

1

2 Monte Carlo simulation of parameter confidence intervals for non-
3 linear regression analysis of biological data using Microsoft Excel

4

5 Ronald J.W. Lambert¹, Ioannis Mytilinaios¹, Luke Maitland¹ and Angus M. Brown²

6

7 ¹Applied Microbiology Group, Cranfield Health Cranfield University Cranfield MK43 0AL,
8 U.K.

9 ²School of Biomedical Sciences, Queens Medical Centre, University of Nottingham,
10 Nottingham, NG7 2UH, UK.

11

12

13 Corresponding author:

14 Ronald J.W. Lambert

15 ¹Applied Microbiology Group, Cranfield Health, Cranfield University
16 Cranfield MK43 0A, U.K.

17

18 Tel +44-1234 758314; rjwlambert@cranfield.ac.uk

19

20 Keywords: Monte Carlo simulation, SOLVER, Excel

21 **Abstract**

22

23 This study describes a method to obtain parameter confidence intervals from the fitting of
24 non-linear functions to experimental data, using the SOLVER and Analysis ToolPaK Add-
25 In of the Microsoft Excel spreadsheet. Previously we have shown that Excel can fit
26 complex multiple functions to biological data, obtaining values equivalent to those returned
27 by more specialized statistical or mathematical software. However, a disadvantage of
28 using the Excel method was the inability to return confidence intervals for the computed
29 parameters or the correlations between them. Using a simple Monte-Carlo procedure
30 within the Excel spreadsheet (without recourse to programming), SOLVER can provide
31 parameter estimates (up to 200 at a time) for multiple 'virtual' data sets, from which the
32 required confidence intervals and correlation coefficients can be obtained. The general
33 utility of the method is exemplified by applying it to the analysis of the growth of *Listeria*
34 *monocytogenes*, the growth inhibition of *Pseudomonas aeruginosa* by chlorhexidine and
35 the further analysis of the electrophysiological data from the compound action potential of
36 the rodent optic nerve.

37

38 **1 Introduction**

39

40 We have previously described the use of the Microsoft Excel spreadsheet to conduct non-
41 linear regression (NLR) analysis of biological data [1]. Direct fitting of the dose response
42 curve, for example, through the use of NLR techniques has been widely advocated, but
43 access to, and comprehension of, commercial software are often at odds with the direct
44 needs of the researcher. We recognised the ease of access and understanding that most
45 researchers have of Microsoft Excel [2], which could allow even those with an elementary
46 understanding of the spreadsheet to conduct relatively sophisticated data analyses,
47 without the expense of purchasing and learning a new statistical or advanced
48 mathematical package.

49

50 The SOLVER Add-In package of Excel allows the user to conduct investigations of non-
51 linear (NL) functions using the minimization of the sum of squares of the errors between
52 the observed and modelled values [1]. We further described the use of the technique for
53 the modelling of multiple Gaussian functions, which described the observed
54 electrophysiological data from the compound action potential of the rodent optic nerve [3].
55 One particular failing of the SOLVER package was the inability to return parameter
56 confidence intervals. It was noted that this requires the use of the Hessian matrix, whose
57 calculation and use would invalidate the aim of making NLR open to anyone [1]. Hence the
58 error analysis of the modelled and observed data was terminated at the calculation of the
59 standard error of the fit.

60

61 Confidence intervals can be calculated from knowledge of the Hessian but can also be
62 estimated using either the Bootstrap technique of re-sampling errors between the
63 modelled fit and the observed, or from Monte-Carlo (MC) simulation [6]. The MC technique
64 uses the standard error of the fit of the non-linear model to the observed data to produce
65 sets of 'virtual' data. These data are modelled using the same non-linear model and a new
66 group of parameters obtained for each virtual set. From the statistical distribution of these
67 parameters, confidence intervals as well as correlation coefficients can be obtained.

68

69 We describe here the use of NLR within the Excel environment and augment our original
70 method with a simple MC analysis, and show its general utility by applying it to analyse the
71 growth of *Listeria monocytogenes*, the growth inhibition of *Pseudomonas aeruginosa* by
72 chlorhexidine and the further analysis of the electrophysiological data from the compound
73 action potential of the rodent optic nerve.

74

75 **2 Computational Methods and Theory**

76

77 For a given data set and a particular model (y_{fit}), the sum of squares of the errors is given
78 by

79
$$SSE = \sum_{i=1}^n (y_i - y_{fit})^2$$

80 In a regression analysis the value of SSE is minimised by changing the parameter values
81 of the model y_{fit} , resulting in the best estimates of these parameters. In linear regression
82 this is solved analytically, but if using non-linear regression this is carried out numerically,
83 based on the input of initial parameter estimates. The square root of the mean of the
84 square of the error (RMSE) is the standard error of the fit. For a given set of conditions the
85 model will return the expected value of y_i , $E(y_i) = \hat{y}_i$. With linear regression, if all the
86 prerequisite conditions are met, then the reported 95% confidence intervals will contain the
87 true value of the regression parameters 95% of the time. With non-linear regression
88 confidence intervals are found using linear approximations and the labelled 95%
89 confidence intervals may not contain the true interval as often.

90

91 If the conditions required for regression are met, e.g. constant variance of error
92 (homoscedasticity), normal distribution of errors, then the RMSE is an unbiased estimator
93 of the standard deviation of the fit. A virtual data set can be calculated by adding random
94 error to the expected value of y ;

95
$$Y'_i = \hat{y}_i + N(0, RMSE)$$

96 This virtual data set can be analysed by NLR to give another set of parameters (the best fit
 97 estimates for this virtual data set).

$$98 \quad SSE = \sum_{i=1}^n (Y'_i - y_{fit})^2$$

99 Another set of virtual data can be generated and the NLR fitting repeated. In the procedure
 100 outlined here the sum of squares of the errors from multiple data sets are summed and the
 101 fitting of m -sets of data are conducted simultaneously

$$102 \quad SSE_{total} = \sum_{j=1}^m \sum_{i=1}^n (Y'_i - y_{fit})^2$$

103 From the m -sets of parameters obtained, frequency analyses of the parameter values are
 104 performed and the 95% confidence intervals obtained from the normal quantiles;
 105 covariance between parameter pairs can be found by calculating the parameters'
 106 correlation coefficient.

