Development of the Birmingham Lung Improvement Studies (BLISS) prognostic score for COPD patients in primary care: data from the Birmingham COPD cohort

ABSTRACT

Introduction

COPD patients in primary care have high rates of hospital admissions. A prognostic score could be used to guide patient management and reduce risk of admission but currently available scores do not perform well enough and are not used in practice.

Methods

Using data from the Birmingham primary care COPD cohort we developed and internally validated a new prognostic score from 25 candidate variables considered important from the literature and a patient-clinician stakeholder group. 1558 patients on COPD registers of 71 GP practices and 331 newly-identified patients identified from a linked case-finding trial were included and their self-reported and clinical data linked to routine hospital episode statistics. The primary outcome was the record of at least one respiratory admission within 2 years of cohort entry (May 2012-June 2014) and the secondary outcome included full follow-up data up to 01/04/2016. The model was developed using backward elimination with $p<0.157$. Fractional polynomials were considered and multiple imputation using chained equations was used for missing data. Discrimination was assessed using the c-statistic and calibration was also assessed. Bootstrapping was used for internal validation and the optimum-adjusted performance statistics were presented.

Results

Median (min, max) follow up was 2.9 years (1.8, 3.8). Of 25 candidate variables, 9 were retained in the final developed model including age, sex, smoking status, CAT score, respiratory admissions in the previous 12m, BMI, diabetes, FEV1Q and FEV1/h_2. After adjustment for optimism, the primary model performed well in predicting 2yr respiratory admissions (c statistic=0.80 (95%CI 0.77, 0.83) and calibration slope 0.88 (0.75, 1.01)). Three further variables were included in the secondary analysis but with similar score performance.

Conclusions

The BLISS score has better performance in predicting respiratory admissions than the scores currently available. All 9 variables are readily available in primary care records or would be easy to collect, and a simple computer programme could calculate the score. Important next steps are external validation, proposing and evaluating a model of use to guide patient management and exploration of the best ways to implement such a score in primary care practice.

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is one of the most common long-term conditions managed in primary care [1,2], and is also one of the most expensive to healthcare systems with a high rate of hospital admissions due to exacerbations of the condition [3,4]. In many countries there are policies and incentives to keep patients out of hospital where possible [5,6], which would benefit both the patients and the health system; however the rates of hospital admissions for this condition have not declined [7] and better strategies are urgently needed.

Prognostic scores (or indices) are often used in medical practice to assess and communicate patient risk and guide the management of individual patients, or stratify care at a practice level [8]. In the case of COPD, there are a large number of proposed prognostic scores [9,10], including the more well-known ones such as the BODE index [11], the DOSE index [12] and the ADO score [13]. These multicomponent scores have been shown to predict prognosis better than single components such as airflow obstruction, especially for predicting mortality where the ADO score has been recently shown to be the best performing score, followed by the BODE index [9]. However, none of these scores are routinely used in practice because of limitations in the development methodology, lack of validation in appropriate populations, impracticality of obtaining the variables or lack of consideration for the most important clinical outcomes. This is particularly relevant for primary care settings, where some of the proposed clinical measures may not be routinely available or practical to measure [9,10].

However, despite the large number of proposed indices, a recent systematic review revealed the lack of suitable prognostic score for predicting hospital admissions, one of the most pertinent outcomes for primary care [10]. The need for a good quality and useful prognostic score is highlighted in the latest UK NICE guidance consultation [14]. In this paper, we present the development of a new prognostic score, the BLISS score, derived from a specifically recruited primary care COPD cohort in the West Midlands region of the UK [15]. This cohort also includes case-found patients from a linked trial [16], and therefore uniquely represents both traditionally diagnosed and newly identified patients.

METHODS

This paper was written in accordance with the TRIPOD statement [17].

Aims and objectives

Development and internal validation of a new prognostic score to predict acute respiratory hospital admissions among COPD patients, for use in primary care, using data from the Birmingham COPD cohort.

Population and setting

The details of the Birmingham COPD cohort have been described in a previous publication.[15]

The cohort comprises three groups of participants: (1) 1558 COPD patients aged 40 years and over identified from the Quality and Outcomes Framework (QOF) COPD registers of 71 UK general practices within the West Midlands region of the UK; (2) 331 newly detected COPD patients aged 40-79 years from 54 of the 71 practices, identified through a linked case-finding trial (i.e. incident cases) [16]; (3) 413 patients with relevant chronic respiratory symptoms but without airflow obstruction (ie symptomatic normals), also recruited through the TargetCOPD trial [16]. This analysis includes prevalent and incident COPD cases only. Baseline assessments took place at cohort entry (31 May 2012 to 25 June 2014) and follow-up assessments took place from 2015-2016, with linked hospital episode data obtained through the Health and Social Care Information Centre (HSCIC) for the period 1 April 2012 to 31 March 2016.

Potential participants to the cohort were invited to take part by their GP, or directly from the investigators if they had provided consent through the trial, with up to two reminders for non-responders. Informed consent was obtained at the initial face-to-face visit.

Candidate variables

A large pool of potential candidate variables were identified from the literature, including variables used in relevant published prognostic scores and variables shown to be individually prognostic. A final set of candidate variables was selected through discussion with a consensus panel of study investigators/clinicians to take into consideration likely contribution to the model, accuracy and practicality in collecting the data in the primary care setting (table 1). The variables were collected from within the cohort study assessments and questionnaires and linked hospital episode statistics.

