

APPENDIX I ADDITIONAL METHODS

Trial design

The Accurate trial lasted from September 2009 until January 2012 and was designed to compare patients' preferences and costs for asthma treatment regimens targeting three levels of asthma control. In 131 rural and urban general practices in the regions of Leiden, Nijmegen and Amsterdam, patients aged 18-50 years with a doctor's diagnosis of asthma were recruited. They were cluster-randomised (at the level of practice) to one of the three interventions: treatment targeted at:

- 1) Controlled asthma: ACQ-score ≤ 0.75
- 2) Partly controlled asthma: ACQ-score between 0.75 and 1.50
- 3) FeNO controlled asthma: ACQ-score ≤ 0.75 and FeNO level ≤ 25 ppb

After obtaining informed consent during information sessions, a blood sample was collected for an allergy test and baseline FeNO level was measured.

During the twelve months follow up, each three months patients visited their practice where their treatment was adjusted (after assessing ACQ-score and spirometry) by practice nurse, assisted by an online treatment algorithm. Close to the date of each practice visit, patients filled out online questionnaires inquiring after symptoms, quality of life, medication adherence and healthcare utilisation. During the study, 41 (6.7%) patients withdrew, and 6 (1.0%) were lost to follow-up; mean intervention duration was 325 days.

Although statistically not significant, there were numerical differences in the annual exacerbation rates: 0.29 (95%CI 0.17 to 0.40) for the controlled asthma strategy (which aimed at an ACQ ≤ 0.75), 0.29 (95%CI 0.15 to 0.43) for the partly controlled asthma strategy

(which aimed at an ACQ of 0.75-1.5) and 0.19 (95%CI 0.11 to 0.29) for the FeNO strategy group (which aimed at an ACQ \leq 0.75 and FENO \leq 25ppb).

The trial was approved by the Medical Ethics Committee of the Leiden University Medical Center and was registered at www.trialregister.nl (NTR 1756). Details of the study protocol have been published elsewhere.¹

Allergy test

At baseline, a venous blood sample was collected from all patients, in which total immunoglobulin E (IgE) and specific IgE against house dust mite, cat, grass- and birch and pollens were determined by radioallergosorbent test (RAST). Total IgE was regarded as positive when values exceeded 100 kU/l, while specific IgE was regarded as positive when exceeding 0.35kU/l.

Spirometry

At baseline a pre- and post-bronchodilator spirometry was performed in each patient at his or her practice. General practitioners and nurse practitioners were trained to perform spirometry according to the guidelines.²

Fractionated exhaled Nitric Oxide (FeNO)

FeNO was measured in all patients according to international guidelines³ at baseline using a NIOX MINO® meter (Aerocrine, Solna, Sweden), with NO values expressed as part per billion (ppb). Values were corrected for smoking status⁴ according to the manufacturer's recommendations.

Statistical analysis

Missing values were multiply imputed by iterative chained equations (ICE), using all predicting covariates except questionnaire data; ten imputed datasets were generated.⁵ In each imputed dataset, backward variable selection was used in 1,000 bootstrapped samples.

Predictors were selected using bootstrapped backward selection (BBS method) with a significance level of 0.157 (Akaike information criterion; AIC), according to Royston and Sauerbrei.⁶ The BIF is defined as the proportion (or percentage) of bootstrap replications in which the variable is selected by fractional polynomial multivariable modelling and may be used as a criterion for the importance of a variable. If the mean bootstrap inclusion fraction (BIF) of a variable was more than 667/1,000 across the ten datasets, the variable was selected as a predictor. Over-optimism was corrected by multiplying the coefficients of selected predictors by a fixed heuristic shrinkage factor which was calculated as follows:⁷

Shrinkage factor = m / D , with m = number of covariates and

D = deviance = null deviance - residual deviance = $-2 \times LL_0 - -2 \times LL_{\text{model}}$, where LL indicates loglikelihood and LL_0 the loglikelihood of an intercept-only or comparison (history only) model.

Net reclassification improvement (NRI)

Addition of spirometry and subsequently FeNO to the history model was assessed by Net Reclassification Improvement (NRI), at cut points of 25 and 50% risk of exacerbation in the next 12 months, which were chosen arbitrarily as the risk within the highest decile in the derivation population (25%) and double this risk (50%). These could reflect clinical decision points at which a practitioner might intervene to reduce the risk of exacerbations, although it should be noted that there are no studies yet demonstrating the predicted risk of exacerbation at which an increase of treatment step (or other intervention) has a net benefit in terms of safety, costs and patient preferences.

