Early Detection of Rapid Cystic Fibrosis Disease Progression Tailored to Point of Care: A Proof-of-Principle Study*

Rhonda Szczesniak, Cole Brokamp, Weiji Su, Gary L. McPhail, John Pestian, John P. Clancy

Abstract— **Slowing cystic fibrosis (CF) lung disease progression is crucial to survival, but point-of-care technologies aimed at early detection—and possibly prevention—of rapid lung function decline are limited. This proof-of-principle study leverages a rich national patient registry and follow-up data on a local CF cohort to build an algorithm and prototype prognostic tool aimed at early detection of rapid lung function decline. The algorithm was developed using a novel longitudinal analysis of lung function (measured as forced expiratory volume in 1 s of % predicted, FEV1). Covariates included clinical and demographic characteristics selected from the registry based on information criterion. Preliminary assessment of algorithm performance suggested excellent predictive accuracy and earlier detection of rapid decline than standard of care being applied at a local center. Graphical displays were presented and evaluated for clinical utility. Predictions from the algorithms and chosen graphical displays were translated into a prototype web application using RShiny and underwent iterative development based on clinician feedback. This paper suggests that the algorithm and its translation could offer a means for earlier detection and treatment of rapid decline, providing clinicians with a viable point-of-care technology to intervene prior to irreversible lung damage.**

I. INTRODUCTION

Cystic fibrosis (CF) is a lethal autosomal disease marked by a progressive loss in lung function, and currently affects nearly 70,000 individuals worldwide [1]. The leading cause of death in CF is respiratory failure [2]; therefore, early detection and treatment of prolonged drops in lung function are essential to survival. These bouts of lung disease progression have been clinically termed "rapid decline" (Fig. 1). Numerous epidemiologic studies by these authors [3] and others [4] [5] [6] have employed various statistical approaches to estimate trajectories of lung function decline. These trajectories exhibit nonlinearity over the lifespan with substantial variation both between subjects and within an individual subject over time. Despite unique approaches, each study has demonstrated that the most severe bouts of rapid decline tend to occur during adolescence and early adulthood. Indeed, recent work suggests that, although there are distinct patterns or phenotypes of rapid decline, those individuals with the highest lung function initially are at risk for the most severe declines early in life [7]. In addition, adults with CF experience rapid decline [8], suggesting that this pervasive event requires clinical monitoring and treatment throughout its entire course.

All authors are with Cincinnati Children's Hospital Medical Center, 3333 Burnet Ave, Cincinnati, OH 45229 USA (correspondence: Rhonda Szczesniak; phone: 513-803-0563; fax: 303-555-5555; e-mail: rhonda.szczesniak@cchmc.org).

While numerous therapeutic advancements and quality improvement initiatives have extended life expectancy, point-of-care algorithms and technologies that harness well-developed epidemiologic findings regarding prediction—as opposed to explanation—of rapid decline within the individual CF patient are limited. Efforts to translate statistical innovations into CF point of care began with spiromteric reference equations [9] [10], which initially were separated by children and adults but have recently been extended through advanced methodology [11]. Home spirometry has been studied, in which patients are monitored for onset of acute respiratory events known as pulmonary exacerbations [12, 13]. A recent diagnostic tool development highlights the feasibility of CF point-of-care technologies for diagnosis [14], but does not address routine care and clinical surveillance that are necessary to treat rapid disease progression.

A state-of-the-art approach was proposed to translate a flexible algorithm to accurately detect rapid lung function decline within the individual patient into a prognostic tool for CF point of care. The approach undertaken in this proof-of-principle study exhibited promises for detection and exhibited clinical utility.

II. ALGORITHM DEVELOPMENT

A. Data Sources

Data from the US Cystic Fibrosis Foundation Patient Registry (CFFPR) were used to develop the longitudinal model and resulting algorithm. The timeframe included data from January 1, 2003, until December 31, 2015, in order to reflect the most modern era of CF care from the available data. This registry has been tracking outcomes on patients with CF for nearly 50 years; detailed descriptions of its contents have been provided [15]. For model development in this study, we utilized forced expiratory volume in 1 s of % predicted (hereafter, FEV1) as a marker of lung function, and included data on patients aged ≥ 6 years in order to obtain reliable pulmonary function from the FEV1 measure. Other
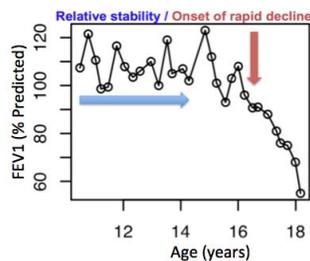

