Development of a risk score to increase detection of severe alpha-1 antitrypsin deficiency

E. Leonard Riley ⬤[1], J. Cory Brunson[2], Soroush Eydgahi[3], Mark L. Brantly ⬤[3] and Jorge E. Lascano[3]

[1]Department of Internal Medicine, Division of Pulmonary, Critical Care, and Sleep Medicine, Kansas City Veterans Affairs Medical Center, Kansas City, MO, USA. [2]Laboratory for Systems Medicine, Division of Pulmonary, Critical Care, and Sleep Medicine, University of Florida College of Medicine, Gainesville, FL, USA. [3]Department of Internal Medicine, Division of Pulmonary, Critical Care, and Sleep Medicine, University of Florida College of Medicine, Gainesville, FL, USA.

Corresponding author: E. Leonard Riley (elmer.riley2@va.gov)

Shareable abstract (@ERSpublications)
Alpha-1 antitrypsin deficiency is an underrecognised genetic disease. Using the National Detection Program database, a medical-history-only prediction model was constructed to risk stratify subjects, and increase testing and detection. https://bit.ly/3r4Rgih

Abstract

Background Alpha-1 antitrypsin deficiency (AATD) is an under-recognised genetic cause of chronic obstructive lung disease, and many fewer cases than estimated have been identified. Can a reported respiratory and hepatic disease history from a large AATD testing database be used to stratify a person's risk of severe AATD?

Methods We analysed data extracted from the AATD National Detection Program. Demographics and medical history were evaluated to predict AATD PI*ZZ genotype. Logistic regression and integer programming models identified predictors and obtained risk scores. These were internally validated on a subset of the data.

Results Out of 301 343 subjects, 1529 (0.5%) had PI*ZZ genotype. Predictors of severe AATD were asthma, bronchitis, emphysema, allergies, bronchiectasis, family history of AATD, cirrhosis, hepatitis and history of abnormal liver function tests. The derived model establishes a subject's risk of severe AATD, and scores ⩾0 had an estimated risk of 0.41%, sensitivity 84.62% and specificity 24.32%. A model simulating guideline recommendations had an estimated risk of 0.51% with a sensitivity of 37.98% and specificity 46.60%. By recommending screening for scores ⩾0, we estimate that more subjects would be screened (75.7% *versus* 53.4%) and detected (84.6% *versus* 58.2%) compared to a guideline-simulated model.

Conclusion This medical history risk model is a useful predictive tool to detect subjects at greater risk of having severe AATD and improves sensitivity of detection. Scores <0 are at lower risk and may need not be screened; testing is recommended for scores ⩾0 and consistent with current guidelines.

Introduction

Alpha-1 antitrypsin deficiency (AATD) is a common yet underdiagnosed genetic disease due to an inherited mutation in the SERPINA1 gene. This mutation produces the misfolded alpha-1 antitrypsin (AAT) protease inhibitor (PI) and predisposes to COPD and liver cirrhosis [1, 2]. This disease is inherited in an autosomal codominant pattern, and the normal AAT allele is termed "M" while the most common abnormal alleles are "Z" and then "S". The allelic homozygous PI*ZZ combination accounts for 95% of persons with severe AATD, and it makes up 0.6–2% of all COPD cases [2–5]. Epidemiological data estimate that there are ~100 000 individuals with severe AATD in the USA, although only ~10 000 have been identified [6]. People ultimately diagnosed with AATD see multiple physicians and experience long delays from symptom onset to diagnosis, which is associated with worsened clinical symptoms and overall survival, underscoring the importance of identifying persons with AATD earlier [7, 8].

While the World Health Organization (WHO), American Thoracic Society (ATS) and European Respiratory Society (ERS) recommend testing all individuals with COPD for AATD, there is low uptake of

this recommendation in clinical practice [9–12]. Failure to fully implement these guidelines is due to inadequate awareness of AATD, unclear test results, cost, testing taking too much time and the belief that testing will not impact clinical care [13, 14]. In clinical practice, some physicians subscribe to a traditional understanding of AATD characteristics, such as younger age and family history of AATD, to guide testing [13]. Strategies have been developed to enhance detection and increase testing prevalence, but there has been minimal improvement in the rate of detection of AATD [5, 15–17].

Clinical prediction models are tools that estimate the prognosis, specific outcomes or an individual's risk for a certain diagnosis; for example, the widely used Well's score predicts the risk of a subject having a pulmonary embolism. These models are most frequently utilised by general practitioners and viewed favourably when guiding testing and avoiding additional costs [18]. At this time, there are no models based only on medical history designed and implemented for the diagnostic prediction of AATD.

