Sweat conductivity for diagnosing cystic fibrosis after positive newborn screening: prospective, diagnostic test accuracy study

Renata Marcos Bedran,1 Cristina Gonçalves Alvim,1 Olívia Gonçalves Sader,1 José Vicente Alves Júnior,2 Fernando Henrique Pereira,2 Daniela Magalhães Nolasco,2 Linjie Zhang,3 Paulo Camargos

ABSTRACT
Objective To assess the accuracy of sweat conductivity among newborns and very young infants.
Design Prospective, population-based, diagnostic test accuracy study.
Setting Public Statewide Newborn Screening Programme where the incidence rate of cystic fibrosis (CF) is 1:11 000.
Patients Newborns and very young infants with positive two-tiered immunoreactive trypsinogen.
Interventions Sweat conductivity and sweat chloride were performed simultaneously, on the same day and facility by independent technicians, with the cut-off values of 80 mmol/L and 60 mmol/L, respectively.
Main outcome measures Sensitivity, specificity, positive and negative predictive values (PPV and NPV), overall accuracy, positive and negative likelihood ratios (+LR, -LR) and post (sweat conductivity (SC)) test probability were calculated to assess SC performance.
Results 1193 participants were included, 68 with and 1108 without CF, and 17 with intermediate values. The mean (SD) age was 48 (19.2) days, ranging from 15 to 90 days. SC yielded sensitivity of 98.5% (95% CI 95.7 to 100), specificity of 99.9% (95% CI 99.7 to 100), PPV of 98.5% (95% CI 95.7 to 100) and NPV of 99.9% (95% CI 99.7 to 100), overall accuracy of 99.8% (95% CI 99.6 to 100), +LR of 1091.7 (95% CI 153.8 to 7744.9) and -LR of 0.01 (95% CI 0.00 to 0.10). After a positive and negative sweat conductivity result, the patient’s probability of CF increases around 350 times and drops to virtually zero, respectively.
Conclusion Sweat conductivity had excellent accuracy in ruling in or ruling out CF after positive two-tiered immunoreactive trypsinogen among newborns and very young infants.

INTRODUCTION
Due to the limited availability and affordability of CFTR gene mutations analysis in resourcelimited settings—where, like in Brazil, only cases with abnormal chloride values undergo DNA testing, after positive two-tier IRT sweat chloride determination is the strategy implemented by several newborn screening programmes for cystic fibrosis (CF) worldwide.1 It offers maximal accuracy but requires a specialised centre and skilled personnel.1 In addition, the time-consuming Gibson and Cooke’s method is still used in those settings instead of the extensively recommended Macroduct system.1

WHAT IS ALREADY KNOWN ON THIS TOPIC
⇒ Sweat conductivity is considered a screening test for cystic fibrosis (CF); accuracy studies are needed to assess its role as an alternative method to rule in/out CF.

WHAT THIS STUDY ADDS
⇒ Sweat conductivity showed an excellent diagnostic accuracy; for instance, positive and negative likelihood ratios were 1091.7 and 0.01, respectively.
⇒ In settings where the incidence rate of CF is around 1:11 000, post-test probability increases 350 times the likelihood of CF after a positive sweat conductivity, while negative results virtually rule out CF.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY
⇒ In settings with a shortage of chloride titration, sweat conductivity can expedite diagnosis and allow early prompt treatment, minimising family and parental distress due to a positive newborn bloodspot screening.
titration of sweat chloride for diagnosing CF in newborns and young infants.

METHODS

Study design
We conducted this prospective, population-based, paired diagnostic test accuracy study between April 2013 and October 2020, following the Standards for Reporting Diagnostic Accuracy guidelines (STARD).15

Participants and predefined inclusion and exclusion criteria
We included newborns and infants younger than 3 months, public Statewide Newborn Screening Programme participants, who presented two IRT positive results (ie, >70 mg/L) in our CF referral centre, located in Belo Horizonte, Brazil. To ensure blinding and to avoid interference of other clinical conditions on the result of SC and chloride titration if they were done on different days, participants underwent both tests on the same day and facility, conducted simultaneously, once only, by independent technicians.

