Automatic Adjudication of Symptom-Based Exacerbations in Bronchiectasis Patients Treated With Azithromycin

Joint Statistical Meetings, Seattle, WA, 12th August 2015

Mark C. Wheldon[1,2]

Alain C. Vandal[1,3], Anne-Cécile Bourien[4], Lata Jayaram[5], Noel Karalus[6], Cecilia Tong[2], Hans Hockey[7], and Conroy 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.

Middlemore Clinical Trials
Future of healthcare

AUT UNIVERSITY HEALTH+ENVIRONMENTAL SCIENCES

EMBRACE Trial

This work is an extension to the trial: Wong, C., et al. *Lancet* (2012) vol. 380, pp. 660–7

Articles

Azithromycin for prevention of exacerbations in non-cystic fibrosis bronchiectasis (EMBRACE): a randomised, double-blind, placebo-controlled trial

Conroy Wong, Lata Jayaram, Noel Karalus, Tam Eaton, Cecilia Tong, Hans Hockey, David Milne, Wendy Fergusson, Christine Tuffery, Paul Sexton, Louanne Storey, Toni Ashton

Lancet 2012; 380: 660–67

See Comment page 627

Department of Respiratory Medicine (C Wong FRACP), Centre for Clinical Research and effective practice (C Tong RN), and Department of Microbiology (L Storey BSc), Middlemore Hospital, Counties Manukau District Health Board, Auckland, New Zealand; Department of Respiratory Medicine, Monash Medical Centre, Clayton, VIC, Australia (L Jayaram FRACP); Department of Respiratory Medicine, Waikato Hospital, Waikato District Health Board, Hamilton, New Zealand (N Karalus FRACP, C Tuffery BN); 85 Totara Road, Whenuapai, Auckland,

Summary

Background Azithromycin is a macrolide antibiotic with anti-inflammatory and immunomodulatory properties. We tested the hypothesis that azithromycin would decrease the frequency of exacerbations, increase lung function, and improve health-related quality of life in patients with non-cystic fibrosis bronchiectasis.

Methods We undertook a randomised, double-blind, placebo-controlled trial at three centres in New Zealand. Between Feb 12, 2008, and Oct 15, 2009, we enrolled patients who were 18 years or older, had had at least one pulmonary exacerbation requiring antibiotic treatment in the past year, and had a diagnosis of bronchiectasis defined by high-resolution CT scan. We randomly assigned patients to receive 500 mg azithromycin or placebo three times a week for 6 months in a 1:1 ratio, with a permuted block size of six and sequential assignment stratified by centre. Participants, research assistants, and investigators were masked to treatment allocation. The coprimary endpoints were rate of event-based exacerbations in the 6-month treatment period, change in forced expiratory volume in 1 s (FEV_1) before bronchodilation, and change in total score on St George's respiratory questionnaire (SGRQ). Analyses were by intention to treat. This study is registered with the Australian New Zealand Clinical Trials Registry, number ACTRN12607000641493.

Findings 71 patients were in the azithromycin group and 70 in the placebo group. The rate of event-based exacerbations was 0·59 per patient in the azithromycin group and 1·57 per patient in the placebo group in the 6-month treatment period (rate ratio 0·38, 95% CI 0·26–0·54; p<0·0001). Prebronchodilator FEV_1 did not change from baseline in the

EMBRACE Design

Objectives	Assess the effect of *azithromycin* on health-related quality of life and lung function in patients 18–80 years with bronchiectasis (diagnosed by CT scan).
Design	Multicenter (3), double-blind, placebo-controlled, parallel group (1:1), 141 pts total.
Intervention	500mg azithromycin capsule vs. placebo, 3 days per week, for six months.
1° Endpoints	i) Rate of Event Based Exacerbations (EBEs) over 6 mo. treatment period; ii) Change in St. George's Respiratory Questionnaire (tot. score); (+ *others*).
2° Endpoints	Symptom scores for: sputum purulence, sputum volume, dyspnoea; (+ *others*).

Exacerbations

- Patients exacerbations and symptom scores recorded prospectively in patient diaries.
- Each patient-day judged *exacerbation* or *no exacerbation*.
- Key symptoms of an exacerbation are
  - Sputum volume
  - Sputum purulence (colour)
  - Dyspnoea (shortness of breath, coughing).
- Two types: Event-based (EBE) and Symptom-based (SBE).
  - Ascertainment of EBE requires contact with clinician.
  - SBE is determined from patient diary data.

