Journal of Cystic Fibrosis 16 (2017) 573 – 578

Journal of Cystic Fibrosis

www.elsevier.com/locate/jcf

Original Article

A multiple reader scoring system for Nasal Potential Difference parameters

George M. Solomon [a,*], Bo Liu [a], Isabelle Sermet-Gaudelus [b,c], Isabelle Fajac [c,d], Michael Wilschanski [e], Francois Vermeulen [f], Steven M. Rowe [a]

[a] Department of Medicine, The Gregory Fleming James Cystic Fibrosis Research Center, University of Alabama at Birmingham, Birmingham, AL, USA
[b] Hôpital Necker-Enfants Malades, France
[c] Université Paris René Descartes, Sorbonne Paris Cité, Paris, France
[d] AP-HP, Hôpital Cochin, Service de Physiologie et Explorations Fonctionnelles, Paris, France
[e] Hadassah Medical Center, Hebrew University, Jerusalem, Israel
[f] Catholic University of Leuven, Leuven, Belgium

Received 14 October 2016; revised 10 March 2017; accepted 10 April 2017
Available online 29 April 2017

Abstract

Background: Nasal Potential Difference (NPD) is a biomarker of CFTR activity used to diagnose CF and monitor experimental therapies. Limited studies have been performed to assess agreement between expert readers of NPD interpretation using a scoring algorithm.
Methods: We developed a standardized scoring algorithm for "interpretability" and "confidence" for PD (potential difference) measures, and sought to determine the degree of agreement on NPD parameters between trained readers.
Results: There was excellent agreement for interpretability between NPD readers for CF and fair agreement for normal tracings but slight agreement of interpretability in indeterminate tracings. Amongst interpretable tracings, excellent correlation of mean scores for Ringer's Baseline PD, $\Delta_{amiloride}$, and $\Delta_{Cl\text{-}free + Isoproterenol}$ was observed. There was slight agreement regarding confidence of the interpretable PD tracings, resulting in divergence of the Ringers and $\Delta_{amiloride}$, and $\Delta Cl_{\text{-free + Isoproterenol}}$ PDs between "high" and "low" confidence CF tracings.
Conclusion: A multi-reader process with adjudication is important for scoring NPDs for diagnosis and in monitoring of CF clinical trials.

Keywords: Nasal Potential Difference; CFTR; Clinical trial outcomes

1. Introduction

Nasal Potential Difference (NPD) measurements have been important for diagnostic evaluations of cystic fibrosis (CF) since the technique was developed almost 40 years ago [1,2]. This method remains important due to the ability to quantify ion channel function in the respiratory epithelium [3–5]. It also provides the only in vivo assay to detect the function of both the epithelial sodium channel (ENaC) [6,7] and cystic fibrosis transmembrane conductance regulator (CFTR) [8]. NPD is a key outcome measure for experimental therapeutics addressing the CF ion transport abnormality, including modulators of CFTR function [9] Ivacaftor [10] and Lumacaftor [11] and other CFTR modulators [12,13] in development; CFTR-gene directed treatments [14–18]; and inhibitors of ENaC [19–22].

While improved NPD methods now allow for the use of NPD in multi-center trials with electronic data capture and blinded interpretation that have improved its reliability [23,24] and has been agreed upon by the US TDN and EU CTN, there has not yet been a proposed standardized interpretation protocol to ensure uniform interpretation within clinical trials. Furthermore, a standardized scoring system allows for a multiple

Abbreviations: NPD, Nasal Potential Difference; PD, potential difference.

* Corresponding author at: University of Alabama at Birmingham, 1900 University Blvd, THT 422, Birmingham, AL 35294, United States.
E-mail address: msolomon@uab.edu (G.M. Solomon).

http://dx.doi.org/10.1016/j.jcf.2017.04.011
1569-1993

reader approach to adjudicate the inclusion of questionable tracings, which could improve the performance of NPD, particularly for smaller early phase and proof-of-concept clinical trials.

In the present study, we investigated the agreement of expert NPD readers using a standardized scoring algorithm originally developed at the Center for CFTR detection at UAB. Six expert NPD readers were trained on this algorithm, which provides a method for quantifying NPD values and assigns a 2-tier approach to rating the quality of NPD tracings: a lower stringency rating determines "interpretability" (i.e. whether a tracing should be included in a dataset), and a higher stringency tier which determines if tracings are "high" or "low confidence" for measurements of sodium and chloride. Here we evaluated a scoring system and measured inter-reader agreement that included expert readers from two continents.

