ONLINE SUPPLEMENT FOR

Inhaled antibiotics improve symptoms of cough and sputum in patients with bronchiectasis: a post-hoc analysis of the AIR-BX studies

Megan L Crichton[1], Mike Lonergan[1], Alan F Barker[2], Oriol Sibila[3], Pieter Goeminne[4], Amelia Shoemark[1], James D Chalmers[1]

1. Division of Molecular and Clinical Medicine, University of Dundee, Ninewells Hospital and Medical School, Dundee, Scotland, UK;
2. Oregon Health & Science University, Portland, OR, USA
3. Department of Respiratory Medicine, Hospital Clinic, Barcelona, Spain.
4. Department of Respiratory Medicine, AZ Nikolaas, Sint-Niklaas, Belgium

Corresponding author: Professor James D Chalmers, University of Dundee, Ninewells Hospital and Medical School, Dundee, DD1 9SY. jchalmers@dundee.ac.uk. Telephone 01382 6600111

Trial	QOL-B question	Difference (95% CI)	p-value	% change in patient response (95% CI)
AIRBX1	29- Congestion	0.04 (-0.21 to 0.29)	0.74	4.2 (-23.2 to 25.5)
	30- Cough	0.09 (-0.14 to 0.32)	0.43	8.8 (-14.5 to 27.4)
	31- Sputum production	0.22 (-0.01 to 0.46)	0.06	20.1 (-1.2 to 36.9)
	32- Sputum colour	0.18 (-0.03 to 0.39)	0.10	16.2 (-3.4 to 32.1)
	33- Shortness of breath	-0.01 (-0.24 to 0.22)	0.94	-0.9 (-26.4 to 19.5)
	34- Wheezing	-0.04 (-0.25 to 0.17)	0.68	-4.5 (-28.8 to 15.3)
	35- Chest pain	-0.07 (-0.19 to 0.06)	0.30	-6.8 (-21.1 to 5.8)
	36- SOB when talking	-0.05 (-0.27 to 0.17)	0.67	-4.9 (-30.7 to 15.9)
	37- nocturnal cough	0.14 (-0.10 to 0.37)	0.26	12.6 (-10.3 to 30.7)
AIRBX2	29- Congestion	0.20 (-0.01 to 0.41)	0.07	17.8 (-1.2 to 33.3)
	30- Cough	**0.32 (0.13 to 0.50)**	**0.001**	27.0 (12 to 39.5)
	31- Sputum production	**0.35 (0.18 to 0.53)**	**<0.0001**	29.7 (16.1 to 41.1)
	32- Sputum colour	**0.37 (0.18 to 0.55)**	**<0.0001**	30.6 (16.7 to 42.1)
	33- Shortness of breath	-0.05 (-0.23 to 0.14)	0.62	-4.8 (-26.0 to 12.8)
	34- Wheezing	-0.08 (-0.26 to 0.09)	0.35	-8.7 (-29.3 to 8.6)
	35- Chest pain	-0.02 (-0.15 to 0.11)	0.77	-2.0 (-16.2 to 10.5)
	36- SOB when talking	0.06 (-0.10 to 0.21)	0.48	5.4 (-10.2 to 18.8)
	37- nocturnal cough	-0.03 (-0.20 to 0.14)	0.75	-2.8 (-21.9 to 13.2)

Table S1 – Treatment response for individual symptoms in AIRBX 1 & 2. Statistically significant effects are highlighted in bold.

QOLB-RSS item	AIRBX1	AIRBX2	Pooled (improved MCID)	Pooled (worsening MCID)
Q29 – Congestion	0.499	**0.035**	**0.015**	0.171
Q30 – Daily cough	-	-	0.074	0.659
Q31 – Sputum production	0.053	**0.02**	**0.003**	0.143
Q32 – Sputum purulence	0.864	**0.004**	**0.01**	0.392
Q33 – Breathlessness on daily activity	-	-	0.687	0.126
Q34 – Wheeze	-	-	0.929	0.16
Q35 – Chest pain	-	-	0.657	0.334
Q36 – Breathlessness on talking	-	-	0.363	0.073
Q37 – Nocturnal cough	-	-	0.327	0.27

Table S2. Chi-squared test determining treatment effect with a MCID response following 4 weeks of treatment. Where a statistically significant effect was observed in the pooled data the analysis was repeated for each trial separately.

Table S3- Trajectories for answers to individual questions.