Exclusion criteria were skin lesions of any aetiology, clinical instability and insufficient sweat quantity.

Test methods and procedures
Data collection and recording were completed immediately after performing SC and chloride determination.

Collection of sweat samples by Macroduct system and SC analysis (index test)
Sweat collection was performed as described by others.3 7 10 11 Briefly, the technician cleaned the patient’s skin before placing the Macroduct coil. As recommended by the manufacturer, the sweat collection time was 30 min.16 Following the manufacturer’s guidelines, after sweat collection, the catheter was separated from the disk and a syringe was connected to one end of the catheter. To measure conductivity, the other end was connected to the Sweat-Chek analyzer.

Collection of sweat samples by Gibson and Cooke’s technique and chloride analysis (reference standard)
Gibson and Cooke’s method was performed according to well-established standards in the contralateral forearm.12 14 After pilocarpine iontophoresis, chloride quantitation was carried out through a chloridometer.12 14

Cut-off points for positive and negative results
For chloride concentration, we adopted the CFF and the reference values recommended by others, that is, results below 30 mmol/L and equal to or greater than 60 mmol/L, were considered negative and positive, respectively, indicating the definitive diagnosis of non-CF and CF.4 12 14 For SC, we followed both the manufacturer’s guidelines16 and the values reported by others.3 7 SC was considered negative when values were equal or lower than 60 mmol/L and positive when they were equal or greater to 80 mmol/L. Sweat chloride values of 30–59 mmol/L and SC values of 60–79 mmol/L were considered intermediate and analysed apart.

Statistics

Analysis
The primary analysis included only participants with positive and negative results for both methods; intermediate results

Figure 1 Flow diagram displaying the study cohort retrieval process. SC, sweat conductivity; ST, sweat test.
were not accounted for at this stage. We conducted descriptive statistics and constructed a 2×2 contingency table to estimate sensitivity, specificity, PPV, NPV, positive and negative likelihood ratios (+LR and -LR) and their respective 95% CI.

Intermediate values are expected to occur with chloride titration and SC. According to STARD guidelines, they can be ignored altogether, be reported but not accounted for or handled as a separate test result category. It is also possible to reclassify all such results as false positives or false negatives, depending on the reference standard result (‘worst-case scenario’) or as true positives and true negatives (‘best-case scenario’). As suggested by STARD guidelines, we conducted secondary analyses to reassess SC performance using both scenarios.

For the calculation of the post (sweat conductivity) test probability, we considered that in our setting, in the last 5 years, the average prevalence (pretest probability) of CF among newborns and young infants was at around 0.2%.

Table 1

<table>
<thead>
<tr>
<th>Age-group</th>
<th>Positive</th>
<th>Negative</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (%)</td>
<td>N (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Newborns</td>
<td>13 (19.1)</td>
<td>161 (14.5)</td>
<td>0.30</td>
</tr>
<tr>
<td>Infants</td>
<td>55 (80.9)</td>
<td>947 (85.5)</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td>0.86</td>
</tr>
<tr>
<td>Female</td>
<td>35 (51.5)</td>
<td>555 (50.1)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>33 (48.5)</td>
<td>546 (49.3)</td>
<td></td>
</tr>
<tr>
<td>Not recorded</td>
<td>0</td>
<td>7 (0.6)</td>
<td></td>
</tr>
<tr>
<td>Region of residence</td>
<td></td>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td>Urban area</td>
<td>8 (11.8)</td>
<td>277 (25.0)</td>
<td></td>
</tr>
<tr>
<td>Rural area</td>
<td>60 (88.2)</td>
<td>831 (75.0)</td>
<td></td>
</tr>
<tr>
<td>Gestational age</td>
<td></td>
<td></td>
<td>0.03</td>
</tr>
<tr>
<td>&lt;27</td>
<td>1 (1.5)</td>
<td>102 (9.2)</td>
<td></td>
</tr>
<tr>
<td>≥27</td>
<td>61 (89.7)</td>
<td>999 (90.2)</td>
<td></td>
</tr>
<tr>
<td>Birth weight (g)</td>
<td></td>
<td></td>
<td>0.21</td>
</tr>
<tr>
<td>&lt;2.500</td>
<td>5 (7.4)</td>
<td>142 (12.8)</td>
<td></td>
</tr>
<tr>
<td>≥2.500</td>
<td>60 (88.2)</td>
<td>954 (86.1)</td>
<td></td>
</tr>
<tr>
<td>Not recorded</td>
<td>3 (4.4)</td>
<td>12 (1.1)</td>
<td></td>
</tr>
<tr>
<td>Weight at the tests day (g)</td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>&lt;2.000</td>
<td>zero</td>
<td>4 (0.4)</td>
<td></td>
</tr>
<tr>
<td>≥2.000</td>
<td>36 (52.9)</td>
<td>1061 (95.8)</td>
<td></td>
</tr>
<tr>
<td>Not recorded</td>
<td>32 (47.1)</td>
<td>43 (3.8)</td>
<td></td>
</tr>
</tbody>
</table>