The University of Florida Alpha-1 Antitrypsin Genetic Laboratory, through the National Detection Program (NDP), has tested >300 000 people throughout the USA and its territories for AATD. Testing is offered free to patients and demographic and clinical information are collected. We hypothesise that evaluating a large database of tested patients may identify certain patient characteristics associated with AATD as well as lead to development of a diagnostic prediction model.

Study design and methods
National Detection Program
The NDP was a study approved by the University of Florida (IRB201801599) to collect data and test individuals for AATD. The study was based at the University of Florida from 2003 until 2014, when it was transferred to GeneAidyx. The programme is available to all healthcare providers and is available throughout the USA and its territories [19]. Healthcare providers are given information about AATD and a blood collection card. An attached questionnaire collects demographic and clinical information including gender, race, date of birth, smoking history (recorded as never, past, current, passive smoke as a child and passive smoke as an adult), hepatic disease history (including cirrhosis, jaundice, hepatitis, liver transplant or history of abnormal liver function test), respiratory disease history (asthma, chronic bronchitis, emphysema, COPD, allergies, bronchiectasis, AATD, family history AATD or family history COPD), and AAT augmentation therapy status. The laboratory testing and protocol performed by the NDP has been described previously [20, 21].

Outcome measures
The primary outcome of the study was a subject having severe AATD based on medical history. Alpha-1 antitrypsin deficiency is defined by the ATS and ERS as a serum AAT level <100 mg·dL^{-1} (<20 µmol·L^{-1}), although a serum level 57 mg·dL^{-1} (11 µmol·L^{-1}) is regarded as a "protective threshold" value and lower values represent severe deficiency [11]. The genotype PI*ZZ accounts for the majority of severely deficient serum AAT levels; therefore, for the purpose of this study we defined "severe AATD" as PI*ZZ genotype, although we also compared models of PI*ZZ or PI*SZ as a composite outcome [21, 22].

Modelling approach
We set out to construct a simple, interpretable clinical decision support tool to predict severe AATD based on variables collected by the NDP and to internally validate this tool on a subset of the data. We separately considered solutions using two sets of possible predictors: medical history only (excluding liver transplantation and self-reported AATD) and medical history together with smoking history. We compared a variety of predictive model families in order to establish performance benchmarks, but we considered only four from which an interpretable tool could be derived: decision rules, decision trees, logistic regression and integer programming. Ultimately, we decided to construct a risk score using either logistic regression or integer programming.

Risk scores
A risk score is a sparse solution to an integer optimisation problem. Given binary predictors $x_1, \ldots, x_p$, the goal is to obtain point values $\beta_0, \beta_1 \ldots, \beta_p$ (an intercept and linear coefficients) that yield the log-odds of the genotype class y. Where i indexes one case with predictor values $x_{1i}, \ldots, x_{pi}$ and $y_i = 1$ encodes an abnormal genotype, then the risk score produces the probability estimate

$$\Pr(y_i = 1) = \frac{1}{1 + \exp\left(-\beta_0 - \sum_{j=1}^{p} \beta_j x_{ji}\right)}$$

that case i has abnormal genotype. The problem is to find point values β_j that maximise some measure of predictive performance subject to two constraints: that they take integer values within pre-specified bounds, and that the number of nonzero values satisfies a pre-specified trade-off against the marginal performance gain. We considered multiple choices of constraints in order to obtain several candidate risk scores.

We used two methods to obtain candidate risk scores. One method uses penalisation to obtain a logistic regression model with a desired number of terms, scales its coefficients to desired bounds and rounds the scaled coefficients to integer point values. This procedure is described in detail by SULLIVAN *et al.* [23], and we refer to it as rounded regularised logistic regression. The other method is an integer programming solution that takes bounds on the number of terms and on their point values as inputs and returns a pool of integer-valued risk scores [24]. The procedure, called FasterRisk, improves upon previous state-of-the-art integer programming solutions.

Data pre-processing

Reported values of passive smoke exposure are infrequent in the data (roughly 8000) and were combined with "never" for the analysis. In cases where multiple samples were submitted for the same patient, only the first sample was used in the analysis, and the repeat tests were removed (n=7745). For comparison, we modelled current screening guidelines from WHO, ATS and ERS as a risk score with one composite item for COPD, chronic bronchitis or emphysema ("guidelines"). Medical and smoking history models were built using the 237 882 (79%) of cases with nonmissing smoking responses.

