Clinical features of primary ciliary dyskinesia in Cyprus with emphasis on lobectomized patients.** Respir Med 2015, **109**(3):347-356.

451. Kouis P, Yiallouros PK, Middleton N, Evans JS, Kyriacou K, Papatheodorou SI: **Prevalence of primary ciliary dyskinesia in consecutive referrals of suspect cases and the transmission electron microscopy detection rate: a systematic review and meta-analysis.** Pediatr Res 2016, .

452. Kouis P, Papatheodorou SI, Yiallouros PK: **Diagnostic accuracy of nasal nitric oxide for establishing diagnosis of primary ciliary dyskinesia: a meta-analysis.** BMC pulmonary medicine 2015, **15**(1):153.

453. Jackson CL, Behan L, Collins SA, Goggin PM, Adam EC, Coles JL, Evans HJ, Harris A, Lackie P, Packham S, Page A, Thompson J, Walker WT, Kuehni C, Lucas JS: **Accuracy of diagnostic testing in primary ciliary dyskinesia.** Eur Respir J 2016, **47**(3):837-848.

454. Boon M, Jorissen M, Jaspers M, Cuppens H, De Boeck K: **Is the sensitivity of primary ciliary dyskinesia detection by ciliary function analysis 100%?** Eur Respir J 2013, **42**(4):1159-1161.

455. Anonymous *Proceedings of the Seminars in Perinatology:* Elsevier; 2006.

456. Kuehni CE, Goutaki M, Carroll M, Lucas JS: **Primary ciliary dyskinesia: the patients grow up.** Eur Respir J 2016, **48**(2):297-300.

457. Bradley JM, Blume SW, Balp MM, Honeybourne D, Elborn JS: **Quality of life and healthcare utilisation in cystic fibrosis: a multicentre study.** Eur Respir J 2013, **41**(3):571-577.

458. Fryback DG, Dasbach EJ, Klein R, Klein BE, Dorn N, Peterson K, Martin PA: **The Beaver Dam Health Outcomes Study: initial catalog of health-state quality factors.** Medical Decision Making 1993, **13**(2):89-102.

459. Stavem K: **Reliability, validity and responsiveness of two multiattribute utility measures in patients with chronic obstructive pulmonary disease.** Quality of Life Research 1999, **8**(1):45-54.

460.

461.

462.

463. Van Hout BA, Al MJ, Gordon GS, Rutten FF: **Costs, effects and C/E-ratios alongside a clinical trial.** Health Econ 1994, **3**(5):309-319.

464. Judgements SV: **Principles for the development of NICE guidance.** London: National Institute for Health and Care Excellence 2008, **478**.

465. Shiroiwa T, Sung Y, Fukuda T, Lang H, Bae S, Tsutani K: **International survey on willingness-to-pay (WTP) for one additional QALY gained: what is the threshold of cost effectiveness?** Health Econ 2010, **19**(4):422-437.

466. Schwabe GC, Hoffmann K, Loges NT, Birker D, Rossier C, De Santi MM, Olbrich H, Fliegauf M, Failly M, Liebers U: **Primary ciliary dyskinesia associated with normal axoneme ultrastructure is caused by DNAH11 mutations.** Hum Mutat 2008, **29**(2):289-298.

467. Olbrich H, Cremers C, Loges NT, Werner C, Nielsen KG, Marthin JK, Philipsen M, Wallmeier J, Pennekamp P, Menchen T: **Loss-of-function GAS8 mutations cause primary ciliary dyskinesia**

and disrupt the nexin-dynein regulatory complex. The American Journal of Human Genetics 2015, **97**(4):546-554.

468. Leigh MW, O'callaghan C, Knowles MR: **The challenges of diagnosing primary ciliary dyskinesia.** Proceedings of the American Thoracic Society 2011, **8**(5):434-437.

469. Behan L, Dunn Galvin A, Rubbo B, Masefield S, Copeland F, Manion M, Rindlisbacher B, Redfern B, Lucas JS: **Diagnosing primary ciliary dyskinesia: an international patient perspective.** Eur Respir J 2016, **48**(4):1096-1107.