EBE Incidence

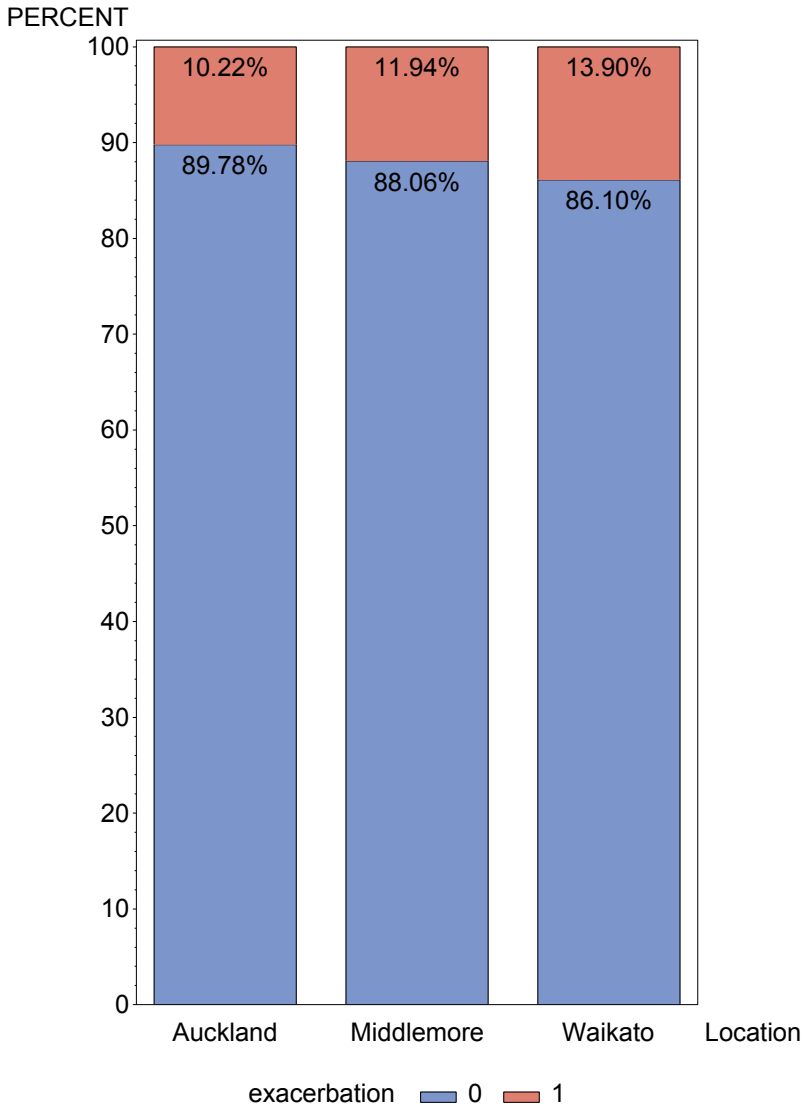

Figure : Diarized days by Event-based exacerbations status for each EMBRACE location, all patient-days.

Goal

Automatic adjudication of symptom-based exacerbations

New Definition of SBEs

- Adjudication of SBEs originally done by manual review of diaries.
- Automation of SBE adjudication presented opportunity to revisit definition.
- New definition of SBE to be based on a prediction rule validated against the clinically adjudicated EBEs.

Validated Prediction Rule

- Build a regression model for EBE_t using symptom scores and EBE at times $t \in [t - \delta, t_0)$.

Data

- Observed $\sim$ 50 000 patient-days observed on 141 patients across 3 centres over 6 months.
- Patients rated severity of
  - Sputum purulence, sputum volume, dyspnoea

  on a validated 5-point scale, 0 “no symptom” $\rightarrow$ 4 “very much”.

Data

- Observed $\sim$ 50 000 patient-days observed on 141 patients across 3 centres over 6 months.
- Patients rated severity of
  - Sputum purulence, sputum volume, dyspnoea

on a validated 5-point scale, 0 "no symptom" $\rightarrow$ 4 "very much".

Table : Example data

			Symptoms			
Pat.	Day	EBE	SP	SV	DY	SBE
1	0	0	1	1	1	$\uparrow$
1	1	1	3	3	4	$\uparrow$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	Adjudicated
1	17	1	2	1	5	$\downarrow$
1	18	0	1	0	0	$\downarrow$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\downarrow$

Goal is a Model for Prediction

- Statistical goal is a model to predict a time-ordered, clustered, binary outcome, $\text{EBE}_{i,t}$.
- Selected GLM, logit link, random intercepts for patient.

$$\text{logit}\,\Pr(\text{EBE}_{i,t} = 1|\cdot) = \mathbf{x}'_{i,t}\boldsymbol{\beta} + z_i b_i + \epsilon_{i,t}$$
$$b_i \sim \text{Normal}(0, \tau^2) \perp\!\!\!\perp \epsilon_{i,t} \sim \text{Normal}(0, \sigma^2)$$

(columns of $\mathbf{X}$ are symptom scores and $\text{EBE}_{i,t-\delta}$)

$$\widehat{\text{EBE}}_{i,t} = \begin{cases} 1 & \text{if } \widehat{\Pr}\big(\text{EBE}_{i,t} \,|\, \cdot\big) > c \\ 0 & \text{if } \widehat{\Pr}\big(\text{EBE}_{i,t} \,|\, \cdot\big) \leq c \end{cases}$$

(c chosen such that sensitivity $=$ specificity)

- Design parameters are: $\mathbf{X}$, δ, c.