Data collection within the cohort study

At cohort entry, participants completed a face-to-face baseline clinical assessment and several self-reported questionnaires including socio-demographic variables (age, sex, ethnicity, smoking status, social contact), disease specific variables (number of exacerbations in the previous 12 months (estimated by courses of steroids and antibiotics taken), presence of chronic bronchitis [18], extent of dyspnoea (MRC scale) [18]) and selected physician-diagnosed conditions. Disease-specific health-related quality of life (HRQL) was measured using the COPD assessment test (CAT)[19] and general health using a 5-point Likert scale. Self-reported exercise levels were reported using the IPAQ-short [20] and exercise capacity measured using the sit-to-stand test[21]. Height was measured to the nearest 0.1 cm using a Leicester height monitor, and weight (to the nearest 0.1 kg) was assessed using the Tanita BC-420SMA body composition scale.

Lung function (FEV_1) was measured using the nddEasy One Spirometer (ndd, Switzerland), administered by researchers trained to ARTP Foundation Spirometry Certificate standard [15] before (max eight blows) and after (max six blows) 400µg salbutamol, stopping when repeatability within 100mls was achieved. The highest recording was then taken. FEV_1% predicted was estimated using the GLI equations [22]. Due to the documented statistical problems with the use of the FEV1% predicted measure, we examined 3 different measures of FEV_1 as potential predictors: FEV1Q, FEV1/height^2 and FEV1% predicted [23]. Bronchial hyper-responsiveness was defined as change between pre & post BD FEV1 >12% and >200ml, OR change between pre & post BD FVC1 >12% and >200ml. The IMD (2010) score was calculated as a measure of deprivation, based on patients' individual postcode [24].

We obtained data on current or main occupation using a questionnaire administered by trained research assistants, who used information on skill content and skill level to assign a 4-digit standard occupational classification (SOC 2010) [25] code using the CASCOT (computer assisted structured coding tool) software.[26] Risk of occupational exposure to vapours, gases, dust and fumes (VGDF) was derived using a job exposure matrix [27], modified for use with SOC 2010 codes.

Use of cardiovascular medications was self-reported by patients.

Outcomes

Data on hospital episodes were obtained from NHS Digital using patient NHS number and linked to the cohort data via a unique study ID. The primary outcome was one or more acute

respiratory admissions during the two year period since entry to the cohort, defined using specific respiratory ICD10 codes (see Appendix 1). As a sensitivity analysis, we developed a prognostic model to predict occurrence of one or more acute respiratory admissions during the full follow-up period from cohort entry until the NHS Digital admissions data was obtained (01/04/2016).

Statistical analyses

Developing the prognostic model

The outcome was modelled using a logistic regression model. Firstly the full model was fitted, including all candidate variables, and then backward elimination performed, with a conservative significance level of 0.157 used [28]. For categorical variables included in the model, the category with the lowest p-value was used to assess the significance level. No variables were forced into the model. Continuous variables remained in their raw form to ensure data were not lost through dichotomisation. Initially a linear trend was assumed, then, where possible, fractional polynomials were considered (set of powers considered: -2, -1, -0.5, natural logarithm, 0.5, 1, 2, 3) with $p<0.001$ indicating the use of a fractional polynomial rather than linear trend. [29] Fractional polynomials were also used for the continuous variables eliminated from the model to check whether they should be included in the fractional polynomial format.

Multiple imputation (using chained equations) was used for all variables considered in the model and auxiliary variables used to aid the imputation. The number of imputed data sets used was equal to the fraction of missing data (64 data sets for 64% missing data). [30]

Assessment of prognostic model performance

Assessment of the fitted model was achieved by estimating calibration and discrimination. A calibration plot was produced by plotting the observed risk against the predicted risk and the calibration slope calculated. To judge discrimination the area under the receiver operating curve was calculated (equivalent to the c-statistic).

Internal validation of the prognostic model

This developed ‘apparent’ model was then internally validated using bootstrap methods. Each imputed dataset was used to generate 100 bootstrapped datasets. Each one of these bootstrapped data sets was then used to develop a prognostic model in the same way as the original model. Estimates of performance (c-statistic and calibration slope) were obtained

from the model fitted using each of the bootstrapped data sets. The estimates obtained from the bootstrapped data sets were averaged and subtracted from the estimates from the original model to estimate optimism and provide optimism-adjusted performance statistics.

Final prognostic model

The optimised adjusted calibration slope was then used as a uniform shrinkage factor. Each of the coefficients from the original apparent model was adjusted by multiplying by the shrinkage factor. The intercept was also adjusted to ensure calibration-in-the-large.

Subsidiary and sensitivity analyses

Although the final model included two different measures of FEV_1, we also considered how the inclusion of only one of the three potential measures would impact on the model performance by evaluating their separate inclusion at the development stage within the apparent model. We also evaluated how well the model would perform on the prevalent cases only.

Sample size calculation

With 267 events for the primary outcome, up to 26 candidate variables could be used, based on the rule of thumb of 10 events per candidate variable. [31]

RESULTS

Characteristics of participants

Of 7176 invited to the cohort, 1558 prevalent and 331 incident participants completed baseline assessments and were included in these analyses [15]. Median follow-up (min, max) was 2.9 years (1.8, 3.8 years), 382 (16%) had a respiratory admission recorded during the study period, and 267 (12%) had a respiratory admission in the primary two year period. Participants with hospitalisations were more likely to be older (70.6 vs 6.7 years, $p<0.001$), male (65% vs 59% $p=0.017$), more deprived (median IMD score 30.7 vs 23.8, $p<0.001$), have lower BMI (mean 28.1 vs 28.9, $p=0.017$), more severe airflow obstruction (mean FEV1 55.2 vs 76.8% predicted), worse dyspnoea (MRC 3-5 74% vs 49%, $p<0.001$)), worse quality of life scores (median CAT score 24 vs 16, $p<0.001$), report previous exacerbations (62% vs 42%, $p<0.001$) and previous hospitalisations (16.0% vs 2.2%, $p<0.001$)), higher rate of VGDF (71% vs 62%, $p=0.001$) and smoking exposure (31.8% vs 26.5% current smokers, $p<0.001$), and have diabetes (24% vs 15%, $p<0.001$) and cardiovascular disease (22% vs 14% with coronary heart disease, $p<0.001$) (Table 2).