Figure 1. Lung function over age for a male CF patient, initially stable then declining nonlinearly over age.

clinical and demographic characteristics included static variables: sex (male or female), genotype (copies of F508del alleles coded as heterozygous, homozygous, or none), birth cohort (a categorical variable defined as birth year < 1981, 1981-1988, 1989-1994, 1995-1998, 1999-2005, ≥ 2006), chronic infection with *Pseudomonas aeruginosa* (Pa, defined as ≥ 4 positive cultures over time), persistent methicillin-resistant *Staphylococcus aureus* (MRSA, defined as ≥ 4 positive cultures over time); time-varying variables: age (in years), low socioeconomic status (defined as having received federal/state insurance). These variables were chosen based on existing CF epidemiologic literature. The local institutional review board approved the study.

B. Specifications and Fitting

A longitudinal model was developed to fit age-related FEV1 progression and account for its nonlinearity by expanding an established method that has been successfully used to monitor markers of renal disease progression [16]. The expanded model used in this paper was presented at the 40th European Cystic Fibrosis Society Meeting (abstract to be published in Fall 2017 in the Journal of Cystic Fibrosis). Let Y_{ij} be a random variable representing the longitudinal process of FEV1 taken on the ith patient at the jth time point ($i = 1, \ldots, N; j = 1, \ldots, n_i$); here, let time be represented by age (in years). The longitudinal model can be expressed as:

$$Y_{ij} = f(t_{ij}) + X_i'\beta + U_i + W_i(t_{ij}) + \varepsilon_{ij}. \quad (1)$$

In Equation (1), the function $f(\cdot)$ is used to depict nonlinear FEV1 progression over time t_{ij}, which is expressed as age (in years); $X_i'\beta$ represent the patient-specific vector of covariates defined previously and their corresponding coefficients; $U_i \sim N(0, \sigma_U^2)$ are random intercepts allowing FEV1 trajectories to be shifted across individual that follow a normal distribution with mean 0 and variance σ_U^2; $W_i(t_{ij})$ are independent realizations of a zero-mean, continuous-time stochastic process known as integrated Brownian motion, representing change in a patient's FEV1 over time that cannot be accounted for with the other terms in the model; $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$ is independent, identically distributed measurement error. The model was fitted using the 'lmenssp' package available in R.

C. Performance

Forecast validation was performed by randomly selecting roughly 20% of patients and masking the last two years of their data. Metrics included mean absolute error (MAE), root mean square error (RMSE), mean absolute % error (MAPE). These were calculated using the actual data during the masked period and the predicted data based on the model. Over the two-year window, overall and h-step ahead forecasts were computed; steps included 0.5 years, 1 year and 2 years. Table I includes performance metrics for this model.

TABLE I.   MODEL PERFORMANCE

Metric	Period[a]			
	Overall	*0.5 years*	*1 year*	*2 years*
MAE	4.71	3.27	3.94	4.53
RMSE	6.76	5.21	5.96	6.68

Metric	Period[a]			
	Overall	*0.5 years*	*1 year*	*2 years*
MAPE	8.24%	5.82%	7.20%	8.81%

Abbreviations: mean absolute error (MAE); mean absolute percentage error (MAPE); root mean-square error (RMSE). [a] Selected based on clinical input.

III. EARLY TRANSLATION TO POINT OF CARE

A. Clinical Identification of Rapid Decline

In this application of the model, there were $N = 27{,}296$ patients "at risk" of rapid decline with a total of $\sum_{i=1}^{N} n_i = 619{,}960$ observed FEV1 data points and $j = 1, \ldots, n_i \le 89$ visits per patient. Age range during follow up was 6 to 83 years. The covariate history on patient i up to time t can be denoted as $H_i(t) = \{x_i, (t_{ij}, y_{ij}): t_{ij} > t\}$.

The first derivative of Equation (1) may be used as an estimate of rate of change in FEV1. A threshold of -1.5% predicted/year was selected based on clinical judgment and graphical inspection. In order to identify periods in which a given patient is at risk of rapid decline based on this threshold, the following probability needs to be estimated.

$$P\left(\frac{d}{dt}Y_i(t) < -1.5 \Big| H_i(t)\right) \quad (2)$$

This probability in Equation (2) corresponds to the risk of rapid decline for patient i based on his or her information history, which includes clinical and demographic data as well as past FEV1. By conditioning probabilities on this information, more accurate risk predictions are expected.