Method Overview

1. Build a "retrospective" prediction model for EBE_t using
   - symptom scores
   - observed EBE status at times $t \in [t - \delta, t_0)$.
2. Convert to a "prospective" model for EBE_t using
   - retrospective design
   - predicted EBE status at times $t \in [t - \delta, t_0)$.
3. Estimate its predictive performance.

Design

- Fixed effect design matrix, $\mathbf{X}$, contains
  - patient-specific symptom scores at a *contemporaneous* time interval, $t \in [-a, t_0]$.
  - patient-specific symptom scores at an *earlier* time interval, $t \in [-c, -b]$, $a < b < c$.
  - EBE status at an earlier time point, $\text{EBE}_{i,t-\delta}$.
- For any given choice of a, b, c, symptom scores can be averaged, or not, over the intervals.

Design

- 72 combinations of δ, a, b, c, and averaging schemes were defined based on existing "by-hand" adjudication methods.

$$\underbrace{2}_{\substack{\text{contemporaneous}\\ \text{window}}} \times \underbrace{4}_{\substack{\text{earlier}\\ \text{window}}} \times \underbrace{3}_{\delta} \times \underbrace{3}_{\text{avg. schemes}} = 72 \text{ models}$$

Design

- 72 combinations of δ, a, b, c, and averaging schemes were defined based on existing "by-hand" adjudication methods.

$$\underbrace{2}_{\substack{\text{contemporaneous}\\ \text{window}}} \times \underbrace{4}_{\substack{\text{earlier}\\ \text{window}}} \times \underbrace{3}_{\delta} \times \underbrace{3}_{\text{avg. schemes}} = 72 \text{ models}$$

- In-sample predictive performance of each compared using AUC.

Table : Best designs within each avg. scheme.

Avg. Scheme	Contemp. Window	Comp. Window	δ	AUC
1	$[-3, 0]$	$[-11, -7]$	5	0.84
2	$[-3, 0]$	$[-11, -7]$	10	0.85
3	$[-3, 0]$	$[-8, -4]$	5	0.97

Selected Model

- Averaging scheme consists in arithmetic average of symptom scores over the interval.
- Includes all two-way interactions between $\text{EBE}_{i,t-5}$ and averaged symptom scores.

$$
\begin{aligned}
\eta_{i,t} &\equiv \text{logit}\,\Pr(\text{EBE}_{i,t} = 1 | \text{EBE}_{i,t-5}, \mathbf{X}_{i,t}, b_i) \\
&= \beta_0 + \text{EBE}_{i,t-5} \times \\
&\qquad \left(\overline{\text{SV}}_{i,\text{cont.}} + \overline{\text{SP}}_{i,\text{cont.}} + \overline{\text{DY}}_{i,\text{cont.}} + \overline{\text{SV}}_{i,\text{earl.}} + \overline{\text{SP}}_{i,\text{earl.}} + \overline{\text{DY}}_{i,\text{earl.}} \right) \\
&\quad + b_i + \epsilon_{i,t}
\end{aligned}
$$

$$
\eta_{i,t} \in (-\infty, \infty), \quad \text{EBE}_{i,t} \in \{0, 1\}, \ t = 0, 1, \ldots, T,
$$

$$
b_i \sim \text{Normal}(0, \tau^2) \perp\!\!\!\perp \epsilon_{i,t} \sim \text{Normal}(0, \sigma^2)
$$

Prospective Prediction

- The selected model is retrospective in that today's prediction depends on earlier observed EBEs.

$$\text{logit}\,\Pr(\text{EBE}_{i,t} = 1 \mid \text{EBE}_{i,t-5}, \cdot) = \mathbf{x}_i'\beta + z_i b_i + \epsilon_{i,t}$$

- We want a prospective model that uses earlier predictions to make today's prediction.

$$\text{logit}\,\Pr(\text{EBE}_{i,t} = 1 \mid \widehat{\text{EBE}}_{i,t-5}, \cdot) = \mathbf{w}_i'\beta + z_i b_i + \epsilon_{i,t}$$

A Model for Prediction

Goal

- Recall that our goal is a model for prediction.
- $\Rightarrow$ propose a model (somehow!).
- Verify it has good predictive power.