470. Shoemark A, Frost E, Dixon M, Ollosson S, Kilpin K, Patel M, Scully J, Rogers AV, Mitchison HM, Bush A: **Accuracy of Immunofluorescence in the Diagnosis of Primary Ciliary Dyskinesia.** American Journal of Respiratory And Critical Care Medicine (ja).

471. Boaretto F, Snijders D, Salvoro C, Spalletta A, Mostacciuolo ML, Collura M, Cazzato S, Girosi D, Silvestri M, Rossi GA: **Diagnosis of Primary Ciliary Dyskinesia by a Targeted Next-Generation Sequencing Panel: Molecular and Clinical Findings in Italian Patients.** The Journal of Molecular Diagnostics 2016, **18**(6):912-922.

472. Kim RH, A. Hall D, Cutz E, Knowles MR, Nelligan KA, Nykamp K, Zariwala MA, Dell SD: **The role of molecular genetic analysis in the diagnosis of primary ciliary dyskinesia.** Annals of the American Thoracic Society 2014, **11**(3):351-359.

473. Leigh MW, Hazucha MJ, Chawla KK, Baker BR, Shapiro AJ, Brown DE, LaVange LM, Horton BJ, Qaqish B, Carson JL: **Standardizing nasal nitric oxide measurement as a test for primary ciliary dyskinesia.** Annals of the American Thoracic Society 2013, **10**(6):574-581.

474. Harris A, Bhullar E, Gove K, Joslin R, Pelling J, Evans HJ, Walker WT, Lucas JS: **Validation of a portable nitric oxide analyzer for screening in primary ciliary dyskinesias.** BMC pulmonary medicine 2014, **14**(1):18.

475. Raidt J, Wallmeier J, Hjeij R, Onnebrink JG, Pennekamp P, Loges NT, Olbrich H, Haffner K, Dougherty GW, Omran H, Werner C: **Ciliary beat pattern and frequency in genetic variants of primary ciliary dyskinesia.** Eur Respir J 2014, **44**(6):1579-1588.

476. Werner C, Lablans M, Ataian M, Raidt J, Wallmeier J, Grosse-Onnebrink J, Kuehni CE, Haarman EG, Leigh MW, Quittner AL, Lucas JS, Hogg C, Witt M, Priftis KN, Yiallouros P, Nielsen KG, Santamaria F, Uckert F, Omran H: **An international registry for primary ciliary dyskinesia.** Eur Respir J 2016, **47**(3):849-859.

477. Kobbernagel HE, Buchvald FF, Haarman EG, Casaulta C, Collins SA, Hogg C, Kuehni CE, Lucas JS, Omran H, Quittner AL: **Study protocol, rationale and recruitment in a European multi-centre randomized controlled trial to determine the efficacy and safety of azithromycin maintenance therapy for 6 months in primary ciliary dyskinesia.** BMC Pulmonary Medicine 2016, **16**(1):104.

478. O'Callaghan C, Rutman A, Williams GM, Hirst RA: **Inner dynein arm defects causing primary ciliary dyskinesia: repeat testing required.** Eur Respir J 2011, **38**(3):603-607.

479. Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen, Jan HP, Bossuyt PM: **Empirical evidence of design-related bias in studies of diagnostic tests.** JAMA 1999, **282**(11):1061-1066.

480. Rutjes AW, Reitsma JB, Vandenbroucke JP, Glas AS, Bossuyt PM: **Case-control and two-gate designs in diagnostic accuracy studies.** Clin Chem 2005, **51**(8):1335-1341.

481. Lin J, Yin W, Smith MC, Song K, Leigh MW, Zariwala MA, Knowles MR, Ostrowski LE, Nicastro D: **Cryo-electron tomography reveals ciliary defects underlying human RSPH1 primary ciliary dyskinesia.** Nature communications 2014, **5**:5727.

482. Berg JS, Evans JP, Leigh MW, Omran H, Bizon C, Mane K, Knowles MR, Weck KE, Zariwala MA: **Next generation massively parallel sequencing of targeted exomes to identify genetic mutations in primary ciliary dyskinesia: implications for application to clinical testing.** Genet Med 2011, **13**(3):218-229.

