Supplementary Digital Content

Pseudomonas aeruginosa eradication therapy and risk of acquiring *Aspergillus* in young children with cystic fibrosis

Sabariah Noor Harun,[1,2] Nicholas H. G. Holford,[2,3] Keith Grimwood,[4,5] Claire E. Wainwright,[6,7] Stefanie Hennig,[2] on behalf of the Australasian Cystic Fibrosis Bronchoalveolar Lavage (ACFBAL) study group

[1]School of Pharmaceutical Sciences, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia;

[2]School of Pharmacy, The University of Queensland, St Lucia, QLD 4072, Australia;

[3]Department of Pharmacology and Clinical Pharmacology, University of Auckland, Auckland, New Zealand.

[4]School of Medicine and Menzies Health Institute Queensland, Griffith University, Southport, QLD 4222, Australia;

[5]Departments of Infectious Diseases and Paediatrics, Gold Coast Health, Southport, QLD 4215, Australia;

[6]Department of Respiratory and Sleep Medicine, Queensland Children's Hospital, South Brisbane, QLD 4101, Australia;

[7]Child Health Research Centre, The University of Queensland, South Brisbane, QLD 4101, Australia.

ABBREVIATIONS

ACFBAL	Australasian Cystic Fibrosis Bronchoalveolar Lavage
BAL	bronchoalveolar lavage
CBA	chocolate bacitracin agar
CF	cystic fibrosis
HBA	horse blood agar
HR	hazard ratio
OFV	objective function value
RTTE	repeated time-to-event
VIF	variance inflation factor
VPC	visual predictive check

Methods

Culture method

In the Australasian Cystic Fibrosis Bronchoalveolar Lavage (ACFBAL) study, the *Pseudomonas aeruginosa* colony count in the bronchoalveolar lavage (BAL) fluid samples was determined by the spread plate method.[1] This involved 500µL of BAL fluid being serially diluted from 10^{-1} to 10^{-5} in sterile phosphate buffered saline for quantitative colony counts and 100µL of undiluted BAL fluid and 100µL from each of the serial dilutions were added to six different selective and non-selective media: (1) horse blood agar (HBA), (2) mannitol salt agar, (3) MacConkey agar, (4) chocolate bacitracin agar (CBA), (5) cetrimide or Pseudomonas agar and (6) *Burkholderia cepacia* agar. Plates were then incubated in air at 37^0C (in 5%CO_2 for CBA plates) and read at 24- and 48-hours. The colony count for BAL fluid was determined by choosing whichever plate contained between 30 and 100 individual colonies. Respiratory bacterial pathogens were identified by standard methods.

Five of eight centres also included Sabouraud dextrose agar with gentamicin routinely as selective media to isolate fungal organisms. The plates were incubated in air at 35^0C for up to 7-days for the primary isolation of moulds. When mould was found growing on either selective or non-selective media (e.g. HBA agar), *Aspergillus* species were identified by their colonial and microscopic appearance.

Cross-sectional analysis at age 5-years

Before conducting univariable and multivariable logistic regression, statistical tests and multicollinearity were examined using the IBM SPSS Statistics 22 package.[2] Multicollinearity between explanatory variables was tested using the variance inflation factor (VIF). If two variables showed a high correlation (VIF >3), the least significant variable was excluded from the model.

The objective function value (OFV) is calculated by NONMEM as a measure of how the model fitted the data. It is a numerical value computed by an extended least squares method and is proportional to minus twice the logarithm likelihood of the data (-2*ln* (likelihood)) given the model.[3] A lower OFV represents a better model to describe the data. Model explanatory variables were included in a stepwise approach and selected when the OFV decrease was significant (*P*-value <0.05). Inclusion ceased when adding a new variable did not result in further significant improvement. The explanatory variables were then removed from the full model one at a time and tested using a stricter criterion (*P*-value $\leq$0.01).

Internal validation of the final model was performed using a nonparametric bootstrap with sample replacement to estimate parameter uncertainty. When the 95th percentile bootstrap (2.5th – 97.5th percentile) confidence interval of the odds ratio included the value of 1, the variable associated with the parameter was considered uninformative and removed from the model.

Longitudinal analysis from enrolment at age <6-months until 5-years of age

Development of repeated time-to-event (RTTE) models

A parametric hazard model was utilised to describe the time to the first and any recurrent *Aspergillus* and *P. aeruginosa* events. The time of an event, *T,* was defined as the time at which an event occurred since study enrolment. The distribution of *T* can be defined by the hazard function. Time-independent (constant) baseline hazard and time-varying baseline hazards were investigated using either a Gompertz or a Weibull hazard model for *T*.[4 5]

Between-subject variability around the hazard was estimated, assuming an exponential distribution for the random effect.

Possible explanatory variables that may influence or predict the changes in hazard were explored by including each explanatory variable in the hazard function. A parameter, β_n, for each of the n explanatory variables, X_n, was estimated using the following equation;

$$h(t) = h_0 * \exp^{\beta_1 X_1 + \beta_2 X_2 \ldots + \beta_n X_n} \qquad \text{Equation 1}$$

where h_0 is the baseline hazard (λ_0), β_n is the coefficient for the explanatory variable, X_n, describing how the hazard varies with the explanatory variable. Exponentiation of the explanatory variable coefficient provides the hazard ratio (HR), which reflects the influence of the explanatory variables relative to the hazard when the explanatory variable is not present.

All event times were treated as interval censored data as the exact *T* of occurrence was unknown. It was assumed that the event occurred some time between the previous negative BAL (t_{pre}) and the time of the positive BAL procedure event (t_{post}). S(*t*) is the survivor function calculated using the following equation;

$$\boldsymbol{S(t) = exp\left(\int_{u=0}^{u=t} h(t).du\right)} \qquad \textit{Equation 1}$$

where S(*t*) is the survivor function calculated from the time-varying hazard h(t). The likelihood for interval-censored *T was* computed as $S(t_{pre}) - S(t_{post})$.

