

SUPPLEMENTARY INFORMATION

Quantifying antibody kinetics and RNA shedding during early-phase SARS-CoV-2 infection

^{1,2,3}*Borremans Benny*, ¹*Gamble Amandine*, ¹*Prager KC*, ¹*Helman Sarah K*, ⁴*McClain Abby M*, ¹*Cox Caitlin*,
^{1,5}*Savage Van*, ¹*Lloyd-Smith James O*

Affiliations

¹Ecology and Evolutionary Biology Department, University of California Los Angeles, United States;
²Interuniversity Institute for Biostatistics and Statistical Bioinformatics (I-BIOSTAT), Universiteit Hasselt,
Belgium; ³Evolutionary Ecology Group, Universiteit Antwerpen, Belgium; ⁴National Marine Mammal
Foundation, San Diego, United States; ⁵Biomathematics Department, University of California Los Angeles,
United States.

bennyborremans@gmail.com; jlloydsmith@ucla.edu

Seroconversion time distributions

We use a stepwise bootstrap procedure to integrate all types of data that contain useful information about the timing of seroconversion of different antibodies in days post symptom onset (dpo). The first step uses only the best available resolution of seroconversion data (i.e. reported for exact days, as opposed to a range of days) to bootstrap a distribution of observed possible seroconversion times. For this first step, bootstrapping is done through repeated random sampling of individual seroconversion times from the dataset. By doing this many times (50,000), a random sampling procedure is simulated, resulting in an estimate of the distribution of seroconversion times. Some studies report seroconversion times for groups of individuals simultaneously. In this case, each individual group member is treated as a separate individual that can be sampled randomly. After this first bootstrapping step, a normal distribution is fit to

the observed sampled distribution, and this fitted normal distribution is used as prior information for the next bootstrapping step.

The following steps includes all data for which the maximum reported seroconversion time range is the next one observed in the data (for up to maximum time range present in a dataset). For example, if a number of results are reported not as an exact time but as a period ranging 3 days (reported as such or as part of a time series), the data included in step 2 consist of results reported as exact days, and results reported as 3-day ranges. Bootstrapping in this case is again done through repeated random sampling of an individual. When that individual has a result reported as an exact time, that time is stored as a bootstrap sample. When that individual has a result reported as a time range, a time within that range is sampled, but importantly the times within that range do not have the same probability of being sampled. This probability is determined by the normal distribution that was estimated after the preceding bootstrapping step. This ensures that the best available data have the largest contribution to the analysis, and data of lower resolution are used while taking into account the information contained in the higher resolution data. This stepwise procedure continues until data of all resolutions (i.e. including the largest reported seroconversion time periods) has been bootstrapped.

Seroconversion times are sometimes reported as a mean time (\pm error) instead of an exact time or time period. In these cases, the standard deviation of time around the mean is calculated (using reported sample size and standard error), and a random time is drawn from this normal distribution.

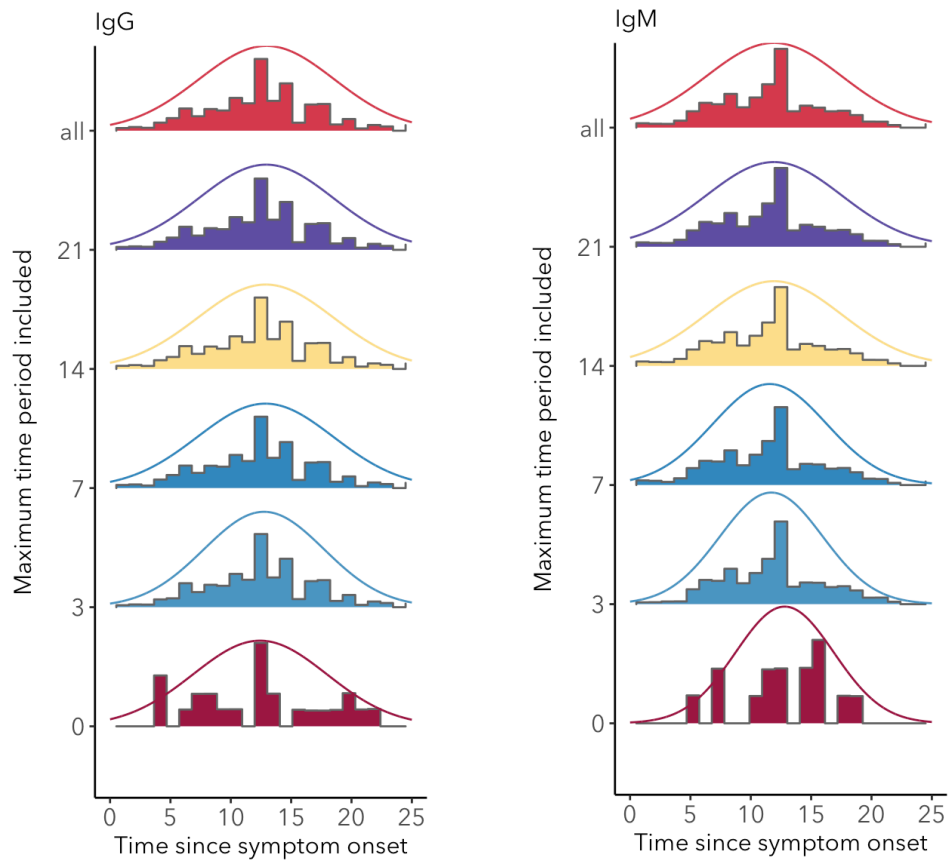


Figure S1. Distributions of IgG and IgM seroconversion times (including all assays) for increasing levels of data inclusion, from exact time data only (time period = 0) to the inclusion of the longest reported time periods (all).