483. Marshall CR, Scherer SW, Zariwala MA, Lau L, Paton TA, Stockley T, Jobling RK, Ray PN, Knowles MR, FORGE Canada Consortium, Hall DA, Dell SD, Kim RH: **Whole-Exome Sequencing and Targeted Copy Number Analysis in Primary Ciliary Dyskinesia.** G3 (Bethesda) 2015, **5**(8):1775-1781.

484. Djakow J, Kramná L, Dušátková L, Uhlík J, Pursiheimo J, Svobodová T, Pohunek P, Cinek O: **An effective combination of sanger and next generation sequencing in diagnostics of primary ciliary dyskinesia.** Pediatr Pulmonol 2016, **51**(5):498-509.

485. King SM, Patel-King RS: **Planaria as a Model System for the Analysis of Ciliary Assembly and Motility.** Cilia: Methods and Protocols 2016, :245-254.

486. Bukowy Z, Ziętkiewicz E, Witt M: **In vitro culturing of ciliary respiratory cells—a model for studies of genetic diseases.** J Appl Genet 2011, **52**(1):39-51.

487. Ostrowski LE, Yin W, Rogers TD, Busalacchi KB, Chua M, O'neal WK, Grubb BR: **Conditional deletion of dnaic1 in a murine model of primary ciliary dyskinesia causes chronic rhinosinusitis.** American journal of respiratory cell and molecular biology 2010, **43**(1):55-63.

488. Lucas JS, Adam EC, Goggin PM, Jackson CL, Powles-Glover N, Patel SH, Humphreys J, Fray MD, Falconnet E, Blouin J: **Static respiratory cilia associated with mutations in Dnahc11/DNAH11: a mouse model of PCD.** Hum Mutat 2012, **33**(3):495-503.

489. Rogers GB, Shaw D, Marsh RL, Carroll MP, Serisier DJ, Bruce KD: **Respiratory microbiota: addressing clinical questions, informing clinical practice.** Thorax 2015, **70**(1):74-81.

490. Alanin MC, Nielsen KG, von Buchwald C, Skov M, Aanaes K, Høiby N, Johansen HK: **A longitudinal study of lung bacterial pathogens in patients with primary ciliary dyskinesia.** Clinical Microbiology and Infection 2015, **21**(12):1093. e1-1093. e7.

491. Anonymous *Proceedings of the BMC Proceedings:* BioMed Central; 2016.

492. Grosse-Onnebrink J, Werner C, Loges NT, Hörmann K, Blum A, Schmidt R, Olbrich H, Omran H: **Effect of TH2 cytokines and interferon gamma on beat frequency of human respiratory cilia.** Pediatr Res 2016, **79**(5):731-735.

493. Yaghi A, Zaman A, Cox G, Dolovich MB: **Ciliary beating is depressed in nasal cilia from chronic obstructive pulmonary disease subjects.** Respir Med 2012, **106**(8):1139-1147.

494. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, Abraham J, Adair T, Aggarwal R, Ahn SY: **Global and regional mortality from 235 causes of death for 20 age groups**

in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. The Lancet 2013, **380**(9859):2095-2128.

495. Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S: **Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010.** The lancet 2013, **380**(9859):2197-2223.

496. Chilvers MA, O'Callaghan C: **Analysis of ciliary beat pattern and beat frequency using digital high speed imaging: comparison with the photomultiplier and photodiode methods.** Thorax 2000, **55**(4):314-317.

497. Werner C, Lablans M, Ataian M, Raidt J, Wallmeier J, Grosse-Onnebrink J, Kuehni CE, Haarman EG, Leigh MW, Quittner AL, Lucas JS, Hogg C, Witt M, Priftis KN, Yiallouros P, Nielsen KG, Santamaria F, Uckert F, Omran H: **An international registry for primary ciliary dyskinesia.** Eur Respir J 2016, **47**(3):849-859.

498. Kobbernagel HE, Buchvald FF, Haarman EG, Casaulta C, Collins SA, Hogg C, Kuehni CE, Lucas JS, Omran H, Quittner AL: **Study protocol, rationale and recruitment in a European multi-centre randomized controlled trial to determine the efficacy and safety of azithromycin maintenance therapy for 6 months in primary ciliary dyskinesia.** BMC pulmonary medicine 2016, **16**(1):104.