Model selection

We followed a three-phase machine learning workflow to optimise, evaluate and select our models. In phase 1, we used cross-validation on a one-sixth sample of the data to compare all predictive model families. In phase 2, we optimised model hyperparameters and compared some of the best-performing families on a four-fifths training set containing the one-sixth sample, again using cross-validation. Finally, in phase 3, we modified the best-performing interpretable families to yield simple decision support tools, fit these to the entire training set and evaluated them on the testing set comprising the remaining fifth of the data. Fitted models were compared using several criteria described in the next section. All samples were stratified by AAT genotype.

Statistical analysis

Key criteria for the risk scores were overall performance, trade-offs with respect to current screening guidelines, discriminability and usability. We assessed overall performance at detecting severe AATD visually using received operating characteristic ROC curves and quantitatively using the area under these curves as well as sensitivity, specificity, risk and screening burden for each score. We report 2×2 contingency tables for each candidate score based on two cut-offs, namely the scores that obtained specificity below and above, but otherwise closest to, that of guidelines. We assessed discriminability visually using histograms and quantitatively as the quotient of the standard deviation of the values of a score by its theoretical range. We assessed usability primarily by the number of items in a score and secondarily by the largest point value of any item in a score, both of which count against. Through discussion based on these criteria, we selected one risk score for each of the two predictor sets.

Software

NDP data were provided in a Microsoft Excel Workbook. The data were pre-processed and analysed in R (v4.2.1) using tidyverse (v1.3.2) and tidymodels (v1.0.0) workflows. FasterRisk was run on Python (v3.8) through R *via* reticulate (v1.26) [24–26].
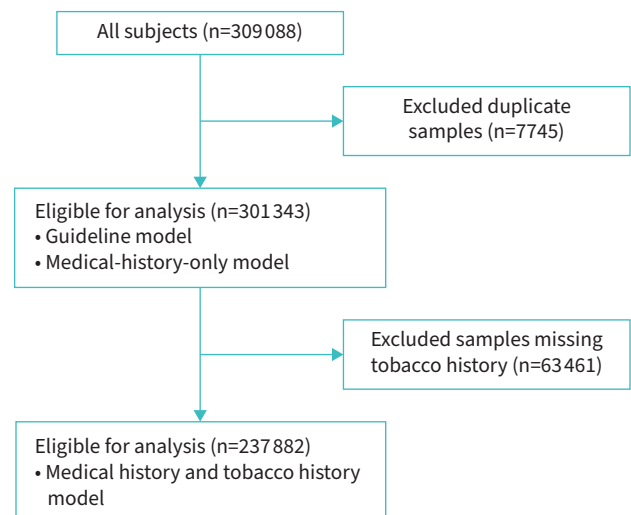
Results

Characteristics of study participants

A total of 301 343 separate individuals submitted a sample and underwent testing (figure 1). Supplementary table S1 describes the subjects' characteristics. In this cohort, the median age was 56.7 years (interquartile range 43.2–66.5 years) and 57.9% were female. The most common comorbidities were COPD (46.3%), asthma (25.4%) and emphysema (15.2%); however, 90.6% and 33.2% had no liver or lung disease, respectively. Smoking was noted in the past (48.6%), current (27.7%) and never (23.7%).

Predicting severe AATD

The testing data comprised 47 000 cases, of which 208 subjects had severe AATD. The models consistently performed better at predicting the PI*ZZ genotype than the composite PI*ZZ/PI*SZ outcome. Throughout model selection, logistic regression and integer programming were the best-performing model families. The final risk scores selected for both sets of predictors had been obtained using FasterRisk.

```
┌────────────────────────────┐
│  All subjects (n=309 088)  │
└────────────────────────────┘
              │                    ┌──────────────────────┐
              ├───────────────────▶│ Excluded duplicate   │
              │                    │ samples (n=7745)     │
              ▼                    └──────────────────────┘
┌────────────────────────────────┐
│ Eligible for analysis (n=301 343) │
│ • Guideline model               │
│ • Medical-history-only model    │
└────────────────────────────────┘
              │                    ┌──────────────────────────┐
              ├───────────────────▶│ Excluded samples missing │
              │                    │ tobacco history (n=63 461)│
              ▼                    └──────────────────────────┘
┌────────────────────────────────┐
│ Eligible for analysis (n=237 882) │
│ • Medical history and tobacco    │
│   history model                  │
└────────────────────────────────┘
```

FIGURE 1 Flowchart showing participant disposition through the study.

Table 1 compares their point values. A subject's score is calculated by adding the point values for those items reported by the subject.

Figure 2 compares the models' ROC curves along with that of guidelines. For each model, the values with specificity on either side of that of guidelines (46.6%) are outlined. Contingency tables show the estimates for the number screened and detected using a model simulating the guidelines, medical history only, and medical history with tobacco history (table 2). Unlike guidelines, neither model achieves roughly equal sensitivity and specificity at some cut-off. Instead, as shown in table 3, any cut-off significantly improves one at some cost to the other, in comparison to guidelines.