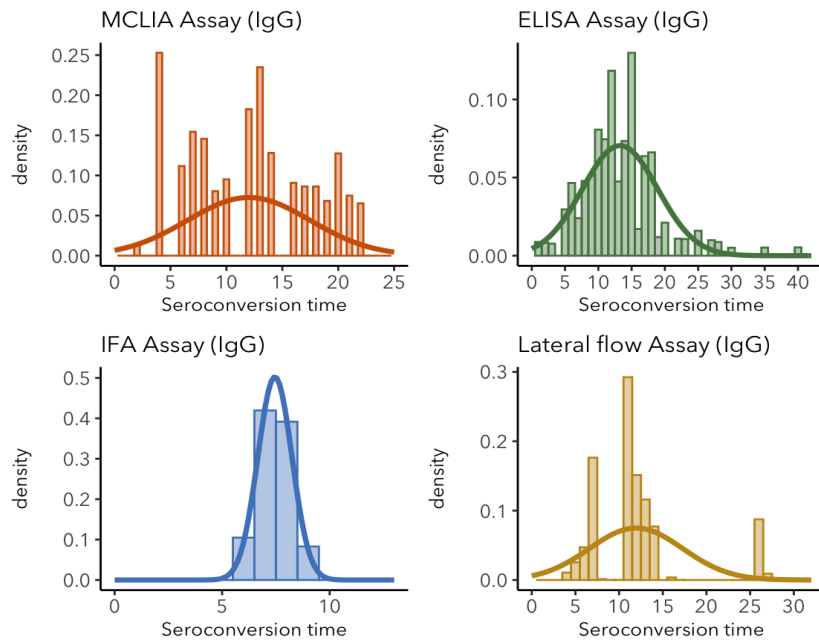


Figure S2. IgG seroconversion time distributions for different assays. Sparse distributions are the result of low sample sizes in the study. These sparse results are included for informative reasons, and should not be interpreted as representative of the real distribution.

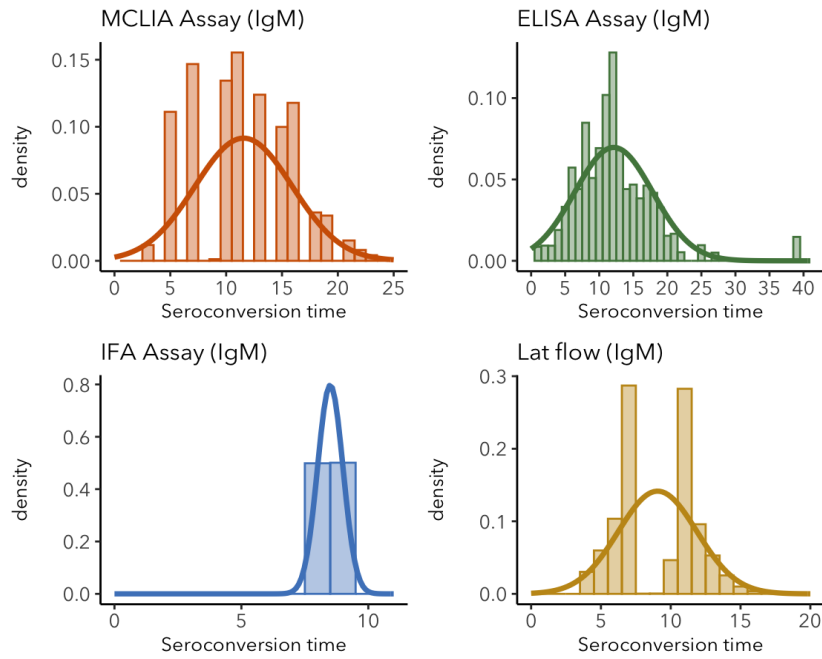


Figure S3. IgM seroconversion time distributions for different assays. Sparse distributions are the result of low sample sizes in the study. These sparse results are included for informative reasons, and should not be interpreted as representative of the real distribution.

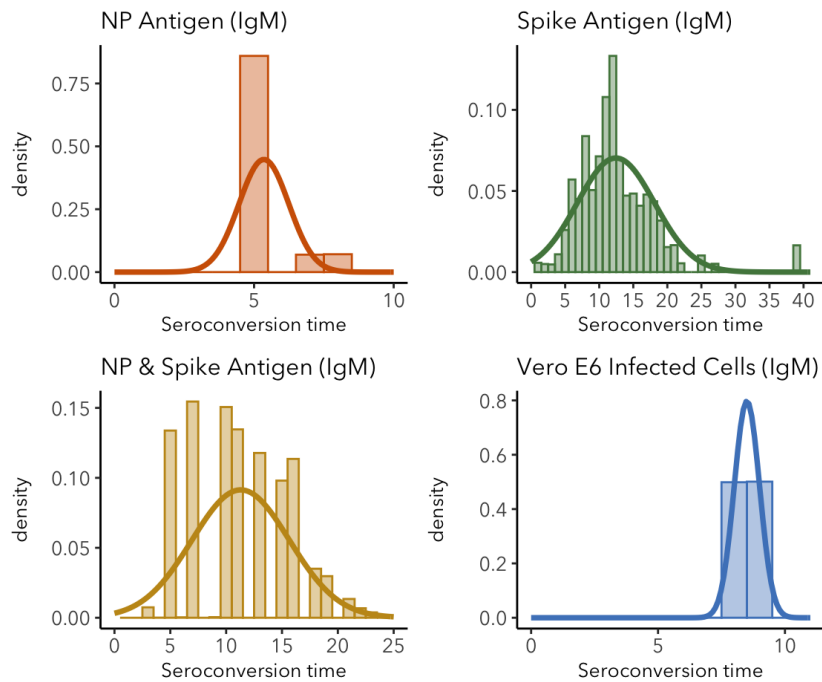


Figure S4. IgG seroconversion time distributions for different target antigens. Sparse distributions are the result of low sample sizes in the study. These sparse results are included for informative reasons, and should not be interpreted as representative of the real distribution.