External validation

External validation of the models was performed in the U-BIOPRED⁸ (Unbiased Biomarkers for the Prediction of Respiratory Disease Outcomes) adult cohort, a multicentre prospective

study, in which adult patients with asthma were recruited in 16 clinical centres in 11 European countries from 2010 until 2014. Three groups within the cohort (severe non-smoking asthma (n=307), smokers and ex-smokers with severe asthma (n=109) and non-smoking patients with mild/moderate asthma (n=88) were merged (Table A5); healthy controls in this study were not used for validation. Outcomes assessed in the validation study were courses of oral corticosteroids, regardless of the place they were prescribed (general practice, emergency department or hospital), as well as increased dose of corticosteroids; both events were only counted if lasted at least three days. Data on the seven predictors and the outcome were complete in 128 patients (25.1%); 309 patients (60.7%) missed 1 value; 54 (10.6%) missed two values and three patients (0.6%) three values. Data were imputed similar to the validation dataset (using ICE, generating 10 datasets); 10 patients missed all baseline variables and were excluded. Overall, 34.9% patients in this population had one or more courses of oral corticosteroids (i.e. a severe exacerbation) during follow-up. The full models were validated (not the risk scores); models were not recalibrated.

References

- 1 Honkoop PJ, Loymans RJ, Termeer EH, et al. Asthma control cost-utility randomized trial evaluation (ACCURATE): the goals of asthma treatment. *BMC Pulm Med*. 2011;11:53.
- 2 Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardisation of spirometry. *Eur Respir J* 2005;26:319-38.
- 3 ATS/ERS recommendations for standardized procedures for the online and offline measurement of exhaled lower respiratory nitric oxide and nasal nitric oxide, 2005. *Am J Respir Crit Care Med*. 2005;171(8):912-30.

- 4 Dressel H, de la Motte D, Reichert J, et al. Exhaled nitric oxide: independent effects of atopy, smoking, respiratory tract infection, gender and height. *Respir Med*. 2008;102(7):962-9.
- 5 White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;30(4):377-99.
- 6 Royston P, Sauerbrei W. *Multivariable Model-building: a pragmatic approach to regression analysis based on fractional polynomials for modelling continuous variables*. West Sussex, UK: John Wiley and Sons, 2008.
- 7 Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JD. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. *Stat Med* 2004;23(16):2567-86.
- 8 Shaw DE, Sousa AR, Fowler SJ et al. Clinical and inflammatory characteristics of the European U-BIOPRED adult severe asthma cohort. *Eur Respir J*. 2015 Nov;46(5):1308-21.

Table A1: mean bootstrap inclusion fraction (BIF) across 10 imputed data sets; variables exceeding a BIF of 66.7 (bold), were included in the model.

Variables	Mean BIF (%)
Asthma Control Questionnaire	76.0
Excessive SABA use	37.7
Inadequate ICS: not prescribed	61.4
Inadequate ICS: poor adherence	24.3
Antidepressants use	16.8
Income	49.0
Current smoking	82.7
Allergen exposure if sensitised	21.9
Body Mass Index	4.8
Chronic sinusitis	69.9
Self-reported food allergy	26.7
Ever hospitalisation for asthma	76.1
Systemic corticosteroids previous year	100.0
selected predictors (bold) extended with:	
FEV1	87.9
FEV1 + FeNO	78.7

Table A2: discrimination, expressed as the 10th, 50th and 90th centiles of the distributions of predicted probabilities of the three models.

	p10	p50	p90
history	3.9	6.9	24.1
history + spirometry	3.9	8.2	28.5
history + spirometry + FeNO	3.8	8.1	29.6

A p50 of 6.9 indicates that 50% of the patients had a risk of $\leq 6.9\%$ of experiencing a severe exacerbation in the following 12 months.