Two-fold Cross-validation

- Split data into a training set and a hold-out set for validation.
- Randomly select 70 percent of the patients and allocate all their observations to the training set.
- The remainder go into the hold-out set.

Sequential Approach

1. Initialize by generating retrospective predictions, $\widehat{\text{EBE}}_{i,1}^{[\text{ret}]}, \ldots,$ $\widehat{\text{EBE}}_{i,5}^{[\text{ret}]}$, using $\widehat{\beta}$ and threshold, $c^{[\text{ret}]}$, from the retrospective model.

Using the training set:

2. Sequentially generate prospective predictions $\widehat{\text{EBE}}_{i,6}^{[\text{pro}]}, \widehat{\text{EBE}}_{i,7}^{[\text{pro}]}, \ldots.$
   - Use "population level" predictions
   - Use $c^{[ret]}$ to threshold the predicted probabilities (we have to because this is the only c currently available).
3. Re-estimate the binary threshold, $c^{[pros]}$, using $\widehat{\text{EBE}}_{i,t}^{[\text{pro}]}$.

Using the hold-out set:

4. Repeat 2 using $c^{[pros]}$.

Results

Dataset	Model	c (used)	c (Opt.)	Sens. (%)	Spec. (%)
Training	Retro.	0.093	0.093	90	92

Results

Dataset	Model	c (used)	c (Opt.)	Sens. (%)	Spec. (%)
Training	Retro.	0.093	0.093	90	92
	Prosp.	0.093	0.048	76	88

Results

Dataset	Model	c (used)	c (Opt.)	Sens. (%)	Spec. (%)
Training	Retro.	0.093	0.093	90	92
	Prosp.	0.093	0.048	76	88
	Prosp.	0.048	0.048	83	83

Results

Dataset	Model	c (used)	c (Opt.)	Sens. (%)	Spec. (%)
Training	Retro.	0.093	0.093	90	92
	Prosp.	0.093	0.048	76	88
	Prosp.	0.048	0.048	83	83
Hold-out	Prosp.	0.048	—	90	79

- Our prospective SBE predictor,
  - misses 10 percent of the EBEs (1 in 10)
  - calls an EBE 21 percent of time there isn't one (1 in 5).
- Relative to using the whole dataset,
  - Estimated sensitivity equal,
  - Specificity is 86 percent ($\downarrow$ 13 percentage points).

Patient-reported Wellbeing

- Ultimately, we're interested in patient wellbeing, and a patient-centred measure that is sensitive to changes in physical state.
- Patients also reported wellbeing each day using a 1–5 scale (SGRQ).
- How is our new definition of SBE associated with wellbeing?
- Do we get "closer" to wellbeing with SBE rel. to EBE?

EBE, SBE, and Wellbeing

- Consider EBE and SBE as "tests" for wellbeing,
  - i.e., sens. $= \Pr(\text{EBE} = 1 \,|\, \text{WB} \leq 2)$.
- Wellbeing is dichotomized between 2 and 3.

bad $\equiv 1$		good $\equiv 0$		
1	2	3	4	5

(so indicator of "bad" wellbeing corresponds to indicator of "bad" EBE/SBE)

Wellbeing and EBE

EBE	wellbeing_leq2		
Frequency / Percent / Col Pct	0	1	Total
0	24274 52.36 83.44	16579 35.76 96.03	40853 88.13
1	4818 10.39 16.56	685 1.48 3.97	5503 11.87
Total	29092 62.76	17264 37.24	46356 100.00

- Spec. = 83%.
- Sens. = 4%.
- Most (96%) episodes of poor wellbeing do not correspond to an EBE.

Wellbeing and SBE

SBE	wellbeing_leq2		
Frequency			
Percent			
Col Pct	0	1	Total
0	17414	14567	31981
	40.18	33.61	73.79
	63.51	91.48	
1	10006	1356	11362
	23.09	3.13	26.21
	36.49	8.52	
Total	27420	15923	43343
	63.26	36.74	100.00

- Well . . . we've doubled the proportion of bad wellbeing days captured!
- Spec. $\downarrow$ 20 percentage points.
- Work in progress!

Summary

- Exacerbations are an outcome of interest in the study of bronchiectasis.
- Ascertainment of event-based exacerbations (EBEs) requires clinical assessment.
- Symptom-based exacerbations (SBEs) are ascertained from patient-reported symptom scores and exacerbation history, coded "by hand".
- We used logistic regression to develop an "automatic" coding scheme; changes in symptoms that are associated with changes in physical state (EBE).
- As a classifier of EBE the performance was quite good (sens. 90%, spec. 79%).
- Unlcear we moved closer to patient-reported wellbeing (SGRQ) ...to be continued.

Bibliography

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. *The Lancet*, 380(9842):660–7.