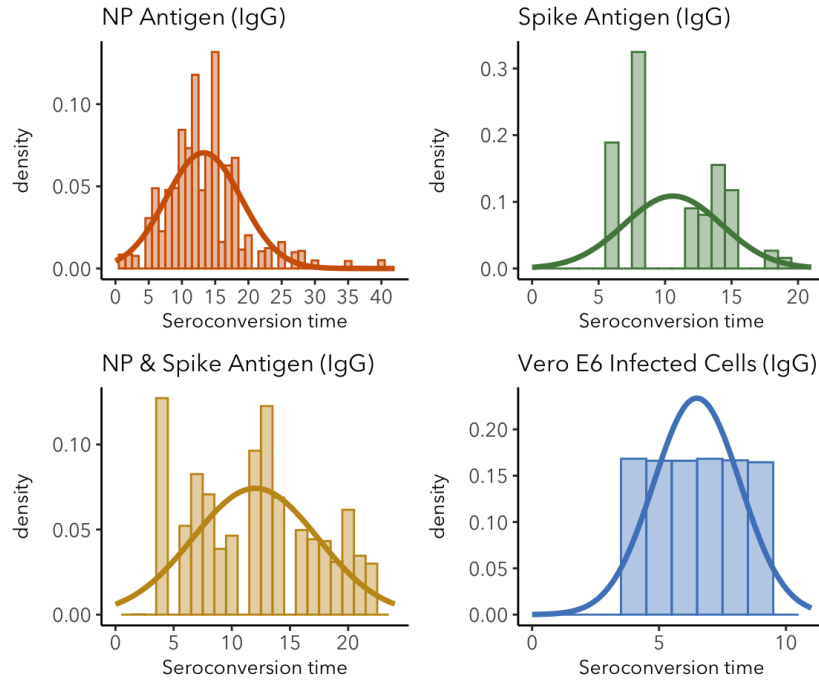


Figure S5. IgM seroconversion time distributions for different target antigens. Sparse distributions are the result of low sample sizes in the study. These sparse results are included for informative reasons, and should not be interpreted as representative of the real distribution.

Table S1. Fitted normal distribution parameters for different assays. SD: standard deviation; “N too low” indicates a sample size too small to compute a mean and SD.

Antibody	Assay	Mean	SD
IgG	ELISA (any antigen)	13.3	5.7
IgG	ELISA-NP	13.3	5.7
IgG	ELISA-Spike	10.5	4.1
IgG	MCLIA	12.0	5.5
IgG	LFIA	11.9	5.3
IgG	IFA	7.5	0.8
IgM	ELISA (any antigen)	12.1	5.7
IgM	ELISA-NP	<i>N too low</i>	<i>N too low</i>
IgM	ELISA-Spike	12.6	5.8
IgM	MCLIA	11.6	4.4
IgM	LFIA	9.1	2.8
IgM	IFA	8.5	0.5
IgG	Mild/Moderate (ELISA-NP)	12.9	5.0
IgG	Severe/Critical (ELISA-NP)	15.5	6.8
IgM	Mild/Moderate (ELISA-Spike)	12.3	5.8
IgM	Severe/Critical (ELISA-Spike)	13.2	5.0