499. Kuehni CE, Goutaki M, Kobbernagel HE: **Hypertonic saline in patients with primary ciliary dyskinesia: on the road to evidence-based treatment for a rare lung disease.** Eur Respir J 2017, **49**(2):10.1183/13993003.02514-2016. Print 2017 Feb.

500. Sims EJ, Mugford M, Clark A, Aitken D, McCormick J, Mehta G, Mehta A, UK Cystic Fibrosis Database Steering Committee: **Economic implications of newborn screening for cystic fibrosis: a cost of illness retrospective cohort study.** The Lancet 2007, **369**(9568):1187-1195.

501. Robson M, Abbott J, Webb K, Dodd M, Walsworth-Bell J: **A cost description of an adult cystic fibrosis unit and cost analyses of different categories of patients.** Thorax 1992, **47**(9):684-689.

APPENDIX I: Publications

Relevant publications in peer reviewed journals

1. **Kouis P**, Yiallouros PK, Middleton N, Evans JS, Kyriacou K, Papatheodorou SI. Prevalence of Primary Ciliary Dyskinesia in consecutive referrals of suspect cases and the Transmission Electron Microscopy detection rate: A systematic review and meta-analysis. *Pediatric Research*, doi:10.1038/pr.2016.263, **2016**

2. Rubbo B, Behan L, Dehlink E, Goutaki M, Hogg C, **Kouis P**, Kuehni CE, Latzin P, Kim N, Norris D, Nyilas S, Price M, Lucas JS.: Proceedings from the COST Action BM1407 Inaugural Conference: BEAT-PCD: Translational research in primary ciliary dyskinesia - bench, bedside, and population perspectives. *BMC Proceedings*. Vol. 10. No. 9. BioMed Central, **2016**

3. **Kouis P**, Papatheodorou SI, Yiallouros PK. Diagnostic accuracy of nasal nitric oxide for establishing diagnosis of primary ciliary dyskinesia: A meta-analysis. *BMC Pulmonary Medicine* 15:153, **2015**

4. Yiallouros PK, **Kouis P**, Middleton N, Nearchou M, Adamidi T, Georgiou A, Eleftheriou A, Ioannou P, Hadjisavvas A, Kyriacou K. Clinical features of primary ciliary dyskinesia in Cyprus with emphasis on lobectomized patients. *Respir Med* 109(3): 347-356, **2015**

Relevant publications in conference proceedings

1. **Kouis P**, Papatheodorou SI, Middleton N, Kyriacou K, Evans JS, Yiallouros PK. A decision analysis approach for the development of an evidence based diagnostic algorithm for Primary Ciliary Dyskinesia. Sept 2017. ERS International Congress 2017, Milan, Italy

2. Yiallouros PK, **Kouis P**, Pirpa P, Michailidou K, Loizidou M, Hadjisavvas A, Kyriacou K. Clinical disease spectrum in RSPH9 Primary Ciliary Dyskinesia patients: a case series. Sept 2017. ERS International Congress 2017, Milan, Italy

3. **Kouis P**, Yiallouros P, Papatheodorou S. Prevalence of primary ciliary dyskinesia among patients referred for specialized testing: A systematic review and meta-analysis. Sept **2016.** ERS International Congress 2016, London, UK

4. Behan L, Quittner AL, Galvin AD, Kuehni CE, Maurer E, Jose A, Haarman AG, Paff T, Rutjes NWP, Werner C, Lofruthe A, Edelbusch C, Heisch L, Kobbernagel HE, Philipsen M, Contreras B, Yiallouros P, **Kouis P**, Giannakou K, Toliopoulou A, Thouvenin G, Tamalet A, O' Neill C, Lucas JS. The development of translated, cross-cultural patient-reported outcome measures for patients with primary ciliary dyskinesia. Sept **2016.** ERS International Congress 2016, London, UK