107

108 **2.1 FITTING THE MODIFIED GOMPERTZ EQUATION TO MICROBIAL GROWTH DATA**

109 The modified Gompertz equation is a standard empirical model for the fitting of microbial
 110 growth data [5].

$$111 \quad \log N(t) = A + C \exp\{-\exp(B(M - t))\} \quad (1)$$

112 Where A is the asymptotic number ($\log \text{cfu ml}^{-1}$) as t tends to negative infinity, $A+C =$
 113 maximum population density as t tends to positive infinity, B is a measure of the slope and
 114 M is the time of maximum slope. From these fitted parameters the growth rate and lag are
 115 calculated, respectively, as

$$116 \quad \text{CB}/\exp(1) \quad (2)$$

$$117 \quad M-1/B \quad (3)$$

118 Data for the growth of *Listeria monocytogenes* at 30°C in growth media (Tryptone Soya
119 broth) containing 9% salt in terms of log cfu ml⁻¹ over a period of 100 hours were obtained.

120 **2.2 FITTING THE LAMBERT-PEARSON MODEL TO MICROBIAL GROWTH INHIBITION DATA**

121 The time taken for a microbial culture to reach a specific optical density (also known as the
122 time to detection, TTD) in the presence of an inhibitor is dependent on the concentration
123 and dose response of that inhibitor. The Lambert-Pearson model (LPM, equation 4) [4]
124 describes the visual growth of a culture as an exponential decay function of the
125 concentration of the applied inhibitor. A plot of the log concentration against the relative
126 rate to detection (RRTD, the ratio of the time to detection of the uninhibited culture, or
127 positive control, to the time to detection of the test culture) gives a characteristic sinusoid,
128 with inflexion at $RRTD = 1/\exp(1)$. A linear extrapolation from this point to the log
129 concentration axis allows the estimation of the minimum inhibitory concentration (MIC,
130 equation 5), and a linear extrapolation to the $RRTD = 1$ axis allows the estimation of the
131 non-inhibitory concentration (NIC, equation 6), the concentration below which normal

132 visual growth is observed even in the presence of the inhibitor.

$$133 \quad RRTD = \begin{cases} \text{if} & [x] = 0, & 1 \\ \text{else if} & [x] < [P_1] \\ \text{then} & \exp\left(-\left(\frac{[x]}{P_1}\right)^{P_2}\right) \\ \text{else if} & \frac{1}{e}(1 - P_2(\ln[x] - \ln P_1)) < 0, & 0 \\ \text{else} & \frac{1}{e}(1 - P_2(\ln[x] - \ln P_1)) \end{cases} \quad (4)$$

134 Where RRTD = relative rate to detection, $[x]$ is the concentration of the given inhibitor, P_1
 135 is the concentration of inhibitor giving a relative inhibition of $1/e$, where e is the exponential
 136 of 1, and P_2 is a slope parameter which has been defined as the dose response due its
 137 similarity with the Hill model.

138 Two biologically important parameters can be obtained from the LPM; the minimum
 139 inhibitory concentration (MIC) and the non-inhibitory concentration (NIC) and are defined,
 140 respectively, as

$$141 \quad MIC = P_1 \exp\left(\frac{1}{P_2}\right) \quad (5)$$

$$142 \quad NIC = P_1 \exp\left(\frac{1-e}{P_2}\right) \quad (6)$$

143 Data from the growth inhibition of *Pseudomonas aeruginosa* (ATCC 15442) in the
 144 presence of chlorhexidine at 37°C was obtained using standard, published, methods
 145 (Lambert and Pearson 2000).

146

147 **2.3 FURTHER ANALYSIS OF COMPOUND ACTION POTENTIAL OF THE RODENT OPTIC** 148 **NERVE**

149 The compound action potential from the rodent optic nerve typically has three peaks
 150 (indicating the presence of three populations of axons with different conduction velocities)
 151 with a rapidly decaying transient or artefact from the initial stimulus. Originally this
 152 phenomenon was modelled using the sum of four Gaussian functions, one for each feature
 153 of the CAP (Brown 2006).

$$154 \quad CAP = \sum_{i=1}^4 \frac{A_i}{w_i \sqrt{\pi/2}} \text{Exp} \left\{ -2 \left(\frac{t-c_i}{w_i} \right)^2 \right\} \quad (7)$$

156 where A_i is the area under the curve, w_i the width at half the maximum amplitude and c_i is
 157 the latency to the maximum amplitude of peak i . This equation models the artefact as a
 158 Gaussian; a secondary model which has some desirable features as described in this
 159 report, models the artefact as a simple decay, modelling from the initial recording time of
 160 1.04 ms.

$$161 \quad CAP = D_1 \text{Exp}(-D_2(t - 1.04)) \sum_{i=1}^3 \frac{A_i}{w_i \sqrt{\pi/2}} \text{Exp} \left\{ -2 \left(\frac{t - c_i}{w_i} \right)^2 \right\} \quad (8)$$

162 Where D_1 and D_2 are parameters describing the decay of the initial stimulus.

163 **3 Program description and sample runs**

164 The method can be split into 3-stages: Stage 1 fits the given NL model to the observed
165 data using SOLVER. This generates the initial best fit parameters and the RMSE value of
166 the fit. Stage 2 uses the RMSE to generate a set of random numbers based on
167 $N(0, \text{RMSE})$, which are added to the predicted data from Stage 1. The NL model is then
168 applied to multiple virtual data sets simultaneously (using SOLVER) to generate multiple
169 values of best-fit parameters. In stage 3 these values are statistically analysed to provide
170 the mean of the best-fit parameters, their standard errors, 95% confidence intervals and
171 parameter correlations. The data can also be used to provide confidence intervals for
172 parameters calculated from the regressed parameters, which are often dependent on the
173 magnitude of the correlation between those parameters.

174 **3.1 NLR: EXCEL ANALYSIS OF THE GROWTH OF LISTERIA MONOCYTOGENES AT 30°C** 175 **IN HIGH SALT MEDIA**

176 **3.1.1 Monte-Carlo: Excel Analysis**

177 Stage 1. The initial part of the procedure generally follows that given previously [3] except
178 in this case the modified Gompertz model (1) was used. The data used and the initial NLR
179 are shown in [Figure 1](#). The sum of the squares of the errors (SSE, Cell E39) was
180 calculated using the inbuilt “SUMXMY2(data range 1,data range 2)”, where the first data
181 range was the modelled values and the second data range was the observed values. Cell
182 E41 divides this by the degrees of freedom (Cell E40) and takes the square root to give
183 the root mean square error (RMSE). This is the standard error of the curve fit. The

184 SOLVER Package was used to minimise this value, by changing the values of the four
185 parameters A, C, B and M. [Figure 2](#) shows a plot of the observed data and the modelled
186 function.