Models were separately fitted to the answers to each question. As there are only four possible answers to each question, uncertainties around these models' estimates were generated by bootstrap resampling the data and a pragmatic and empirical approach taken to model selection. The data from the two trials was pooled for simplicity and because preliminary analyses identified no differences between them. Each model contained an intercept term and a random effect for individual. The full model for each question also contained four terms: short (visit 3,4,6) and long (study duration) term placebo effects, affecting both arms equally, and equivalent drug effects that only affected the people on AZLI. 15 sub-models were fitted, containing each combination of the four treatment terms. Terms were considered statistically significant where their 95% confidence interval from bootstrapping did not cross zero. The model for each question, among those for which all parameters were statistically significant, that contained the largest number of parameters was considered best. Where this produced multiple candidates, parameters that were significant in the full model, or failing that, those with the largest absolute parameter value (effect size), were favoured.

			Placebo effect, affecting both arms equally		Drug effect on AZLI population only	
Q	model	Intercept	short term (visit 3,4,6)	long term (study duration)	short term (visit 3,4,6)	long term (study duration)
29	full	**1.40 (1.31, 1.48)**	0.06 (-0.03, 0.15)	**0.20 (0.10, 0.30)**	0.06 (-0.09, 0.21)	-0.06 (-0.19, 0.07)
	best	**1.40 (1.31, 1.48)**	**0.09 (0.02, 0.16)**	**0.17 (0.10, 0.25)**	-	-
30	full	**1.09 (1.02, 1.15)**	0.05 (-0.04, 0.13)	**0.16 (0.06, 0.26)**	0.06 (-0.07, 0.19)	**0.14 (0.02, 0.28)**
	best	**1.09 (1.02, 1.15)**	-	**0.19 (0.11, 0.27)**	**0.11 (0.01, 0.20)**	**0.11 (0.007, 0.23)**
31	full	**1.00 (0.94, 1.07)**	**0.09 (0.01, 0.17)**	**0.16 (0.07, 0.24)**	0.09 (-0.03, 0.21)	**0.14 (0.01, 0.26)**
	best	**1.00 (0.94, 1.07)**	**0.14 (0.08, 0.20)**	**0.13 (0.05,0.21)**	-	**0.19 (0.09, 0.29)**
32	full	**1.39 (1.33, 1.46)**	-0.01 (-0.09, 0.06)	**0.12 (0.04, 0.21)**	**0.27 (0.16, 0.39)**	-0.03 (-0.16, 0.09)
	best	**1.39 (1.33, 1.46)**	-	**0.11 (0.05, 0.16)**	**0.25 (0.17, 0.33)**	-
33	full	**1.48 (1.39, 1.57)**	0.06 (-0.01, 0.13)	**0.10 (0.02, 0.18)**	-0.07 (-0.17, 0.03)	-0.03 (-0.14, 0.08)
	best	**1.48 (1.39, 1.57**)	**0.07 (0.01, 0.13)**	**0.09 (0.02, 0.15)**	**-0.09 (-0.17, -0.002)**	-
34	full	**1.98 (1.90, 2.06)**	0.05 (-0.01, 0.12)	**0.11 (0.03, 0.19)**	-0.05 (-0.15, 0.05)	-0.03 (-0.13,0.08)
	best	**2.01 (1.93, 2.08)**	**0.11 (0.06, 0.17)**	-	**-0.11 (-0.20, -0.02)**	**0.08 (0.01, 0.15)**
35	full	**2.55 (2.48, 2.61)**	0.01 (-0.03, 0.06)	**0.09 (0.03, 0.16)**	-0.01 (-0.10, 0.07)	-0.03 (-0.11, 0.05)
	best	**2.55 (2.48, 2.61)**	-	**0.08 (0.04, 0.12)**	-	-
36	full	**2.05 (1.98, 2.13)**	0.05 (-0.01,0.11)	0.04 (-0.03, 0.11)	-0.03 (-0.13, 0.05)	-0.02 (-0.12, 0.08)
	best	**2.05 (1.98, 2.13)**	**0.05 (0.004, 0.10)**	-	-	-
37	full	**1.96 (1.88, 2.03)**	0.05 (-0.02,0.13)	0.08 (-0.01, 0.17)	0.03 (-0.09, 0.14)	-0.003 (-0.12, 0.11)
	best	**1.96 (1.88, 2.03)**	**0.07 (0.01,0.12)**	**0.07 (0.01, 0.14)**	-	-

Table S3: Models of answers to each question. Each cell contains the parameter estimate with the 95% confidence interval estimated by bootstrapping. Bold type indicates terms where the 95% CI does not include zero.