2. Methods

2.1. Development of the scoring system

The UAB Center for CFTR Detection (CCD) has previously developed a scoring algorithm for use in multi-center clinical trials of CFTR modulators employing NPD as an outcome for CFTR function. This scoring system was reviewed by the 2 co-directors of the CCD and 4 additional expert readers in the Therapeutics Development Network (TDN) and the European Clinical Trials Network (CTN).

This scoring system was developed to assess both a stringent criteria of "interpretability" based on completeness of the tracing protocol, biological plausibility of the response data, and appropriate control responses, which enhances and standardizes the current quality standards for tracings using the current standard operating procedures employed by the Therapeutics Development Network. In addition, we analyzed a less stringent "confidence" score that reflects more subtle abnormalities in the tracing that do not alter its ability to be analyzed. As measures of sodium and chloride transport are derived from different time parts of the tracing, each component was scored separately. The criteria for rating "interpretability" and "confidence" are detailed in Table 1.

Table 1
Criteria for interpretability and confidence of NPD analysis.

Intepretability — "Intepretable" = absence of all of the following
1. Missing portion of tracing
2. Incomplete tracing
3. >1 mV shift in the last 30 s of each perfusion tracing
4. Displaced catheter without recovery to pre-displacement value
5. Biologically implausible values (especially positive charge)
6. Poorly responsive tracing (<3 mV variability of PD with amiloride and/or ATP)
Confidence — "High" = absence of all of the following
1. Excessive noise or artifact that interferes with PD (last 30 s of each solution tracing that is <1 mV)
2. Large break or shift (≥30 s duration) without recovery to pre-shift values
3. Catheter displacement that is resolved to the pre-displacement value

The resultant scoring system was approved by the review committee.

2.2. Analysis of tracings

After approval each committee member was acquainted with the scoring algorithm in an in-person training session as well as provided with review materials that detail the scoring system. Examples of "interpretable" tracings are detailed in Fig. 1A (high confidence) and Fig. 1B (low confidence). Examples of uninterpretable tracings are shown in Fig. 1C–D. The original lab chart files corresponding to these examples are shown in Supplementary Fig. 1.

2.3. Analysis of the scoring system

After appropriate training, each reader was assigned blinded tracings to determine the correlation of qualitative scores and quantitative values for key NPD parameters including: Ringer's Baseline potential difference (PD), $\Delta_{Amiloride}$ PD, and $\Delta_{Cl\text{-}free + Isoproterenol}$ PD. These key PD measures (Ringer's Baseline potential difference (PD), $\Delta_{Amiloride}$ PD, and $\Delta_{Cl\text{-}free + Isoproterenol}$ PD) were quantified by the mean of the last 10 s of the perfusion as previously described [25]. Each reader was assigned 40 single-nostril CF, 40 single-nostril non-CF, and 20 indeterminate (non-diagnostic) tracings for CFTR function. Tracings were chosen from recent clinical trial databases and our database of diagnostic NPDs sent for interpretive over-read to our center. Since the scoring system was developed for individual nostril tracings, the readers analyzed tracings in this manner. All NPDs were collected using the Therapeutics Development Network standard Operating Procedure (NPD SOP 528.0), and conformed to general quality standards in the document at that time. Each NPD was selected at random and blinded by study staff before review by expert readers. The diagnosis of "CF" was defined by the presence of sweat chloride ≥60 meq/L and/or 2 disease-causing mutations on CFTR genetic analysis [26]. "Indeterminate" was defined by the presence of a questioned CF clinical diagnosis with indeterminate sweat chloride values (40–60 meq/L) and <2 CFTR causing mutations on CFTR genetic analysis. We grouped level of agreement into 3 categories: "complete" (all 6 readers agree), "moderate" (4–5 of 6 readers agree), and "poor" (only 3 readers agree).

2.4. Statistical analysis

After unblinding at the CCD, kappa statistics (κ) for inter-reader comparisons were calculated for "interpretability" and "confidence" of the ENaC-mediated portion of the tracing and the CFTR-mediated portions of the tracing. Significant correlation was determined when $p < 0.05$ (ANOVA). Comparative statistics were calculated using SPSS 13.0 (IBM Corporation, Armonk, NY) and Graph Pad Prism 6.0 (GraphPad Software, Inc., La Jolla, CA). The level of agreement was assessed as previously described [27].