Detection probability

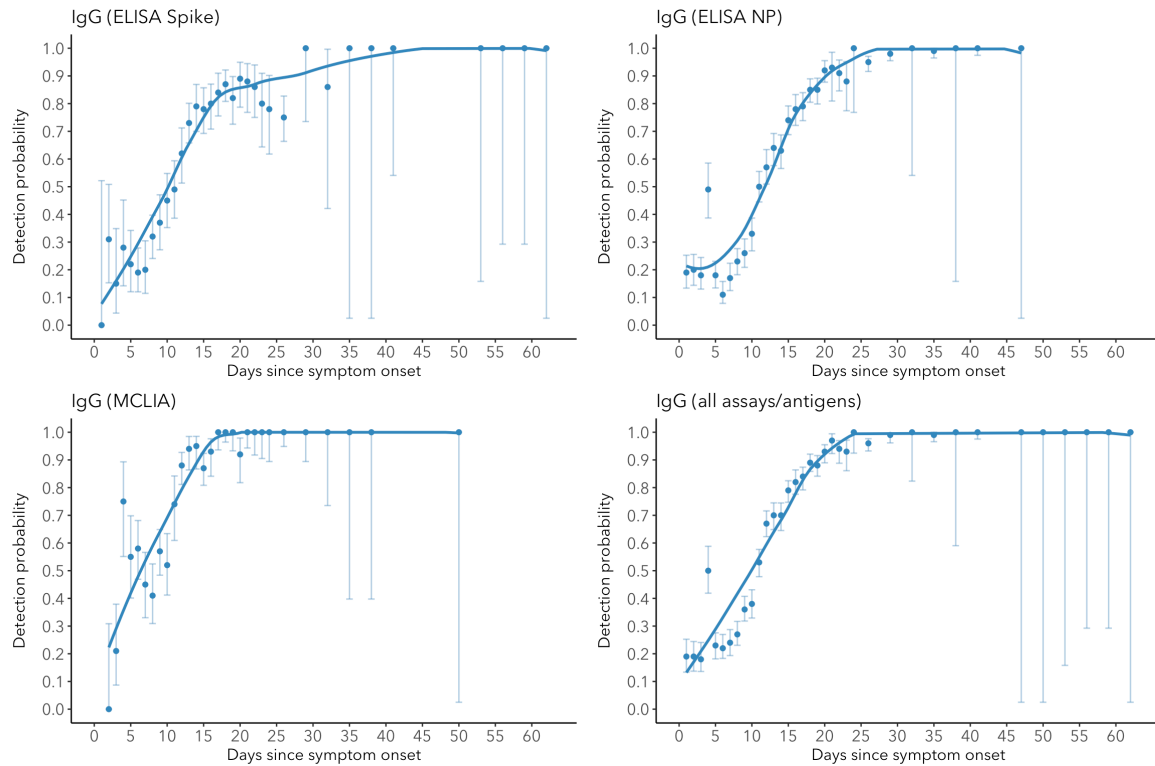


Figure S6. IgG detection probability for different assays/antigens.

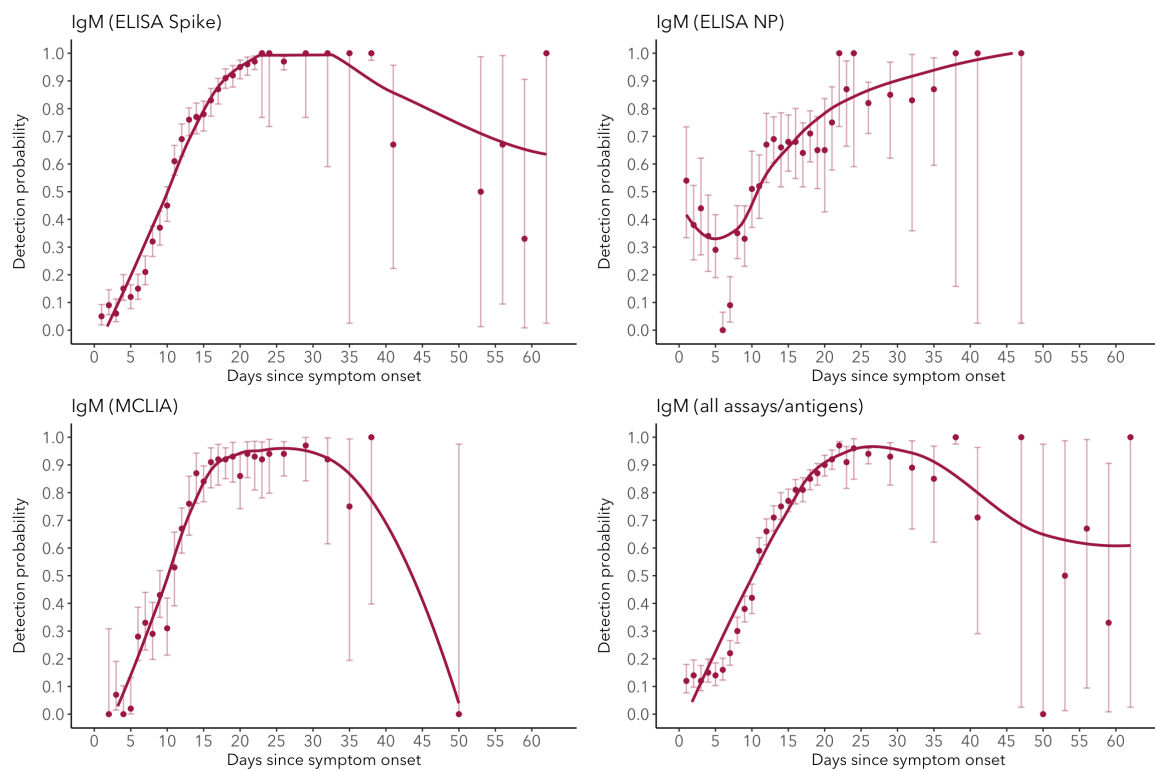


Figure S7. IgM detection probability for different assays/antigens.

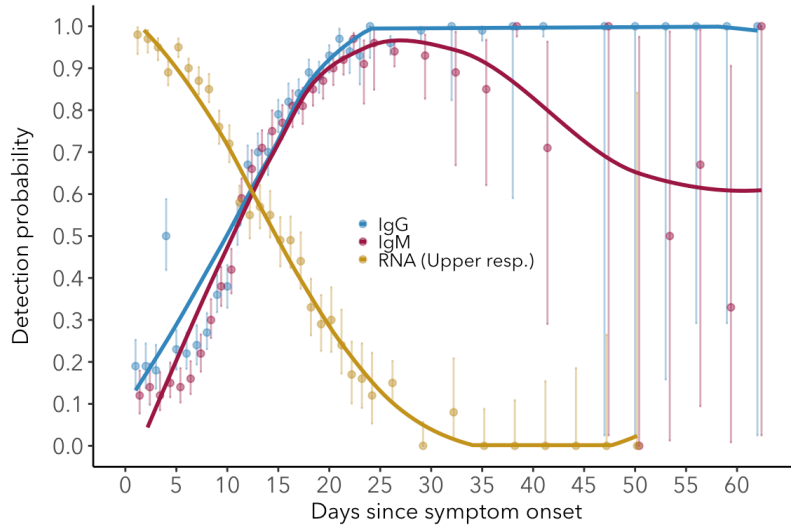


Figure S8. Detection probability for IgG (ELISA-NP), IgM (ELISA-Spike) and RNA in upper respiratory tract samples.