5. Kyriacou K, **Kouis P**, Nearchou M, Pirpa P, Hadjisavvas A, Papatheodorou SI, and Yiallouros P. Primary Ciliary Dyskinesia: An update with molecular correlation. Sept **2016**. XXXI International Congress of the International Academy of Pathology and 28th Congress of the European Society of Pathology, Cologne, Germany

6. **Kouis P**, Yiallouros PK, Papatheodorou SI, Kyriacou K. The Detection Rate of Transmission Electron Microscopy in Primary Ciliary Dyskinesia: A Systematic Review and Meta-analysis. July **2016**. Ultrapath Conference XVIII, Lisbon, Portugal

7. Kyriacou K, Yiallouros PK, **Kouis P**, Papatheodorou SI, Omran H. Better Experimental Screening and Treatment for Primary Ciliary Dyskinesia: The FP7 BESTCILIA project. July **2016.** Ultrapath Conference XVIII, Lisbon, Portugal

8. Yiallouros P, **Kouis P,** Middleton N, Nearchou M, Adamidi T, Georgiou A, Eleftheriou A, Ioannou P, Hadjisavvas A, Kyriacou K. Clinical features of primary ciliary dyskinesia in Cyprus with emphasis on lobectomized patients. Apr **2016**, BEAT-PCD Training School - Second Young Researchers' Meeting on Primary Ciliary Dyskinesia, Paris, France

9. **Kouis P**, Papatheodorou S, Yiallouros P. Diagnostic accuracy of nasal nitric oxide measurement for establishing diagnosis of primary ciliary dyskinesia: A systematic review and meta-analysis. Sept **2015.** ERS International Congress 2015, Amsterdam, Netherlands

Appendix II

List of Abbreviations

ACER – Average Cost Effectiveness Ratio

ATP - Adenosine triphosphate

ATS – American Thoracic Society

AUC – Area under Curve

BH – Breath Hold

CBA – Cost Benefit Analysis

CBP – Cilia Beat Pattern

CBF – Cilia Beat Frequency

CEA – Cost Effectiveness Analysis

CEAC – Cost Effectiveness Analysis Curve

CEF – Cost Effectiveness Frontier

CF - Cystic Fibrosis

CFTR – Cystic Fibrosis Transmembrane conductance Regulator

CMA – Cost Minimization Analysis

COPD – Chronic Obstructive Pulmonary Disease

CP – Central Pair

CT – Computed Tomography

CUA – Cost Utility Analysis

CXR – Chest X-ray

DA – Dynein Arms

DOR – Diagnostic Odds Ratio

ER – Exhalation against Resistance

ERS – European Respiratory Society

FEV1 – Forced Expiratory Volume in one second

FN – False Negative

FP – False Positive

FVC – Forced Vital Capacity

HKSJ - Hartung, Knapp, Sidik and Jonkman

HR – Hazard Ratio

HRQoL – Health Related Quality of Life

HSROC – Hierarchical Summary Receiver Operator Curve

HSVM – High Speed Video Microscopy

ICER – Incremental Cost Effectiveness Ratio

IDA – Inner Dynein Arms

LR – Likelihood Ratio

MCC – Mucociliary Clearance

MeSH – Medical Subject Headings

MTD – Microtubular disorganization

nNO – nasal Nitric Oxide

NGS – Next Generation Sequencing

NO – Nitric Oxide

NOS – Nitric Oxide Synthase

NPV – Negative Predictive Value

NU – Normal Ultrastructure

ODA – Outer Dynein Arms

ODA/IDA – Combined Outer and Inner Dynein Arms

OR – Odds Ratio

PCD – Primary Ciliary Dyskinesia

PPV – Positive Predictive Value

QALYs – Quality Adjusted Life Years

RCT – Randomized Clinical Trial

RD – Risk Difference

ROC – Receiver Operator Curve

ROI – Region of Interest

RR – Relative Risk

RS – Radial Spoke

SAVA – Sisson Ammons Video Analysis

SE – Standard Error

SG – Standard Gamble

TB – Tidal Breathing

TEM – Transmission Electron Microscopy

TN – True Negative

TP – True Positive

TTO - Time Trade Off

VAS – Visual Analog Scale

VC – Velum Closure

WFA – Whole Field Analysis

WTP – Willingness to Pay