



OPEN

## A systematic review of animal and human data comparing the nasal potential difference test between cystic fibrosis and control

Cathalijn H. C. Leenaars<sup>1</sup>✉, Frans R. Stafleu<sup>2</sup>, Christine Häger<sup>1</sup>, Hendrik Nieraad<sup>1</sup> & André Bleich<sup>1</sup>

The nasal potential difference test (nPD) is an electrophysiological measurement which is altered in patients and animal models with cystic fibrosis (CF). Because protocols and outcomes vary substantially between laboratories, there are concerns over its validity and precision. We performed a systematic literature review (SR) of the nPD to answer the following review questions: A. Is the nasal potential difference similarly affected in CF patients and animal models?”, and B. “Is the nPD in human patients and animal models of CF similarly affected by various changes in the experimental set-up?”. The review protocol was preregistered on PROSPERO (CRD42021236047). We searched PubMed and Embase with comprehensive search strings. Two independent reviewers screened all references for inclusion and extracted all data. Included were studies about CF which described in vivo nPD measurements in separate CF and control groups. Risk of bias was assessed, and three meta-analyses were performed. We included 130 references describing nPD values for CF and control subjects, which confirmed substantial variation in the experimental design and nPD outcome between groups. The meta-analyses showed a clear difference in baseline nPD values between CF and control subjects, both in animals and in humans. However, baseline nPD values were, on average, lower in animal than in human studies. Reporting of experimental details was poor for both animal and human studies, and urgently needs to improve to ensure reproducibility of experiments within and between species.

**Keywords** Animal-to-human translation, Cystic fibrosis, Nasal potential difference, Electrophysiology, Predictive value, Reproducibility

### Abbreviations

CF	Cystic fibrosis
CFTR	Cystic fibrosis transmembrane conductance regulator
CI	Confidence interval
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
IQR	Interquartile range
nPD	Nasal potential difference
RoB	Risk of bias
SD	Standard deviation
SE	Standard error
SOP	Standard operating procedure
SR	Systematic literature review

In its basics, the nasal potential difference test (nPD) is an electrophysiological measurement showing polarisation of the nasal epithelium. In cystic fibrosis (CF), the epithelia are hyperpolarised. Because of the relative ease of measuring potential differences in the nose, and in light of the limitations of the preceding diagnostic technique (i.e. sweat tests), the nPD was further developed as a diagnostic tool, a process which started well before the *CFTR*-gene was discovered. In this process, several laboratories developed standard operating procedures

<sup>1</sup>Institute for Laboratory Animal Science, Hannover Medical School, Hannover, Germany. <sup>2</sup>Department of Animals in Science and Society-Human-Animal Relationship, Utrecht University, Utrecht, The Netherlands. ✉email: Leenaars.Cathalijn@mh-hannover.de

(SOPs). The SOPs were extended to include buffers flowing over the epithelium, which alter the ion flow over the membrane and thereby the potential difference. The use of these buffers, which affect the CFTR and other channels, increased the diagnostic sensitivity of the nPD for CF<sup>1,2</sup>.

Because of the association of nPD with disease severity<sup>3</sup>, the nPD can also be measured in (pre)clinical trials as a proxy of disease severity, both in CF patients and in animal models. Besides, as it reflects the function of the CFTR, it can be used in mechanistic studies. However, while there are standard operating procedures, exact protocols vary between laboratories, and this can in turn affect the measured membrane potential. Consequently, scientists have expressed their concern over the observed variation in protocol and outcomes<sup>4,5</sup>.

The nPD has thus been used for different purposes and in multiple species, and there are valid concerns about its validity and precision<sup>4–7</sup>. Therefore, we performed a systematic literature review (SR) of the nPD. Our SR was designed to answer multiple review questions. In this paper, we describe the relevant data for two of them; first, “is the nasal potential difference similarly affected in CF patients and animal models?”, and second, “is the nPD in human patients and animal models of CF similarly affected by various changes in the experimental set-up?”. This SR of the nPD test follows several narrative reviews of e.g. the nPD in CF patients<sup>8,9</sup>.

Besides being a valuable tool for anyone involved in translational nPD study design, by comparing the same outcome between animal models and CF patients, this SR contributes to the ongoing analysis of translational value of animal studies. Preceding work has shown that the predictive value of animal experiments for humans ranges from 0 to 100%<sup>10</sup>, with relevant differences in average predictive values between medical fields<sup>11</sup>. Straight-forward analyses of several potentially relevant factors could not predict the translational value, which means that part of the animal studies that are currently performed cannot be reproduced in humans and thus have limited informative value. While in-depth analyses of combinations of factors may be of future benefit<sup>12</sup>, further in-depth quantitative analyses of specific tests and interventions in animals and humans seem crucial. To the best of our knowledge, our SR is one of the first comparing a specific test; the nPD, between animals and humans.

## Materials and methods

### Protocol

This review followed a protocol that was preregistered on PROSPERO (CRD42021236047) on 05 March 2021. Protocol refinements (i.e. more detailed operationalisations of our planned work) are described in the respective sections below. Except for additional data extraction and analyses, there were no protocol violations. This publication describes the review process (methods), and the data relating to the first and third review question from the protocol, focussing on comparisons between cystic fibrosis and control. Reporting follows the PRISMA guidelines. Other papers will describe the data relating to potential treatment effects (protocol review question 2), the relative sensitivity of the nPD measure compared to other outcome measures in animal models (protocol review question 4), a detailed analysis of study reporting quality, and text frequency analyses of the abstracts of the included studies. This review has also been used as a case study to compare several review tools (manuscript under review).

### Search

PubMed and Embase (via Ovid) were searched on 23 March 2021, without restrictions for publication date or language. Separate search strings were used for: animals and humans, CF, and the nPD. Strings comprised both title-abstract-keyword terms as well as thesaurus (MeSH/Emtree) terms. The comprehensive animal filters, mentioning all potentially-relevant species in single and plural as <[tiab]> or <.ti,ab,kw.> terms, as well as all relevant thesaurus terms in the appropriate syntax, have been published elsewhere<sup>13</sup>. The other search strings are provided in Table 1.

The reference lists of relevant reviews and included studies were screened by two independent reviewers (FS & CL) to retrieve additional eligible studies.

### Screening

Screening was performed in two distinct phases: title-abstract screening and full text screening, following the criteria in Table 2. Screening was performed by two independent reviewers (FS & CL) for both phases, in two separate projects in Rayyan<sup>14</sup>. Discrepancies could all be resolved by discussions between the reviewers.

### Data extraction

Data were extracted by two independent reviewers (FS & HN or FS & CL) in Covidence<sup>15</sup>, in two distinct phases. In the first phase, basic data on study design and included population were extracted. This allowed us to categorise the included studies into the CF versus control comparisons included in this publication, and the treatment studies to be published at a later stage. In the second phase, all data presented in this publication were extracted for the relevant subset. Discrepancies were resolved by discussion between the reviewers. When standard deviations (SDs) were provided, they were converted to Standard Errors (SEs). When values were not provided in text or tables, digital rulers were used.

It became clear from the first data extractions that many alternative nPD measures are reported. The nowadays common low chloride nPD values are less sensitive to infection and mild trauma, and more stable and reliable in distinguishing CF from non-CF<sup>1</sup>. However, many of the older studies only report baseline measures. To include data from as many studies as possible, we decided to separately extract both baseline and low chloride measurements.

