Automatic Adjudication of Exacerbations in Respiratory Patients

M.C. Wheldon[1,2], A.C. Vandal[1,3], A. Bourien[4], L. Jayaram[5], N. Karalus[6], C. Tong[2], H. Hockey[7], C. Wong[3]

[1]Auckland University of Technology, [2]Middlemore Clinical Trials, [3]Counties Manukau Health, [4]ENSAI France, [5]Monash Health Sleep Lab, [6]Waikato District Health Board, [7]Biometrics Matters Ltd.

Summary

- Bronchiectasis patients move in and out of a state of exacerbation. There are two kinds: event-based exacerbations (EBEs) and symptom-based exacerbations (SBEs).
- Ascertainment of EBEs requires contact with a clinician. SBE status is determined by adjudication of patient-recorded symptom scores (ordinal, 0–4) for sputum volume, sputum purulence and dyspnoea.
- Daily symptom diaries kept by 140 bronchiectasis patients over a 6 month period studied by Wong et al. (2012) were manually adjudicated. EBEs and wellbeing (St. George's Resp. Q'aire) were also recorded.
- Manual adjudication of SBEs is labour intensive and based on a complicated scoring rule.
- A new definition of SBE is proposed based on a prediction rule validated against clinically adjudicated EBEs. The prediction rule is derived by regressing current EBE status on current and previous symptom scores and previous EBE status.

Discussion

Conclusions

- A new definition of SBE was proposed which can be automatically adjudicated with useful precision.
- SBE was no more associated with patient-reported wellbeing than clinically adjudicated EBE.
- Future work will investigate joint validation against patient-reported wellbeing and EBE.

Limitations

- Assigned equal loss for both types of misclassification; differential loss could improve predictive performance.
- Error in the predicted EBE state, $\mathrm{Var}\left(\widehat{\mathrm{EBE}}_{i,t}\right)$, was not propagated.

References

Bonney, G. E. (1987). Logistic regression for dependent binary observations. *Biometrics*, 43(4): 951–973.

Wong, C., Jayaram, L., Karalus, N., Eaton, T. Tong, C., Hockey, H., Milne, D., Fergusson, W., Tuffery, C., Sexton, P., Storey, L., and Ashton, T. (2012). Azithromycin for prevention of exacerbations in non-cystic fibrosis bronchiectasis (EMBRACE): a randomised, double-blind, placebo-controlled trial. *Lancet,* 380(9842): 660–7.

Acknowledgements

Presented at the IMS New Researchers' Conference, Univ. of Washington, Seattle, WA, August 2015.

Travel support came from an AUT Faculty of Health and Environmental Sciences Travel Award.

Correspondence to Mark C Wheldon, Dept. of Biostatistics & Epidemiology, Auckland University of Technology, Private Bag 92006, Auckland 1142, New Zealand. Email: mwheldon@aut.ac.nz.

Results

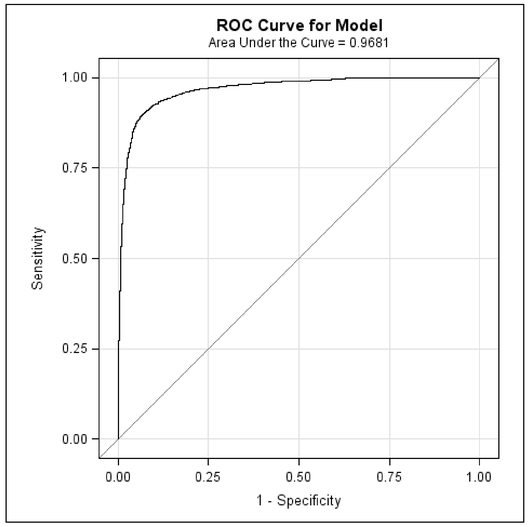

Figure. ROC curve for the *retrospective* model.
AUC = 0.97
Dichot'n threshold = 0.09
Sens. = spec. = 0.91
Event predictions from this model were included as fixed effects in the *prospective* model.