Table S4 and S5. Impact of baseline symptoms on overall QOL-B response.

Baseline symptom groups were associated with different responses in the overall QOL-B RSS scores from baseline to visit 4 as shown below in table S4.

Table S4 - ***Effect of baseline symptoms on changes to total QOL-B RSS at week 4***

	Overall effect on QOLB	
Symptom	Effect estimate (95% CI)	p-value
Q31- sputum production		
1 - 2 (severe)	4.82 (1.12 to 8.53)	**0.011**
3 - 4 (mild)	-2.61 (-7.58 to 2.37)	0.305
Q32 colour		
1 - 2 (severe)	5.02 (1.19 to 8.86)	**0.01**
3 - 4 (mild)	-0.78 (-5.88 to 4.33)	0.766
Q34 Wheezing		
1 - 2 (severe)	-0.66 (-8.06 to 6.74)	0.861
3 - 4 (mild)	3.74 (0.50 to 6.97)	**0.024**
Q36 SOB when talking		
1 - 2 (severe)	0.66 (-7.67 to 8.99)	0.876
3 - 4 (mild)	3.23 (0.12 to 6.33)	**0.042**
Q37 Woken due to cough		
1 - 2 (severe)	-0.33 (-7.01 to 6.35)	0.923
3 - 4 (mild)	3.62 (0.27 to 6.97)	**0.034**

Nine sets of 194 mixed models were fitted to the data from visits 2-7 of both trials together looking at the effect of each question separately. These models allowed a common drift over time for all individuals, and potentially different long and short term treatment effects on placebo and AZLI and an interaction between each of these and the answer to the symptom question at baseline. The estimates for the effects of the baseline questions are shown in table S4 below, the other parameter values are uninteresting. In each case the full model contains all the parameters, the best model is the one with the lowest AICc, and the average is a model averaged result, that uses the AICc weight for each model. This stage was necessary because there would be over 10^{15} models if each combination of all these terms for all questions were considered.

Q	model	effect of baseline question answer	interaction of baseline question answer with:			
			Placebo effect, affecting both arms equally		Drug effect on AZLI population only	
			long term (study duration)	short term (visit 3,4,6)	long term (study duration)	short term (visit 3,4,6)
29	full	12.86 (0.66)	-3.34 (0.74)	-1.23 (0.71)	0.41 (1.08)	-1.02 (1.09)
	best	12.86 (0.66)	-3.17 (0.58)	-1.67 (0.54)	-	-
	average	12.86 (0.66)	-3.23 (0.63)	-1.32 (0.78)	-0.46 (0.83)	0.02 (0.36)
	p	***<0.001***	***<0.001***	*0.09*	*0.58*	*0.95*
30	full	13.74 (0.87)	-3.95 (0.90)	-0.90 (0.85)	0.01 (1.33)	-1.25 (1.32)