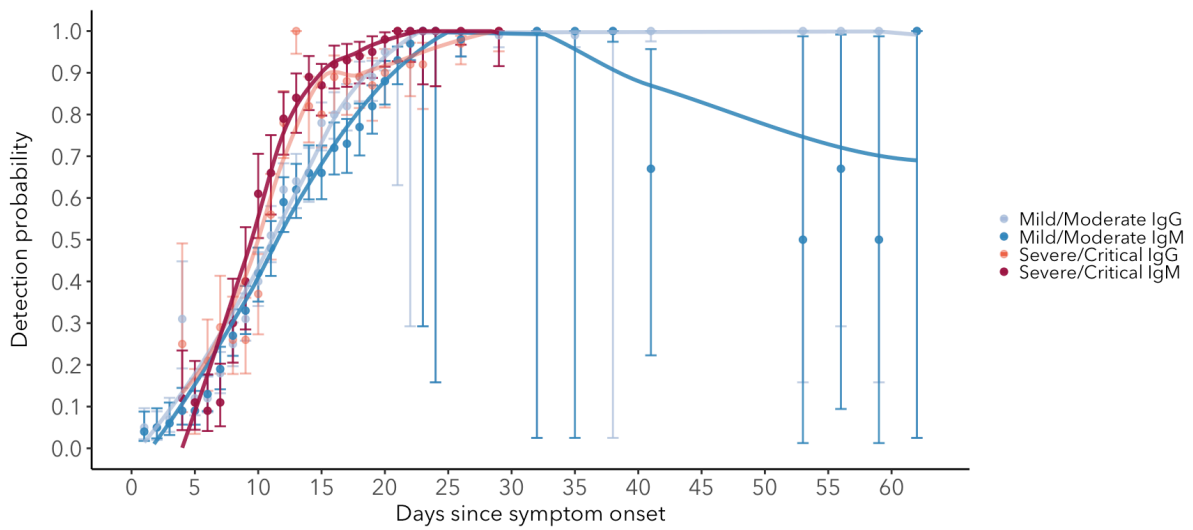


Figure S9. Detection probability for Mild/Moderate and Severe/Critical cases.

Table S2. IgG - ELISA detection probability

N: interpolated sample size.

Day after symptom onset	IgG	
	Percentage positive	N
1	19	181
2	19	220
3	18	228
4	50	143
5	23	323
6	22	379
7	24	336
8	27	430
9	36	467
10	38	364
11	53	413
12	67	413
13	70	368
14	70	346
15	79	448
16	82	311
17	84	335
18	89	410
19	88	314
20	93	274
21	97	111
22	94	160
23	93	100
24	100	47
26	96	361
29	99	225
32	100	19
35	99	161
38	100	7
41	100	148
47	100	1
50	100	1
53	100	2
56	100	3
59	100	3
62	100	1

Table S3. IgM – ELISA detection probability

N: interpolated sample size.

Day after symptom onset	IgM	
	Percentage positive	N
1	12	173
2	14	212
3	12	224
4	15	312
5	14	299
6	16	357
7	22	352
8	30	391
9	38	430
10	42	344
11	59	422
12	66	411
13	71	405
14	75	321
15	77	429
16	81	330
17	81	332
18	85	404
19	87	307
20	90	265
21	92	251
22	97	262
23	91	67
24	96	45
25	94	329
26	93	56
27	89	19
28	85	20
29	100	148
30	71	7
47	100	1
50	0	1
53	50	2
56	67	3
59	33	3
62	100	1

Table S4. RNA – Upper respiratory tract det. prob.
N: interpolated sample size.

RNA upper respiratory tract		
Day after symptom onset	Percentage positive	N
1	98	106
2	97	182
3	95	300
4	89	484
5	95	398
6	90	433
7	87	432
8	85	448
9	76	460
10	72	413
11	58	543
12	55	383
13	57	370
14	55	324
15	49	292
16	49	268
17	44	220
18	33	198
19	29	190
20	30	145
21	24	132
22	17	109
23	16	96
24	12	67
26	15	222
29	0	63
32	8	39
35	0	40
38	0	32
41	0	22
44	0	18
47	0	12
50	0	2

Table S5. RNA – Lower respiratory tract det. prob.
N: interpolated sample size.

RNA lower respiratory tract		
Day after symptom onset	Percentage positive	N
2	100	13
3	100	50
4	100	161
5	100	123
6	100	134
7	100	127
8	100	125
9	95	114
10	92	107
11	93	170
12	85	88
13	96	81
14	89	72
15	91	64
16	92	59
17	71	52
18	91	47
19	95	40
20	9	31
21	70	23
22	78	77
23	100	6
24	0	9
26	61	18
29	0	2

Table S6. RNA – Faeces detection probability

N: interpolated sample size.

Day after symptom onset	RNA faeces	
	Percentage positive	N
3	100	6
4	85	13
5	88	16
6	91	66
7	75	81
8	91	87
9	87	82
10	88	73
11	90	77
12	88	74
13	89	73
14	90	62
15	87	52
16	80	51
17	80	44
18	77	47
19	61	41
20	43	35
21	37	30
22	27	22
23	39	28
24	14	14
26	16	43
29	12	26
32	0	10
35	0	6
38	0	6
41	0	6
44	0	6
47	0	2

Titer kinetics

We estimate important early-phase titer kinetic characteristics (increase rate and start timing, peak titer level and timing) by fitting a Gompertz growth rate function to different subsets of data. We choose the following three-parameter Gompertz function,

$$y(t) = ae^{-be^{-ct}},$$

as its parameters correspond with relevant titer kinetic characteristics, asymptote a (peak titer level), displacement parameter b (start of the increase phase) and growth rate c (increase rate).