*χ² or Fisher’s exact test.

We also constructed the receiver operator characteristic (ROC) curve to estimate the area under the curve (AUC). Finally, we assessed the agreement between index and reference tests through Kappa coefficient.

Analyses were performed using SPSS software, V.23 (SPSS, USA).

Sample size

Sample size calculation was based on the classical equation, that is, N=4Zα²P(1−P)/W², in which Z returns 1.96; P, a proportion that equals the sensitivity or specificity estimates and W, the width of the 95% CI.

According to the published literature, the average sensitivity for SC is 98%; accepting a margin of error within 0.06 for the 95% CI, we would need a minimum of 83 patients with CF. Regarding specificity, the literature reports an average specificity of 99%; adopting a width of 0.02 for the 95% CI would require at least 380 individuals without CF. Thus, the final sample size would need at least 463 subjects.

RESULTS

Participants

During the study period, 1752,587 newborns were screened. A total of 2631 newborns and young infants had two positive IRT tests. According to predefined criteria, we excluded 1750,999 and considered ineligible 1295 participants; therefore, 1193 were recruited for the study. Among them, 17 subjects (17/1193, 1.4%) presented intermediate results for ST, SC or both tests (figure 1).

Table 1 compares the general characteristics of 1176 participants with positive and negative ST and SC.

There were no significant differences between the two groups, except that the CF group had a higher prevalence in rural areas and a lower proportion of preterm birth. (Gestational age<37 weeks).

The median age of the patients was 5 days old at the first IRT and 15 days of age for those in which the first dosage was higher than 70 mg/L, while the mean (SD) and median (IQR) age of the participants were 48 (19.2) and 43 days old (15-90), respectively. In the CF group, the mean and median values were 100.2 (21.8) mmol/L and 98.1 (66–242) mmol/L for sweat chloride and 103.9 (14.3) mmol/L and 106.0 (25–131) mmol/L for SC. In the non-CF group, the mean and median values for sweat chloride and conductivity were 10.9 (4.4) mmol/L and 10.2 (2.4–29.7) mmol/L and 29.2 (6.0) and 29 (11-103), respectively (data are not shown in table 1).

Test results and estimates: diagnostic accuracy of SC

Table 2 shows the results of ST and SC among the 1176 patients with no intermediate values to chloride analysis and conductometry.

Sweat conductivity yielded a sensitivity of 98.5% (95% CI 95.7 to 100), a specificity of 99.9% (95% CI 99.7 to 100), a PPV of 98.3% (95% CI 95.7 to 100), a NPV of 99.9% (95% CI 99.7 to 100) and an overall accuracy of 99.8% (95% CI 99.6 to 100). Furthermore, the positive and negative LR were 1091.7 (95% CI 153.8 to 7744.9) and 0.01 (95% CI 0.0 to 0.1), respectively.