Primary analysis predicting acute respiratory admissions in a 2-year period

For the primary analysis, of 25 candidate predictors, 9 were retained in the final developed model (table 3), including age, sex, smoking status, CAT score, previous respiratory admission, BMI, self-report of a diagnosis of diabetes and two different measures of obstruction. After adjusting for optimism (using a uniform shrinkage factor of 0.869), the prediction model was able to discriminate between COPD participants with and without a respiratory admission with a c-statistic of 0.80 (0.77, 0.83) (table 4). There was also good agreement between observed and predicted probabilities with a calibration slope of 0.88 (0.75, 1.01) (fig 1).

Sensitivity analysis predicting acute respiratory admissions for the full follow-up

We repeated the analysis using the full follow-up period. An additional 3 variables were retained in the model (antibiotic/steroid prescription in the last 12 months, bronchial hyper-responsiveness and self-report of a diagnosis of heart failure) were retained in the final developed model (Table S1). After adjusting for optimism (shrinkage factor 0.877) the c-statistic was similar at 0.80 (0.78, 0.83), again with good agreement between observed and expected probabilities (fig S1).

Further sensitivity analyses

At the initial model development stage (the apparent model), we explored the use of only one measure of airflow obstruction. With only FEVQ included the c-statistic was 0.77 (0.74, 0.80); with only FEV1/h_2 the c-statistic was 0.78 (0.75, 0.81) and with only FEV1% predicted, the c-statistic was 0.78 (0.75, 0.81). Including only prevalent cases resulted in a c-statistic of 0.76 (95%CI: 0.73 to 0.79). These results were not adjusted for optimism.

Examples of the application of this score (see Table 3 for equation)

Example 1 – A 70 year old male, who is a current smoker, has a BMI of 20, and has had a respiratory related hospitalisation in the previous 12 months. His obstruction is measured as 0.25 (FEV1/h_2) and 8 (FEV1q). His disease specific HRQL category is 35, and he does not have diabetes. He has a predicted risk of 83.4% of having a respiratory related hospitalisation in the next two years. *Interpretation*: If 1000 people with the same risk factors are followed for two years, 834 would have a respiratory related hospitalisation.

Example 2 – A 60 year old female, who has never smoked, has a BMI of 25, and has not had a previous respiratory related hospitalisation in the previous 12 months. Her obstruction is measured as 0.1 (FEV1/h_2) and 4 (FEV1q). Her disease specific HRQL category is 20, and she has diabetes. She has a predicted risk of 13.5% of having a respiratory related hospitalisation within two years. *Interpretation*: If 1000 people with the same risk factors are followed for two years, 135 would have a respiratory related hospitalisation.

DISCUSSION

Key findings

Although there are more than 27 proposed prognostic models and scores in the published literature evaluated for use in predicting exacerbations of COPD, none of these are suitable for use in practice because of limitations in the methodology of their development or validation, inadequate performance in predicting hospital admissions (indicating that further variables are needed) or impracticality in measuring some components in primary care. [10] We have used data from a unique primary care COPD cohort to develop a novel prognostic score for primary care, considering all the potential predictors from previously published scores and other prognostic factors likely to be important. Using best practice methodology we have produced the BLISS score, which has good discriminative ability and good calibration and better performance than any previously published scores in predicting risk of respiratory admissions [10]. There are 9 variables, all readily available in primary care records or easy to collect and each of the components has been shown to be individually associated with increased risk of admission and therefore not a surprising inclusion. Age and respiratory admission in the previous 12 months were strong predictors in the model, which is consistent with other evidence [32,33]. BMI is known to have a non-linear relationship with poor prognosis [34], which may also explain its non-linear function in our score.

Comparison with existing literature

The BLISS score has many variables in common with other prognostic scores for COPD. Airflow obstruction is the most commonly found variable, followed by previous exacerbations, age, smoking, COPD-specific quality of life, BMI and sex [10]. The most commonly known scores have been developed to predict other outcomes such as mortality or health-related quality of life [9]. Of these, the BODE index contains two of the BLISS score variables (BMI, obstruction) but dyspnoea and exercise capacity rather than the CAT score. [11] The DOSE index contains three of the BLISS index variables (obstruction, smoking status and exacerbations) [12] and the ADO score contains age and obstruction in common [13], but both also contain dyspnoea as well. It is likely that the CAT score and the MRC score measure similar dimensions (impact of breathlessness) and they are frequently used as alternatives to each other [35]. A number of scores also include comorbidities [36-38] although none identify diabetes as a single predictive component. Very few scores have been developed within a primary care setting. However, the most relevant comparative score is

probably that produced by Bertens et al [36] which aimed to predict exacerbations (described by steroid use or hospitalisation) in a 2-year period among COPD patients in primary care. This score, containing 4 variables (previous exacerbations, FEV1% predicted, pack years of smoking and presence of vascular disease), was derived within a primary care cohort of COPD patients aged 65 years and over from 51 general practices in the Netherlands, and validated in a cohort aged 50 years and over. Although having good discrimination and good calibration in the derivation cohort (c=0.75) it had moderate discrimination in the validation cohort (c-statistic of 0.66 (95% CI: 0.62–0.71), considered a more limited range of candidate variables and defined exacerbations more broadly as those requiring courses of antibiotics or steroids or hospital admissions.

Limitations of this study

Although the BLISS score is based on 9 readily available or easily obtainable components, the non-linear nature of several of them makes it more difficult to understand and compute than a simple points-based score. However, most GP systems have an inbuilt facility to calculate such scores, or a simple programme in Excel could do this.

Many of the components are based on self-report, which for comorbidities may not be as accurate as data available in routine GP records, although is unlikely to be systematically biased.