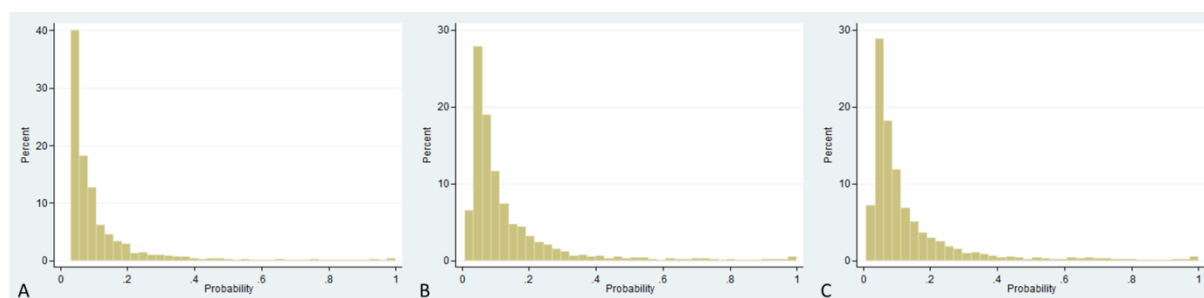


Figure A1: Percentage of the patients (Y-axis) versus the predicted probabilities (x-axis) for models based on history (A), history + spirometry (B) and history + spirometry + FeNO (C).

Net reclassification improvement

Table A3: Net Reclassification Improvement tables for cut-off values 25 and 50%.

	adding spirometry		adding FeNO	
	<25%	>=25%	<25%	>=25%
<25%	43	4	41	3
>=25%	1	32	2	35
<25%	487	19	487	5
>=25%	5	20	5	34
	NRI _{events} 3.8% NRI _{non events} 2.8%		NRI _{events} 1.3% NRI _{non events} 0.0%	
	total NRI 0.011 p 0.71		total NRI 0.013 p 0.66	
	adding spirometry		adding FeNO	
	<50%	>=50%	<50%	>=50%
<50%	57	8	56	2
>=50%	0	15	1	22
<50%	526	3	526	0
>=50%	0	2	1	4
	NRI _{events} 10.0% NRI _{non events} 0.6%		NRI _{events} 1.3% NRI _{non events} -0.2%	
	total NRI 0.094 p 0.01		total NRI 0.014 p 0.51	

Spirometry was added to the history model, FeNO was added to the history + spirometry model. For the 25% risk cut-off, in the patients with exacerbations, 43 patients remained in the low (<25%) risk group after adding spirometry to the model. Four patients were reclassified correctly from the low to the high-risk (>=25%) group and one incorrectly from the high risk to the low risk group after adding spirometry; 32 patients remained unchanged in the high-risk group. In patients without events nineteen were reclassified incorrectly to the high-risk group and five correctly to the low risk group. Total NRI 0.011 shows the difference of the percentages of the patients correctly classified in the events and the non-events groups.

Risk score

Table A4: scores and predicted probabilities for the low, medium and high risk group for exacerbations in the population, for the history model.

score	probability	risk	exacerbation		total
			no	yes	
<= 3	<= 8.9	low	326	23	349
4 - 5	11.6 – 14.9	intermediate	118	12	130
>= 6	>= 18.9	high	87	45	132
total			531	80	611

External validation

Table A5: comparison of ACCURATE¹ and U-BIOPRED² cohorts

	ACCURATE (n = 611) (derivation)	U-BIOPRED* (n=504) (validation)
Sex (% female)	68.4	59.5
Mean age (yrs, SD; range)	39.4 (9.1; 17-55)	50.1 (14.4; 18-79)
BMI (kg/m ² ,SD; range)	26.4 (5.4; 13.0-56.8)	28.5 (6.1; 17.8-53.6)
Current smokers (%)	14.1	8.3
Previous smokers (%)	32.9	25.1
Chronic sinusitis (%)	11.0	35.3
ACQ-5 score (SD; range)	1.01 (0.95; 0.0-5.6)	1.99 (1.24; 0.0-5.8)
Hospitalized for asthma (%)	11.7	54.2
Exacerbation previous year (%)*	11.8	55.6
FEV1 (% predicted, SD; range)	91.3 (15.4; 36.8-147)	69.0 (22.2; 18.4-132.3)
FeNO (ppb, SD; range)	26 (28; 5-297)	37 (32.6; 2-207)
Total exacerbations during follow-up (%)	13.1	34.9

*Six patients in the severe non-smoking and 4 in the severe (ex)smoking group had no baseline data and were not used in the analysis. Data derived from Shaw *et al.*² Underlined items were predictors in the derivation model

References

- 1 Honkoop PJ, Loymans RJ, Termeer EH, et al. Asthma control cost-utility randomized trial evaluation (ACCURATE): the goals of asthma treatment. *BMC Pulm Med.* 2011;11:53.
- 2 Shaw DE, Sousa AR, Fowler SJ et al. Clinical and inflammatory characteristics of the European U-BIOPRED adult severe asthma cohort. *Eur Respir J.* 2015 Nov;46(5):1308-21.