Post-test probability for positive and negative LR were 68.6 and 0.002, respectively, meaning that the patient’s probability of CF increases from 0.2% to almost 70% (ie, around 350 times more) and drops from 0.2% to virtually zero, after a positive and negative SC result, respectively.
respectively. For the best-case scenario scenario, where true positives=67, false-positives=8, false-negatives=11 and true negatives=1107, sensitivity, specificity, PPV and NPV and positive and negative LR were 89.3% (95% CI 80.0 to 95.2), 99.0% (95% CI 98.2 to 99.3), 85.9% (95% CI 77.0 to 91.6), 99.2% (95% CI 98.6 to 99.6), 90.8 (95% CI 50.1 to 164.3) and 0.1 (95% CI 0.0 to 0.2), respectively. For the best-case scenario, where true positives=74, false-positives=1, false-negatives=1 and true negatives=1117, sensitivity, specificity, PPV and NPV and positive and negative LR were 98.6% (95% CI 99.5 to 99.9), 99.9% (95% CI 99.5 to 100), 98.6% (95% CI 91.2 to 99.8), 99.9% (95% CI 99.3 to 99.9), 1.103,0 (95% CI 155.4 to 7,825), 0.01 (95% CI 0.0 to 0.1), respectively.

DISCUSSION

Chloride titration and SC are two sides of the same coin. An increase in chloride anion reflects a proportionate increase in the total electrolyte concentration, which is the basis of analysis by electrical conductivity.3

To date, this is the first study that demonstrated the excellent accuracy of SC as an alternative diagnostic test for CF among 1193 newborns and very young infants—the target age group of the NBS—compared with the classical ST.

All accuracy estimates obtained in the primary and secondary analyses evidenced the discriminatory diagnostic power of SC, specially the very high positive LR and very low negative LR. Positive and negative LR have unequivocal clinical relevance and are independent of disease prevalence offering a more practical way of interpreting SC accuracy. The higher the positive LR, the better the test result for ruling in the disease; conversely, the lower the negative LR, the better the test result is for ruling out the disease.3,4 Irrespective of the way that primary and secondary analyses were done, all accuracy indices for SC remained relatively stable. Moreover, post-test probability values mean that the likelihood of CF increases 350 times after a positive SC, while negative results virtually rule out CF.

The ROC curve is a useful summary of the overall accuracy of any test. The very large AUC (98.8%) and the near-perfect agreement between index and reference tests (Kappa=0.99) provide additional evidence for the accuracy of SC. The rate of the intermediate ST (1.0%) and SC (0.7%) results also demonstrated the comparability between them.

As the mean and median age of the included participants were lower than 50 days when they underwent ST and SC, the first clinical appointment was accomplished before completing the second month of life irrespective of their gestational age, birth weight and weight at the testing day. These data are comparable to those described by Leung et al reporting a mean age of 2.58 (SD 0.69) months at initial examination among 231 American CF children diagnosed exclusively through classical ST.19

We searched studies on SC, disregarding investigations carried out with its related Nanoduct, and found only one prospective study including newborns.1 That study assessed 534 subjects and had similar estimates for SC, that is, 100% sensitivity, 96.2% specificity, 83.3% PPV, 100% NPV, 31.7% positive LR, and zero negative LR.1 Other studies in different age groups found good to high estimates for sensitivity (83.3%–100%),

Table 3 Sweat chloride and SC results among the 17 (out of 1193) participants with intermediate results

<table>
<thead>
<tr>
<th>Sweat chloride</th>
<th>Positive</th>
<th>Intermediate</th>
<th>Negative</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
<td>N (%)</td>
</tr>
<tr>
<td>Sweat conductivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>zero</td>
<td>3 (25)</td>
<td>zero</td>
<td>3 (17.6)</td>
</tr>
<tr>
<td>Intermediate</td>
<td>4 (100)</td>
<td>3 (25)</td>
<td>1 (100)</td>
<td>8 (47.1)</td>
</tr>
<tr>
<td>Negative</td>
<td>zero</td>
<td>6 (50)</td>
<td>zero</td>
<td>6 (35.3)</td>
</tr>
<tr>
<td>Total</td>
<td>4 (100)</td>
<td>12 (100)</td>
<td>1 (100)</td>
<td>17 (100)</td>
</tr>
</tbody>
</table>

SC, sweat conductivity.
specifcity (97.3%–100%), PPV (90.9%–100%) and NPV (94.7%–100%).1 3 5 7 10 11 To note, the present study had the narrowest width of the 95%CI for all estimates, demonstrating the highest statistical precision.