In the UK, the CAT score is less commonly recorded than the MRC score (which is required for QOF), although it is suggested to be collected during annual reviews [39] and appears to be more useful for prognosis and would not be difficult to collect and record as it consists of only eight questions on a Likert scale.

There has been considerable debate about the value of including non-modifiable factors such as age in COPD prognostic scores. However, excluding important predictive factors such as age and sex would lead to confounding and biased estimates of the remaining predictors, producing a score which performs badly. The aim of a prognostic score should be to predict risk accurately; the role of the clinician is to then use the score to guide their management, which can address the factors which are modifiable.

The inclusion of two different measures of FEV1 may be considered unusual. Due to controversies surrounding the best potential measure [23], we considered three different possibilities and allowed the statistical approach to determine which was more useful. The best combination included both FEV1/height$_2$ and FEV1Q. These capture slightly different

dimensions where the FEV1/height$_2$ standardises for a person's size, and the FEV1Q is an index of the number of turnovers of a nominal lower limit of lung function remaining, and takes into account some sex and size differences in lung function [23]. This is consistent with another study suggesting that FEV1Q and FEV1/h$_2$ were the best measures to use [40]. However, our sensitivity analyses showed that including only one at a time reduced the overall performance of the score a little, although of the 3 single measures, the traditional FEV1% predicted performed the best.

Most of the included participants had 2 years of follow-up data which provided the primary outcome. Our secondary analyses included full follow-up data (median 2.9 years), an extra 115 events and a further 3 variables although these contributed less to the model than the original variables and were probably included due to the increased statistical power available. Most previous studies have follow-up limited to one year [10].

Finally, it is possible that our population does not truly represent primary care as we included those who were case-found, and also those who were prepared to take part in a research study who would be more likely to have milder disease than the average of primary care [15]. Indeed the score performed slightly less well amongst prevalent-only cases, although this was not statistically significant.

Implications for research and practice

Although we have performed internal validation, before the score should be used, further external validation in relevant primary care datasets is important. Further work with primary care clinicians is also needed to understand the reasons for lack of uptake of such scores in practice, and then using this information to propose and test a practical use for the score in guiding or stratifying patient management. [41]

It is possible that a whole practice COPD population could be stratified by 2-year risk of admission, and then the greatest resources directed towards those at greatest risk. Trials which test this approach are needed. A further use might be to guide individual patient management. Within the GOLD guidelines, the new ABCD matrix includes one dimension which relates to exacerbation risk [35]. At the moment, that exacerbation risk is defined by number of previous exacerbations. Perhaps the BLISS score could be used as a better marker of future risk? However it would be important to decide how to categorise level of risk within the BLISS score, and how many cut-points it should have.

There are now many such scores available, and given our rigorous approach and the fact that the score has both good discrimination and calibration, now it is time to move to the next phase and test its utility in practice rather than developing new scores.

Conclusions

Using robust methodology and a COPD patient cohort which represents primary care, we have developed and internally validated a new prognostic score which performs very well in predicting respiratory admissions within a two-year timeframe. The components are easy to collect and the score performs better than any other published score. The next steps are to test its application in practice and identify how best to implement its use in a real life primary care setting.

REFERENCES

1. Halbert RJ, Natoli JL, Gano A, et al. Global burden of COPD: systematic review and meta-analysis. The European Respiratory Journal 2006;28(3):523-32.
2. Quality and Outcomes Framework 2013/14. Available from: http://www.hscic.gov.uk/catalogue/PUB15751
3. Mahon J, Bartlett C, Isojarvi J, Edwards M, Glanville J. York Health Consortium YH. Literature review: the economic costs of lung disease and the cost-effectiveness of policy and service interventions. 2017 Oct 2017. https://www.brit-thoracic.org.uk/document-library/research-and-innovation/documents/literature-review-the-economiccosts-of-lung-disease/
4. McLean S, Hoogendoorn M, Hoogenveen RT, Feenstra TL, Wild S, Simpson CR, et al. Projecting the COPD population and costs in England and Scotland: 2011 to 2030. Sci Rep. 2016;6:31893
5. Department of Health (England). Improving the health and well-being of people with long-term conditions. Department of Health; 2010.
6. Department of Health (England). An outcomes strategy for people with chronic obstructive pulmonary disease (COPD) and asthma in England. 2011
7. Hospital Episode Statistics, Admitted Patient Care, England - 2013-14. Available from: http://www.hscic.gov.uk
8. Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ [Internet] 2009;338(7709):1487–1490. Available from: http://www.bmj.com/cgi/doi/10.1136/bmj.b606
9. Guerra B, Haile SR, Lamprecht B, et al. Large-scale external validation and comparison of prognostic models: An application to chronic obstructive pulmonary disease. BMC Med 2018;16(1):1–13
10. Guerra B, Gaveikaite V, Bianchi C, et al. Prediction models for exacerbations in patients with COPD. Eur Respir Rev 2017; 26: 160061 [https://doi.org/10.1183/16000617.0061-2016].
11. Celli B, Cote C, Marin JM, et al. The body-mass index, airflow obstruction, dyspnea, and exercise capacity index in chronic obstructive pulmonary disease. N Engl J Med [Internet] 2004;350(10):1005–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14999112