Baseline measures were preferentially extracted without flow or with only buffer, but if these data were not available, measurements in the presence of amiloride were extracted. Low or zero chloride data were preferentially

Pubmed	
nPD	Membrane potentials [MeSH:NoExp] OR ((Nasa*[tiab] OR naso*[tiab] OR membran*[tiab] OR transmembran*[tiab]) AND (Potential[tiab] OR potentials[tiab] OR voltage[tiab] OR voltages[tiab] OR current[tiab] OR currents[tiab]) AND (Difference[tiab] OR differences[tiab] OR change[tiab] OR changes[tiab] OR alteration[tiab] OR alterations[tiab] OR variance[tiab])))
CF	Cystic fibrosis [MeSH] OR Mice, Inbred CFTR [MeSH] OR Cystic Fibrosis Transmembrane Conductance Regulator [MeSH] OR (cystic[tiab] AND (fibrosis[tiab] OR fibroses[tiab] OR fibrotic[tiab])) OR Mucoviscidos* [tiab] OR Mucoviscoid* [tiab] OR Mukoviszid* [tiab] OR CFTR [tiab] OR Fibrocystic Disease [tiab] OR Fibrocystic Diseases [tiab] OR Mckusick [tiab] OR CFRD [tiab] OR "pancreas cystic disease" [tiab] OR muco-patient* [tiab] OR muko-patient* [tiab] OR (CF [tiab] AND (lung [tiab] OR lungs [tiab] OR pulmonary [tiab] OR ABPA [tiab] OR mucus [tiab] OR liver [tiab] OR livers [tiab] OR steatosis [tiab] OR cirrhosis [tiab] OR cirrhotic [tiab] OR meconium ileus[tiab] OR gastrointestinal [tiab] OR intestine [tiab] OR intestines [tiab] OR intestinal [tiab] OR duodenum [tiab] OR jejunum [tiab] OR colon [tiab] OR caecum [tiab] OR DIOS [tiab] OR ((sweat [tiab] OR eccrine [tiab] OR apocrine [tiab] OR salivary [tiab] OR parotid [tiab] OR sublingual [tiab] OR submandibular [tiab] OR sub-lingual [tiab] OR sub-mandibular [tiab] OR von Ebner [tiab])) AND (gland [tiab] OR glands [tiab])) OR ((Paranasal [tiab] OR Para-nasal [tiab] OR frontal [tiab] OR ethmoidal [tiab] OR maxillary [tiab] OR sphenoidal [tiab]) AND (sinus [tiab] OR sinuses [tiab])) OR pancreas [tiab] OR pancreatic [tiab]))
Human	clinical study [pt] OR clinical trial [tiab] OR intervention study [tiab] OR "clinical studies as topic" [MeSH] OR first in man [tiab] OR proof of concept [tiab] OR randomized [tiab] OR placebo [tiab] OR drug therapy [sh] OR randomly [tiab] OR trial [tiab] OR groups [tiab] OR multicenter study[pt] OR "Multicenter Studies as Topic" [Mesh]
Embase	
nPD	exp potential difference/OR (exp nose/AND exp electrical parameters/) OR ((Nasa* OR naso* OR membran* or transmembran*) AND (Potential OR potentials OR voltage OR voltages OR current OR currents) AND (Difference OR differences or change OR changes OR alteration OR alterations OR variance)).ti,ab,kw
CF	Cystic fibrosis/OR cystic fibrosis transmembrane conductance regulator/OR (cystic adj2 fibros*).ti,ab,kw. OR fibrocystic diseas*.ti,ab,kw. OR (mucovisc* or Mukoviszidose).ti,ab,kw. OR CFRD.ti,ab,kw. OR muco-patient*.ti,ab,kw. OR muko-patient*.ti,ab,kw. OR pancreas cystic disease.ti,ab,kw. OR pancreas fibrocystic disease.ti,ab,kw. OR pancreas fibrosis.ti,ab,kw. OR pancreatic cystic disease.ti,ab,kw. OR pancreatic fibrosis.ti,ab,kw. OR ((CF adj30 (lung OR liver OR stomach OR intestines OR pulmonary OR meconium ileus OR gastrointestinal OR intestine OR intestines OR intestinal OR pancreas OR pancreatic OR ((sweat OR eccrine OR apocrine OR salivary OR parotid OR sublingual OR submandibular OR von Ebner) adj2 (gland OR glands)) OR ((Paranasal OR frontal OR ethmoidal OR maxillary OR sphenoidal) adj2 (sinus OR sinusses))).ti,ab,kw
Human	exp clinical trial/OR clinical study/OR human subject.ti,ab,kw. OR clinical drug trial.ti,ab,kw. OR major clinical trial.ti,ab,kw. OR trial, clinical.ti,ab,kw. OR clinical study.ti,ab,kw. OR phase 1 clinical trial.ti,ab,kw. OR phase 2 clinical trial.ti,ab,kw. OR phase 3 clinical trial.ti,ab,kw. OR clinical trial, controlled.ti,ab,kw. OR clinical trial, phase 1.ti,ab,kw. OR clinical trial, phase 2.ti,ab,kw. OR clinical trial, phase 3.ti,ab,kw. OR clinical trials.ti,ab,kw. OR clinical trial, phase I.ti,ab,kw. OR clinical trial, phase II.ti,ab,kw. OR clinical trial, phase III.ti,ab,kw. OR intervention study.ti,ab,kw

**Table 1.** Search strategy.

Exclusion criteria	Screening phase	
	Title-abstract	Full text
Study not about CF	X	X
Study not in vivo <sup>a</sup>	X	X
No nPD measured	X	X
No untreated control group (either with or without CF) <sup>b</sup> present		X
No full peer-reviewed publication <sup>c</sup>		X

**Table 2.** Exclusion criteria. There were no restrictions for publication date or language. <sup>a</sup>Ex vivo, in vitro and in silico models were excluded. <sup>b</sup>Included were both studies that compared CF with healthy controls at baseline, and studies that compared treated with untreated CF. <sup>c</sup>Conference proceedings and short communications lacking a detailed description of the methods were excluded because we planned analyses of the experimental set-ups.

extracted in combination with the additional stimulation by forskolin or isoproterenol, but when data with these chloride channel activators were lacking, low/no chloride buffer data without these activators were extracted.

When a paper described multiple comparisons, we only included one of these comparisons to prevent dependency of data. We preferentially included the first one performed, or when a cross-over design was used, the first one presented. For example, De Wachter et al.<sup>16</sup> compared two different nPD methods in CF and control subjects. We included the first described into our analyses; the one with the subcutaneous agar-filled needle as the reference electrode. Similarly, for Duperrex et al.<sup>17</sup>, values from the inferior turbinate were included. We only deviated from this method when a later-presented sample was substantially larger, as e.g. the third protocol described in<sup>1</sup>.

When repeated baseline measures of the same condition were performed in the same session, we took the last measurement of this condition. When repeated treatment effects were separately presented, we selected the last one within the “treatment” period.

### Risk of bias assessments

Risk of bias (RoB) was assessed independently by 2 reviewers using a list of questions based on the SYRCLE RoB tool<sup>18</sup> and the Cochrane Collaboration's tool<sup>19</sup>. Besides, compliance with all items from the ARRIVE<sup>20</sup> and

CONSORT<sup>21</sup> guidelines was assessed as a reflection of reporting quality. All discrepancies were resolved by discussion between these reviewers, the third reviewer specified in the protocol was not necessary. In this paper, we summarise the RoB for the papers included in these CF versus control analyses following the SYRCLE and Cochrane tools. RoB for treatment effects has been published with those findings<sup>22</sup>.

The SYRCLE and Cochrane RoB tools consists of separate domains for studies included in systematic reviews. The tools are both focussed on the comparison of treatment versus control. This makes them perfectly applicable for the data we described in our parallel publication of this SR, but less so for the CF-control comparisons described in this one. As the CF or control genotype is present from birth, the risk of bias associated with treatment allocation is irrelevant. For this publication we restrict RoB reporting to the relevant domains, as summarised in Table 3. Findings will be presented in summarised form per domain and per study. RoB and reporting quality will be published in more detail at a later stage.

### Analysis

Data were exported from Covidence to Excel. Data were checked and cleaned in Excel. Cleaning comprised manual reintroduction of vanished decimal separators (due to language settings that were not fully compatible with Covidence export), conversion of all flow rates to mL/min, splitting of ranges for age and weight to minimum and maximum values, selecting median (outcomes) or minimum (n) values for reported ranges, calculating the median for other measures occasionally reported as a range, estimating SDs from InterQuartile Ranges (IQRs, four values from one study, using the formula from the Cochrane handbook<sup>23</sup>:  $[SD \approx (q_3 - q_1)/1.35]$ ), ranges (26 values from 8 studies, using the formulas suggested for the respective sample size from Hozo et al.<sup>24</sup> and Confidence Intervals (CIs, two values from one study, using the formula from the Cochrane handbook<sup>23</sup>:  $[SE = (\text{upper limit} - \text{lower limit})/3.92]$ ) and converting them to SEs, adding columns for categorised variables and risk of bias counts per study, and adapting variable names, text and factor variables to be compatible with R.