Roughly half of the PI*ZZ cases were flagged by the guidelines and final models, medical history only (score $\geqslant 0$) and medical history with tobacco history (score $\geqslant -1$), while most of the remaining PI*ZZ cases were flagged by both final models, but missed by guidelines (supplementary figure S1). The most significant pattern among those remaining is that several were flagged by guidelines and by the medical history only model, but not by the medical and smoking history model, whereas none were missed by the medical history only model.

The medical history only model, with a cut-off score of $\geqslant 0$, was estimated to flag 75.7% (35 586 out of 47 000) subjects for testing and detect 84.6% (176 out of 208) of severe AATD cases, while missing 15.4% (32 out of 208) cases. The guideline model flagged 53.4% (25 106 out of 47 000) subjects, detected 58.2% (121 out of 208) cases and missed 41.8% (87 out of 208) cases. By increasing the cut-off to a score of $\geqslant 1$, the testing burden would be reduced to 20.3% (9564 out of 47 000) of subjects at the cost of detecting only 38.0% (79 out of 208) of cases. Figure 3 is an alluvial plot of how PI*ZZ cases were treated by guidelines and the medical history model scored at either $\geqslant 0$ or $\geqslant 1$.

The medical history only model with a cut-off score of $\geqslant 0$ was ultimately selected due to the similar sensitivity and specificity without missing any additional severe AATD subjects when compared to medical and smoking history model (table 4). The point totals range from -2 to 13 with the relative risks and proportions of people with this score or higher that have severe AATD in table 5. The fourth column is the estimated share of people who would be screened at each threshold, a measure of screening burden.

Discussion

This is the first study to describe a diagnostic prediction model for severe AATD using only a subject's clinical history. It is derived and internally validated on >300 000 patients tested for AATD, making this the largest assessment of AATD testing reported to date. The proposed risk score suggests three clinically relevant risk categories: individuals with scores <0 are at low risk and may not be indicated for screening; patients with scores of 0 are at moderate relative risk and would be indicated for screening; and patients with scores >0 are at a greater relative risk and would be strongly indicated for screening. Adding the

TABLE 1 Candidate models and associated points values for each clinical characteristic; note the difference in area under the receiver operating characteristic curve (AUROC) between the candidates with and without smoking history

	Asthma	Bronchitis	Emphysema	COPD	Allergies	Bronchiectasis	AATD (fam.)	COPD (fam.)	Cirrhosis	Jaundice	Hepatitis	Abnormal LFT	Past smoking	Current smoking	AUROC
Medical	−1	3	2	0	−1	2	10	0	5	0	3	2	NA	NA	0.606
Medical + smoking	−1	0	2	0	−1	2	8	0	4	0	2	1	−1	−3	0.654

AATD: alpha-1 antitrypsin deficiency; fam.: family history; LFT: liver function test; NA: not applicable.
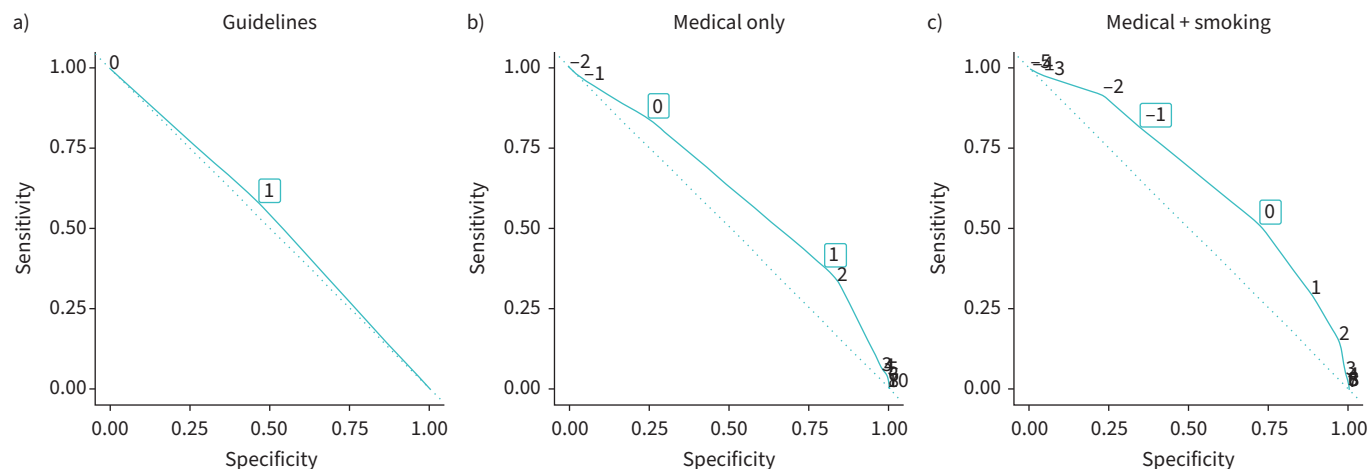
FIGURE 2 Receiver operating characteristic curves (ROC) of the model for a) current screening recommendations and the two final models b) medical history only and c) medical history plus smoking history. Suggested screening threshold values are outlined.