187

188 *Stage 2. Generation of random numbers with a distribution of $N(0, RMSE)$:* (It is assumed
189 the user has installed the Analysis ToolPak Add-In). On a separate worksheet, the
190 Random Number Generator was used to generate an array of 50 columns of 33 random
191 numbers using the RMSE of the fitted model as the standard deviation ([Figure 3](#)). The
192 number of columns used is set by the maximum number of values that SOLVER can
193 handle (200) divided by the number of parameters in the model. A random seed number of
194 2 was used for illustrative purposes as the use of this seed number will allow any reader to
195 recreate the exact procedure carried out here.

196

197 *Generation of virtual data:* The random data was added to the modelled log cfu ml⁻¹ data to
198 produce a set of 50 virtual observed data. On the spreadsheet these were conveniently
199 located below the random number array.

200

201 *Stage 3. Fitting multiple models simultaneously using SOLVER:* Below each set of virtual
202 data the modified Gompertz model was entered ([Figure 4](#)). The regression parameters
203 were placed below this: the initial parameters for each set used the parameters from the
204 initial model fit, although it is advised to check that the parameters do not represent a local
205 rather than the global minimum). The SSE between the virtual data and the modelled data
206 was calculated per data set (Cells C115 to AZ115). The SSE from each data set was

207 summed and this total value placed in cell C117; SOLVER was used to minimise this total
208 SSE by changing all the 200 parameters concurrently. This procedure gave 50 sets of
209 modelled parameters per run. A target SSE of 50 times the SSE obtained from the initial fit
210 (cell B118) was used to monitor the progress of the fitting.
211 After the minimization procedure, from the fitted parameters the mean and standard
212 deviations were found. The 95% confidence intervals were calculated from the 95%
213 percentiles using the syntax '=PERCENTILE(array, 0.025)', and '=PERCENTILE(array,
214 0.975)'. The correlation coefficient was found using the "CORREL(data range 1, data range
215 2)" function. [Tables 1a](#) and [1b](#) give the results of this MC analysis.

216

217 **3.1.2 Calculation of Biological Parameters**

218 The growth rate, maximum population density and the lag before the onset of growth are
219 important biological parameters and have to be calculated from the parameters obtained
220 from the fitting of the modified Gompertz. However, a singular problem is the calculation of
221 the confidence interval of the calculated parameter. For example the MPD = A + C, but the
222 variance is given by

$$Var(A + C) = Var(A) + Var(C) + 2Cov(AC)$$

223 Knowledge of the correlation between parameters allows the covariance to be calculated.
224 However in more complex cases such as the calculation of the growth rate (given by
225 BC/e), the calculation of the confidence interval becomes complex. The confidence
226 intervals for these biological parameters, however, can be estimated from the parameter
227 data of the MC analysis. For each parameter set the particular biological parameter was

228 calculated, giving 101 values (including the original fitting parameters).From these values
229 the required percentiles were calculated. In this particular case a re-parameterized
230 version of the modified Gompertz [5] was used to show that the ranges obtained by
231 running a non-linear analysis using JMP were equivalent to those obtained directly from
232 the MC analysis ([Table 2](#)).
233

234 4 Samples of Program Runs

235 4.1 INHIBITION OF PSEUDOMONAS AERUGINOSA BY CHLORHEXIDINE

236 The LPM (Eqn. 4) can be written in Excel as a nested series of IF-statements (e.g. see
237 [Figure 6](#)). Initial estimates for the regression parameters can be obtained from an analysis
238 of a plot of the chlorhexidine concentration against the RRTD ([Figure 5](#)). Using SOLVER
239 estimates for parameters P1 and P2 and the RMSE were obtained. The fitted parameters
240 and the RMSE were used to prime the MC analysis. The results of the Excel MC analysis
241 are given in [Table 3](#) and [Figure 6](#) shows the spreadsheet used.

242

243 A table (81x100) of random numbers based on $N(0, RMSE)$ was produced (cells C4 to
244 CX84). These random numbers were added to the modelled values (cells C88 to CX168),
245 and the non-linear fitting repeated (C172 to CX252) by regressing all 200 parameters at
246 once (Cells C255 to CX256). [Figure 6](#) shows a portion of the calculation performed. Cell
247 C259 sums the SSE for each regression performed (cells C258 to CX258). Cells B262 to
248 CX262 and C263 to CX263 calculate the MIC and NIC values from the regressed
249 parameters respectively.

250

251 4.2 FURTHER ANALYSIS OF COMPOUND ACTION POTENTIAL OF THE RODENT OPTIC

252 NERVE

253 The stimulus-evoked compound action potential from the mouse optic nerve was
254 successfully modelled using multiple Gaussian functions (Brown 2006). The original work
255 used four Gaussian functions to simulate the three peaks of the CAP and the brief stimulus
256 artefact (Eqn.7). This model was set up in Excel and the 16 parameters regressed. The
257 sum of squares obtained was 0.09987; the parameter values for Peaks 1,2 and 3 were
258 essentially identical to those published (nb., a typographical error in the publication gave
259 the area of peak 1 as 7.663, whereas it should have read 0.633). The estimated parameter
260 values for the artefact were, however, different from those published. The peak area found
261 in this analysis was 5.577 vs. 7.180 found previously and a calculated amplitude of 30.8
262 vs. 38.042.