All analyses were performed in R<sup>25</sup> via RStudio, using the following packages: readxl<sup>26</sup>, dplyr<sup>27</sup>, ggplot2<sup>28</sup>, tm<sup>29</sup>, tidyR<sup>30</sup>, crosstable<sup>31</sup>, meta<sup>32</sup> and metafor<sup>33</sup>. Two papers described both animal and human data<sup>4,34</sup>. Because we performed all analyses at the paper level, these two papers were excluded from all analyses other than the bibliographic ones. Comparisons of fractions of animal and human papers were made with Wilcoxon's rank sum test.

Some authors report the potential as is (i.e., negative at baseline, e.g.,<sup>35</sup>). Others plot negative values in the positive [e.g.,<sup>36</sup>, or reverse the potential in reporting [e.g.,<sup>37</sup>. Consequently, for baseline potential differences, we analysed absolute values. This should result in meaningful analyses with interpretable outcomes. We calculated a simple unweighted Spearman's correlation coefficient for absolute within-study nPD values in CF versus control, using base R, prior to our planned analyses.

Responses to low or zero chloride solutions were reported as absolute values with and without reversion, but also as absolute and percent changes compared to baseline. As these values can cross zero, analysing absolute values would not result in meaningful analyses with interpretable outcomes. Thus, we analysed absolute difference scores between the CF and the control values. As these data were extracted in addition to the protocol, we did not design these analyses upfront.

The following weighted meta-analyses were performed:

1. A planned overall translational meta-analysis for the baseline difference in nPD between cystic fibrosis and control without treatment, comparing human patients and animal models as subgroups. This analysis used standardised mean differences and the "metacont" function from the meta package<sup>32</sup>.
2. an additional (unplanned) overall translational meta-analysis for the difference in nPD between cystic fibrosis and control after low or no chloride buffer perfusion, comparing human patients and animal models as subgroups. This analysis used the "metagen" function from the meta package<sup>32</sup>.
3. a planned meta-regression analysis assessing the effects of buffer perfusion speed and species (mouse or human) on the baseline nPD. This analysis used the "rma" function from the metafor package ref<sup>33</sup>. As a rule

Bias domain	Type of bias	Operationalisation*
Baseline characteristics	Selection bias	Where the CF and control subjects comparable at baseline, at least for age and sex?
Housing	Performance bias	Was housing of animals mixed over the cages or adequately randomised/counterbalanced?
Blinding	Performance bias	Were all involved in contact with the subjects adequately blinded?
Outcome assessment	Detection bias	Was the order of measuring CF and control subjects adequately randomised/counterbalanced?
Blinding	Detection bias	Were those performing the nPD and those analysing the traces adequately blinded?
Missing outcome data	Attrition bias	Were missing outcome data transparently reported and equally distributed over CF and control groups?
Selective outcome reporting ("cherry picking")	Reporting bias	Was a protocol posted before study start, and were all planned outcomes reported?
Other	Other bias	Were there no other reasons for concerns of biassing the results?

**Table 3.** Analysed risk of bias domains (based on<sup>18</sup> and<sup>23</sup>) described in this publication. \*Yes reflects a low risk of bias, no reflects a high risk of bias, unclear reflects insufficient information to be sure.

of thumb, regression analyses should not be performed if there are fewer than 10 observations for a specific (level of a factorial) variable. Therefore, our protocol specified restriction of meta-regression to variables for which we had data from at least 10 papers. In line with this idea, but not explicitly described in our protocol, we excluded studies of pigs ( $k=2$  included papers) and rats ( $k=4$  included papers) from this meta-regression.

All meta-analyses used a frequentist (classical) random effects model with standardised mean differences. Heterogeneity between the studies was assessed using the  $I^2$  statistic. Publication bias was assessed with funnel plots and trim&fill analyses.

## Results

### Study flow and included literature sample

The flow of studies from search to included is summarised in Fig. 1.

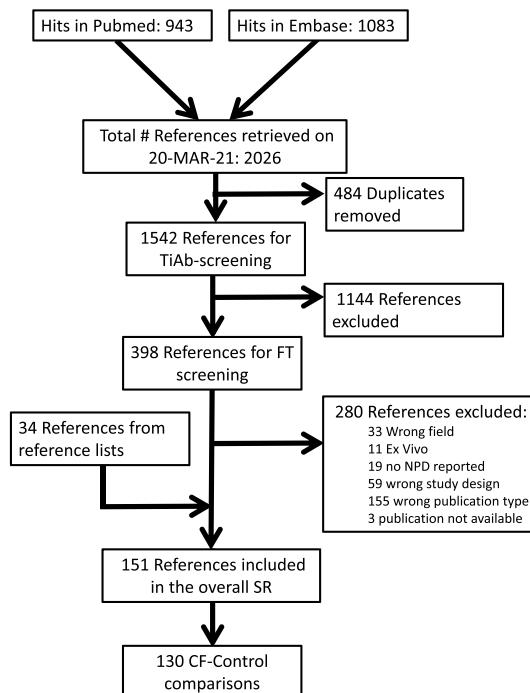
The 130 papers describing baseline CF-control comparisons comprise 55 papers about animal studies, 73 about human studies, and 2 describing both. Of the 55 animal studies, 13 reported multiple relevant experimental groups. Of the 73 human studies, 16 reported multiple relevant experimental groups. The number of relevant CF-control comparisons per paper was 2 for 28 papers (12 animal papers, 14 human papers and 2 papers describing both an animal and a human CF-control comparison), and 3 for the remaining 3. In animal SRs, it is common practice to extract data from different experiments within a single reference, and even from different groups analysed within a single experiment, and analyse them as if the data were independent. Following that practise, we could in total have included 164 ( $99 + 28 \times 2 + 3 \times 3$ ) comparisons. However, particularly for this baseline comparison, we were worried about issues with correlated data. To prevent these, data were only extracted and analysed for the first-described experimental CF-control comparison meeting the inclusion criteria.

The included nPD animal studies were published from 1992 through 2020; the human studies from 1981 through 2016 (Fig. 2). Animal studies were on average published more recently than human studies (Wilcoxon rank sum test:  $W=2540$ ,  $p=0.01$ ).

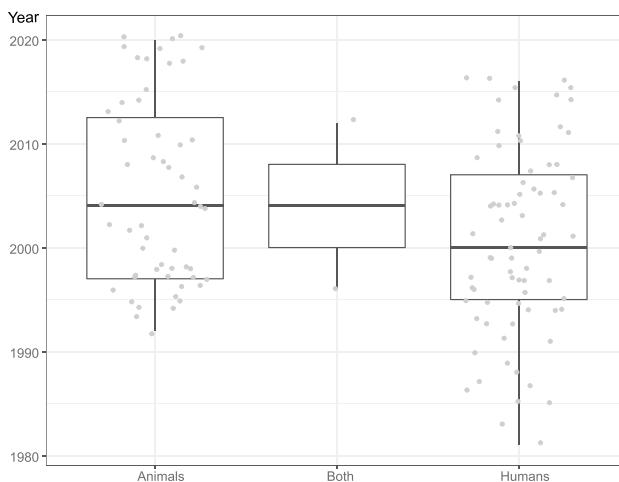
The 4 most popular journals were the American journal of respiratory and critical care medicine ( $k=2$  animal and  $k=12$  human papers); the American journal of respiratory cell and molecular biology ( $k=4$  animal and  $k=3$  human papers); Human gene therapy ( $k=4$  animal and  $k=2$  human papers); and Thorax ( $k=1$  animal and  $k=5$  human papers).