The parametric RTTE analysis was performed using NONMEM v7.4.1[6] and Perl speaks NONMEM (PsN) version 4.1.0.[7] and Wings for NONMEM 742 [8] Model selection was based on the comparison of the OFV between models, bootstrap confidence intervals for parameter estimates, and biological plausibility. The improvement in the fit was measured by a decrease in the OFV generated by NONMEM. The difference in OFV between two hierarchical models is approximately X^2 distributed and can be tested for significance with $X^2_{1,0.05} = 3.84$.[9] [10] Explanatory variables were included in a stepwise approach. The stepwise analysis approach was

repeated using joint model for both types of events. The final joint model of repeated–time to *Aspergillus* and *P. aeruginosa* events was evaluated with a nonparametric bootstrap to assess parameter imprecision and a Kaplan-Meier visual predictive check (VPC) to assess the predictive performance.[11]

Model estimation, selection and evaluation

Parameters were estimated using the FOCE method (ADVAN13 TOL=9 NSIG=3) in NONMEM to obtain maximum likelihood estimates of time-to-event parameters. Model building ceased when adding a new variable did not result in further significant improvement.

Event times for the same 80 children from the ACFBAL study data[1] were simulated using NONMEM with 1000 replicates of the model, where simulated *P. aeruginosa* events influenced the occurrence of AF events as identified in the final model.

To evaluate the predictive performance of the model throughout model building, Kaplan-Meier VPCs using RStudio software was utilized (version 1.1.456, RStudio, Inc., Boston, MA, http://www.rstudio.com/)[12]. The Surv and survfit functions in RStudio were used to obtain Kaplan-Meier non-parametric estimates of the survivor function for both the observed and simulated events. The 5%, 50% and 95% iles for the observed event survivor function were calculated by Surv and survfit. The percentiles of the simulated event survivor function were obtained by interpolation of each simulated survivor function at defined time points and calculation of the percentiles from the resulting distribution of survivor function values at these time points.

RESULTS

Table E1 Multivariable model of risk factors for positive *Aspergillus* BAL culture at age 5-years

2nd Multivariable step	ΔOFV	OR (95%CI)
Base model + number of *P. aeruginosa* eradication therapy courses received before BAL positive cultures for *Aspergillus* [b] +		
Female	-3.02	0.45 (0.29,1.07)
Meconium ileus	-0.48	1.34 (0.38,1.36)
Children in BAL arm of ACFBAL study	-0.42	1.52 (0.56,4.14)
Received gentamicin	-4E-05	1.00 (0.29,1.24)
Received anti-staphylococcal prophylaxis until their first birthday	-0.74	0.79 (0.19,3.32)
Body Mass Index z-score at BAL	-0.03	0.97 (0.67,1.01)
Cumulative dosage of intravenous tobramycin (not associated with eradication therapy) (mg)	-0.35	1.01 (1.00,1.01)
Cumulative dosage of intravenous tobramycin received (mg)[a]	-0.43	1.02 (0.95,1.09)
Cumulative dosage of inhaled tobramycin received (mg)[a]	-0.07	0.99 (0.98,1.00)
Cumulative dosage of intravenous and oral antibiotics (not associated with eradication therapy) (mg)[c]	-0.65	1.00 (1.00,1.00)
Minimum annual temperature at geographic region at baseline (^{0}C)	-0.99	0.91 (0.83,0.98)
Maximum annual temperature at geographic region at baseline (^{0}C)	-4.56[d]	1.11 (1.01,1.20)
Number of 'all-cause' CF-related hospitalisations before age 5-years BAL	-0.03	1.01 (0.98,1.02)
Number of pulmonary exacerbations requiring hospitalisation before BAL	-1.99	0.94 (0.88,1.01)
3rd multivariable step [e]		
Number of *P. aeruginosa* eradication therapy courses received before BAL positive cultures for *Aspergillus*	-14.55	1.61 (1.23,2.12)**

[a] combined with an anti-pseudomonal beta-lactam antibiotic; [b] model carried forward from the univariable step; [c] antibiotics other than tobramycin and not associated with eradication therapy (e.g. intravenous cefuroxime, oral macrolides, amoxicillin-clavulanate and co-trimoxazole) [d] excluded from the final model during backward elimination step as p>0.0; [e] Final multivariable model**Bootstrap results (95% confidence interval).

Abbreviation: ACFBAL, Australasian Cystic Fibrosis Bronchoalveolar Lavage study; BAL, bronchoalveolar lavage; ^{0}C, degree Celsius; IQR, interquartile range; mg, milligram; n, number of children contributing data, ΔOFV, change in objective function value.

Table E2: Multivariable model selection; the objective function (OVF) and the change in OFV (ΔOFV) associated with a tested explanatory variable are shown.

Model Number	Model Description	OFV	ΔOFV	df	*P*-value
1	**Full Model**	10260.31	0.00	.	.
2	Remove *Asp* Gompertz step 2	10260.39	0.08	1	0.775
3	Remove *Asp* Gompertz step 4	10260.67	0.36	1	0.551
4	Remove *Asp* Gompertz step 5	10261.58	1.27	1	0.261
5	Remove *Asp* Gompertz step 3	10263.55	3.24	3	0.357
6	Remove effect of PNA at enrollment on *Pa* hazard	10264.08	3.77	4	0.438
7	Remove effect of PNA at enrollment on *Asp* hazard	10267.36	7.05	7	0.423
8	Remove *Asp* Gompertz step 0	10272.35	12.04	1	0.000520
9	Remove *Asp* Gompertz step 5	10292.12	31.81	1	1.705E-08
10	**Reduced model (Final)**	10292.95	32.64	7	3.084E-05