Parameter fitting is done using a Bayesian Monte Carlo Markov Chain approach, with Metropolis-Hastings acceptance/rejection algorithm. Proposed values are generated according to a normal distribution with the mean centered on the previously accepted values and standard deviation optimized to obtain an overall acceptance rate between 0.2 and 0.5. Growth rate c is estimated at the population level (i.e. all individuals are assumed to have the same growth rate), while asymptote a and displacement b are estimated at the individual level. Uninformative uniform prior distributions are used for all parameters. Six parallel chains with different starting values are run for up to 2,000,000 iterations, of which the last 250,000 are used for analysis in order to ensure the use of converged values.

All R code for model fitting can be found in the accompanying R markdown file.

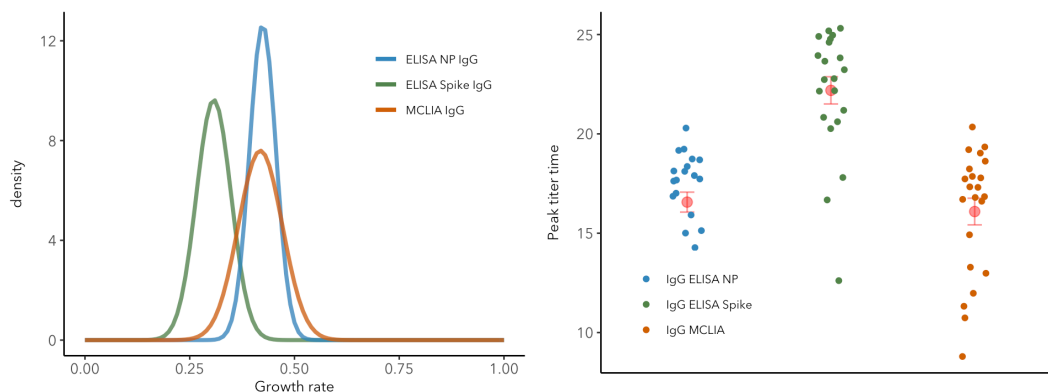


Figure S10. Posterior growth rates (normal distributions fitted to the posteriors) and estimated peak titer times (with mean \pm SE) for IgG ELISA-NP, IgG ELISA-Spike and IgG MCLIA.

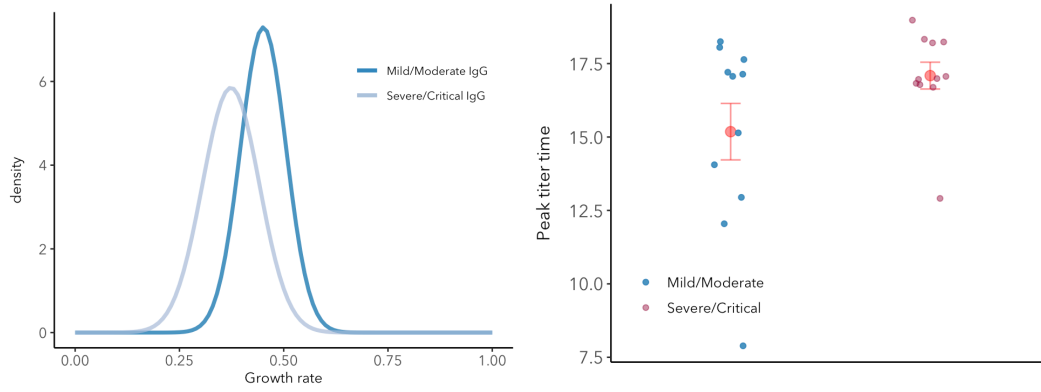


Figure S11. Posterior growth rates (normal distributions fitted to the posteriors) and estimated peak titer times (with mean \pm SE) for IgG Mild/Moderate and IgG Severe/Critical cases.

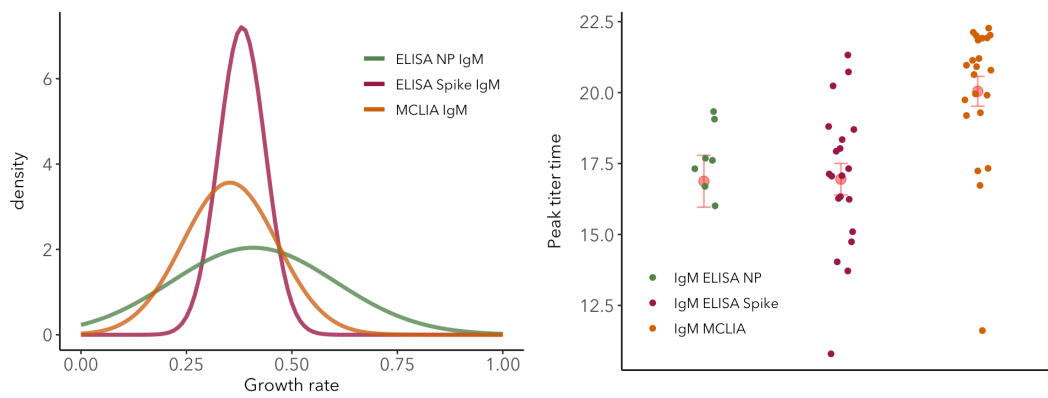


Figure S12. Posterior growth rates (normal distributions fitted to the posteriors) and estimated peak titer times (with mean \pm SE) for IgM ELISA-NP, IgM ELISA-Spike and IgM MCLIA.