The language of most papers ( $k=126$ , including all animal papers) was English. Three human papers were in French, 1 was in German. We extracted the countries in which the protocols for the included studies had been submitted to ethical review. These data were available for  $k=81$  of the 130 included papers. The majority of the studies was approved in the USA ( $k=18$  animal and  $k=16$  human studies). Other countries with more than 5 approved studies were Belgium (overall  $k=7$ ), France ( $k=7$ ), Germany ( $k=6$ ) and Canada ( $k=6$ ).

A simple text frequency analysis of the titles of the included studies showed no relevant differences in the main topics covered between included animal and human studies (data not shown).



**Figure 1.** Reference flow.



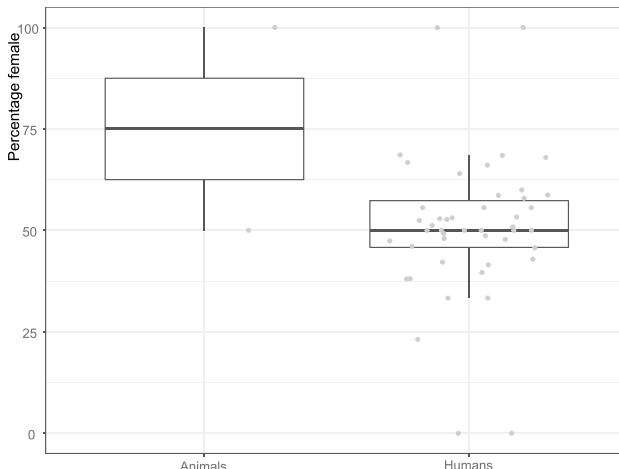
**Figure 2.** Publication dates by study population.

### Population

Because we performed all analyses at the paper level, the two papers describing both animal and human studies<sup>5,34</sup> were excluded from further analyses. The animal species studied were mice ( $k=49$ ), rats ( $k=4$ ) and pigs ( $k=2$ ). Sex/gender was reported in  $k=68$  (53.1%) of the included papers. Most of these analysed both males and females ( $k=63$ ), 2 exclusively analysed males, 3 exclusively analysed females. Animal papers were responsible for the main part of underreporting, only two out of the 55 animal papers reported sex. In the studies of both genders /sexes, the percentage female varied from 11.8 to 89.0 (Fig. 3), with median values for CF and control groups around 50%.

The most studied mutation was the p.F508del mutation ( $k=22$  animal and  $k=9$  human papers). Multiple mutations were analysed in 2 (out of 55) animal and 22 (out of 73) human papers. The next commonly studied mutation was the *CFTR* knock-out ( $k=23$  animal papers). Other mutations (e.g. G551D and G542X) were specifically studied in fewer than 5 of the included papers. Information about the analysed mutation(s) was not reported in 1 animal and 39 human papers. While many other CF animal models exist<sup>38</sup>, all included animal studies described genetic animal models. Disease severity was (briefly) mentioned in 30 human and 2 animal papers. In humans, it varied from “not severe” to “frequently hospitalised”. In animals, one mentioned that the disease was not severe<sup>39</sup>, another described strain differences for CF disease severity<sup>36</sup>.

Some information on age or weight of the subjects was reported in 27 (out of 55) animal and 60 (out of 73) human studies. Animal (mouse) ages ranged from 0 to 56 weeks, human ages from 0 to 70 years. Mouse weight varied from 15 to 53g. The immune status was only reported in 7 animal papers, all specific pathogen-free. Prevention of intestinal obstruction (a common problem in CF animal models) was described in 25 animal papers and targeted diet ( $k=2$  liquid,  $k=1$  other), laxatives ( $k=1$ ), a combination of diet and laxative ( $k=1$ ) and genetic correction of the gut ( $k=2$ ).



**Figure 3.** Percentage female by species.

Of the human papers, 16 mentioned something about ongoing background treatments (e.g. “no regular medications” or “pancreatic supplements and antibiotic treatment”). None of the animal papers described co-medication.

### Experimental design

None of the papers included in this part of the review reported protocol registration. Note that registered protocols were more common in the treatment vs. control comparisons.

Part of the methods were referenced by 33 (out of 55) animal and 53 (out of 73) human papers, explaining part of the underreporting of experimental details described below. However, several authors only half referenced the methods (e.g. “a modification of ...” or “adapted from ...”<sup>36,40</sup>), without details on what was modified and adapted, which limits reproducibility.

Total numbers of study subjects were reported by 41 (out of 55) animal and 70 (out of 73) human papers, and ranged from 2 to 175 for animal and from 8 to 260 for human papers. On average, human studies were significantly larger than animal studies ( $W = 784$ ,  $p = 0.001$ , Fig. 4).

CF animals were mainly compared to heterozygous animals ( $k = 32$ ). The use of control littermates (wildtype, heterozygous or both) were only specified in  $k = 9$  papers. CF patients were mostly compared to healthy controls ( $k = 63$ ), with or without the exclusion of specific diseases. Repeated nPD measurements were reported in  $k = 5$  animal papers and in  $k = 15$  human papers. Animals were measured up to 5 times, humans up to 6 times. Time between repeated measurements varied from a day to nearly 2 months.

Only two papers provided information on the time the tests were performed;  $k = 1$  animal study described performing the tests “throughout the day”<sup>35</sup>, one human paper “in the afternoon”<sup>41</sup>. For studies of nocturnal rodents, the light cycle can be reversed. Only  $k = 5$  (out of 55) animal studies provided some information about the light cycle, which was normal ( $k = 2$ ), 12:12 ( $k = 2$ ) and reversed ( $k = 1$ ). Only 3 animal (out of 128) papers reported laboratory temperature, which was 21 °C ( $k = 2$ ) or 22 °C ( $k = 1$ ). None of the included papers reported laboratory humidity.

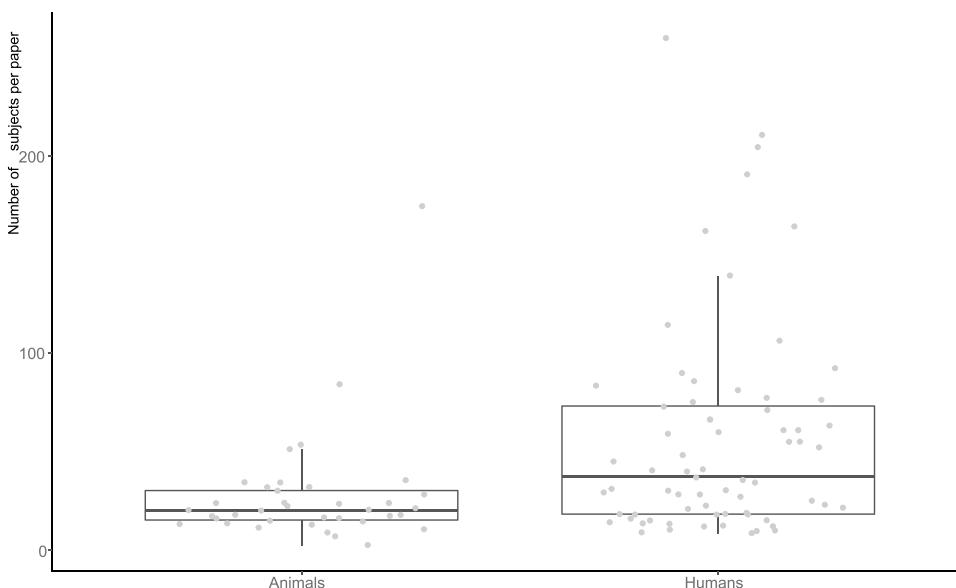
The number of subjects leaving the study before the last nPD (“DropOuts”) were only mentioned by two human papers (which had none, or at least 4).

### nPD technique

For animals, the nPD can only be performed under anaesthesia. The anaesthetics used were reported in  $k = 33$  (out of 55) animal papers. The most commonly used combination anaesthetic was ketamine-xylazine. Other combinations and single anaesthetics were used in fewer than 3 papers each.