B. Preliminary Comparison to Actual Standard of Care

Prior to the algorithm described in this paper, a local center study was conducted to develop and implement a systematic algorithm specific to rapid decline in pediatric patients. Those patients whose peak FEV1 in the prior 3 months was not within 10% predicted of their highest FEV1 in the prior 12 months were classified as high-risk "Red Zone" patients. Modifiable risk factors for this degree of FEV1 decline included untreated/newly identified infectious organisms, gaps in or failure to prescribe pulmonary therapies, gastroesophageal reflux disease, unrecognized allergic bronchopulmonary aspergillosis and infrequent clinic follow up.

The local Red Zone algorithm to identify rapid decline was compared to the algorithm described in this paper with respect to age at which rapid decline was first identified using a retrospective analysis of 124 patients (age range: 6-22.3 years) who had received care based on the Red Zone algorithm at Cincinnati Children's Hospital Medical Center (2012-2015). Age at which the algorithm first estimated probability of rapid decline to be ≥ 0.80 was considered high risk. For those 120 patients who were classified as high risk based on both approaches, the proposed algorithm identified rapid decline an average of 0.65 years (95% CI: 0.41-0.89), or roughly 8 months, earlier than the local Red Zone algorithm (paired t-test, P < 0.0001).

C. Integrating Stakeholder Feedback

The algorithm and graphical displays were refined based on individual and group consultations with clinician through an iterative process (Fig. 2).
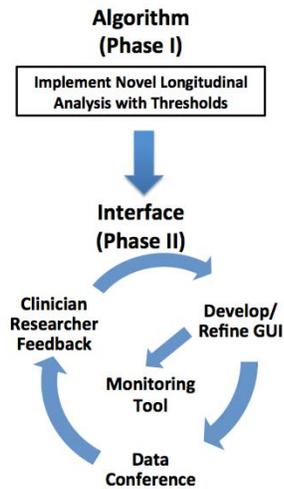

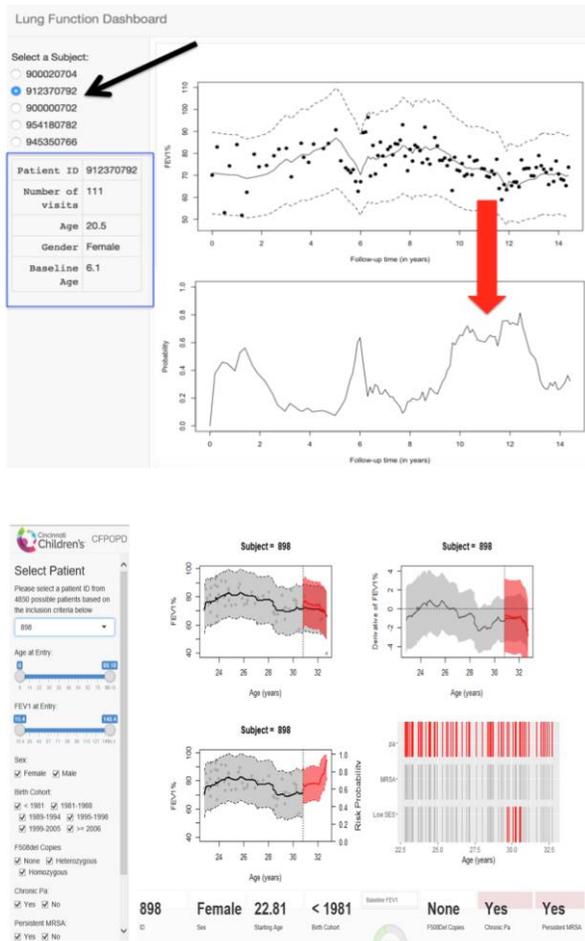Figure 2. Process for point-of-care prototype to detect rapid decline in CF patients.

Figure 3. Rapid decline dashboard screenshots from initial (upper) and revised (lower) prototypes based on clinician feedback. See text for additional explanation.

Initial presentation was based on a case series of patients using a graphic platform with FEV1 trends and predicted probabilities (Fig. 3, upper panel). In each display, the black dots are the observed FEV1 data from the patient. The application shown in the upper panel allows the user to select particular patient displays by clicking on the de-identified patient ID (see black diagonal arrow on the left). Information is displayed about the patient (blue box). This patient is female, has had 111 pulmonary function tests (denoted by the number of visits), and was 6.1 years old for her first test (zero on the x-axis). Her FEV1 time course is shown, along with 95% confidence bands. Her probability of rapid decline is also displayed with 95% confidence bands in the lower graph in this panel and varies according to her individual FEV1 time course shown above it. The probability is highest (~0.80 or 80%) during 10-12 years of follow-up (at age 16.1-18.1 years); this risk is reflected in her actual FEV1 course during that time (see vertical red arrow). Risk here was based on currently accrued data.