12. Jones RC, Donaldson GC, Chavannes NH, et al. Derivation and validation of a composite index of severity in chronic obstructive pulmonary disease: The DOSE index. Am J Respir Crit Care Med 2009;180(12):1189–1195.
13. Puhan MA,Hansel NN, Sobradillo P, et al. Large-scale international validation of the ADO index in subjects with COPD: an individual subject data analysis of 10 cohorts. BMJ Open 2012;2:e002152.doi:10.1136/bmjopen-2012-002152
14. https://www.nice.org.uk/guidance/indevelopment/gid-ng10026/documents
15. Adab P, Fitzmaurice DA, Dickens AP, et al. Cohort Profile: The Birmingham Chronic Obstructive Pulmonary Disease (COPD) Cohort Study. Int J Epidemiol [Internet] 2016;46(1):dyv350. Available from: https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dyv350
16. Jordan RE, Adab P, Sitch A, et al. Targeted case finding for chronic obstructive pulmonary disease versus routine practice in primary care (TargetCOPD): a cluster-randomised controlled trial. Lancet Respir Med [Internet] 2017;4(9):720–730. Available from: http://dx.doi.org/10.1016/S2213-2600(16)30149-7
17. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. Ann Intern Med 2017;162(1):55–63.
18. Fletcher CM. Standardised questionnaire on respiratory symptoms: a statement prepared and approved by the MRC Committee on the Aetiology of Chronic Bronchitis (MRC breathlessness score). British Medical Journal 1960;(2):1665.
19. Jones PW, Harding G, Berry P, et al. Development and first validation of the COPD Assessment Test. The European Respiratory Journal 2009;34(3):648-54.
20. Craig CL, Marshall AL, SjostromM et al. International physical activity questionnaire: 12-country reliability and validity. Medicine and science in sports and exercise 2003;35:1381–95.
21. Ozalevli S, Ozden A, Itil O, Akkoclu A. Comparison of the Sitto-Stand Test with 6 min walk test in patients with chronic obstructive pulmonary disease. Respir Med 2007;101:286–93.
22. Quanjer PH, Stanojevic S, Cole TJ, et al. ERS Global Lung Function Initiative. Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. Eur Respir J 2012; 40: 1324–43.

23. Miller et al. New concepts for expressing forced expiratory volume in 1 s arising from survival analysisEur Respir J 2010; 35: 873–882. DOI: 10.1183/09031936.00025809FEV1/h2
24. Department for Communities and Local Government. English Indices of Deprivation 2010. http://data gov uk/dataset/index-of-multiple-deprivation 2010 [cited 2015 Apr 14]
25. SOC 2010. https://www.ons.gov.uk/methodology/classificationsandstandards/standardoccupationalclassificationsoc/soc2010
26. Warwick Institute for Employment Research. CASCOT: Computer Assisted Structured Coding Tool. http://www2 warwick ac uk/fac/soc/ier/software/cascot/ 2014
27. Sadhra SS, Kurmi OP, Chambers H, et al. Development of an occupational airborne chemical exposure matrix. Occupational Medicine 2016;66(5):358-64.
28. Sauerbrei W. The use of resampling methods to simplify regression models in medical statistics. Journal of the Royal Statistical Society Series C-Applied Statistics. 1999;48:313-29.
29. Thangaratinam S, Allotey J, Marlin N, Mol BW, Von Dadelszen P, Ganzevoort W, Akkermans J, Ahmed A, Daniels J, Deeks J, Ismail K. Development and validation of Prediction models for Risks of complications in Early-onset Pre-eclampsia (PREP): a prospective cohort study. Health Technol Assess. 2017.
30. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Statistics in Medicine. 2011;30(4):377-99Peduzzi
31. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996 Dec 1;49(12):1373-9.
32. Motegi T, Jones RC, Ishii T, et al. A comparison of three multidimensional indices of COPD severity as predictors of future exacerbations. Int J COPD 2013; 8: 259–271.
33. Hurst JR, Vestbo J, Anzueto A, et al. Susceptib ility to exacerbation in chronic obstructive pulmonary disease. N Engl J Med 2010; 363: 1128–1138.
34. Guo Y, Zhang T, Wang Z, et al. Body mass index and mortality in chronic obstructive pulmonary disease: A dose–response meta-analysis. Liu. J, ed. Medicine. 2016;95(28):e4225. doi:10.1097/MD.0000000000004225.
35. Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease. GOLD, 2018.

https://goldcopd.org/wp-content/uploads/2017/11/GOLD-2018-v6.0-FINAL-revised-20-Nov_WMS.pdf

36. Bertens LC, Reitsma JB, Moons KG et al. Development and validation of a model to predict the risk of exacerbations in chronic obstructive pulmonary disease. Int J Chron Obstruct Pulmon Dis 2013;8:493–99.
37. Almagro P, Barreiro B, Ochoa de Echaguen A, et al. Risk factors for hospital readmission in patients with chronic obstructive pulmonary disease. Respiration 2006; 73: 311–317.
38. Amalakuhan B, Kiljanek L, Parvathaneni A, et al. A prediction model for COPD readmissions: catching up, catching our breath, and improving a national problem. J Community Hosp Intern Med Perspect 2012; 2: 9915.
39. NICE quality standards. https://www.nice.org.uk/guidance/qs10
40. Tang-Hsiu Huang, Tzuen-Ren Hsiue, Sheng-Hsiang Lin, Xin-Ming Liao, Po-Lan Su, Chiung-Zuei Chen European Respiratory Journal 2018 51: 1700577; DOI: 10.1183/13993003.00577-2017
41. van Dijk WD, van den Bemt L, van den Haak-Rongen S, Bischoff E, van Weel C, in ‘t Veen JCCM, Schermer TRJ. Multidimensional prognostic indices for use in COPD patient care. A systematic review. Respiratory Research 2011, 12:151 http://respiratory-research.com/content/12/1/151