model to current guidelines of testing all persons with COPD has the potential to increase testing of subjects more likely to have severe deficiency and decrease the burden of testing subjects with a lower risk. Secondly, this approach could address some of those barriers experienced by healthcare providers in clinical practice.

GREULICH et al. [13] demonstrated from a survey of healthcare providers that cost and time resources limit AAT testing. The described AATD diagnostic prediction model provides a minimum-cost, evidence-based approach to stratify subjects for testing, which is viewed favourably by primary care providers [18]. In addition, it was shown that providers believe patients with AATD are only seen by specialists [13]. Primary care providers are chiefly responsible for diagnosing and managing subjects with COPD and are subsequently a source of under-diagnosis and delayed diagnosis of AATD [27–30]. A strength of this model is that it provides an actionable plan to screen subjects, which primary-care providers value from such models [18]. Lastly, the prediction model lacks age, race, gender and tobacco use variables, which are factors known to introduce bias into the selection of subjects for AAT testing [10, 13].

	Normal	Abnormal
Guidelines		
No	21 807 (46.40)	87 (0.19)
Yes	24 985 (53.16)	121 (0.26)
Medical (⩾2)		
No	11 382 (24.22)	32 (0.07)
Yes	35 410 (75.34)	176 (0.37)
Medical (⩾3)		
No	37 307 (79.38)	129 (0.27)
Yes	9485 (20.18)	79 (0.17)
Medical + smoking (⩾−1)		
No	16 175 (34.41)	38 (0.08)
Yes	30 617 (65.14)	170 (0.36)
Medical + smoking (⩾0)		
No	33 712 (71.73)	101 (0.21)
Yes	13 080 (27.83)	107 (0.23)

TABLE 2 Contingency tables for final models on the testing data at thresholds that obtain specificity on either side of guidelines specificity

Data are presented as n (%).

| TABLE 3 Thresholds for the two final models and their estimated sensitivity and specificity			
	Risk	Sensitivity	Specificity
Guidelines	0.51	58.17	46.60
Medical (⩾2)	0.41	84.62	24.32
Medical (⩾3)	0.55	37.98	79.73
Medical + smoking (⩾−1)	0.45	81.73	34.57
Medical + smoking (⩾0)	0.66	51.44	72.05
Data are presented as %.			

A number of detection strategies have been implemented, although there are differences and strengths from the proposed prediction model [31, 32]. These strategies focus on identifying a singular entity, such as a diagnosis of COPD, use of bronchodilators or airflow obstruction on pulmonary function tests (PFTs) to establish a subject's risk [5, 15, 33, 34]. In contrast, the prediction model allows for the inclusion of multiple factors to individualise risk of a severe genotype. Additionally, the prediction model is portable and may be paired with PFT software, medical documentation, electronic clinical reminders, a questionnaire linked to appointment reminders, when checking-in for an appointment, mobile applications or an interactive web-based calculator [35].

In addition, the model demonstrated novel findings. The presence of asthma and allergies alone reduced the estimated probability of severe AATD, which diverges from our current understanding. EDEN et al. [36] reviewed >1000 subjects from the AATD Registry, and found that asthma and allergies were reported in 38% and 29%, respectively. Similarly, CAMPOS et al. [27] showed not only that a prior diagnosis of asthma was reported in 43.3% of subjects registered with AlphaNet, but also that asthma was an increasing reason for AATD testing. It is important to recognise that people with AATD have a higher prevalence of asthma and asthma-like features, yet the pathomechanisms connecting these two processes is unclear and continues to be investigated [37]. The negative contributions of these factors, amid additional, stronger predictors, may reflect that their co-occurrence with AATD is largely due to their having comorbidities in common. Similarly, bronchiectasis has been considered an association with AATD and current guidelines recommend testing for AATD in persons with unexplained bronchiectasis. In this study, bronchiectasis was found to have an increased risk of severe AATD, although CARRETO et al. [38] had a low detection rate of AATD when screening persons with bronchiectasis at two centres in the UK. This also raises the question of co-occurrence of AATD with bronchiectasis and the impact of geography and prevalence of AATD detection rates.