ET – eradication therapy, *Pa* – *P. aeruginosa*, *Asp* – *Aspergillus*, PNA – postnatal age, yellow highlights – non-significant factors can be removed, green highlights – removal of potentially significant factor, should remain in the model, df – degree of freedom

full model – included a 5-step Gompertz baseline hazard model for *Aspergillus* (at age 0.5, 1,2,3 &4) an effect of after ET, an effect of PNA on the hazard for *Aspergillus* and *P*a, an effect of a 2nd, 3rd and 4th event for either event;

final model – a 1-step (at age 1-year) Gompertz baseline hazard model for *Aspergillus*, an effect of ET after 2.5-months of ET on *Aspergillus* hazard, influence of 2nd and 3rd *Aspergillus* event, a Gompertz baseline hazard model for *Pa*, an effect of ET after 2.5 months of ET on *Pa*, influence of 2nd, 3rd and 4th *Pa* event

The joint RTTE model of Aspergillus and Pseudomonas aeruginosa

Figure E1A shows the distribution of the 53-observed surveillance and symptomatic *Aspergillus* event times. The distribution of 89 observed surveillance and symptomatic *P. aeruginosa* event times is shown in Figure E1B. The distribution of the symptomatic *P. aeruginosa* events suggested a non-constant hazard, which was confirmed by RTTE modelling.

A full model (model 1 in Table E2) was developed, which included time since study entry to distinguish both the baseline hazard and the Gompertz hazard (Table E3). The baseline hazard was separated at 0.5-years into early and late values. The Gompertz hazard was separated at 1-year intervals. The post-natal age at study entry was a non-time varying explanatory variable. Time varying explanatory variables predicting the hazard were based on the interval following

eradication therapy until the next BAL. An initial reduced model (not shown) was obtained from the full model by fixing parameters whose 95% bootstrap confidence interval in the full model included 0 to the non-influential value of 0 (Table E4). A second reduced model (model 547) was obtained in a similar fashion based on the confidence intervals in the initial reduced model. This second reduced model confidence intervals indicated the post-natal age effect should be removed and this led to the final model (model 10 in Table E2; parameters shown in Table E5).

Table E3: Full model parameter estimates from original and bootstrap distribution (model 1 in Table E2)

Parameter	Description	**Units**	**Original**	**Bootstrap average**	**2.5% ile**	**97.5% ile**	**RSE**
Asp _BHAZ_EARLY	*Asp* baseline hazard <1y	1/kiloy	78.2	138	4.35	200	350%
Asp _BHAZ_LATE	*Asp* baseline hazard >=1y	1/kiloy	0.000210	0.000201	0.00000210	0.000812	124%
BNAF2	*Asp* effect of 1st and 2nd *Asp* event on *Asp* hazard	.	13.5	15.0	12.7	21.3	14%
BNAF3	*Asp* effect of 3rd *Asp* event on *Asp* hazard	.	12.3	13.6	11.1	20.8	27%
Asp _BGOMP EARLY	*Asp* Gompertz hazard <0.5y	1/y	-8720.0	-56636	-448525	-490	237%
Asp _BGOMP_LATE1	*Asp* Gompertz hazard >=0.5y and <1y	1/y	-7780	-10534454	-11592500	-2267	757%
Asp _BGOMP_LATE2	*Asp* Gompertz hazard >=1y and <2y	1/y	-0.291	-0.977	-4.87	1.88	-216%
Asp _BGOMP_LATE3	*Asp* Gompertz hazard >=2y and <3y	1/y	0.901	0.818	-0.237	1.92	78%
Asp _BGOMP_LATE4	*Asp* Gompertz hazard >=3y and <4y	1/y	-0.254	-0.633	-2.90	0.381	167%
Asp _BGOMP_LATE5	*Asp* Gompertz hazard >=4y	1/y	-0.351	-0.678	-2.247	0.106	96%
Asp _BPA	*Asp* effect after *Pa* eradication therapy	.	1.70	2.23	0.598	6.408	82%
Asp _BAGE0	*Asp* effect of postnatal age at baseline	1/y	5.62	6.25	0.748	15.58	66%
Pa _BHAZ	*Pa* baseline hazard <1y	1/y	0.0427	0.0536	0.000427	0.194	96%
BNPA2	*Pa* effect of 2nd PA event on *Pa* hazard	1/y	5.43	6.09	4.17	10.6	29%
BNPA3	*Pa* effect of 3rd PA event on *Pa* hazard	.	7.68	8.23	4.92	13.5	28%
Pa _BGOMP	*Pa* Gompertz hazard	1/y	-1.09	-1.16	-1.76	-0.599	28%
Pa _BPA	*Pa* effect after *Pa* eradication therapy	.	-1.51	-32.9	-89.9	0.090	589%
Pa _BAGE0	*Pa* effect of postnatal age at baseline	1/y	3.12	3.49	-0.10	7.64	60%

Asp, Aspergillus; Pa, Pseudomonas aeruginosa; yellow highlights – 95% bootstrap confidence interval in the model included 0 to the non-influential value of 0 and can be removed,

Table E4: Initial reduced model parameter estimates from original and bootstrap distribution (model not shown Table E2)