Table 1 Candidate variables and data source

Description	Form of variable	Data source
Demographics		
BMI	Categorical	Cohort assessment data
Age	Continuous	Cohort self-report data: questionnaires
Sex	Binary	Cohort self-report data: questionnaires
Ethnicity	Categorical	Cohort self-report data: questionnaires
COPD specific risk factors		
Obstruction—FEV1 % predicted	Continuous	Cohort assessment data
Obstruction—FEV1Q*	Continuous	Cohort assessment data
Obstruction—FEV1/height$_2$	Continuous	Cohort assessment data
Bronchial hyper-responsiveness	Binary	Cohort assessment data
Dyspnoea—MRC scale	Categorical	Cohort self-report data: questionnaires
Disease specific HRQL – CAT	Continuous	Cohort self-report data: questionnaires
Previous respiratory hospitalisations (12 months prior to baseline)	Binary	NHS Digital
Course of antibiotics/steroids within last 12 months	Binary	Cohort self-report data: questionnaires
Chronic cough and/or chronic phlegm for 3 or more months of the year	Binary	Cohort self-report data: questionnaires
Other risk factors		
Smoking	Categorical	Cohort self-report data: questionnaires
Social isolation	Binary	Cohort self-report data: questionnaires
Exercise capacity—sit to stand test	Continuous	Cohort assessment data
Physical activity—IPAQ	Categorical	Cohort self-report data: questionnaires
History of CVD	Binary	Cohort self-report data: questionnaires
Medication for CVD	Binary	Cohort self-report data: questionnaires
Heart failure	Binary	Cohort self-report: questionnaires
Asthma	Binary	Cohort self-report data: questionnaires
Depression	Binary	Cohort self-report data: questionnaires
Diabetes	Binary	Cohort self-report data: questionnaires
Any cancer	Binary	Cohort self-report data: questionnaires
Osteoporosis	Binary	Cohort self-report data: questionnaires

*See Miller et al [23] for calculation

Table 2: Baseline characteristics of BLISS cohort study participants by respiratory hospitalisation during study period.

Variable	Total population (N =2,305)	No hospitalisation (N = 1,923)	Respiratory hospitalisation (N = 382)	P-value[1]
Demographics				
Age, Mean (SD)	67.4 (9.9)	66.7 (9.7)	70.6 (10.2)	<0.001
Sex (Male)	1,382 (60)	1,132 (59)	250 (65)	0.017
Ethnicity				
White British	1,918 (83.2)	1,601(83.3)	317(83.0)	0.525
Asian	53 (2.3)	47(2.4)	6(1.6)	
African/Caribbean	23 (1.0)	20(1.0)	3(0.8)	
Mixed	13 (0.6)	10(0.5)	3(0.8)	
Other	120 (5.2)	94(4.9)	26(6.8)	
Unclear/Missing	178 (7.7)	151(7.9)	27(7.1)	
Deprivation (IMD), median [IQR]	25.0 [14.4 to 41.4]	23.8 [14.1 to 39.7]	30.7 [17.1 to 45.1]	<0.001
BMI, mean (SD)	28.7 (5.8)	28.9 (5.5)	28.1 (6.7)	0.017
COPD specific				
Obstruction (FEV1% predicted) , median [IQR]	73.7 [56.7 to 88.8]	76.8 [60.9 to 90.9]	55.2 [38.9 to 72.2]	<0.001
Obstruction (FEV1 Q) , median [IQR]	0 [0 to 2]	0 [0 to 2]	2 [0 to 4]	<0.001
Obstruction (FEV1/h2) , mean (SD)	0.69 (0.27)	0.73 (0.26)	0.52 (0.23)	<0.001
Bronchial hyper-responsiveness	155 (6.7)	121 (6.3)	34 (8.9)	0.063
Dyspnoea				
Grade 1 – 2	1,014 (47)	920 (51)	94 (26)	<0.001
Grade 3 – 5	1,157 (53)	895 (49)	262 (74)	
Disease specific HRQL Categories, median [IQR]	17 [11 to 24]	16 [10 to 23]	24 [18 to 31]	<0.001
Respiratory hospitalisation in the previous 12m[2] (Count)				
0	2,202 (95.5)	1,881 (97.8)	321 (84)	<0.001
1	78 (3.4)	37 (1.9)	41 (10.7)	
2+	25 (1.1)	5 (0.3)	20 (5.2)	
Previous hospitalisation[2] (Binary)	103 (4.5)	42 (2.2)	61 (16.0)	<0.001
Antibiotics/Steroids	1,040 (45)	802 (42)	238 (62)	<0.001
Chronic cough and/or chronic phlegm	1,332 (58)	1,072 (56)	260 (68)	<0.001
VGDF exposure	2,239 (63)	1,152 (62)	263 (71)	0.001
Other risk factors				

Smoking				
Never Smoker	268 (12.6)	246 (13.8)	22 (6.3)	<0.001
Current Smoker	583 (27.4)	472 (26.5)	111 (31.8)	
Ex Smoker	1,279 (60.0)	1,063 (59.7)	216 (61.9)	
Social isolation	130 (6)	100 (6)	30 (8)	0.048
Exercise capacity (Sit to stand test)	19 [15 to 23]	20 [15 to 24]	16 [13 to 20]	<0.001
Physical activity (IPAQ)				
Low Activity	730 (41)	580 (39)	150 (55)	<0.001
Moderate Activity	595 (34)	515 (34)	80 (30)	
High Activity	439 (25)	398 (27)	41 (15)	
General health (Likert scale)				
1	171 (7.8)	159 (8.7)	12 (3.4)	<0.001
2	848 (38.7)	769 (42.0)	79 (22.1)	
3	926 (42.3)	745 (40.6)	181 (50.7)	
4	215 (9.8)	143 (7.8)	72 (20.2)	
5	30 (1.4)	17 (0.9)	13 (3.6)	
Asthma	811 (40)	669 (39)	142 (44)	0.071
Depression	479 (24)	409 (24)	70 (22)	0.445
Diabetes	330 (16)	252 (15)	78 (24)	<0.001
Cancer	266 (13)	223 (13)	43 (13)	0.884
Osteoporosis	156 (8)	124 (8)	32 (11)	0.075
Cardiovascular disease related				
Coronary heart disease	30 (15)	235 (14)	73 (22)	<0.001
Heart failure	158 (8)	116 (7)	42 (13)	<0.001
Medication	1,152 (50)	924 (48)	228 (60)	<0.001

Values are Number (percentage) unless specified.