	best	13.74 (0.87	-4.33 (0.62)	-	-	-1.92 (0.87)
	average	13.74 (0.87	-4.17 (0.74)	-0.66 (0.82)	-0.02 (0.44)	-0.97 (1.14)
	p	***<0.001***	***<0.001***	*0.42*	*0.96*	*0.39*
31	full	12.63 (0.97)	-3.52 (0.99)	-0.85 (0.94)	0.57 (1.43)	-2.66 (1.41)
	best	12.63 (0.97)	-3.67 (0.67)	-	-	-3.07 (0.92)
	average	12.63 (0.97)	-3.52 (0.78)	-0.55 (0.89)	0.07 (0.56)	-2.49 (1.39)
	p	***<0.001***	***<0.001***	*0.53*	*0.90*	*0.07*
32	full	6.90 (1.02)	-2.06 (0.98)	-2.23 (0.93)	-2.27 (1.41)	0.13 (1.38)
	best	6.90 (1.02)	-3.10 (0.74)	-2.17 (0.69)	-	-
	average	6.86 (1.03)	-2.63 (1.20)	-1.90 (1.04)	-0.98 (1.47)	-0.43 (1.08)
	p	***<0.001***	***0.03***	*0.68*	*0.51*	*0.69*
33	full	11.22 (0.65)	-1.24 (0.74)	-1.67 (0.71)	2.42 (1.04)	-1.88 (1.03)
	best	11.22 (0.65)	-2.44 (0.47)	-	-	-
	average	11.22 (0.65)	-2.07 (0.83)	-0.57 (0.83)	-0.39 (0.96)	0.66 (1.12)
	p	***<0.001***	***0.01***	*0.50*	*0.69*	*0.56*
34	full	12.59 (0.79)	-3.49 (0.83)	-0.60 (0.80)	-0.97 (1.21)	-0.24 (1.20)
	best	12.59 (0.80)	-4.34 (0.53)	-	-	-
	average	12.59 (0.80)	-4.05 (0.65)	-0.20 (0.47)	-0.16 (0.55)	-0.39 (0.73)
	p	***<0.001***	***<0.001***	*0.68*	*0.77*	*0.59*
35	full	12.46 (1.09)	-4.83 (1.19)	0.89 (1.13)	0.65 (1.57)	-1.05 (1.55)
	best	12.46 (1.09)	-4.29 (0.69)	-	-	-
	average	12.46 (1.09)	-4.33 (0.78)	0.11 (0.48)	0.01 (0.42)	-0.07 (0.55)
	p	***<0.001***	***<0.001***	*0.82*	*0.90*	*0.99*
36	full	11.92 (0.83)	-2.86 (0.89)	0.77 (0.85)	-0.68 (1.27)	0.74 (1.26)
	best	11.92 (0.83)	-3.17 (0.68)	1.11 (0.63)	-	-
	average	11.92 (0.83)	-2.88 (0.73)	0.56 (0.72)	-0.06 (0.44)	0.29 (0.71)
	p	***<0.001***	***<0.001***	*0.44*	*0.88*	*0.68*
37	full	12.80 (0.76)	-5.01 (0.83)	-0.56 (0.79)	2.40 (1.17)	-0.61 (1.16)
	best	12.80 (0.76)	-4.84 (0.76)	-0.84 (0.58)	2.02 (0.92)	-
	average	12.80 (0.76)	-4.74 (0.86)	-0.37 (0.61)	1.29 (1.35)	-0.15 (0.72)
	p	***<0.001***	***<0.001***	*0.54*	*0.34*	*0.83*
total score	full	0.93 (0.03)	-0.26 (0.04)	-0.06 (0.04)	-0.01 (0.06)	-0.02 (0.06)
	best	0.93 (0.03)	-0.26 (0.03)	-0.07 (0.03)	-	-
	average	0.93 (0.03)	-0.27 (0.03)	-0.05 (0.04)	-0.002 (0.017)	-0.02 (0.04)
	p	***<0.001***	***<0.001***	*0.23*	*0.92*	*0.63*

Table S5: Parameter estimates in models of the effect of each question's answer at visit 1 on QOLB-RSS scores at visits 2-7. Each full model contained the terms shown plus separate short and long term effects of treatment, allowing these to be different for placebo and AZLI. The data from the two trials was combined, and an intercept plus a single common drift term included along with individual as a random effect.

The 25 question-based parameters that had p-values less than 0.5 were put into a model along with the four treatment terms and the models nested within it refitted. At this stage, all the individual question main effects had values close to 3, so they were replaced with the total score at visit 1, and the model fitting was redone. Separate terms were included for the interaction between being on the trial and the answer to each question, as well as with the total score at baseline. The main effects for the baseline score and the short term (visit=3,4,6) and longer term (duration of study) placebo effects were required to be in all of the 98,304 models. Table S5 shows the results:

parameter		estimate (SE) in models:			p
		full (A)	best (B)	averaged	
	Intercept	4.75 (1.62)	4.76 (1.58)	4.88 (1.62)	**0.003**
	drift	0.015 (0.009)	0.015 (0.009)	0.009 (0.01)	0.35
	baseline score	0.93 (0.03)	0.93 (0.03)	0.93 (0.03)	**<0.0001**
on trial (study duration)	main (placebo) effect	21.12 (2.42)	20.18 (1.97)	20.5 (2.32)	**<0.0001**
	drug effect	-0.09 (0.98)	-	-0.02 (0.51)	0.97
	interaction with Q29	-0.12 (0.67)	-	-0.16 (0.45)	0.73
	interaction with Q30	-1.51 (0.63)	-1.63 (0.54)	-1.48 (0.69)	0.03
	interaction with Q31	-0.43 (0.68)	-	-0.18 (0.48)	0.71
	interaction with Q32	-3.42 (0.52)	-3.39 (0.48)	-3.37 (0.49)	**<0.0001**
	interaction with Q33	0.43 (0.59)	-	0.10 (0.31)	0.75
	interaction with Q34	-2.31 (0.55)	-2.11 (0.49)	-2.20 (0.52)	**<0.0001**
	interaction with Q35	-1.63 (0.63)	-1.53 (0.60)	-1.40 (0.74)	0.06
	interaction with Q36	-0.25 (0.57)	-	-0.02 (0.30)	0.94
	interaction with Q37	-1.33 (0.65)	-1.40 (0.48)	-1.33 (0.65)	**0.04**
being treated (visit 3,4,6-short term)	Main (placebo) effect	2.70 (2.17)	3.91 (0.94)	3.44 (1.57)	**0.03**
	drug effect	4.69 (1.39)	4.64 (1.14)	4.48 (1.37)	**0.001**
	interaction with total baseline score	0.04 (0.08)	-	0.003 (0.034)	0.93
	interaction with Q29	-1.70 (0.87)	-1.44 (0.45)	-1.37 (0.72)	0.06
	interaction with Q33	-0.25 (0.74)	-	0.05 (0.31)	0.88
	interaction with Q37	-0.28 (0.83)	-	-0.05 (0.43)	0.90
	interaction with Q31 and on AZLI	-2.31 (0.93)	-2.36 (0.83)	-2.18 (1.04)	**0.04**
	AICc	20799.0	20784.6		
	ΔAICc	+14.4	0		
	AICc weight	0.00002	0.03		