263
264 The standard error of the fit (0.02986) was used to produce an array (96 x124) of normally
265 distributed random numbers. These values were added to the modelled data (on a
266 separate Excel sheet) to produce 96 virtual data sets (Cells DH5 to GY128), [Figure 7](#). (nb
267 columns and rows have been 'hidden' to show the full sheet). The model was added to
268 each cell M5 to DD128, the parameters were placed below each set of modelled data:
269 cells M132 to DD143. The calculated SSE between the modelled data and their respective
270 virtual data set was placed in cells M145 to DD145. The sum of these SSE was calculated
271 in Cell L146. Due to the SOLVER limit of 200 parameters, the MC analysis had to be done
272 in batches of 16. The results of the Monte-Carlo analysis (96 runs; 6 runs of 16) are given

273 in [Table 4a](#). The values obtained for the parameters are very similar to that previously
274 published, apart for the values for the artefact. The confidence interval for the area of the
275 artefact ranged from 4 to 10.7 and all correlations between parameters A4, w4 and c4 had
276 magnitudes greater than 0.994, suggesting that the model was over-parameterized.

277

278 A NLR analysis using the JMP statistical package gave an estimate for A4 of 5.5788 (95%
279 CI of 2.595 – 21.835), and correlations greater than 0.994 between A4, w4 and c4.

280

281 In a second Excel MC study, the artefact was modelled by a simple exponential decay,
282 replacing the Gaussian for the artefact by the function $D_1 \text{Exp}(-D_2(t-1.04))$. A similar MC
283 analysis was undertaken; the parameter estimates found for the three principal peaks
284 were relatively unchanged, but the parameter estimates for the artefact now had narrower
285 confidence intervals and the correlation between D1 and D2 = 0.246 ([Table 4b](#))

286 **4.3 CAVEATS TO USING MC ANALYSIS WITHIN EXCEL**

287 The MC analysis within Excel is initially primed using the parameter estimates from the
288 initial SOLVER minimisation procedure. It is possible that the estimates relate not to a
289 global minimum but to a local minimum, especially if there is a high degree of correlation
290 between given parameters (as was observed in the first analysis of the CAP data). One
291 method to overcome any such possibility is to use different initial parameter estimates for
292 each virtual dataset. This can be done using Excel's "RANDBETWEEN(a,b)" function,
293 where the user generates a scaled value of a parameter between two integer extrema (a
294 and b), and rescales to accommodate the desired magnitude of the initial parameter.

295

296 **4.4 COMPARISON OF THE EXCEL NLR AND MC ANALYSIS WITH *MATHEMATICA***

297 The non-linear regression capability of *JMP* was used to fit the three sets of sample data;
298 Tables 1, 2, 3, and 4c give the parameters, and their confidence intervals obtained using
299 this sophisticated software; Table 1b also compares the parameter correlations obtained
300 between the Excel MC and JMP NLR analysis. A small program was written to conduct a
301 Monte-Carlo analysis within *Mathematica* (Version 8) with the subsequent analysis of
302 10,000 virtual data sets (approx 2 to 5 minutes per 10,000 runs) for each of the fitted
303 models. The results of these MC analyses are also given in Tables 1a, 3, and 4d. The NLR
304 analysis using *JMP* essentially gave the same parameter estimates as that from the NLR
305 Excel analysis; the confidence intervals calculated using the Excel MC technique closely
306 agree with those calculated using the Hessian method within *JMP*. A comparison of the
307 results of the MC analysis of *Mathematica* also compares well with the Excel output. It
308 should be noted that the confidence intervals obtained from the MC analysis and the direct
309 NLR analysis using *JMP* do not completely agree with each other and this simply reflects
310 the differences in the techniques used. Each is equally correct.

311

312 The comparison between the parameter estimates and their confidence intervals as
313 obtained from the Excel MC analysis with either the direct NLR or the MC analyses using
314 *JMP* or *Mathematica* demonstrates the capacity of Excel to produce results equivalent to
315 those from more sophisticated packages.

316 **5 Hardware and software specification**

317

318 The method was carried out on a basic desktop computer with an AMD Phenom 9750
319 Quad core processor (2.4GHz), using Microsoft Excel 2007 under Windows 7. Non linear
320 regression comparisons were carried out using the JMP (4.0.4) Statistical Software (SAS
321 Institute Inc, Cary NC) and Monte-Carlo comparisons were carried out using *Mathematica*
322 Version 8.0.0.0 (Wolfram Research Inc, Champaign IL, USA).

323

324

325

326 **6 Program availability**

327 Spreadsheets with the worked examples are available from the author on request. The
328 *Mathematica* coding used to produce the NLR fits and Monte-Carlo simulations are also
329 available from the corresponding author.

330

331 Acknowledgments: we would like to acknowledge the financial support given by the Food
332 Standards Agency (UK) to Mr Ioannis Mytilinaios (post graduate scholarship PG 1025).

7 References

334

335 [1] A.M. Brown, A step-by-step guide to non-linear regression analysis of experimental
336 data using a Microsoft Excel spreadsheet. *Comp. Meth. Prog. Biomed.* 65 (2001) 191-200.

337 [2] w.w.w. Microsoft.com, Microsoft, Redmond, WA, USA.

338 [3] A.M. Brown, A non-linear regression analysis program for describing
339 electrophysiological data with multiple functions using Microsoft Excel, *Comp. Meth. Prog.*
340 *Biomed.* 83 (2006) 51 – 57.

341 [4] R.J.W. Lambert, and J. Pearson, Susceptibility testing: accurate and reproducible
342 minimum inhibitory concentration (MIC) and non-inhibitory concentration (NIC) values. *J.*
343 *Appl. Microbiol.*(2000) **88**, 784-790.

344

345 [5] M.H. Zwietering, I. Jongenburger, F.M. Rombouts and K. Van't Riet, Modeling the
346 Bacterial growth curve. *Appl. Environ. Microbiol.* (1990) **56**, 1875 – 1881.

347

348 [6] C. Robert and G. Casella, Monte-Carlo Statistical Methods. (2004) Springer texts in
349 Statistics, Springer (New York, NY USA), 2nd Edition, ISBN 0-387-21239-6.