1: P-value obtained from t-test, Mann-whitney U test, or chi-squared test. 2: Hospitalisation for respiratory related problem in previous 12 months obtained from Hospital episode statistics . IQR: Inter-quartile range.

Table 3: Final multivariable model for risk of respiratory hospitalisation within two years for participants with chronic obstructive pulmonary disease

Variable	OR (95%CI)	β coefficients
FEV1/h_2	0.131 (0.069 - 0.252)	-2.02928538
Disease specific HRQL Categories	1.045 (1.028 - 1.062)	0.04386269
Sex (Male)	1.402 (1.076 - 1.827)	0.33805492
Previous 12 month respiratory hospitalisation	3.882 (2.591 - 5.816)	1.35624740
Diabetes	1.550 (1.103 - 2.179)	0.43830211
None Smoker	Reference	-
Current Smoker	1.691 (0.979 - 2.921)	0.52533606
Ex Smoker	1.687 (1.022 - 2.782)	0.52265980
Fractional polynomial transformation		
$(BMI/10)_3$		-0.13840965
$BMI/10)_3$ x ln(BMI/10)		0.08650160
Age/10		-5.74991978
(Age/10) x ln(Age/10)		2.02527365
(FEV1Q + 0.000005)/100		0.53161365
((FEV1Q + 0.000005)/100) x ln((FEV1Q + 0.000005)/100)		-1.09589609
Constant		10.95740000

Risk score = 10.957 – 2.029FEV1h2 + 0.044CAT + 0.338Male + 1.356previous hospitalisation + 0.438Diabetes + 0.525Current smoker + 0.523Ex smoker – 0.138((bmi/10)3)+0.087(bmi/10)3ln(bmi/10) – 5.750 Age/10 + 2.025 Age/10 ln(Age/10) +0.532 ((FEV1Q+ 0.000005)/100) – 1.096((FEV1Q+ 0.000005)/100)ln((FEV1Q+ 0.000005)/100).

Note: ln= natural logarithm

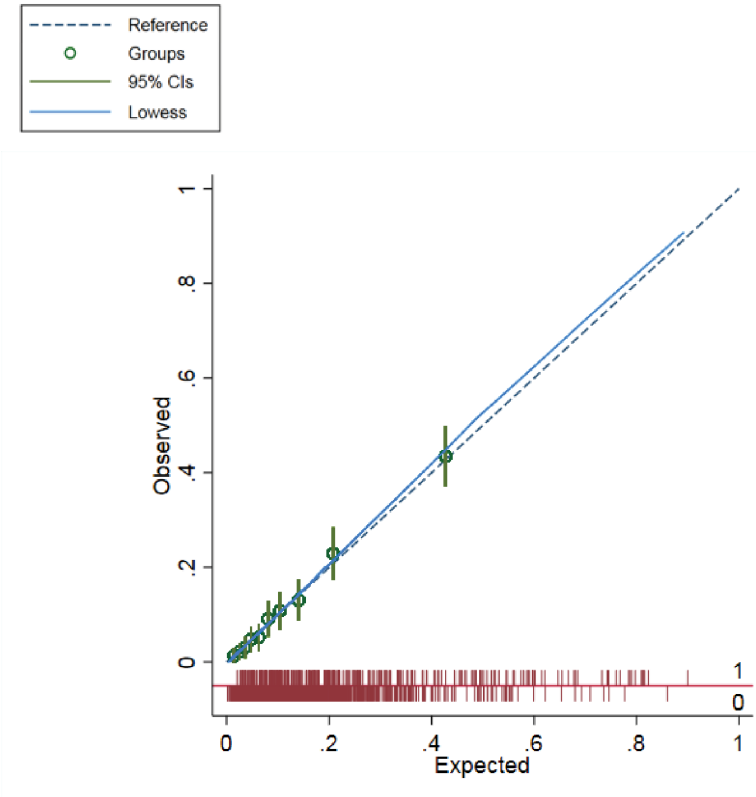
All variables are coded as binary (0 for absence of presence of a risk factor), except for FEV1/h_2, FEV1Q, HRQL, BMI, and Age. The value 10.957 is the intercept, and the other numbers reflect the estimated coefficients for the predictors, indicating their contribution to the risk. The regression coefficients represent the log odds ratio for a change in 1 unit in the corresponding predictor. The predicted risk of hospitalisation is $1/(1+e^{-riskscore})$.

Table 4: Model diagnostics (with 95% CI)

Measure	Apparent Performance[1]	Test Performance[2]	Average Optimism[3]	Optimism corrected[4]
C-Statistic [5]	0.79 (0.76 to 0.82)	0.80 (0.77 to 0.80)	-0.0123	0.80 (0.77 to 0.83)
Calibration slope	1.00 (0.87 to 1.13)	0.88 (0.76 to 0.99)	0.1216	0.88 (0.75 to 1.01)

[1]: Refers to performance estimated from imputation datasets that were used to develop prediction model
[2]: Determined by developing model in each bootstrapped sample (100 samples with replacement), calculating performance (bootstrap performance), and applying bootstrap model in original imputed dataset.
[3]: Average difference between model performance in bootstrap data and original imputation data
[4]: Subtracting optimism from apparent performance
[5]: Probability that for any randomly selected pair of patients with diagnosed COPD with and without respiratory hospitalisation, the patient with respiratory hospitalisation had higher predicted risk. A value of 0.5 represents no discrimination and 1.00 represents perfect discrimination.