Table S6: parameters in models of total score that allow effects from answers to individual questions at visit 1. The baseline QOL-B score is out of 100, while each question is scored 1-4, complicating the comparison of their effect sizes.

The intercepts for the models of different questions differ, reflecting the pattern of baseline symptoms among the participants, but the placebo estimated effects are remarkably similar. This suggests that such reported improvements were distributed across the range of symptoms. There seems to be more variability in the reported effects of AZLI itself, with answers to some questions (30,31,32,34) being more affected than the remaining ones.

Table S7- Effect of baseline symptoms on changes in FEV_1 at week 4

	Overall effect on FEV_1	
Symptom	Effect estimate (95% CI)	p-value
Overall effect	-0.01 (-0.12 to 0.09)	0.85
Q29- congestion		
1	-0.113 (-0.32 to 0.09)	0.277
2	0.036 (-0.14 to 0.22)	0.696
3	0.035 (-0.19 to 0.26)	0.754
4	-0.096 (-0.28 to 0.08	0.296
Q30- cough		
1	-0.095 (-0.26 to 0.07)	0.266
2	-0.072 (-0.27 to 0.13)	0.483
3	0.083 (-0.09 to 0.25)	0.339
4	-0.147 (-0.46 to 0.17)	0.348
Q31- sputum production		
1	-0.156 (-0.33 to 0.02)	0.082
2	0.11 (-0.16 to 0.18)	0.902
3	0.114 (-0.10 to 0.33)	0.289
4	0.036 (-0.56 to 0.64)	0.906
Q32 colour		
1	-0.274 (-0.46 to 0.02)	0.07
2	0.108 (-0.05 to 0.27)	0.181
3	-0.085 (-0.28 to 0.11)	0.382
4	0.022 (-0.39 to 0.44)	0.916
Q33 daily activity		
1	0.038 (-0.18 to 0.25)	0.729
2	-0.049 (-0.25 to 0.16)	0.637
3	0.018 (-0.16 to 0.20)	0.842
4	-0.098 (-0.35 to 0.15)	0.437
Q34 Wheezing		
1	-0.029 (-0.51 to 0.45)	0.904
2	-0.066 (-0.25 to 0.12)	0.484
3	0.063 (-0.09 to 0.22)	0.419
4	-0.065 (-0.27 to 0.14)	0.532
Q35 chest pain		
1	-0.332 (-1.1 to 0.44)	0.396
2	0.164 (-0.08 to 0.41)	0.195
3	0.105 (-0.10 to 0.31)	0.327

4	-0.064 (-0.19 to 0.07)	0.331
Q36 SOB when talking		
1	0.085 (-0.30 to 0.47)	0.661
2	-0.018 (-0.28 to 0.25)	0.896
3	-0.032 (-0.19 to 0.13)	0.69
4	-0.004 (-0.19 to 0.18)	0.968
Q37 Woken due to cough		
1	0.338 (-0.01 to 0.68)	0.053
2	0.046 (-0.18 to 0.27)	0.69
3	0.03 (-0.16 to 0.16)	0.968
4	-0.124 (-0.32 to 0.07)	0.213

Table S7. No effect of AZLI on FEV_1 change from baseline to week 4 regardless of baseline symptoms. As no effect was seen for any individual symptoms no further analysis was performed.