As a result of multiple determinants, a glaring gap is evident in the median age of survival of patients with CF living in affluent and non-affluent countries. For instance, the life expectancy of individuals with CF in Canada, the USA and European Union countries are in the mid-40s or above, while only half of that in Brazil and even lower in some countries of the former Soviet Union.20 It is widely recognised that individuals with CF in developed nations are diagnosed earlier than their counterparts in developing ones and are likely to receive sooner standard CF care because of organisational factors, appropriate healthcare infrastructure and human resources.20 Providing the cheaper and simple SC test in rural or remote settings could contribute to speeding up the diagnosis, bridging at least some of these critical gaps.

Then, in settings with a relative shortage of chloridometers, implementing an SC network in close cooperation with CF referral centres labs could reduce their burden with negative sweat chloride testing after NBS and contribute to the earlier diagnosis of CF. For instance, the CF Referral Centre labs must retest sweat samples through a chloridometer from either positive or negative SC results chosen randomly through a rigorous and periodical schedule. As shown, around 80% of our studied newborns and young infants resided in the countryside, that is, between 30 and 470 miles from our CF Centre, the sole facility for the whole State, which surface area is like that of France. A similar scenario seems to apply in similar settings worldwide.

We should pinpoint the limitations of our study. First, even though chloride titration is still considered a gold standard, it would be better if we had used a more robust gold standard, that is, the combination of sweat chloride plus DNA testing. That strategy would undoubtedly avoid the rate of intermediate results and would guide CFTR modulator therapies. Second, we recruited 72 out of 83 patients with CF, that is, 11 fewer than would be needed. However, the present work has enough power to assess the diagnostic accuracy of SC, given that the final sample size (n=1193) is much larger than that previously calculated (n=463). Finally, in terms of size and age group, this unparalleled sample may also have mitigated the ineligibility of the 1295 potential patients who could not simultaneously undergo both tests, one of the inclusion criteria.

In conclusion, SC demonstrated excellent reliability after two-tier IRT NBS. Therefore, it might be considered an alternative diagnostic tool for CF, especially in poor health infrastructure settings, because it contributes to early treatment and minimises family and parental distress due to a positive NBS.

Acknowledgements The authors acknowledge Roberto Vagner Puglia Ladeira, the CF referral centre laboratory head, for supporting the study as well as Dr José Diniz Ribeiro, Dr Fabiola Vilac Adde, Dr José Nélio Januário, Dr Cláudia Ribeiro de Andrade and Dr Clésio Gontijo do Amaral for their helpful inputs on the preliminary version of this manuscript.

Contributors RMB verified the underlying data, contributed to the conception, writing, revision and editing of the manuscript. CGA contributed to the manuscript’s conception, writing, editing and critical revision. JV Aj and FHP contributed to data collection, analysis, validation and interpretation, verified the underlying data and contributed to writing the manuscript. LZ contributed to the manuscript’s conception, writing, revision and editing. PC conceived the study and its design, had primary responsibility for protocol development, contributed to the investigation, methodology, project administration, validation and verified the underlying data; conception, writing, editing and revising of the manuscript. PC is the guarantor. All authors confirm that they had full access to all the data in the study and accept responsibility for submitting it for publication.

Funding This work was supported by the Brazilian Council for Research and Technological Development, grant number 486201/2013-9, to PC.

Competing interests None declared.

Patient consent for publication Consent obtained from parent(s)/guardian(s).

Ethics approval The Institutional Review Board of the Federal University of Minas Gerais approved the research protocol (reference CAAE 21958014.1.000.5149).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. All data relevant to the study are included in the article. Additional individual patient data are available in the deidentified format on request to the corresponding author.

ORCID iDs Linjie Zhang http://orcid.org/0000-0001-5150-5840
Paulo Camargos http://orcid.org/0000-0003-4731-291X

REFERENCES