Also of note, active and past history of smoking reduced the probability of severe AATD. AATD is classically considered a cause of emphysema that is out of proportion to a subject's tobacco use, and a model that includes smoking history and favours nonsmokers would align with this. However, we caution
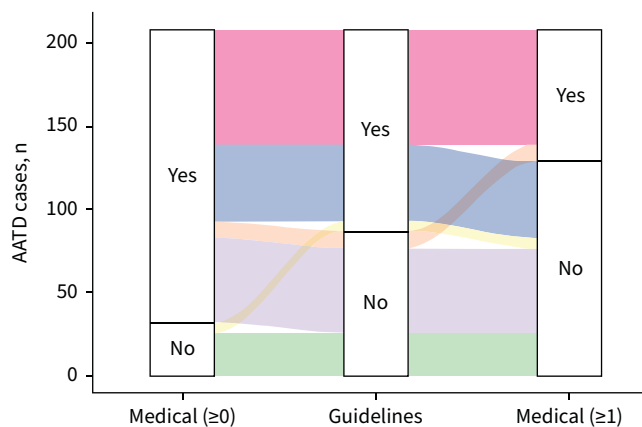
FIGURE 3 Alluvial plot of how all PI*ZZ cases were treated by a model simulating current guidelines and medical history model (scores ⩾0 and ⩾1).

TABLE 4 Final alpha-1 antitrypsin deficiency (AATD) diagnostic prediction model

	Score (points)
Asthma	−1
Bronchitis	3
Emphysema	2
Allergies	−1
Bronchiectasis	2
Family history of AATD	10
Cirrhosis	5
Hepatitis	3
History of abnormal LFTs	2

LFTs: liver function tests.

against including a smoking history in a clinical decision support tool, for two reasons: its inclusion introduces a bias against testing persons who use(d) tobacco; and persons who use tobacco who are ultimately diagnosed with AATD have higher cessation rates [39, 40]. Moreover, the medical history only model performed similarly overall without missing any additional severe AATD cases.

This study and the prediction model have a number of limitations. First, we recognise that a prediction model that includes additional medical history variables, *e.g.* recurrent respiratory infections and oxygen use, PFT measurements or chest imaging findings could produce more accurate and precise results. However, we believe that it would come at the cost of an additional cognitive burden, time to complete the questions and further testing with additional expenses [18]. Second, while the ROC of 0.606 is poor, it is a substantial improvement over our model of current guidelines. Moreover, if used as in table 5, our model would detect more severe AATD cases than guidelines while providing more specific expectations and risk estimates in each case. Third, there are chances for missed populations, such as clinically unaffected subjects who would not produce a score high enough to recommend AAT testing. We note that the same would be true for the ATS, ERS and WHO recommendations. Furthermore, the model does not address issues with heterozygotes, such as PI*MZ, who are also predisposed to disease [41], although heterozygotes with symptomatic disease, *e.g.* bronchitis or emphysema, would score high enough to warrant testing and matches guideline recommendations. Fourth, the data are not without potential biases. Most conspicuously, predictors were self-reported and subject to recall bias and interpretation. Also, the selection of subjects for AATD testing was at the discretion of the healthcare provider; thereby introducing selection bias. Regrettably, there currently exists no other large AATD database to avoid these biases, nor would it be readily feasible to construct to a large enough database, of prospectively and consecutively tested subjects to meet this request. Fifth, our comparison model, which approximates guidelines put forth by leading organisations, does not reflect the real-world processes that currently lead some, but far from

TABLE 5 Alpha-1 antitrypsin deficiency diagnostic prediction model score risk table

	Risk, %	Relative risk	Population (at or above this score), %	Risk assessment
−2	0.22	0.50	100.0	Low
−1	0.30	0.68	95.49	
0	0.41	0.92	75.71	Moderate
1	0.55	1.25	20.35	Greater
2	0.75	1.69	16.15	
3	1.01	2.29	1.89	
4	1.37	3.09	1.38	
5	1.85	4.18	0.83	
6	2.49	5.64	0.26	
7	3.36	7.58	0.20	
8	4.50	10.17	0.04	
9	6.01	13.58	0.04	
10	7.99	18.05	0.02	
11	10.54	23.82	0.00	
≥13	17.84	40.30	0.00	

all, at-risk patients to be tested for AATD. Besides professional guidelines, these processes are subject to access to care, confounding comorbidities and clinician judgement, among other factors. Lastly, the model was internally validated on a portion of the NDP's data. Validation on an external dataset would provide a better estimate of the model's performance, stronger demonstration of its usefulness and indications for further improvement.