Retrospective Model Fit

- The best δ was 5 days. The best performing summarization scheme for the symptoms, $\boldsymbol{X}_{i,t}$, was an average of scores over two windows:
  - Current window: [-3, 0] days
  - Comparison window: [-8, -4] days.
- A very well-fitting retrospective model was found, AUC = 0.97, sens. = spec. = 0.91

$$\mathbb{E}\ \mathrm{logit}\ \Pr\left(\mathrm{EBE}_{i,t}|\ \cdot\right) = f\left(\boldsymbol{X}_{i,t\in[-3,0]}, \boldsymbol{X}_{i,t\in[-8,-4]}, \mathrm{EBE}_{i,t-5}\right)\boldsymbol{\beta}$$

EBE Prediction Performance

Dataset	*Model*	*t'hold (used)*	*t'hold (opt.)*	*Sens.*	*Spec.*
Training	Retro.	0.093	0.093	0.91	0.91
	Prosp.	0.093	0.048	0.76	0.88
	Prosp.	0.048	0.048	0.83	0.83
Hold-out	Prosp.	0.048		0.90	0.79

Validation Against Wellbeing

- Association between wellbeing and EBE was weak: sens. = 0.04, spec. = 0.83; most days of bad wellbeing are not on days with EBE.
- Association between wellbeing and our new definition of SBE was also weak: sens. = 0.64, spec. = 0.09; specificity increased at large expense to sensitivity.

Method

Overview

1. Build a *retrospective* prediction model for $\mathrm{EBE}(t_0)$ using observed symptom scores and observed EBEs at times $t \in (t_0 - \tau, t_0]$.
2. Convert to a *prospective* model for $\mathrm{EBE}(t_0)$ using observed symptom scores and $\widehat{\mathrm{EBE}}_t$ at times $t \in (t_0 - \tau, t_0]$.
3. Estimate predictive performance using cross-validation.

Retrospective Prediction Model

- EBE status at current time, t, dependent on current and past symptom scores *and* EBE status at $t - \delta$, $0 < \delta \leq \tau$.
- Used generalized linear mixed model with logit link, rand. intercepts for patient (a 'regressive logistic' with random effect, Bonney, 1987) to predict the time-ordered, clustered, binary outcome, EBE, estimated with ML:

$$\mathrm{logit}\ \Pr\left(\mathrm{EBE}_{i,t}|\ \cdot\right) = f\left(\boldsymbol{X}_{i,t}, \mathrm{EBE}_{i,t-\delta}\right)\boldsymbol{\beta} + \boldsymbol{Z}_i b_i + \varepsilon_{i,t}$$

- Entries $\{x_{i,t}\}$ in $\boldsymbol{X}_{i,t}$ are diarized symptom scores at times $t \in (t_0 - \tau, t_0]$, $f(\cdot)$ indicates an averaging scheme, $b_i \sim \mathrm{Normal}(0, \tau^2)$, $\varepsilon_{i,t} \sim \mathrm{Normal}(0, \sigma^2)$.
- Used area under the ROC curve to search among models defined by $\{f, \delta, \boldsymbol{X}_{i,t}\}$
- Binary prediction by dichotomizing at the threshold where sens. = spec.

Prospective Prediction Model

- Fitted prospective model by sequentially fitting the retrospective model with predicted EBE in place of observed:

$$\mathrm{logit}\ \Pr\left(\mathrm{EBE}_{i,t}|\ \cdot\right) = f\left(\boldsymbol{X}_{i,t}, \widehat{\mathrm{EBE}}_{i,t-\delta}\right)\boldsymbol{\beta} + \boldsymbol{Z}_i b_i + \varepsilon_{i,t}$$

- Re-estimated the dichotomization threshold.
- Two-fold cross-validation used to estimate predictive performance.

Wellbeing

- Assessed daily from St. George's Resp. Q'aire total score, dichotomized (1, 2 = "bad"; 3, 4, 5 = "good").
- Dichotomization threshold determined by regressing dichotomized wellbeing under each threshold on symptom scores, and selecting that which led to a model with best prediction performance.