Parameter	Description	**Units**	**Original**	**Bootstrap average**	**2.5% ile**	**97.5% ile**	**RSE**
Asp _BHAZ_EARLY	*Asp* baseline hazard <1y	1/kiloy	0.553	141	0.00576	437	729%
Asp _BHAZ_LATE	*Asp* baseline hazard >=1y	1/kiloy	0.00055	0.173	0.00000550	0.316	667%
BNAF2	*Asp* effect of 1st and 2nd *Asp* event on *Asp* hazard	.	13.60	13.76	7.94	18.2	16%
BNAF3	*Asp* effect of 3rd *Asp* event on *Asp* hazard	.	12.30	15.83	3.35	17.8	221%
Asp _BGOMP EARLY	*Asp* Gompertz hazard <0.5y	1/y	0	FIXED	.	.	.
*Asp*_BGOMP_LATE1	*Asp* Gompertz hazard >=0.5y and <1y	1/y	-40.6	-41.6	-72.0	-14.30	58%
*Asp*_BGOMP_LATE2	*Asp* Gompertz hazard >=1y and <2y	1/y	0.897	-0.400	-0.689	4.68	3456%
*Asp*_BGOMP_LATE3	*Asp* Gompertz hazard >=2y and <3y	1/y	1.06	1.41	0.54	3.44	59%
*Asp*_BGOMP_LATE4	*Asp* Gompertz hazard >=3y and <4y	1/y	0	FIXED	.	.	.
*Asp*_BGOMP_LATE5	*Asp* Gompertz hazard >=4y	1/y	0	FIXED	.	.	.
Asp _BPA	*Asp* effect after *Pa* eradication therapy	.	0	FIXED	.	.	.
Asp _BAGE0	*Asp* effect of postnatal age at baseline	1/y	0	FIXED	.	.	.
Pa _BHAZ	*Pa* baseline hazard	1/y	0.0951	0.119	0.00524	0.324	71%
BNPA2	*Pa* effect of 2nd *Pa* event on *Pa* hazard	1/y	4.93	5.25	3.95	8.27	19%
BNPA3	*Pa* effect of 3rd *Pa* event on *Pa* hazard	.	6.94	7.29	5.88	10.0	14%
Pa _BGOMP	*Pa* Gompertz hazard	.	-0.918	-0.993	-1.66	-0.510	32%
Pa _BPA	*Pa* effect after *Pa* eradication therapy	.	-1.93	-31.0	-453.9	-0.288	444%
Pa _BAGE0	*Pa* effect of postnatal age at baseline	1/y	0	FIXED	.	.	.

Asp, Aspergillus; Pa, Pseudomonas aeruginosa; yellow highlights – 95% bootstrap confidence interval in the model included 0 to the non-influential value of 0 and can be removed

Table E5: Second reduced model and final model parameter estimates from original and bootstrap distribution (model 10 in Table 2)

Parameter	Description	**Units**	**Original**	**Bootstrap average**	**2.5% ile**	**97.5% ile**	**RSE**
Asp _BHAZ_EARLY	*Asp* baseline hazard <1y	1/kiloy	78.9	80.2	68.0	95.4	11%
Asp _BHAZ_LATE	*Asp* baseline hazard >=1y	1/kiloy	0.000253	0.000364	0.000200	0.00129	80%
BNAF2	*Asp* effect of 1st and 2nd *Asp* event on *Asp* hazard	.	13.5	13.51	12.20	14.42	4%
BNAF3	*Asp* effect of 3rd *Asp* event on *Asp* hazard	.	13.3	13.41	11.70	14.54	11%
*Asp*_ BGOMP EARLY	*Asp* Gompertz hazard <0.5y	1/y	-6470	-6129	-10653	-1884	42%
*Asp*_BGOMP_LATE1	*Asp* Gompertz hazard >=0.5y and <1y	1/y	-6140	-6796	-12135	-2019	43%
*Asp*_BGOMP_LATE2	*Asp* Gompertz hazard >=1y and <2y	1/y	0	FIXED	.	.	.
*Asp*_BGOMP_LATE3	*Asp* Gompertz hazard >=2y and <3y	1/y	0	FIXED	.	.	.
*Asp*_BGOMP_LATE4	*Asp* Gompertz hazard >=3y and <4y	1/y	0	FIXED	.	.	.
*Asp*_BGOMP_LATE5	*Asp* Gompertz hazard >=4y	1/y	0	FIXED	.	.	.
Asp _BPA	*Asp* effect after *Pa* eradication therapy	.	1.01	1.00	0.375	1.69	38%
Asp _BAGE0	*Asp* effect of postnatal age at baseline	1/y	3.18	2.57	-0.265	6.00	62%
*Pa*_BHAZ	*Pa* baseline hazard	1/y	0.095	0.118	0.0134	0.325	72%
BNPA2	*Pa* effect of 2nd PA event on *Pa* hazard	1/y	4.93	5.26	3.92	7.12	18%
BNPA3	*Pa* effect of 3rd *Pa* event on *Pa* hazard	.	6.94	7.35	5.81	9.68	14%
Pa _BGOMP	*Pa* Gompertz hazard	1/y	-0.918	-1.02	-1.71	-0.513	33%
Pa _BPA	*Pa* effect after *Pa* eradication therapy	.	-1.93	-7.75	-85.4	-0.242	250%
Pa _BAGE0	*Pa* effect of postnatal age at baseline	1/y	0	FIXED	.	.	.

Asp, Aspergillus; Pa, Pseudomonas aeruginosa; yellow highlights – 95% bootstrap confidence interval in the model included 0 to the non-influential value of 0 and can be removed

Evaluation of Kaplan-Meier VPCs

Internal evaluation using Kaplan Meier VPCs showed good agreement between the observed data survivor function and the median simulated survivor function for *P. aeruginosa* events. (Figure E2). For *Aspergillus* events the model predicted median survivor function was within the observed survivor function 95% confidence interval but tended to under-predict the hazard at later times (Figure E3).

Application of the RTTE models

The final RTTE models was used to evaluate the probability of having recurrent *Aspergillus* and *P. aeruginosa* events within the first 8-years of life for young children with cystic fibrosis (CF) under different hypothetical clinical scenarios. The evaluated probability of having an *Aspergillus* event within the first 8-years of life in children receiving eradication therapy for *P. aeruginosa* varied at different times during the first 5-years of life. The probability after 5-years of age of having an initial positive *Aspergillus* culture in a child who had not undergone eradication therapy was predicted to be 5.3% compared to 39.8% for a child that received this therapy for the first time during the first-year of life and 26.1% for a child that received eradication therapy after age 4-years (Figure E4) Figure E4 illustrates that mainly eradication therapy impacts on the probability of a *Aspergillus* event rather than time.