Information about the measuring electrode was reported in 32 (out of 55) animal and 51 (out of 73) human papers. Commonly used measuring electrodes were various Foley catheters, (pulled) PE-10 and PE-20 tubing, umbilical catheters and double-lumen silicone rubber tubes. Information about the bio-electrical connection was reported in 19 (out of 55) animal and 45 (out of 73) human papers and comprised continuous flow of the baseline buffer, agar bridges or electrode cream.

Information about the placement of the measuring electrode was provided in  $k = 37$  (out of 55) animal and  $k = 58$  (out of 73) human papers. For the animal papers, this information was mainly “in a nostril”, with or without a depth. For human papers, terms used in more than 5 papers were “under the inferior turbinate” and “floor of the nasal cavity/nose”. Testing of the placement was described for  $k = 5$  animal and  $k = 29$  human papers. Most



**Figure 4.** Total number of subjects for which data were extracted per paper.

of this placement testing comprised advancement or withdrawal of the measuring electrode up to the maximal potential difference.

Information about the placement of the reference electrode was provided in  $k=29$  (out of 55) animal and  $k=52$  (out of 73) human papers. In humans, the reference electrode was most commonly placed in (subcutaneously) or on (abraded skin) the forearm. In animals, it was mainly placed subcutaneously. If the location was specified, it was mostly the abdomen or a paw.

The perfusion solutions were reported for  $k=51$  (out of 55) animal and 57 (out of 73) human papers. Most frequently used were various solutions with Amiloride, different buffers with no or low chloride, and different types of Ringer. The only noticeable difference between animal and human studies was the use of Isoproterenol, which was more common in animal studies ( $k=29$  versus  $k=6$ ). Duration of the perfusion was reported for  $k=32$  (out of 55) animal and  $k=35$  (out of 73) human papers. It ranged from 0.2 to 47.5 min for animals, and from 1.2 to 35.0 min for human subjects. Information about the perfusion device was reported in 16 (out of 55) animal and 15 (out of 73) human papers, and comprised various types of peristaltic, microperfusion and syringe pumps.

Information about the baseline buffer composition was reported in 47 (out of 55) animal and 52 (out of 73) human papers. It comprised exact concentrations, but standard (HEPES-buffered) Ringers and Krebs solutions were also regularly reported. The perfusion speed was reported in 32 (out of 55) animal and 39 (out of 73) human papers. The mostly used units to describe the flow were mL/min ( $k=2$  animal and  $k=32$  human papers),  $\mu$ L/min ( $k=15$  animal and  $k=2$  human papers) and mL/h ( $k=9$  animal and  $k=3$  human papers). After conversion to mL/min, flow ranged from 0.00001 to 0.1 for mice, and from 0.05 to 5.0 for humans.

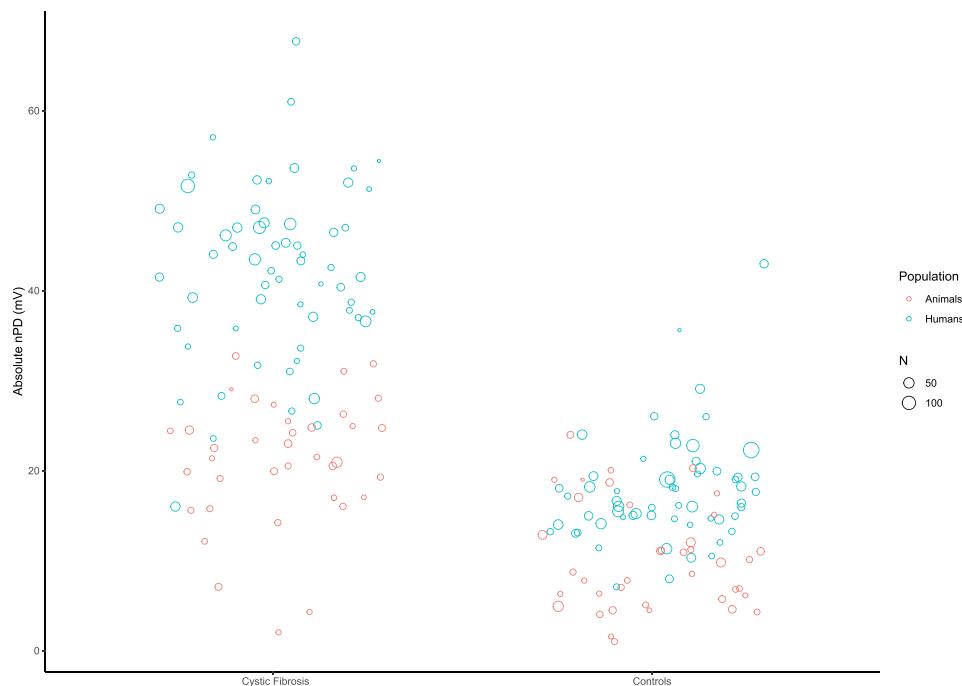
### nPD outcomes in CF and control

Overall, values were estimated (from a provided graph/range/CI/IQR) for at least one of the outcomes (baseline or low chloride nPD) or their associated parameters (SEs or  $n$ ) for 48 out of 128 papers. Because baseline nPD values are all negative, but some authors reverse the sign, we analysed absolute values for baseline values and difference scores for low chloride nPDs.

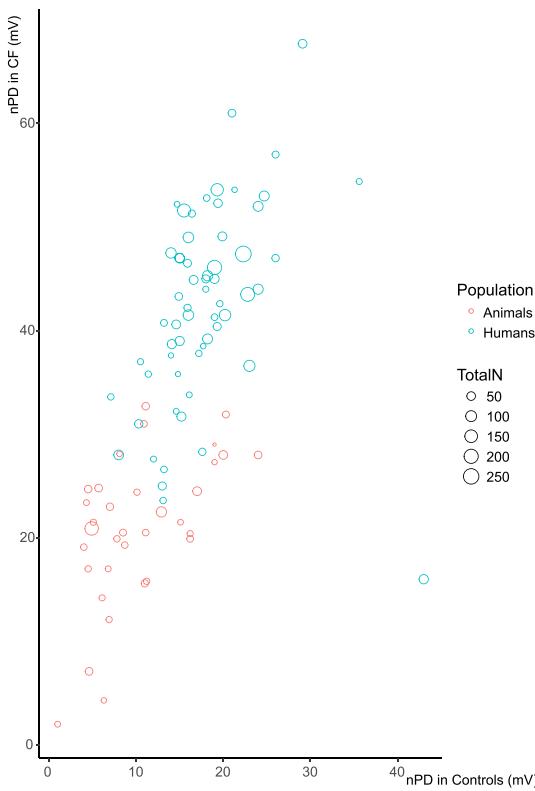
Baseline nPD data were available for 43 (out of 55) animal and 62 (out of 73) human papers. For animals, absolute control nPDs ranged from 1.0 to 24.0 mV, with a median value of 9.0mV, and absolute CF nPDs from 2.0 to 32.7mV, with a median value of 21.5mV. For humans, absolute control nPDs ranged from 7.1 to 43.0mV, with a median value of 16.9mV, and absolute CF nPDs from 16.0 to 67.7mV, with a median value of 43.0 (Fig. 5).

A clear within-study correlation between CF and control values (Fig. 6,  $r(103)=0.69$ ,  $p<0.01$ , uncorrected for study size) suggests that while there is a convincing and reproducible difference in the nPD values between CF and control (Fig. 5), the values heavily depend on the local protocols, and the animal protocols on average result in lower values than the human protocols.