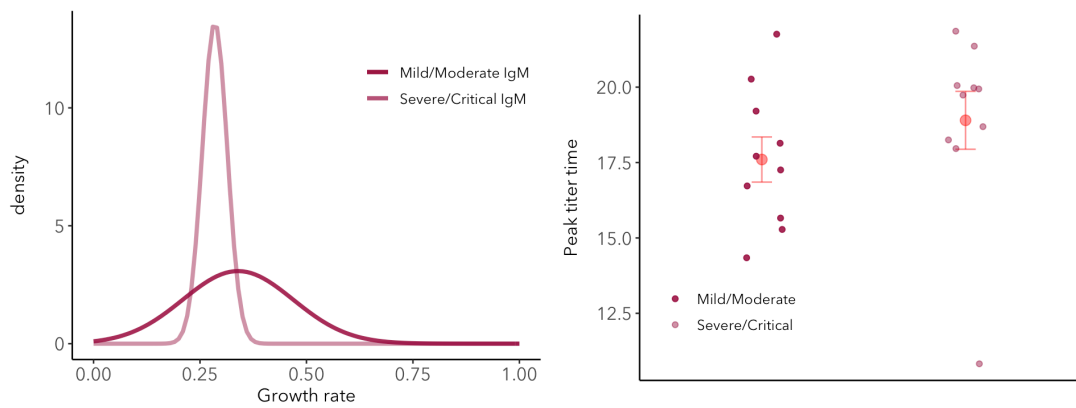
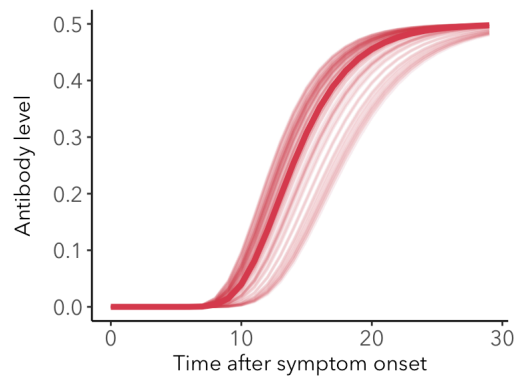
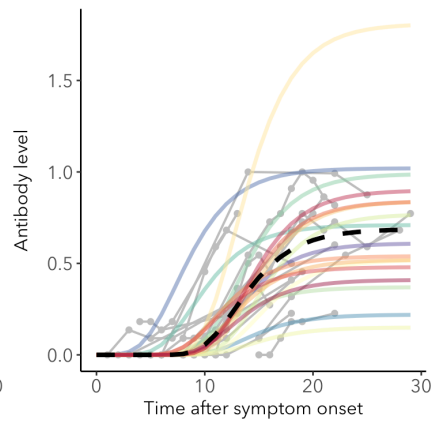
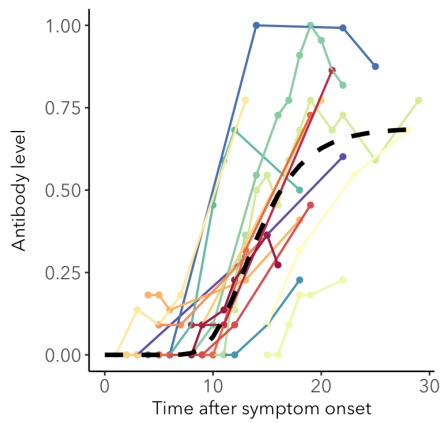
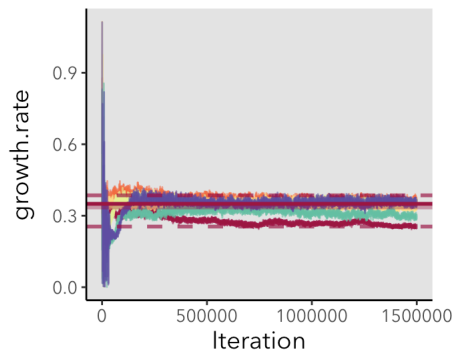
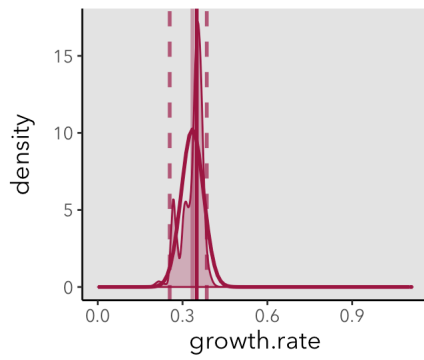
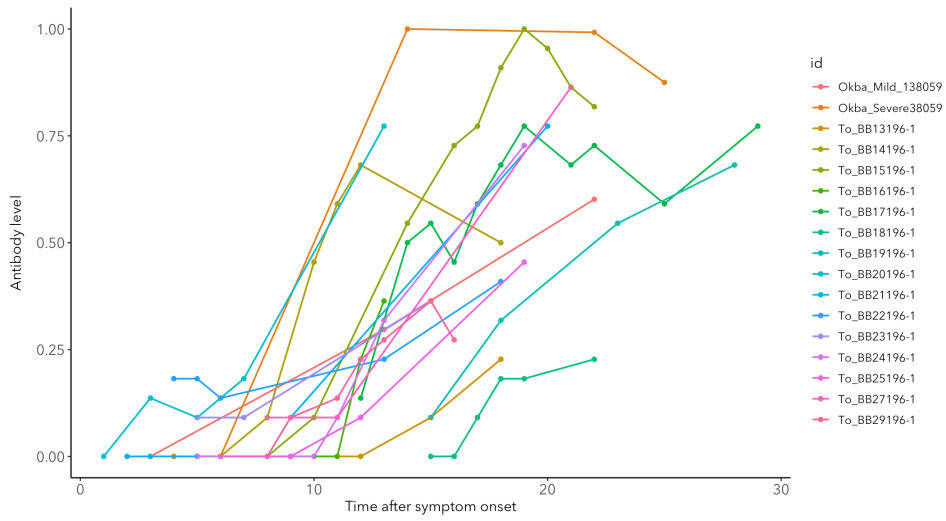