350

351

352 Tables

353

354 **Table 1a.** Excel MC (100 iterations), JMP (non-linear regression) and *Mathematica* MC
 355 (10,000 iterations) analyses for the fitting of the modified Gompertz equation to *Listeria*
 356 *monocytogenes* growth data

Method	Parameter	Estimate	StdErr	LCL	UCL
Excel MC	A	3.948	0.0278	3.906	3.999
	C	4.896	0.0520	4.801	4.977
	B	0.095	0.0033	0.090	0.101
	M	30.318	0.2617	29.766	30.792
JMP NLR	A	3.951	0.0279	3.892	4.008
	C	4.900	0.0486	4.802	5.000
	B	0.095	0.0033	0.089	0.102
	M	30.341	0.2858	29.738	30.937
<i>Mathematica</i> MC	A	3.95	0.0279	3.895	4.0048
	C	4.9003	0.0481	4.807	4.995
	B	0.095	0.0033	0.089	0.102
	M	30.340	0.2833	29.776	30.882

357 Units: A and C: $\log_{10}\text{cfu ml}^{-1}$; B: 1/hr M: hr

358 **Table 1b. Parameter Correlation Table Obtained using Excel MC**
 359 **and JMP NLR analysis (brackets)**

	A	C	B	M
A	1			
C	-0.697 (-0.675)	1		
B	0.347 (0.357)	-0.658 (-0.650)	1	
M	0.546 (0.499)	-0.135 (-0.088)	0.0122 (0.161)	1

360

361

362 **Table 2. Calculation of Biological Parameters from the modified Gompertz using**
 363 **Excel MC, compared to the JMP NLR fitting of the re-parameterised modified**
 364 **Gompertz equation for the growth of *Listeria monocytogenes* at 30°C in high salt**
 365 **(9%).**

Method	Parameter	Estimate	StdErr	LCL	UCL
Excel MC	Growth rate	0.172	0.005	0.162	0.182
	Lag	19.806	0.470	18.733	20.589
	MPD	8.845	0.039	8.765	8.918
JMP NLR	Growth rate	0.171	0.005	0.162	0.182
	Lag	19.829	0.499	18.765	20.878
	MPD	8.850	0.036	8.778	8.924

366 Units: Growth rate: $\log_{10} \text{ cfu ml}^{-1} \text{ hr}^{-1}$; lag: hrs; MPD $\log_{10} \text{ cfu ml}^{-1}$
 367

368 **Table 3. Excel MC (100 iterations), JMP (non-linear regression) and *Mathematica***
 369 **MC(10,000 iterations) analyses for the fitting of the LPM to the inhibition of**
 370 ***Pseudomonas aeruginosa* in the presence of chlorhexidine.**

Method	Parameter	Estimate	StdErr	LCL	UCL
Excel MC	P1	7.265	0.093	7.109	7.455
	P2	1.150	0.022	1.107	1.187
JMP NLR	P1	7.257	0.085	7.092	7.430
	P2	1.149	0.019	1.114	1.185
<i>Mathematica</i> MC	P1	7.257	0.085	7.095	7.427
	P2	1.149	0.019	1.113	1.186
Excel MC	MIC	17.34	0.428	16.61	18.18
	NIC	1.630	0.044	1.550	1.709
<i>Mathematica</i> MC	MIC	17.333	0.372	16.62	18.09
	NIC	1.626	0.037	1.554	1.698

371 Units; P1, MIC and NIC (mg/l); P2 dimensionless.
 372

373

374

Table 4a. Excel MC (96 iterations) analysis of CAP data (Eqn. 7)

Peak	Parameter	Estimate	Stdev	LCL	UCL
Pk1	A1	0.633	0.007	0.620	0.648
	w1	0.243	0.002	0.238	0.247
	c1	1.397	0.001	1.394	1.398
Pk2	A2	1.552	0.017	1.517	1.579
	w2	0.418	0.004	0.410	0.423
	c2	1.875	0.001	1.872	1.877
Pk3	A3	1.288	0.013	1.262	1.313
	w3	0.609	0.006	0.598	0.621
	c3	2.566	0.004	2.559	2.574
Artefact	A4	8.214	2.675	4.002	10.697
	w4	0.152	0.009	0.135	0.162
	c4	0.883	0.021	0.863	0.920

375

376

Table 4b: Excel MC analysis (96 iterations) of CAP data (Eqn. 8)

Peak	Parameter	Estimate	Stdev	LCL	UCL
Pk1	A1	0.621	0.008	0.606	0.636
	w1	0.238	0.002	0.234	0.243
	c1	1.397	0.001	1.394	1.399
Pk2	A2	1.564	0.019	1.522	1.595
	w2	0.421	0.005	0.411	0.427
	c2	1.874	0.002	1.871	1.877
Pk3	A3	1.283	0.015	1.254	1.313
	w3	0.607	0.007	0.596	0.621
	c3	2.568	0.004	2.559	2.577
Artefact	D1	4.996	0.034	4.928	5.058
	D2	30.906	0.194	30.447	31.455

377

378

379

380

Table 4c: JMP NLR analysis of CAP data (Eqn. 8)

Peak	Parameter	Estimate	ApproxStdErr	LCL	UCL
Pk1	A1	0.620	0.009	0.603	0.637
	w1	0.238	0.003	0.233	0.243
	c1	1.396	0.001	1.394	1.399
Pk2	A2	1.566	0.019	1.529	1.603
	w2	0.421	0.004	0.413	0.430
	c2	1.874	0.002	1.871	1.877
Pk3	A3	1.281	0.015	1.252	1.312
	w3	0.607	0.007	0.593	0.621
	c3	2.568	0.004	2.560	2.576
Artefact	D1	4.997	0.034	4.931	5.063
	D2	30.907	0.414	30.124	31.712

381

382

383

Table 4d: *Mathematica* MC (10,000 iterations) analysis of CAP data (Eqn. 8)

Peak	Parameter	Estimate	ApproxStdErr	LCL	UCL
Pk1	A1	0.621	0.008	0.603	0.636
	w1	0.238	0.002	0.233	0.242
	c1	1.397	0.001	1.394	1.399
Pk2	A2	1.564	0.018	1.530	1.601
	w2	0.421	0.004	0.413	0.430
	c2	1.874	0.016	1.871	1.877
Pk3	A3	1.282	0.015	1.253	1.312
	w3	0.607	0.007	0.594	0.621
	c3	2.568	0.004	2.560	2.576
Artefact	D1	4.997	0.034	4.931	5.063
	D2	30.925	0.413	30.141	31.752

384

385 **Legends to Figures**

386

387 Figure 1. Non-linear regression analysis of the observed growth data for *Listeria*

388 *monocytogenes* at 30°C in 9% salt from an initial cellular density of 7.586×10^3 cfu ml⁻¹.

389 The growth was monitored over a 90 hour period. The observed numbers as their decimal
390 log were modelled using the standard modified Gompertz equation [5].