A meta-analysis of the nPD in CF versus control, subgrouped by animal vs. human studies, showed a clear overall difference between CF and control (Overall SMD: 0.56; 95%-CI: 0.46/0.66;  $z=-10.79$ ;  $p<0.0001$ ), and a trend-level difference between animals and humans ( $Q=3.48$ ;  $df=1$ ;  $p=0.062$ ). However, heterogeneity was high ( $I^2=50.4\%$  overall, 63.1% for animals and 38.5% for humans) and small study effects were observed. In the



**Figure 5.** Baseline nPD values in animals and humans with and without CF.



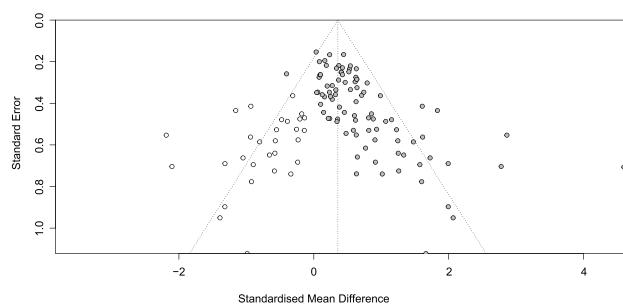
**Figure 6.** Correlation of baseline CF and control nPD values.

trim-and-fill analysis, 30 studies were added, and the overall difference between CF and control became smaller (Overall SMD: 0.36; 95%-CI: 0.21/0.50;  $z = 4.89$ ;  $p < 0.0001$ , Fig. 7).

Low chloride nPD data were available for 38 (out of 55) animal and 47 (out of 73) human papers. For animals, absolute CF-control difference scores ranged from 0.7 to 24.8 mV, with a median value of 10.9 mV. For humans, difference scores ranged from 1.6 to 36.4 mV, with a median value of 16.0 mV. The meta-analysis shows an overall difference between CF and control ( $z = 16.6$ ;  $P < 0.0001$ ), and between the animal and human subgroups ( $Q = 16.1$ ;  $p < 0.0001$ ). Note that heterogeneity was high;  $I^2 = 95.4\%$  (95% CI 94.7–96.0%). Besides, small study effects were present. In the trim and fill analysis 19 studies were added and the overall difference between CF and control became smaller.

### Meta-regression

The only relevant variables with sufficient data ( $k \geq 10$ ;  $k \geq 10$  per category for categorical variables) to analyse in the planned meta-regression were buffer flow rate and species. Note that while data from rats ( $k = 4$ ) and pigs ( $k = 2$ ) were included in all analyses described above, they were excluded from the meta-analyses. The results from the meta-regression are shown in table 4.



**Figure 7.** Funnel plot. Open circles reflect added studies.

Variable	Range/values	Estimate (SE) in mV	p-value
(Intercept)		9.73 (0.90)	<0.0001
Buffer flow rate (mL/min)	0.00001 to 5.0	0.38 (0.40)	0.35
Species	Mice/humans	8.26 (1.64)	<0.0001

**Table 4.** Results from the meta-regression.

### RoB

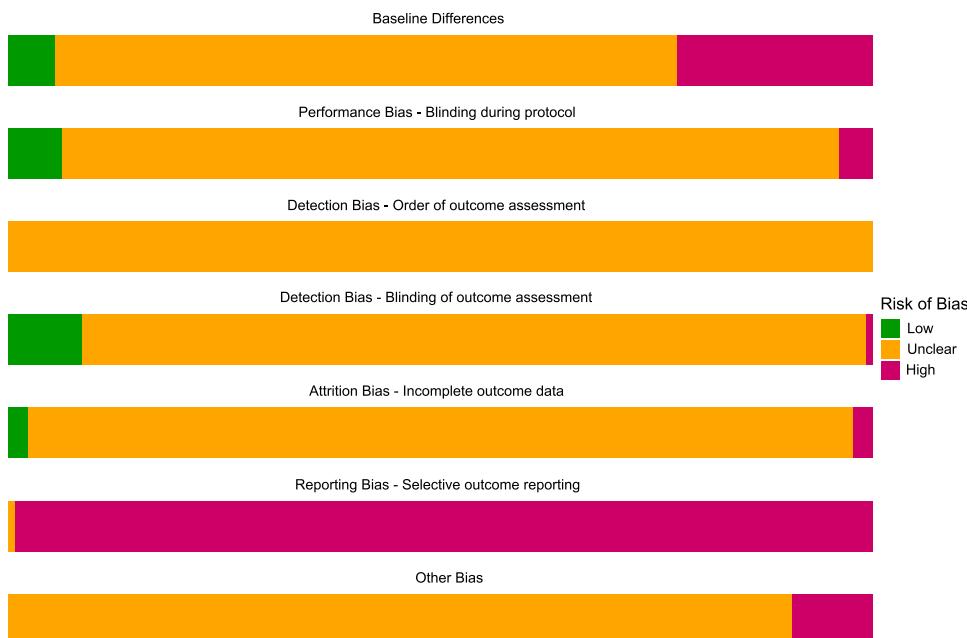
The median number of RoB items scored “unclear” per study was 7 for animal and 5 for human papers, and the number ranged from 2 to 7. Human studies scored significantly fewer “unclears”, reflecting more complete reporting, than animal studies ( $W = 3667$ ,  $p < 0.001$ ). The median number of RoB items scored “high” per study was 1 for both animal and human papers. The median number of RoB items scored “low” per study was 0 for both animal and human papers. The scores of the included studies per item are shown in Fig. 8, except for the item “housing”, which is unique to animal studies. Risk of bias due to non-random or unbalanced housing was unclear for all  $k=55$  included animal studies. More detailed presentation of our analyses of RoB and reporting quality will be presented in a separate publication.

### Conclusions

This SR summarises 130 papers comparing baseline nPD values between CF and control. It confirmed substantial variation in the experimental design between groups, as previously described<sup>42</sup>. These variations may contribute to differences in the nPD outcome value. Consistent with the observation that the human nPD is a reproducible test when performed in a single centre, and extending this observation to animal studies, we showed a clear correlation between CF and control values within studies. Our study further confirmed a clear difference in baseline nPD values between CF and control, both in animals and in humans. However, baseline nPD values were, on average, lower in animal than in human studies.

To the best of our knowledge, this is the first SR of the nPD, and one of the few SRs directly comparing a specific outcome between animals and humans. Other systematic animal-human comparisons focussed on dental outcomes<sup>43–45</sup> and haemodynamic effects (blood pressure, heart rate)<sup>46</sup>. Experimental design was also previously compared between animal and human studies in a systematic manner<sup>47</sup>. We here provide a full overview of the use of the nPD up to our search date (23 March 2021), which showed relevant differences between animal and human studies.

While this search date could be seen as a limitation, SRs generally take a long time to complete, with a median of 66 weeks and extremes up to 186 from start to completion<sup>48</sup>. Full review updates are crucial if the newly available literature could alter the conclusions of a review. For the here-presented baseline comparisons of CF versus control in animals and humans we would not expect new data to change the conclusions, but additional data could potentially allow us to perform more meaningful analyses of factors affecting the nPD. Thus, we tested if

**Figure 8.** Summary of the risk of bias.

an update would be informative, by a brief analysis of an updated PubMed search performed on 26 May 2023, as described in the discussion of our preceding manuscript of treatment effects<sup>22</sup>. Only five new references met our inclusion criteria for the here-presented CF-control comparisons<sup>49–53</sup>. Based on this small number, particularly with the observed low level of reporting experimental details (discussed below), we do not expect a full update to allow for more meaningful analyses. A formal update would delay publication of these findings further, especially with peer review currently lasting over 6 months. Thus, we do not think an update is currently warranted.

In general, reproducibility of scientific experiments depends on knowledge of the experimental design details. For this SR, reporting of experimental details was worse than anything we came across before (the authors have worked on over 20 other SRs). Variables of specific concern are disease severity (only reported in 32 papers), background treatments (only reported in 16 human papers), time of testing (only reported in very broad terms in two papers) and the number of subjects not completing the experiments (only reported in 4 papers). Besides, outcome measures varied between studies. Furthermore, specific to animal studies, details about husbandry (light cycle) and the tested sexes were frequently lacking. The poor reporting hinders reproducibility, but it also limited the analyses we could perform. To ensure reproducibility of experiments within and between laboratories and species, we strongly encourage everyone reporting nPD experiments to not only follow the arrive guidelines, but also report in full all details about the used voltmeters, electrodes and other equipment used.

Concerning data synthesis, the limited reporting of experimental details prevented us to include additional variables in our meta-regression. In the other part of this review, where we were planning to synthesise treatment effects, the limited reporting even prevented any relevant analysis from being performed<sup>22</sup>. We hope that reporting of experimental detail will improve over time, which would allow for more meaningful analyses in future systematic reviews.