Comments on this prototype indicated the importance of forecasting, which has been added, and is still being used to further refine the prototype (Fig. 3, lower panel). In this display, the patient/subject selection has been converted to a drop-down menu and there is also a feature in which the user can type the ID. Age and FEV1 at onset also have toggle switches, which could be used to isolate predictions to normative data relative to the patient of interest. There are also check boxes for categorical measures. Snapshot information is available from the lower portion of the dashboard. This display shows data and predictions for a female CF patient whose age at entry was about 23 years; she was born in an older cohort, and has no F508del alleles. For a brief period, her SES was low. Moving clockwise from the upper left graph of this panel, the gray portion depicts observed and fitted FEV1 data that were used to build the model, while the red shaded portion shows predictions for the held-out data used to assess forecast performance. The dark gray dots are the observed FEV1 data; the black line is the model-based fitted curve; the red line is the forecasted curve assuming that the last two years of data are unknown. Because the data are split into development and forecast data, the confidence bands appear disjoint. The graph to the right shows the rate of change in FEV1 for this patient and has the corresponding confidence bands in gray and red. In finalized displays, the gray dots representing observed FEV1 will not be available in the two-year forecasting window but are shown here for validation purposes. Her rate of change in FEV1 shows clinically relevant declines after 28 years of age. Algorithm results suggest that this patient is at increased risk of rapid decline during the forecast period, and that she frequently cultured positive for Pa infection and had low SES for a period of time during follow up.

D. Implementation

The algorithm has been implemented as an online prototype application based on preliminary stakeholder feedback. Currently, the online posting displays individual patient trajectories, along with model forecasting results (source: http://cfpopd.amazon-shiny.duckdns.org/#section-cf-predicted-lung-function). The application may take a

moment to load. Toggles are provided for filtering continuous variables and selections can be made through check boxes for continuous variables. This version will be expanded to include predictive probability displays as in the lower panel of Fig. 3 and normative data by having a look-up table with algorithm-based predictions underlying the web application. These results will be updated daily using electronic health record integration at the local center level.

IV. CONCLUSION

This proof-of-principle study shows that clinical surveillance data can be harnessed to provide effective monitoring tools for clinicians treating patients with CF. Preliminary assessment of the algorithm indicates that it could improve detection of rapid decline, compared to the currently employed approach.

Future prospective studies are needed to understand how delivery of interventions to treat or prevent rapid decline could be more efficacious if given according to the algorithm proposed here, compared to current care. Later phases of these studies may also require additional standardization across centers, as studies of the CFFPR suggest that prescribing patterns vary [17]. In addition, further study of the selected threshold (-1.5% predicted/year) may be needed, although this initial assessment with clinician researcher feedback indicates its relevance to care. Lowering this threshold, which would correspond to earlier identification and potentially an increased number of patients qualifying for additional care, could pose challenges to clinical systems. Increasing this threshold would yield later identifications and possibly less sensitive detection, compared to the approach currently used.

Mixed methods studies are also needed to optimize the prognostic value of this monitoring approach. A focus group study involving clinical care teams at the center level could be used to formally evaluate the utility of the web application. This approach could be extended to include patients and their caregivers. Monitoring tools could be tailored to individual use, empowering the patient in the point-of-care setting.

REFERENCES

References

1 Farrell, P.M.: 'The prevalence of cystic fibrosis in the European Union', J Cyst Fibros, 2008, 7, (5), pp. 450-453
2 Foundation, C.F.: 'Cystic Fibrosis Foundation Patient Registry', in Editor (Ed.)^(Eds.): 'Book Cystic Fibrosis Foundation Patient Registry' (Cystic Fibrosis Foundation, 2016, edn.), pp.