391 Figure 2. Plot of the growth data of *L. monocytogenes* at 30°C in 9% salt (symbols) with
392 the fitted modified-Gompertz model (line).

393 Figure 3. A portion of the (50 x 33) random number array generated using the random
394 number feature of Excel's Analysis Addin with a distribution of N(0, 0.09047), using a
395 random seed number of 2.

396 Figure 4. A portion of the (50 x 33) NLR array; column B reproduces the NLR fitting of the
397 original data, with cells B111 to B114 reproducing the regressed parameters. Well C77
398 shows the syntax used for the formula, which is reproduced over the array. Cells C115 to
399 AZ115 calculate the SSE between the modelled data and the respective virtual data set.
400 Cell B117 sums all the 50 individual SSE values. This value is then minimised using the
401 SOLVER utility.

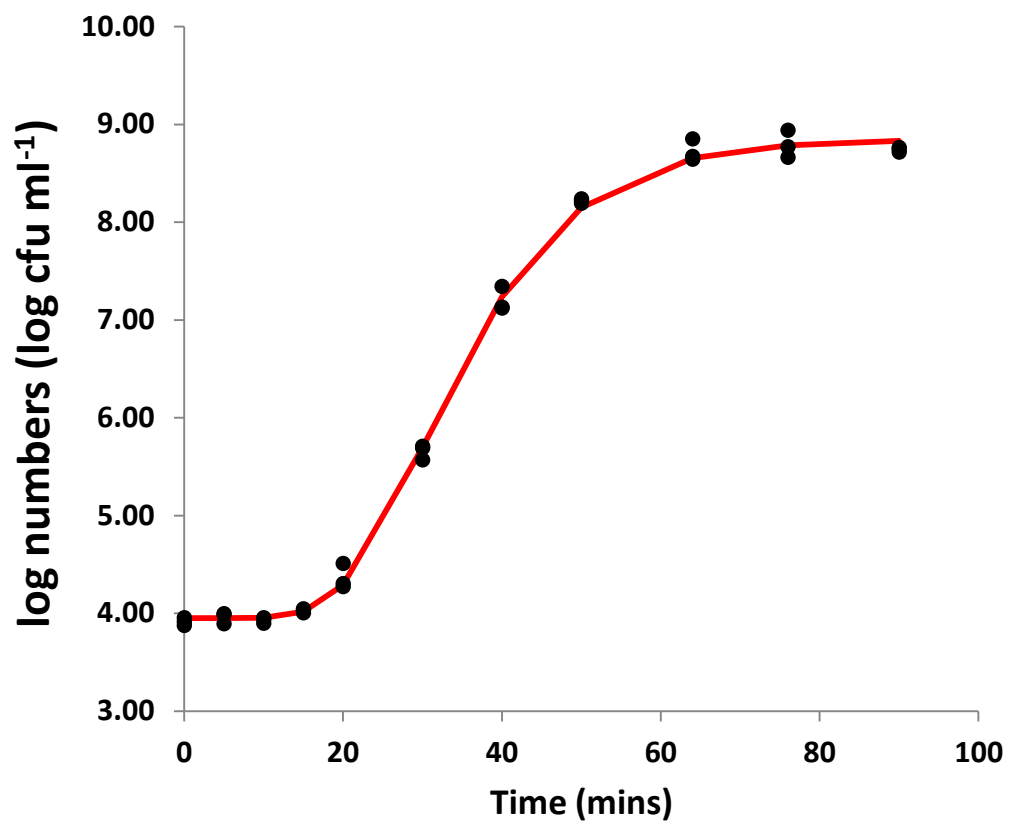
402 Figure 5. A plot of chlorhexidine concentration (mg l⁻¹) against the observed relative rate to
403 detection for the growth of *Pseudomonas aeruginosa* (ATCC 15442) in TSB at 37°C
404 (symbols) and the fitted NLR model (line).

405 Figure 6. Spreadsheet used for the Excel MC analysis of the fitting of the LPM to the data
406 for the growth inhibition of *Ps.aeruginosa* by chlorhexidine.

407 Figure 7. Spreadsheet used for the Excel MC analysis of the fitting of a multiple Gaussian
408 function (Eqn. 7) to CAP data.
409

412

Page 28



413

414

415

416 Figure 2

417

	A	B	C	D	E	F	G	H
1								
2	RMSE	0.090747						
3								
4								
5		no	1	2	3	4	5	6
6		1	-0.27177	0.112094	0.057959	0.009818	0.116198	0.022917399
7		2	-0.0408	0.073162	0.027441	0.01999	-0.00484	-0.016152197
8		3	0.013121	0.005853	-0.04004	-0.12229	0.121177	0.08340568
9		4	-0.13812	-0.14519	-0.03006	0.063873	-0.0022	-0.025016443
10		5	0.074846	0.033108	0.040228	-0.01283	-0.03762	-0.09728479
11		6	-0.11262	-0.0679	0.155139	0.038026	0.108661	0.031230421
12		7	-0.03568	-0.14927	0.053359	-0.12548	-0.07223	0.0336726
13		8	0.154248	0.00085	-0.0747	-0.14346	-0.09552	0.118928897
14		9	0.084781	-0.06085	-0.0096	-0.10539	0.011188	0.04305846
15		10	-0.05888	0.084385	-0.10152	0.116863	0.113325	-0.030171924
16		11	-0.06138	0.070147	0.207505	0.022917	0.031732	0.042074346
17		12	-0.12809	0.049439	0.01095	-0.09689	0.029607	0.245178551
18		13	-0.00068	-0.00305	0.086133	-0.05195	0.056966	0.036319875
19		14	-0.07968	0.219385	0.176514	0.026766	-0.08762	0.099635975
20		15	0.127024	0.04708	0.056282	0.072245	0.004506	0.005210754