Besides limitations due to poor reporting, five further limitations restrict the reliability of our summary findings. First, several of the included papers describe datasets that overlap to some extent, second, our search may have missed relevant references, third, the results may be confounded, fourth, small study effects were present, and fifth, risk of bias in the included studies was substantial.

We tried to prevent overlap in included data sets by excluding data that fully overlapped; the previously reported control data from<sup>54</sup>, which fully overlap with two other included publications<sup>55,56</sup>, were thus excluded from all analyses. With partial overlap we would have lost information by excluding summary values, and we could not always be fully aware of the (extent of the) overlap anyhow. There is one example where overlap is certainly relevant: four of the pigs described in Chen et al. 2010<sup>57</sup> were also part of the analyses by Rogers et al.<sup>58</sup>.

The sensitivity of our search may be disputed, because “snowballing” retrieved 34 additional references. An informal analysis of the (indexing) terms used in these references shows that several of the additional references did not mention the nasal potential difference in title, abstract, or author-provided keywords, and that the databases did not add index terms for potential differences for these records either. However, most added references were missed by our searches because our human search string was developed to retrieve translational studies, focussing on randomised clinical trials. For future literature studies it needs to be extended; terms reflecting different human study types (case-control studies and diagnostic studies for example) need to be added.

The animal-human comparison could be confounded by publication date, as we found a clear difference between the mean publication date of the animal and human studies. Besides, we cannot rule out confounding for variables which were poorly reported, for example sex and time of testing. While most of these variables theoretically do not affect the nPD, caution with interpretation of the findings in this review is warranted.

For future literature analyses on nPD, we recommend extended data extraction to allow for further analyses of the experimental set-ups, e.g. adding the warming of the perfusion solutions<sup>59</sup>, used voltmeters, used electrodes, and tilt for rodent studies. These were not formally analysed for this SR, but our casual observations during extraction suggests that these data were not reported less than the ones we analysed.

In this review, we found that nPD outcome effect sizes are not fully comparable between animals and humans. However, the nPD and the CF animal models show resemblance with human CF patients, indicating construct validity of the animal models. To ensure reproducibility of experiments within and between species, there is a huge need to improve reporting.

## Data availability

All used search strings are publicly available, either within this this manuscript or in one of the references. All search results, screening results and extracted data are available on OSF (<https://osf.io/fkyhc/>).

Received: 8 September 2023; Accepted: 23 April 2024

Published online: 26 April 2024

## References

- Middleton, P. G., Geddes, D. M. & Alton, E. W. Protocols for in vivo measurement of the ion transport defects in cystic fibrosis nasal epithelium. *Eur. Respir. J.* **7**, 2050–2056 (1994).
- Naehrlich, L. *et al.* Nasal potential difference measurements in diagnosis of cystic fibrosis: An international survey. *J. Cyst. Fibros.* **13**, 24–28 (2014).
- Fajac, I. *et al.* Relationships between nasal potential difference and respiratory function in adults with cystic fibrosis. *Eur. Respir. J.* **12**, 1295–1300 (1998).
- Leonard, A. *et al.* Comparative variability of nasal potential difference measurements in human and mice, healthy or carrying two severe CFTR mutations. *J. Cyst. Fibros.* **11**, S62 (2012).
- Leonard, A., Vermeulen, F., De Boeck, K., Leal, T. & Lebecque, P. Nasal potential measurements: Repeatability and reproducibility of TCS in patients homozygous for the F508del mutation and healthy controls. *J. Cyst. Fibros.* **11**, S68 (2012).
- Kyrilli, S. *et al.* Insights into the variability of nasal potential difference, a biomarker of CFTR activity. *J. Cyst. Fibros.* **19**, 620–626 (2020).