3 Szczesniak, R.D., McPhail, G.L., Duan, L.L., Macaluso, M., Amin, R.S., and Clancy, J.P.: 'A semiparametric approach to estimate rapid lung function decline in cystic fibrosis', Ann Epidemiol, 2013, 23, (12), pp. 771-777
4 Konstan, M.W., Morgan, W.J., Butler, S.M., Pasta, D.J., Craib, M.L., Silva, S.J., Stokes, D.C., Wohl, M.E., Wagener, J.S., Regelmann, W.E., Johnson, C.A., Scientific Advisory, G., the, I., and Coordinators of the Epidemiologic Study of Cystic, F.: 'Risk factors for rate of decline in forced expiratory volume in one second in children and adolescents with cystic fibrosis', J Pediatr, 2007, 151, (2), pp. 134-139, 139 e131
5 Liou, T.G., Elkin, E.P., Pasta, D.J., Jacobs, J.R., Konstan, M.W., Morgan, W.J., and Wagener, J.S.: 'Year-to-year changes in lung function in individuals with cystic fibrosis', J Cyst Fibros, 2010, 9, (4), pp. 250-256
6 Vandenbranden, S.L., McMullen, A., Schechter, M.S., Pasta, D.J., Michaelis, R.L., Konstan, M.W., Wagener, J.S., Morgan, W.J., McColley, S.A., Investigators, and Coordinators of the Epidemiologic Study of Cystic, F.: 'Lung function decline from adolescence to young adulthood in cystic fibrosis', Pediatr Pulmonol, 2012, 47, (2), pp. 135-143
7 Szczesniak, R.D., Li, D., Su, W., Brokamp, C., Pestian, J., Seid, M., and Clancy, J.P.: 'Phenotypes of Rapid Cystic Fibrosis Lung Disease Progression during Adolescence and Young Adulthood', Am J Respir Crit Care Med, 2017, 196, (4), pp. 471-478
8 Konstan, M.W., Wagener, J.S., Vandevanter, D.R., Pasta, D.J., Yegin, A., Rasouliyan, L., and Morgan, W.J.: 'Risk factors for rate of decline in FEV1 in adults with cystic fibrosis', J Cyst Fibros, 2012, 11, (5), pp. 405-411
9 Wang, X., Dockery, D.W., Wypij, D., Fay, M.E., and Ferris, B.G., Jr.: 'Pulmonary function between 6 and 18 years of age', Pediatr Pulmonol, 1993, 15, (2), pp. 75-88
10 Hankinson, J.L., Odencrantz, J.R., and Fedan, K.B.: 'Spirometric reference values from a sample of the general U.S. population', Am J Respir Crit Care Med, 1999, 159, (1), pp. 179-187
11 Stanojevic, S., Stocks, J., Bountziouka, V., Aurora, P., Kirkby, J., Bourke, S., Carr, S.B., Gunn, E., Prasad, A., and Rosenfeld, M.: 'The impact of switching to the new global lung function initiative equations on spirometry results in the UK CF Registry', Journal of Cystic Fibrosis, 2014, 13, (3), pp. 319-327
12 Lechtzin, N., West, N., Allgood, S., Wilhelm, E., Khan, U., Mayer-Hamblett, N., Aitken, M.L., Ramsey, B.W., Boyle, M.P., Mogayzel, P.J., Jr., and Goss, C.H.: 'Rationale and design of a randomized trial of home electronic symptom and lung function monitoring to detect cystic fibrosis pulmonary exacerbations: the early intervention in cystic fibrosis exacerbation (eICE) trial', Contemp Clin Trials, 2013, 36, (2), pp. 460-469
13 Lechtzin, N., Mayer-Hamblett, N., West, N.E., Allgood, S., Wilhelm, E., Khan, U., Aitken, M.L., Ramsey, B.W., Boyle, M.P., Mogayzel, P.J., Jr., Gibson, R.L., Orenstein, D., Milla, C., Clancy, J.P., Antony, V., Goss, C.H., and e, I.C.E.S.T.: 'Home Monitoring in CF to Identify and Treat Acute Pulmonary Exacerbations: eICE Study Results', Am J Respir Crit Care Med, 2017
14 Zhang, C., Kim, J.P., Creer, M., Yang, J., and Liu, Z.: 'A smartphone-based chloridometer for point-of-care diagnostics of cystic fibrosis', Biosens Bioelectron, 2017, 97, pp. 164-168
15 Knapp, E.A., Fink, A.K., Goss, C.H., Sewall, A., Ostrenga, J., Dowd, C., Elbert, A., Petren, K.M., and Marshall, B.C.: 'The Cystic Fibrosis Foundation Patient Registry. Design and Methods of a National Observational Disease Registry', Ann Am Thorac Soc, 2016, 13, (7), pp. 1173-1179
16 Diggle, P.J., Sousa, I., and Asar, O.: 'Real-time monitoring of progression towards renal failure in primary care patients', Biostatistics, 2015, 16, (3), pp. 522-536
17 VanDyke, R.D., McPhail, G.L., Huang, B., Fenchel, M.C., Amin, R.S., Carle, A.C., Chini, B.A., and Seid, M.: 'Inhaled tobramycin effectively reduces FEV1 decline in cystic fibrosis. An instrumental variables analysis', Ann Am Thorac Soc, 2013, 10, (3), pp. 205-212