Figure 1: Assessing calibration in original data of the prediction of respiratory hospital admissions within 2 years



Red lines indicate individual respiratory admission events

Table S1: Final multivariable model for respiratory hospitalisation risk for participants with chronic obstructive pulmonary disease from cohort entry until 01/04/2016

Variable	OR (95%CI)	β coefficients
Obstruction, FEV1/h2	0.086 (0.047 - 0.158)	-2.45280977
Obstruction, FEV1Q	0.997 (0.993 - 1.000)	-0.00326171
Disease specific HRQL Categories	1.044 (1.028 - 1.060)	0.04280823
Gender (Male)	1.483 (1.170 - 1.878)	0.3938612
Previous 12 month respiratory hospitalisation	4.133 (2.732 - 6.252)	1.41900155
Antibiotic/Steroid use	1.282 (1.013 - 1.624)	0.24877863
Bronchial Hyper-responsiveness	0.596 (0.300 - 1.184)	-0.51735871
Diabetes	1.718 (1.269 - 2.324)	0.54087298
Heart Failure	1.403 (0.960 - 2.049)	0.33826103
None-smoker	Reference	-
Current Smoker	2.153 (1.320 - 3.511)	0.76692595
Ex Smoker	1.885 (1.201 - 2.958)	0.63383119
Fractional polynomial transformed		
(bmi/10)^2	-	-0.77622761
(bmi/10)^2 x ln(bmi/10)	-	0.44991177
Age/10	-	-6.68969089
Age/10 x ln(Age/10)	-	2.37664856
Constant	-	14.5633

Table S2: Model diagnostics (with 95% CI)

Measure	Apparent Performance[1]	Test Performance[2]	Average Optimism[3]	Optimism corrected[4]
C-Statistic [5]	0.80 (0.78 to 0.83)	0.80 (0.79 to 0.81)	0.0007	0.80 (0.78 to 0.83)
Calibration slope	1.00 (0.88 to 1.11)	0.88 (0.78 to 0.99)	0.122	0.88 (0.76 to 0.99)

[1]: Refers to performance estimated from imputation datasets that were used to develop prediction model

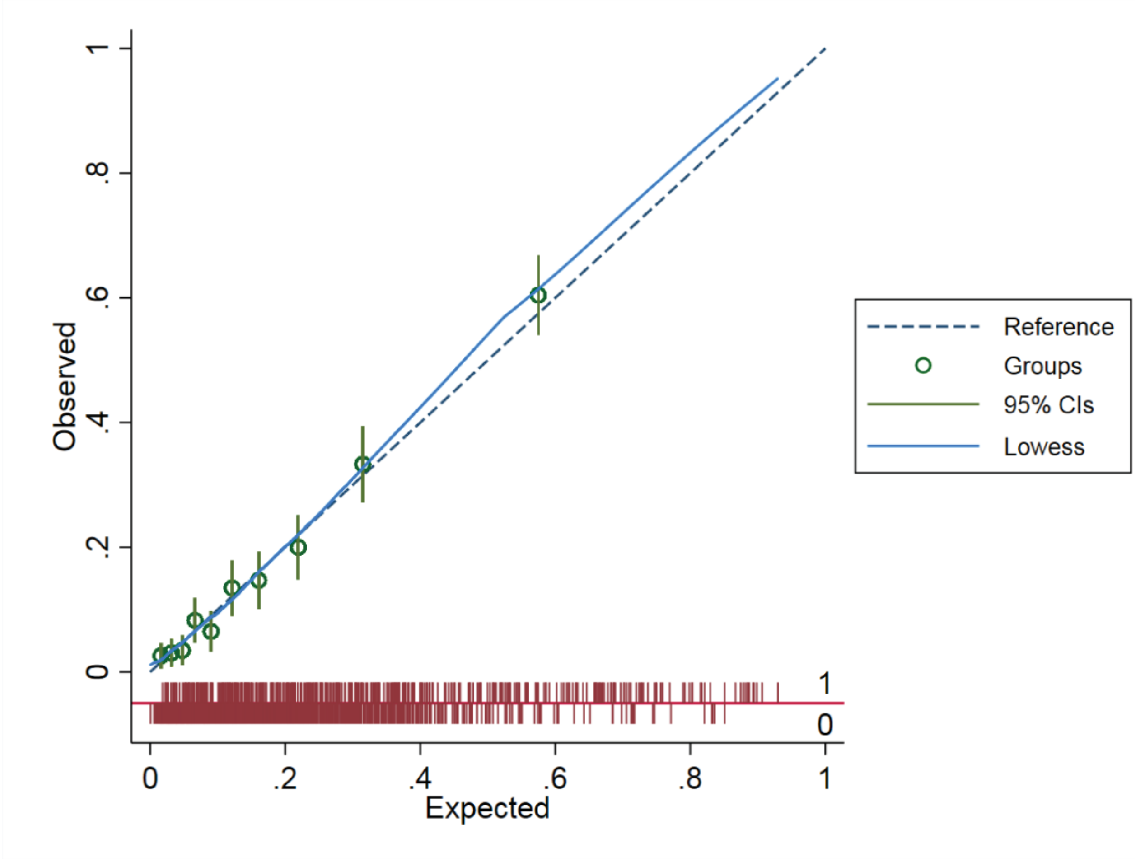
[2]: Determined by developing model in each bootstrapped sample (100 samples with replacement), calculating performance (bootstrap performance), and applying bootstrap model in original imputed dataset.

[3]: Average difference between model performance in bootstrap data and original imputation data

[4]: Subtracting optimism from apparent performance

[5]: Probability that for any randomly selected pair of patients with diagnosed COPD with and without respiratory hospitalisation, the patient with respiratory hospitalisation had higher predicted risk. A value of 0.5 represents no discrimination and 1.00 represents perfect discrimination.

Figure S1: Assessing calibration in original data for respiratory hospitalisation over full study period



Red lines indicate individual respiratory admission events

APPENDIX 1

List of ICD10 codes used to define respiratory hospital admissions

J00-06, J09-18, J20-22, J39.3, J39.8, J39.9, J40-47, J60-70, J80-86, J90-98, R05, R06.0, R06.2, R06.5, R09.2, R09.3