7. Solomon, G. M. *et al.* Standardized measurement of nasal membrane transepithelial potential difference (NPD). *J. Vis. Exp.* <https://doi.org/10.3791/57006-v> (2018).
8. Milanowski, A. & Piotrowski, R. Nasal potential difference measurement in diagnosis and monitoring of cystic fibrosis patients [Polish]. *Pediatria Polska* **81**, 320–322 (2006).
9. Sands, D. Transepithelial nasal potential difference (NPD) measurements in cystic fibrosis (CF). *Medycyna wieku rozwojowego* **17**, 13–17 (2013).
10. Leenaars, C. H. C. *et al.* Animal to human translation: A systematic scoping review of reported concordance rates. *J. Transl. Med.* **17**, 223. <https://doi.org/10.1186/s12967-019-1976-2> (2019).
11. Van de Wall, G. *et al.* Comparing Translational success rates across medical research fields: A combined analysis of literature and clinical trial data. *ALTEX* (2023) (Accepted for publication).
12. Leenaars, C. H. C., Teerenstra, S., Meijboom, F. L. B. & Bleich, A. Methodical advances in reproducibility research: A proof of concept qualitative comparative analysis of reproducing animal data in humans. *J. Neurosci. Methods* **397**, 109931. <https://doi.org/10.1016/j.jneurometh.2023.109931> (2023).
13. van der Mierden, S. *et al.* Laboratory animals search filter for different literature databases: PubMed, Embase, Web of Science and PsycINFO. *Lab. Anim.* **56**, 279–286. <https://doi.org/10.1177/00236772211045485> (2022).
14. Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A. Rayyan—a web and mobile app for systematic reviews. *Syst. Rev.* **5**, 210. <https://doi.org/10.1186/s13643-016-0384-4> (2016).
15. Covidence systematic review software, <http://www.covidence.org>.
16. De Wachter, E., De Schutter, I., Meulemans, A., Buyl, R. & Malfroot, A. A semi-blinded study comparing 2 methods of measuring nasal potential difference: Subcutaneous needle versus dermal abrasion. *J. Cyst. Fibros.* **15**, 60–66. <https://doi.org/10.1016/j.jcf.2015.06.007> (2016).
17. Duperrex, O. *et al.* A new device for in vivo measurement of nasal transepithelial potential difference in cystic fibrosis patients and normal subjects. *Eur. Respir. J.* **10**, 1631–1636 (1997).
18. Hooijmans, C. R. *et al.* SYRCLE's risk of bias tool for animal studies. *BMC Med. Res. Methodol.* **14**, 43. <https://doi.org/10.1186/1471-2288-14-43> (2014).
19. Higgins, J. P. T., Savović, J., Page, M. J., Elbers, R. G. & Sterne, J. A. C. *Cochrane Handbook for Systematic Reviews of Interventions* (2022).
20. Kilkenny, C., Browne, W. J., Cuthill, I. C., Emerson, M. & Altman, D. G. Improving bioscience research reporting: The ARRIVE guidelines for reporting animal research. *PLoS Biol.* **8**, e1000412. <https://doi.org/10.1371/journal.pbio.1000412> (2010).
21. Begg, C. *et al.* Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* **276**, 637–639. <https://doi.org/10.1001/jama.276.8.637> (1996).
22. Leenaars, C., Hager, C., Stafleu, F., Nieraad, H. & Bleich, A. A systematic review of the effect of cystic fibrosis treatments on the nasal potential difference test in animals and humans. *Diagnostics* <https://doi.org/10.3390/diagnostics13193098> (2023).
23. Higgins, J. P. T., Li, T. & Deeks, J. J. *Cochrane Handbook for Systematic Reviews of Interventions* (2022).
24. Hozo, S. P., Djulbegovic, B. & Hozo, I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med. Res. Methodol.* **5**, 13. <https://doi.org/10.1186/1471-2288-5-13> (2005).
25. R\_Core\_Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2021).
26. Wickham, H. & Bryan, J. *readxl: Read Excel Files*. R package version 1.3.1. (2019).
27. Wickham, H., François, R., Henry, L. & Müller, K. *dplyr: A Grammar of Data Manipulation*. R package version 1.0.3., 2021).
28. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
29. Feinerer, I. & Hornik, K. *Text Mining Package*. R package version 0.7-8. (2020).
30. Wickham, H. & Girlich, M. *tidy: Tidy Messy Data*. R package version 1.2.0. (2022).
31. Chaitiel, D. *crosstable: Crosstables for Descriptive Analyses*. R package version 0.5.0. (2022).
32. Balduzzi, S., Rücker, G. & Schwarzer, G. How to perform a meta-analysis with R: A practical tutorial. *Evid.-Based Ment. Health* **22**, 153–160 (2019).
33. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **36**, 1–48. <https://doi.org/10.18637/jss.v036.i03> (2010).
34. McLachlan, G. *et al.* Laboratory and clinical studies in support of cystic fibrosis gene therapy using pCMV-CFTR-DOTAP. *Gene Ther.* **3**, 1113–1123 (1996).
35. Beumer, W. *et al.* Evaluation of eluforsen, a novel RNA oligonucleotide for restoration of CFTR function in in vitro and murine models of p.Phe508del cystic fibrosis. *PLoS ONE* **14**, e0219182 (2019).
36. Rozmahel, R. *et al.* Incomplete rescue of cystic fibrosis transmembrane conductance regulator deficient mice by the human CFTR cDNA. *Hum. Mol. Genet.* **6**, 1153–1162. <https://doi.org/10.1093/hmg/6.7.1153> (1997).
37. Smith, S. N. *et al.* Bioelectric characteristics of exon 10 insertional cystic fibrosis mouse: Comparison with humans. *Am. J. Physiol.* **268**, C297–307. <https://doi.org/10.1152/ajpcell.1995.268.2.C297> (1995).
38. Leenaars, C. H. *et al.* Animal models for cystic fibrosis: A systematic search and mapping review of the literature. Part 2: Nongenetic models. *Lab. Anim.* **55**, 307–316. <https://doi.org/10.1177/0023677221990688> (2021).
39. van Doorninck, J. H. *et al.* A mouse model for the cystic fibrosis delta F508 mutation. *EMBO J.* **14**, 4403–4411. <https://doi.org/10.1002/j.1460-2075.1995.tb00119.x> (1995).
40. Robinson, E. *et al.* Lipid nanoparticle-delivered chemically modified mRNA restores chloride secretion in cystic fibrosis. *Mol. Ther.* **26**, 2034–2046. <https://doi.org/10.1016/j.ymthe.2018.05.014> (2018).
41. Southern, K. W. *et al.* A modified technique for measurement of nasal transepithelial potential difference in infants. *J. Pediatr.* **139**, 353–358 (2001).
42. Saussereau, E. L. *et al.* Characterization of nasal potential difference in cftr knockout and F508del-CFTR mice. *PLoS ONE* **8**, e57317. <https://doi.org/10.1371/journal.pone.0057317> (2013).
43. Faggion, C. M. Jr., Chambrone, L., Gondim, V., Schmitter, M. & Tu, Y. K. Comparison of the effects of treatment of peri-implant infection in animal and human studies: Systematic review and meta-analysis. *Clin. Oral Implants Res.* **21**, 137–147. <https://doi.org/10.1111/j.1600-0501.2009.01753.x> (2010).
44. Valles, C. *et al.* Influence of subcrestal implant placement compared with equicrestal position on the peri-implant hard and soft tissues around platform-switched implants: A systematic review and meta-analysis. *Clin. Oral Investig.* **22**, 555–570. <https://doi.org/10.1007/s00784-017-2301-1> (2018).
45. Yen, C. C., Tu, Y. K., Chen, T. H. & Lu, H. K. Comparison of treatment effects of guided tissue regeneration on infrabony lesions between animal and human studies: A systematic review and meta-analysis. *J. Periodontal Res.* **49**, 415–424. <https://doi.org/10.1111/jre.12130> (2014).
46. Sultan, S. R., Millar, S. A., England, T. J. & O'Sullivan, S. E. A systematic review and meta-analysis of the haemodynamic effects of cannabidiol. *Front. Pharmacol.* **8**, 81. <https://doi.org/10.3389/fphar.2017.00081> (2017).
47. Leenaars, C. *et al.* A systematic review comparing experimental design of animal and human methotrexate efficacy studies for rheumatoid arthritis: Lessons for the translational value of animal studies. *Animals* <https://doi.org/10.3390/ani10061047> (2020).
48. Borah, R., Brown, A. W., Capers, P. L. & Kaiser, K. A. Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. *BMJ Open* **7**, e012545. <https://doi.org/10.1136/bmjopen-2016-012545> (2017).

49. Cmielewski, P., Delhove, J., Donnelley, M. & Parsons, D. Assessment of lentiviral vector mediated CFTR correction in mice using an improved rapid in vivo nasal potential difference measurement protocol. *Front. Pharmacol.* **12**, 714452. <https://doi.org/10.3389/fphar.2021.714452> (2021).
50. Sermet-Gaudelus, I., Nguyen-Khoa, T., Hatton, A., Hayes, K. & Pranke, I. Sweat chloride testing and nasal potential difference (NPD) are primary outcome parameters in treatment with cystic fibrosis transmembrane conductance regulator (CFTR) modulators. *J. Pers. Med.* <https://doi.org/10.3390/jpm11080729> (2021).
51. Reyne, N. *et al.* Single-dose lentiviral mediated gene therapy recovers CFTR function in cystic fibrosis knockout rats. *Front. Pharmacol.* **12**, 682299. <https://doi.org/10.3389/fphar.2021.682299> (2021).
52. Noel, S. *et al.* Correlating genotype with phenotype using CFTR-mediated whole-cell Cl<sup>-</sup> currents in human nasal epithelial cells. *J. Physiol.* **600**, 1515–1531. <https://doi.org/10.1111/jp282143> (2022).
53. Piotrowski-Daspit, A. S. *et al.* In vivo correction of cystic fibrosis mediated by PNA nanoparticles. *Sci. Adv.* **8**, eab00522. <https://doi.org/10.1126/sciadv.ab00522> (2022).
54. Leal, T. *et al.* Airway ion transport impacts on disease presentation and severity in cystic fibrosis. *Clin. Biochem.* **41**, 764–772. <https://doi.org/10.1016/j.clinbiochem.2008.03.013> (2008).
55. Fajac, I. *et al.* Nasal airway ion transport is linked to the cystic fibrosis phenotype in adult patients. *Thorax* **59**, 971–976. <https://doi.org/10.1136/thx.2003.020933> (2004).
56. Leal, T., Lebacq, J., Lebecque, P., Cumps, J. & Wallemacq, P. Modified method to measure nasal potential difference. *Clin. Chem. Lab. Med.* **41**, 61–67. <https://doi.org/10.1515/cclm.2003.011> (2003).
57. Chen, X. *et al.* Nanomolar potency aminophenyltriazine CFTR activator reverses corneal epithelial injury in a mouse model of dry eye. *J. Ocul. Pharmacol. Ther.* **36**, 147–153. <https://doi.org/10.1089/jop.2019.0087> (2020).
58. Rogers, C. S. *et al.* Disruption of the CFTR gene produces a model of cystic fibrosis in newborn pigs. *Science* **321**, 1837–1841. <https://doi.org/10.1126/science.1163600> (2008).
59. Boyle, M. P. *et al.* A multicenter study of the effect of solution temperature on nasal potential difference measurements. *Chest* **124**, 482–489. <https://doi.org/10.1378/chest.124.2.482> (2003).

## Acknowledgements

The authors kindly acknowledge Antje Munder for advising on the review protocol as a CF expert.

## Author contributions

CL and AB acquired the grant to perform this work and designed this study. CL wrote the protocol and performed the searches. CL, FS, and HN screened the search results for inclusion. FS, HN and CH extracted the data. CL performed the analyses and drafted the manuscript. All authors revised the manuscript and approved of the final version.

## Funding

Open Access funding enabled and organized by Projekt DEAL. This research was funded by the BMBF, grant number 01KC1904. During grant review, the BMBF asked for changes in the review design which we incorporated. Publication of the review results was a condition of the call. Otherwise, the BMBF had no role in the collection, analysis, and interpretation of data, or in writing the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to C.H.C.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024