Interpretation

This study describes a novel diagnostic prediction model that uses only a subject's medical history in order to identify those with a higher risk of having severe AATD. Compared to a model simulating guideline recommendation, the diagnostic model would improve sensitivity, testing and detection. This modelling approach could augment current guidelines and help increase the diagnosis of severe AATD.

References

1   Stoller JK, Aboussouan LS. A review of α1-antitrypsin deficiency. *Am J Respir Crit Care Med* 2012; 185: 246–259.
2   Greene CM, Marciniak SJ, Teckman J, *et al.* α1-Antitrypsin deficiency. *Nat Rev Dis Primers* 2016; 2: 16051.
3   O'Brien ML, Buist NRM, Murphey WH. Neonatal screening for alpha$_1$-antitrypsin deficiency. *J Pediatr* 1978; 92: 1006–1010.
4   Sveger T. Liver disease in alpha$_1$-antitrypsin deficiency detected by screening of 200,000 infants. *N Engl J Med* 1976; 294: 1316–1321.
5   Rahaghi FF, Sandhaus RA, Brantly ML, *et al.* The prevalence of alpha-1 antitrypsin deficiency among patients found to have airflow obstruction. *COPD* 2012; 9: 352–358.
6   Blanco I, Bueno P, Diego I, *et al.* Alpha-1 antitrypsin Pi*Z gene frequency and Pi*ZZ genotype numbers worldwide: an update. *Int J Chron Obstruct Pulmon Dis* 2017; 12: 561–569.
7   Tejwani V, Nowacki AS, Fye E, *et al.* The impact of delayed diagnosis of alpha-1 antitrypsin deficiency: the association between diagnostic delay and worsened clinical status. *Respir Care* 2019; 64: 915–922.
8   Meischl T, Schmid-Scherzer K, Loeffler-Ragg J, *et al.* The impact of diagnostic delay on survival in alpha-1-antitrypsin deficiency: results from the Austrian Alpha-1 Lung Registry. *Eur Respir J* 2021; 58: Suppl. 65, PA3302.
9   Soriano JB, Lucas SJ, Jones R, *et al.* Trends of testing for and diagnosis of α_1-antitrypsin deficiency in the UK: more testing is needed. *Eur Respir J* 2018; 52: 1800360.
10 Calle Rubio M, Soriano JB, López-Campos JL, *et al.* Testing for alpha-1 antitrypsin in COPD in outpatient respiratory clinics in Spain: a multilevel, cross-sectional analysis of the EPOCONSUL study. *PLoS One* 2018; 13: e0198777.
11 American Thoracic Society, European Respiratory Society. American Thoracic Society/European Respiratory Society statement: standards for the diagnosis and management of individuals with alpha1 antitrypsin deficiency. *Am J Respir Crit Care Med* 2003; 168: 818–900.
12 Alpha$_1$-antitrypsin deficiency: memorandum from a WHO meeting. *Bull World Health Organ* 1997; 75: 397–415.
13 Greulich T, Ottaviani S, Bals R, *et al.* Alpha1-antitrypsin deficiency – diagnostic testing and disease awareness in Germany and Italy. *Respir Med* 2013; 107: 1400–1408.
14 Taliercio RM, Chatburn RL, Stoller JK. Knowledge of alpha-1 antitrypsin deficiency among internal medicine house officers and respiratory therapists: results of a survey. *Respir Care* 2010; 55: 322–327.
15 Jain A, McCarthy K, Xu M, *et al.* Impact of a clinical decision support system in an electronic health record to enhance detection of α1-antitrypsin deficiency. *Chest* 2011; 140: 198–204.
16 Rahaghi F, Ortega I, Rahaghi N, *et al.* Physician alert suggesting alpha-1 antitrypsin deficiency testing in pulmonary function test (PFT) results. *COPD* 2009; 6: 26–30.
17 Campos M, Hagenlocker B, Martinez N, *et al.* Impact of an electronic medical record clinical reminder to improve detection of COPD and alpha-1 antitrypsin deficiency in the Veterans Administration (VA) system. *Am J Respir Crit Care Med* 2011; 183: A5356.
18 Kennedy G, Gallego B. Clinical prediction rules: a systematic review of healthcare provider opinions and preferences. *Int J Med Inform* 2019; 123: 1–10.

19 Tonelli AR, Rouhani F, Brantly ML. United States targeted detection program for alpha-1 antitrypsin deficiency. *In*: Wanner A, Sandhaus RA, eds. Alpha-1 Antitrypsin: Role in Health and Disease. Cham, Springer International Publishing, 2016; pp. 85–98.