NMTRAN code for final joint model

```
;; DESCRIPTION: Joint PA and Asp RTTE model
$PROBLEM    RTTE TIME TO EVENT DATA

$INPUT ID PID TIME DV MDV EVID FLAG NAF NPA RTTE
CUMTOB CUMTOBR IVTOBRA DURIVTX AGE SEX HT WT SCR
TNUPE ET NETC BMI SAB HIB NSA NHI STENOB MAC NMAC
CUMAC CUERYT GENTA NGENTA MAXTPB MINTPB MEANTP STATE
$DATA PAAF40_AFPA_ET.csv IGNORE=#


; FLAG = 4 -> AF events
; FLAG = 5 -> PA events

$THETA
(0,78.9)        ; AF_BHAZ_EARLY 1/kiloy: AF Baseline hazard before 1 y
(0,0.000253)    ; AF_BHAZ_LATE 1/kiloy: AF Baseline hazard >= 1y
                ; AF effect with second and third events
13.5            ; BNAF2 SECOND EVENT
13.3            ; BNAF3 THIRD EVENT
                ; AF Gompertz hazard change
-6470.          ; AF_BGOMP_EARLY 1/y before 0.5 y
-6140.          ; AF_BGOMP_LATE1 1/y >=0.5 and <1 y
                ; Eradication therapy effect during ET on AF
0 FIX           ; AF_BET (not used in this model)
                ; Eradication therapy effect AFTER ET on AF
1.01            ; AF_BPA
                ; AF Baseline postnatal age effect (not used in the final reduced model)
;3.18           ; AF_BAGE0 1/PNAY SINCE BIRTH

(0,0.095)       ; PA_BHAZ 1/Y, PA Baseline hazard
                ; PA effect with second combined with third PA events
4.93            ; BNPA2 SECOND  & Third EVENT
                ; PA effect with fourth PA event
6.94            ; BNPA3 fourth EVENT
                ; PA Gompertz hazard
-0.918          ; PA_BGOMP 1/y
                ; Eradication therapy effect during ET on PA
0 FIX           ; PA_BET (not used in this model)
                ; Eradication therapy effect AFTER ET on PA
-1.93           ; PA_BPA

$OMEGA 0 FIX ;PPV_BHAZ INTER-INDIVIDUAL VARIABILITY IN BASELINE HAZARD

$SUBROUTINE ADVAN13 TOL=9
```

```
$MODEL
  COMP=HAZAF     ; CUMULATIVE HAZARD COMPARTMENT OF AF
  COMP=HAZPA     ; CUMULATIVE HAZARD COMPARTMENT OF PA

$PK

IF (NEWIND.LE.1) THEN
  LN2=LOG(2)
  ETDUR=2.5/12          ; 2.5 MONTHS of standard ET therapy after a PA event
  SURSTARTAFEST=1
  OLDCHZAFEST=0         ; SET FOR CUMULATIVE HAZARDS FROM PREVIOUS TO
                        CURRENT OBSERVATION FOR AF
  SURSTARTPAEST=1
  OLDCHZPAEST=0         ; SET FOR CUMULATIVE HAZARDS FROM PREVIOUS TO
                        CURRENT OBSERVATION FOR PA
  ENDETEST=0
  EVTTIMEAFEST=0
  EVTTIMEPAEST=0
  WASPAEVTEST=0
  AGE0=AGE
  BLK=-1
ENDIF

IF (ICALL.EQ.4) THEN
  IF (NEWIND.LE.1) THEN
    OLDCHZAFSIM=0
    OLDCHZPASIM=0
    ENDETSIM=0
    EVTTIMEAFSIM=0
    EVTTIMEPASIM=0
    WASPAEVTSIM=0
    NAFX=0
    NPAX=0
    CALL RANDOM(2,R)
    UAF=R               ; UNIFORM RANDOM NUMBER FOR Asp EVENT;
    CALL RANDOM(2,R)
    UPA=R               ; UNIFORM RANDOM NUMBER FOR PA EVENT;
  BLK=-2
  ENDIF
ENDIF

;------------------------------ Aspergillus Model ------------------------------

; AF Baseline hazard before and after 1 year

IF (TIME.LT.1) THEN
```

```
  GRP_AFBHAZ=AF_BHAZ_EARLY/1000          ;1/ky -> 1/y
ELSE
  GRP_AFBHAZ=AF_BHAZ_LATE/1000           ;1/ky -> 1/y
ENDIF

AFBHAZ=GRP_AFBHAZ*EXP(PPV_BHAZ)

; AF Gompertz hazard change

IF (TIME.LT.0.5) THEN
  GRP_GOMPAF=AF_BGOMP
ELSE
  IF (TIME.LT.1) THEN
    GRP_GOMPAF=AF_BGOMP_LATE1
  ELSE
    IF (TIME.LT.2) THEN
      GRP_GOMPAF=AF_BGOMP_LATE2
    ELSE
      IF (TIME.LT.3) THEN
        GRP_GOMPAF=AF_BGOMP_LATE3
      ELSE
        IF (TIME.LT.4) THEN
          GRP_GOMPAF=AF_BGOMP_LATE4
        ELSE
          GRP_GOMPAF=AF_BGOMP_LATE5
        ENDIF
      ENDIF
    ENDIF
  ENDIF
ENDIF
GOMPAF= GRP_GOMPAF

; AF effect with second and third events

IF (NAF.EQ.0) THEN          ; NAF=NUMBER OF AF EVENTS
  BNAF=0
ELSE
  IF (NAF.EQ.1) THEN
    BNAF=BNAF2
  ELSE
    BNAF=BNAF3
  ENDIF
ENDIF

;------------------------------ PA Model -----------------------------
; PA Baseline hazard
```

```
PABHAZ=PA_BHAZ*EXP(PPV_BHAZ)

; PA effect with second, third & fourth events
IF (NPA.EQ.0) THEN                      ; NPA=NUMBER OF PA EVENTS
  BNPA=0
ELSE
  IF (NPA.EQ.1.OR.NPA.EQ.2) THEN        ; 2nd and 3rd event
    BNPA=BNPA2
  ELSE