Figure 3

... X ✓ fx =C\$111+C\$112*EXP(-1*EXP(C\$113*(C\$114-\$A77)))								
	A	B	C	D	E	F	G	
76	Time/hr	MOD	1	2	3	4	5	
77	0	3.951	=C\$111+C	3.94	3.97	3.92	3.97	
78	0	3.951	3.91	3.94	3.97	3.92	3.97	
79	0	3.951	3.91	3.94	3.97	3.92	3.97	
80	5	3.951	3.91	3.94	3.97	3.92	3.97	
81	5	3.951	3.91	3.94	3.97	3.92	3.97	
82	5	3.951	3.91	3.94	3.97	3.92	3.97	
83	10	3.955	3.91	3.95	3.98	3.92	3.98	
84	10	3.955	3.91	3.95	3.98	3.92	3.98	
85	10	3.955	3.91	3.95	3.98	3.92	3.98	
86	15	4.017	3.98	4.03	4.06	4.00	4.04	
87	15	4.017	3.98	4.03	4.06	4.00	4.04	
88	15	4.017	3.98	4.03	4.06	4.00	4.04	
89	20	4.288	4.29	4.33	4.35	4.30	4.32	
90	20	4.288	4.29	4.33	4.35	4.30	4.32	
91	20	4.288	4.29	4.33	4.35	4.30	4.32	
92	30	5.695	5.77	5.69	5.71	5.69	5.74	
93	30	5.695	5.77	5.69	5.71	5.69	5.74	
94	30	5.695	5.77	5.69	5.71	5.69	5.74	
95	40	7.238	7.32	7.18	7.18	7.21	7.26	
96	40	7.238	7.32	7.18	7.18	7.21	7.26	
97	40	7.238	7.32	7.18	7.18	7.21	7.26	
98	50	8.150	8.20	8.11	8.09	8.13	8.15	
99	50	8.150	8.20	8.11	8.09	8.13	8.15	
100	50	8.150	8.20	8.11	8.09	8.13	8.15	
101	64	8.655	8.67	8.66	8.62	8.67	8.64	
102	64	8.655	8.67	8.66	8.62	8.67	8.64	
103	64	8.655	8.67	8.66	8.62	8.67	8.64	
104	76	8.787	8.79	8.82	8.77	8.81	8.76	
105	76	8.787	8.79	8.82	8.77	8.81	8.76	
106	76	8.787	8.79	8.82	8.77	8.81	8.76	
107	90	8.833	8.83	8.88	8.83	8.87	8.81	
108	90	8.833	8.83	8.88	8.83	8.87	8.81	
109	90	8.833	8.83	8.88	8.83	8.87	8.81	
110								
111	A	3.951	3.907	3.944	3.973	3.917	3.973	
112	C	4.900	4.940	4.959	4.876	4.973	4.852	
113	B	0.095	0.097	0.090	0.090	0.092	0.095	
114	M	30.341	29.726	30.463	30.383	30.321	30.115	
115	sse	0.238816	0.253706	0.254936	0.303308	0.211335	0.115958	0.
116								
117	Tot SSE	11.55302						
118	Targ	11.94079						
119								