20 Brantly M. Laboratory diagnosis of α1AT deficiency. *In*: Crystal RG, ed. Alpha 1-Antitrypsin Deficiency: Biology, Pathogenesis, Clinical Manifestations, Therapy. 1st Edn. New York, Marcel Dekker, 1995; pp. 211–226.

21 Brantly M, Crystal RG. Alpha 1-antitrypsin genotypes and phenotypes. *In*: Crystal RG, ed. Alpha 1-Antitrypsin Deficiency: Biology, Pathogenesis, Clinical Manifestations, Therapy. 1st Edn. New York, Marcel Dekker, 1995; pp. 55–60.

22 Ferrarotti I, Thun GA, Zorzetto M, *et al.* Serum levels and genotype distribution of α1-antitrypsin in the general population. *Thorax* 2012; 67: 669–674.

23 Sullivan LM, Massaro JM, D'Agostino RB Sr. Presentation of multivariate data for clinical use: the Framingham Study risk score functions. *Stat Med* 2004; 23: 1631–1660.

24 Liu J, Zhong C, Li B, *et al.* FasterRisk: fast and accurate interpretable risk scores. *arXiv* 2022; preprint [https://doi.org/10.48550/arXiv.2210.05846].

25 Wickham H, Averick M, Bryan J, *et al.* Welcome to the Tidyverse. *J Open Source Softw* 2019; 4: 1686.

26 Kuhn M, Wickham H. Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles. 2020. https://www.tidymodels.org. Date last accessed: 21 March 2023.

27 Campos MA, Wanner A, Zhang G, *et al.* Trends in the diagnosis of symptomatic patients with α1-antitrypsin deficiency between 1968 and 2003. *Chest* 2005; 128: 1179–1186.

28 Stoller JK, Sandhaus RA, Turino G, *et al.* Delay in diagnosis of α1-antitrypsin deficiency: a continuing problem. *Chest* 2005; 128: 1989–1994.

29 Sandhaus R, Turnio G, Brantly M, *et al.* The diagnosis and management of alpha-1 antitrypsin deficiency in the adult. *Chronic Obstr Pulm Dis* 2016; 3: 668–682.

30 Barr RG, Celli BR, Martinez FJ, *et al.* Physician and patient perceptions in COPD: the COPD Resource Network Needs Assessment Survey. *Am J Med* 2005; 118: 1415.

31 Brantly M, Campos M, Davis AM, *et al.* Detection of alpha-1 antitrypsin deficiency: the past, present and future. *Orphanet J Rare Dis* 2020; 15: 96.

32 Stoller JK. Detecting alpha-1 antitrypsin deficiency. *Ann Am Thorac Soc* 2016: 13: Suppl. 4, S317–S325.

33 Lam SW, Strange C, Brantly ML, *et al.* A novel detection method to identify individuals with alpha-1 antitrypsin deficiency: linking prescription of COPD medications with the patient-facing electronic medical record. *Chronic Obstr Pulm Dis* 2022; 9: 26–33.

34 Bals R, Koczulla R, Kotke V, *et al.* Identification of individuals with alpha-1-antitrypsin deficiency by a targeted screening program. *Respir Med* 2007; 101: 1708–1714.

35 DeSalle R, Wikins J, Kennett R. A kiosk survey of perception, attitudes and knowledge (PAK) of Australians concerning microbes, antibiotics, probiotics and hygiene. *Health Promot J Austr* 2022; 33: 838–851.

36 Eden E, Hammel J, Rouhani FN, *et al.* Asthma features in severe α1-antitrypsin deficiency: experience of the National Heart, Lung, and Blood Institute registry. *Chest* 2003; 123: 765–771.

37 Pini L, Paoletti G, Heffler E, *et al.* Alpha1-antitrypsin deficiency and asthma. *Curr Opin Allergy Clin Immunol* 2021; 21: 46–51.

38 Carreto L, Morrison M, Donovan J, *et al.* Utility of routine screening for alpha-1 antitrypsin deficiency in patients with bronchiectasis. *Thorax* 2020; 75: 592–593.

39 Carpenter MJ, Strange C, Jones Y, *et al.* Does genetic testing result in behavioral health change? Changes in smoking behavior following testing for alpha-1 antitrypsin deficiency. *Ann Behav Med* 2007; 33: 22–28.

40 Nakanishi T, Forgetta V, Handa T, *et al.* The undiagnosed disease burden associated with alpha-1 antitrypsin deficiency genotypes. *Eur Respir J* 2020; 56: 2001441.

41 Ghosh AJ, Hobbs BD, Moll M, *et al.* Alpha-1 antitrypsin MZ heterozygosity is an endotype of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2022; 205: 313–323.