    BNPA=BNPA3                          ; fourth event
  ENDIF
ENDIF

IF (ICALL.EQ.4) THEN
  ENDET=ENDETSIM
  STARTET=EVTTIMEPASIM
  WASPAEVT=WASPAEVTSIM
ELSE
  STARTET=EVTTIMEPAEST
  ENDET=ENDETEST
  WASPAEVT=WASPAEVTEST
ENDIF

; ET effects
IF (ENDET.GT.0.AND.TIME.LE.ENDET) THEN          ; DURING ET (ET=2)
  AFBETAET=AF_BET
  PABETAET=PA_BET
ELSE      ; OUTSIDE ET
  IF (WASPAEVT.EQ.1) THEN                       ; AFTER ET (ET=2)
    PABETAET=PA_BPA
    AFBETAET=AF_BPA
  ELSE                                  ; NO PA EVENT THEREFORE NO EFFECT OF
                                        PRECEDING PA EVENT ON AF OR PA
                                        HAZARD (ET=0)
    PABETAET=0                          ; OUTSIDE ET
    AFBETAET=0
  ENDIF
ENDIF


IF (ICALL.EQ.4) THEN
  IF (TIME.GT.ENDET) ENDETSIM=0     ; RESET ENDET
ELSE
  IF (TIME.GT.ENDET) ENDETEST=0     ; RESET ENDET
ENDIF
```

```
GOMPPA=PA_BGOMP

; Baseline post-natal age effect - not used in the final reduced model
; AFAGE=AF_BAGE0*AGE0


$DES

DADT(1)=AFBHAZ*EXP(BNAF+GOMPAF*T+AFBETAET)
DADT(2)=PABHAZ*EXP(BNPA+GOMPPA*T+PABETAET)

$ERROR

;------------TIME TO EVENT FOR AF ----------------------------------------------------------------
SURAFEST=EXP(-(A(1)-OLDCHZAFEST))

IF (FLAG.EQ.4.AND.MDV.EQ.1) THEN          ; START OF CENSORING INTERVAL
  SURSTARTAFEST=SURAFEST                  ; START OF AF EVENT INTERVAL
ELSE
  SURSTARTAFEST=SURSTARTAFEST
ENDIF

;FLAG=4 ALWAYS PRECEDES FLAG=5 AND MDV IS ZERO ONLY AT TIMES OF BAL EVENTS
;DV MAY BE 1 (FOR ESTIMATION OR SIMULATION) OR 0 (FOR SIMULATION) IN THE FLAG=4 RECORD
;WASPAEVT (was a PA event) IS SET TO ZERO BEFORE FLAG=5 RECORD WHEN PA EVENT MIGHT BE OBSERVED OR SIMULATED

IF (FLAG.EQ.4.AND.MDV.EQ.0) THEN
  WASPAEVTEST=0
  WASPAEVTSIM=0
ENDIF

IF (FLAG.EQ.4.AND.MDV.EQ.0.AND.DV.EQ.0) THEN          ; RIGHT CENSORED, NO EVENT FOR AF
  Y=SURAFEST
  SURSTARTAFEST=SURAFEST                  ; START OF AF EVENT INTERVAL
ELSE
  SURSTARTAFEST=SURSTARTAFEST
ENDIF

IF (FLAG.EQ.4.AND.MDV.EQ.0.AND.DV.EQ.1) THEN      ; LIKELIHOOD OF EVENT
  Y=SURSTARTAFEST - SURAFEST
  OLDCHZAFEST=A(1)
  SURSTARTAFEST=1                         ; START OF AF EVENT INTERVAL
```

```
ELSE
  OLDCHZAFEST=OLDCHZAFEST
  SURSTARTAFEST=SURSTARTAFEST
ENDIF
;------------TIME TO EVENT FOR PA ----------------------------------------------------------------------
SURPAEST=EXP(-(A(2)-OLDCHZPAEST))

IF (FLAG.EQ.5.AND.MDV.EQ.1) THEN            ; START OF CENSORING INTERVAL
  SURSTARTPAEST=SURPAEST                    ; START OF PA EVENT INTERVAL
ELSE
  SURSTARTPAEST=SURSTARTPAEST
ENDIF

IF (FLAG.EQ.5.AND.MDV.EQ.0.AND.DV.EQ.0) THEN    ; RIGHT CENSORED, NO EVENT
FOR PA
  Y=SURPAEST
  SURSTARTPAEST=SURPAEST                    ; START OF PA EVENT INTERVAL
ELSE
  SURSTARTPAEST=SURSTARTPAEST
ENDIF

IF (FLAG.EQ.5.AND.MDV.EQ.0.AND.DV.EQ.1) THEN     ; LIKELIHOOD OF EVENT
  Y=SURSTARTPAEST - SURPAEST
  OLDCHZPAEST=A(2)
  EVTTIMEPAEST=TIME
  ENDETEST=EVTTIMEPAEST+ETDUR
  SURSTARTPAEST=1                           ; START OF PA EVENT INTERVAL
  WASPAEVTEST=1
ELSE
  OLDCHZPAEST=OLDCHZPAEST
  SURSTARTPAEST=SURSTARTPAEST
ENDIF

; ---------------------FOR RTTE SIMULATION -------------------------------------

IF (ICALL.EQ.4) THEN
  SURAFSIM=EXP(-(A(1)-OLDCHZAFSIM))
  SURPASIM=EXP(-(A(2)-OLDCHZPASIM))

 ; FLAG REMAINS THE SAME AS IN THE ORIGINAL DATA
 ; DVID IS THE VALUE OF FLAG THAT IS CREATED FOR USE IN THE SIMULATION
OUTPUT
  DVID=FLAG

  IF (FLAG.EQ.1) THEN                       ; FIRST RECORD
BLK=0                                       ; used for debugging only
```

```
    DV=0
    MDV=1
  ENDIF
  IF (FLAG.EQ.4.OR.FLAG.EQ.-4) THEN          ; CHECK FOR AF EVENT
    IF (EVID.EQ.0) THEN                      ; SIMULATED EVENT POSSIBLE
      IF (SURAFSIM.LT.UAF) THEN        ; SIMULATED INTERVAL CENSORED AF EV
BLK=1
        RTTE=1
        NAFX=NAFX+1                    ; COUNT SIMULATED EVENTS
        DV=1
        MDV=0
                                    ; SAVE CUMULATIVE HAZARD FOR THIS EVENT
        OLDCHZAFSIM=A(1)