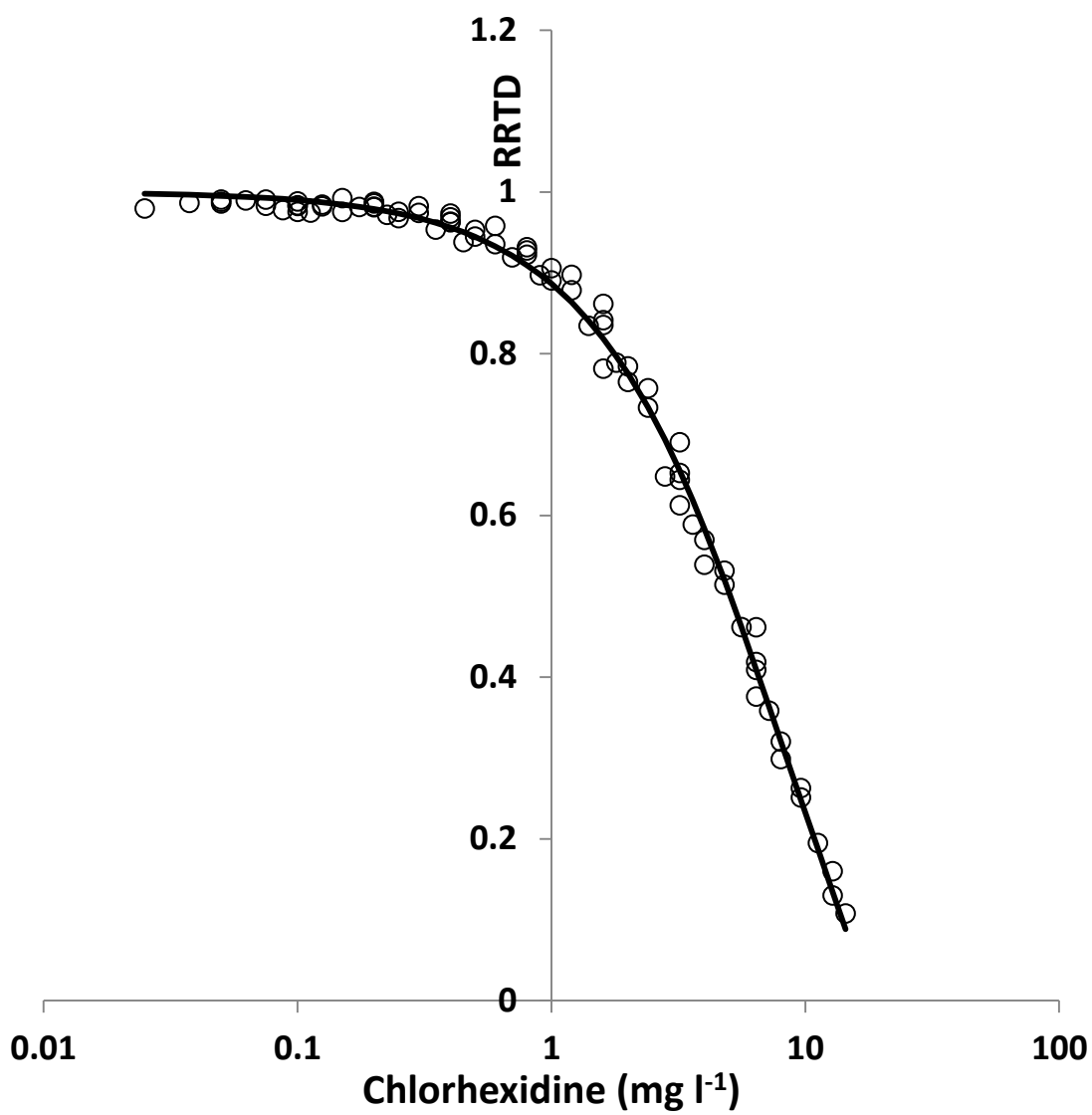
423

424

425

426 Figure 4

427



428

429

430

431 Figure 5

432

	A	B	C	D	E	F	CW	CX
1								
2	RMSE	0.0218361						
3		obs	1	2	3	4	99	100
4		1	-0.0660	0.0035	-0.0189	0.0191	0.0114	0.0003
5		2	-0.0229	0.0358	-0.0234	0.0286	0.0112	-0.0121
6		3	-0.0035	-0.0160	0.0043	0.0018	0.0172	0.0013
82		79	-0.0021	0.0490	0.0411	0.0336	-0.0062	0.0169
83		80	0.0488	-0.0065	-0.0240	0.0045	-0.0028	0.0234
84		81	-0.0254	-0.0056	-0.0149	-0.0001	0.0126	-0.0125
85								
86								
87	model	Obs	1	2	3	4	99	100
88	1	1	0.93399	1.00350	0.98109	1.01907	1.01144	1.00026
89	0.99764	2	0.97477	1.03347	0.97421	1.02628	1.00886	0.98558
166	0.12804	79	0.12594	0.17699	0.16915	0.16167	0.12181	0.14490
167	0.12804	80	0.17683	0.12153	0.10406	0.13250	0.12528	0.15140
168	0.07826	81	0.05288	0.07265	0.06337	0.07815	0.09086	0.06571
169								
170								
171	Conc	model	1	2	3	4	99	100
172	0	1	=IF(\$A172=0,1,IF	1	1	1	1	1
173	0.0375	0.99764	0.99797	0.99771	0.99761	0.99721	0.99773	0.99758
174	0.05	0.99672	0.99716	0.99681	0.99668	0.99616	0.99684	0.99663
251	12.8	0.12804	0.12633	0.12472	0.12574	0.13950	0.12756	0.12732
252	14.4	0.07826	0.07540	0.07465	0.07601	0.09119	0.07749	0.07773
253								
254		Mod	1	2	3	4	99	100
255	P1	7.1213	7.3217	7.2238	7.2140	7.3341	7.2728	7.2288
256	P2	1.0796	1.1754	1.1554	1.1478	1.1148	1.1556	1.1444
257								
258		SSE	0.044169068	0.05177825	0.045781245	0.040551295	0.03294538	0.04006947
259		Tot SSE	3.78178					
260		Target	2.18361					
261								
262	MIC	17.9824	17.1433	17.1647	17.2400	17.9861	17.2793	17.3200
263	NIC	1.4498	1.6972	1.6327	1.6145	1.5701	1.6441	1.6107

Figure 6

$=(M\$132/(M\$133*\text{SQRT}(\text{PI}()/2)))*\text{EXP}(-2*((\$C5-M\$134)^2/M\$133^2))+((M\$135/(M\$136*\text{SQRT}(\text{PI}()/2)))*\text{EXP}(-2*((\$C5-M\$137)^2/M\$136^2)))+(M\$138/(M\$139*\text{SQRT}(\text{PI}()/2)))*\text{EXP}(-2*((\$C5-M\$140)^2/M\$139^2))$														
	J	K	L	M	N	DC	DD	DE	DF	DG	DH	DI	GX	GY
1														
2		MC analysis												
3				Modelled Data							Virtual data sets			
4		time	Modelled	MC1	MC2	MC95	MC96		time	Modelled	MC1	MC2	MC95	MC96
5		1.04	4.9713675	4.96509	4.94832	4.99872			1.04	4.9713675	4.8811	4.97615	4.94275	4.98347
6		1.06	2.8469856	2.83839	2.83586	2.79671	2.81468		1.06	2.8469856	2.84638	2.80143	2.81095	2.84357
126		3.46	0.0226895	0.02001	0.02097	0.0236	0.02328		3.46	0.0226895	0.0555	0.00527	0.06119	0.01604
127		3.48	0.0186681	0.01635	0.01718	0.01946	0.01919		3.48	0.0186681	0.02206	0.01389	0.01855	-0.03865
128		3.5	0.0152933	0.0133	0.01401	0.01598	0.01575		3.5	0.0152933	-0.04361	0.00629	0.00143	-0.02886
129														
130				Parameter Estimates										
131		Parameter	MOD	MC1	MC2	MC95	MC96							
132	Pk1	A1	0.6335031	0.63389	0.62065	0.63325	0.6421							
133		w1	0.2428981	0.24216	0.2401	0.24353	0.24118							
134		c1	1.3963352	1.39583	1.39741	1.39594	1.3978							
135	Pk3	A2	1.5527126	1.5761	1.57711	1.54235	1.536							
136		w2	0.4182147	0.42184	0.42183	0.418	0.41376							
137		c2	1.8745919	1.87693	1.87453	1.8731	1.8741							
138	pK2	A3	1.2868928	1.26079	1.26732	1.30255	1.29881							
139		w3	0.6086458	0.59643	0.60091	0.61345	0.61288							
140		c3	2.5666363	2.57194	2.57021	2.56324	2.56287							
141	Artefact	A4	5.5773766	4.58378	5.3543	10.6411	10.641							
142		w4	0.14447	0.14202	0.14304	0.15886	0.158							
143		c4	0.90182	0.91032	0.90432	0.8665	0.86763							
144														
145		SSE	0.09986	0.08157	0.07409	0.09013	0.09645							
146		Tot SSE	9.759271											

Figure 7



DSPACE

<https://dspace.org/>

Monte Carlo simulation of parameter confidence intervals for non-linear regression analysis of biological data using Microsoft Excel.

Lambert, Ronald J. W.; Mytilinaios, Ioannis; Maitland, Luke; Brown, Angus M.

2012-06-18T00:00:00Z

Elsevier Science B.V., Amsterdam.

<http://dx.doi.org/10.1016/j.cmpb.2011.05.009>

Ronald J.W. Lambert, Ioannis Mytilinaios, Luke Maitland and Angus M. Brown. Monte Carlo simulation of parameter confidence intervals for non-linear regression analysis of biological data using Microsoft Excel. Computer Methods and Programs in Biomedicine, Volume 107, Issue 2, August 2012, Pages 155-163.