                        ; START OF NEXT INTERVAL BECAUSE THIS IS A BAL EVENT
        CALL RANDOM (2,R)              ; RESAMPLE SO EVENTS ARE INDEPENDENT
        UAF=R
                                     ; ENDET=TIME+ETDUR  FOR PA EVENT ONLY
        EVTTIMEAFSIM=TIME
      ELSE
        IF (FLAG.EQ.-4) THEN
          IF (TIME.GT.0.AND.NAFX.EQ.0.OR.TIME.GT.EVTTIMEAFSIM) THEN
; NOT HAD AN EVENT YET OR EVENT NOT AT THIS TIME -> CENSORED EVENT
BLK=2
            DVID=4
            DV=0
            MDV=0
                              ; NEXT RECORDS FOR COMPLETNESS BUT HAS NO
                             EFFECT BECAUSE THIS IS THE LAST SIMULATED EVENT
            OLDCHZAFSIM=A(1)
                        ; START OF NEXT INTERVAL BECAUSE THIS IS A BAL EVENT
            RTTE=1
          ELSE                             ; NO EVENT -> DUMMY RECORD
BLK=3
            DV=-FLAG
            MDV=1
            OLDCHZAFSIM=OLDCHZAFSIM
            RTTE=0
          ENDIF
        ELSE                               ; NO EVENT -> DUMMY RECORD
BLK=4
          DV=-FLAG
          MDV=1
          OLDCHZAFSIM=OLDCHZAFSIM
          RTTE=0
        ENDIF                              ; FLAG -4?
      ENDIF                                ; SURAF < UAF?
```

```
    ELSE                    ; START OF CENSORING INTERVAL DUMMY RECORD
BLK=5
      DV=FLAG
      MDV=1
      OLDCHZAFSIM=OLDCHZAFSIM
      RTTE=0
    ENDIF                   ; EVID=0
  ENDIF                     ; AF EVENT FLAG

  IF (FLAG.EQ.5.OR.FLAG.EQ.-5) THEN       ; CHECK FOR PA EVENT
    IF (EVID.EQ.0) THEN                   ; SIMULATED EVENT POSSIBLE
      IF (SURPASIM.LT.UPA) THEN           ; SIMULATED INTERVAL CENSORED
                                          PA EVENT
BLK=6
        RTTE=2
        NPAX=NPAX+1                 ; COUNT SIMULATED EVENTS
        DV=1
        MDV=0
                                ; SAVE CUMULATIVE HAZARD FOR THIS EVENT
        OLDCHZPASIM=A(2)
                        ; START OF NEXT INTERVAL BECAUSE THIS IS A BAL EVENT
        CALL RANDOM (2,R)           ; RESAMPLE SO EVENTS ARE INDEPENDENT
        UPA=R
        EVTTIMEPASIM=TIME
        ENDETSIM=EVTTIMEPASIM+ETDUR
        WASPAEVTSIM=1               ; THIS IS SET TO ZERO IN THE DEFAULT CASE
                                    AT EACH BAL FOR PA
      ELSE
        IF (FLAG.EQ.-5) THEN
          IF (TIME.GT.0.AND.NPAX.EQ.0.OR.TIME.GT.EVTTIMEPASIM) THEN
                                    ; NOT HAD AN EVENT YET OR EVENT NOT AT
                                    THIS TIME -> CENSORED EVENT
BLK=7
            DVID=5
            DV=0
            MDV=0
                                ; NEXT RECORDS FOR COMPLETNESS BUT HAS NO
                                EFFECT BECAUSE THIS IS THE LAST SIMULATED
                                EVENT
            OLDCHZPASIM=A(1)
                        ; START OF NEXT INTERVAL BECAUSE THIS IS A BAL
EVENT
            RTTE=2
          ELSE              ; NO EVENT -> DUMMY RECORD
BLK=8
            DV=-FLAG
```

```
            MDV=1
            OLDCHZPASIM=OLDCHZPASIM
          ENDIF
        ELSE                ; NO EVENT -> DUMMY RECORD
BLK=9
          DV=-FLAG
          MDV=1
          RTTE=0
          OLDCHZPASIM=OLDCHZPASIM
        ENDIF               ; FLAG -5?
      ENDIF                 ; SURPA < UPA?
    ELSE                    ; START OF CENSORING INTERVAL -> DUMMY RECORD
BLK=10
      DV=FLAG
      MDV=1
      OLDCHZPASIM=OLDCHZPASIM
      RTTE=0
    ENDIF                   ; EVID=0?
  ENDIF                     ; PA EVENT FLAG?
  NAF=NAFX
  NPA=NPAX
  SURAF=SURAFSIM
  SURPA=SURPASIM
ELSE
  SURAF=SURAFEST
  SURPA=SURPAEST
ENDIF                       ; ICALL=4?

;-----------------------------------------
REP=IREP

$ESTIMATION METHOD=ZERO LIKE MAX=9990 NSIG=3 SIGL=9 PRINT=1
MSFO=MSFB014
$COVARIANCE

; $SIMULATION (839928948) (841743102 UNIFORM)
; REWIND                    ; reuse original data set
; ONLYSIM  NOPRED NSUB=1000

$TABLE REP ID EVID DVID NAF NPA TIME DV MDV RTTE FLAG SURAF SURPA BLK
NOAPPEND ONEHEADER NOPRINT FILE=npctab.dat
```

Figure legends

Figure E1. Density distributions with a kernel density smooth (dashed lines) for *Aspergillus* (A) and *Pseudomonas aeruginosa* (B) event times. Distribution of 'surveillance' (black) and 'symptomatic' (grey) events for both *Aspergillus* and *P. aeruginosa* positive culture are shown over the first 5-years of life.

Figure E2. Kaplan-Meier plots showing the *P. aeruginosa* survivor function (probability of no *Pseudomonas aeruginosa* detection) throughout the first 5-years of life for the first four *P. aeruginosa* events. The observed data survivor function is described by a solid black line with 95% confidence intervals (CIs) (dashed black lines). These are overlaid with the median extrapolated survivor function (solid blue line) and 95% CI (shaded area)

Figure E3. Kaplan-Meier plots showing the *Aspergillus* survivor function (probability of no *P. aeruginosa* detection) throughout the first 5-years of life for the first four AF events. Legend is the same as Figure E2.

Figure E4. Probability of having an initial *Aspergillus* event: (A) a child who never received eradication therapy for *P. aeruginosa* infection, (B) a child who received eradication therapy for the first time before 1-year of age, and (C) a child who received the first eradication therapy after age 4-years (solid black line). Black line presents the median expected probability and dotted lines the present 95% confidence intervals. These are generated through simulation from the final model and visualise an extrapolation up to 8-years of age.

Figure E5: Illustrative explanation of the model fitting for the hazards time course of two models; hazard model of *P. aeruginosa* events (A) and *Aspergillus* events (B), the effect of post-ET hazard on *P. aeruginosa* events (C) and the effect of post-ET hazard on *Aspergillus* event (D).

What is known: (A&B) The event occurred between time zero and time 6. Eradication therapy (ET) is from 0 to 2.5. The likelihood of the event is determined by the risk (cumulative hazard) (dotted red line). The best fit has the highest likelihood. (C) *P. aeruginosa* events are detected at year 1 and year 2. (D)*Aspergillus* events are detected at year 1, 2, and 3.

Models: Two models for the hazard time course are shown. (A&B) The constant hazard (solid red line) does not change with time. The ET hazard changes after time 2.5. The constant hazard model has one parameter (baseline hazard). The ET model has two parameters (baseline hazard and hazard after ET). In this example, both models have the same risk (cumulative hazard) at time 6 and thus the same likelihood. This shows the shapes of the hazard and risk under the null hypothesis. (C&D) Two models for the hazard time course are shown. The hazard changes with time and number of events. The ET hazard changes at 2.5 months following a *P. aeruginosa* event until the next event.

Model Fitting: (A&B) The parameters for each model can be estimated by maximising the likelihood. The ET model has a higher likelihood than the constant hazard model. The difference is significant even though the ET model has one extra parameter. Thus, the actual event data means the green curves for the hazard and risk describe the hazard time course better. The null hypothesis can be rejected. (C) The baseline hazard is initially low and decreases slowly with time (Gompertz hazard). The hazard ratio increases similarly after the first event and second event. The post-ET effect reduces the hazard ratio until the next event. (D) The hazard is initially high and decreases

rapidly with time (Gompertz hazard). The hazard ratio increases after the first event and slightly less after the second event. The post-ET effect increases the hazard ratio until the next event.

References

1. Wainwright CE, Vidmar S, Armstrong DS, *et al.* Effect of bronchoalveolar lavage-directed therapy on *Pseudomonas aeruginosa* infection and structural lung injury in children with cystic fibrosis: a randomized trial. *JAMA* 2011;306:163-71.

2. IBM SPSS Statistics for Windows Version 22.0 [program]. Armonk, NY: IBM Corp, 2013.

3. Sheiner LB, Beal SL. Some suggestions for measuring predictive performance. *J Pharmacokinet Biopharm* 1981;9:503-12.

4. Holford N. A time to event tutorial for pharmacometricians. *CPT Pharmacometrics Syst Pharmacol* 2013;2:e43.

5. Collett D. *Modelling survival data in medical research*. Boca Raton, Fla: Chapman & Hall/CRC, 2003.

6. NONMEM User's Guides (1989-2009) [program]: Ellicott City, MD Icon development Solutions 2009.

7. Lindbom L, Ribbing J, Jonsson EN. Perl-speaks-NONMEM (PsN)--a Perl module for NONMEM related programming. *Comput Methods Programs Biomed* 2004;75:85-94.

8. Holford NHG. Wings for NONMEM Version 742 for NONMEM 7.4, 7.3 and 7.2: http://wfn.sourceforge.net, 2017.

9. Wahlby U, Jonsson EN, Karlsson MO. Comparison of stepwise covariate model building strategies in population pharmacokinetic-pharmacodynamic analysis. *AAPS PharmSci* 2002;4:E27.

10. Wahlby U, Jonsson EN, Karlsson MO. Assessment of actual significance levels for covariate effects in NONMEM. *J Pharmacokinet Pharmacodyn* 2001;28:231-52.

11. Nguyen TH, Mouksassi MS, Holford N, *et al.* Model Evaluation of Continuous Data Pharmacometric Models: Metrics and Graphics. *CPT Pharmacometrics Syst Pharmacol* 2017;6:87-109.

12. RStudio: Integrated development environment for R [program], 2015 http://www.rstudio.com/.