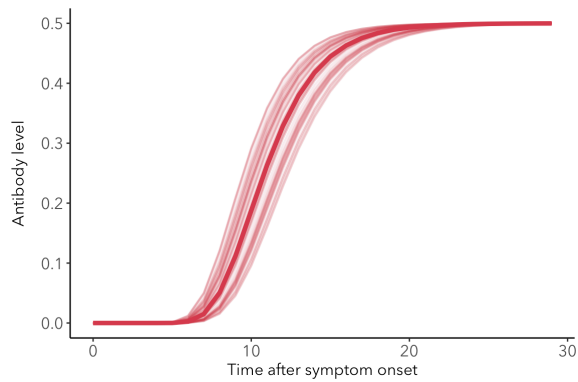
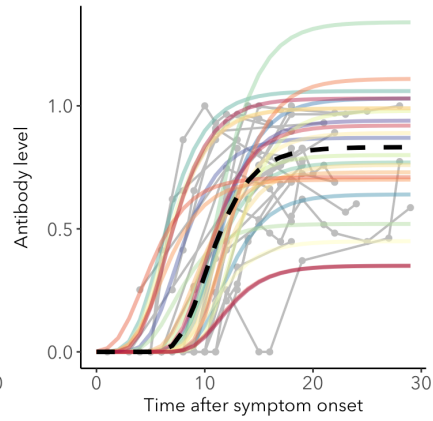
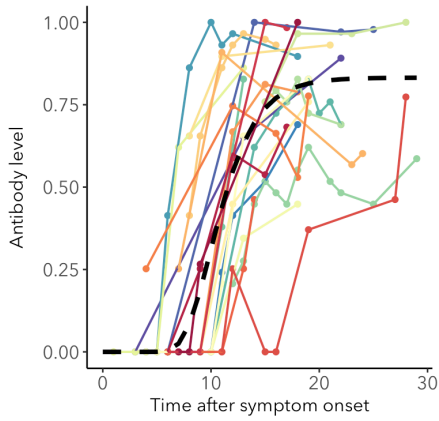
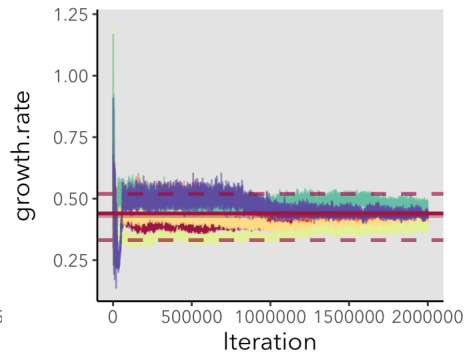
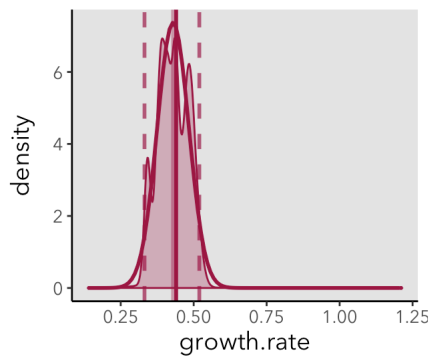
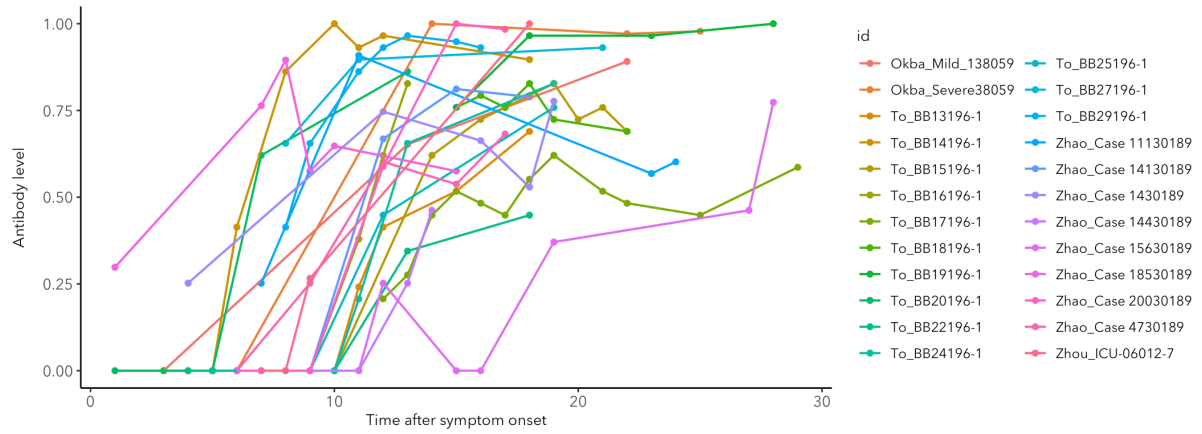
Figure S13. Posterior growth rates (normal distributions fitted to the posteriors) and estimated peak titer times (with mean \pm SE) for IgM Mild/Moderate and IgM Severe/Critical cases.

The pages below show the antibody data used for the different antibody/assay/antigen datasets. The IDs shown in the figures correspond with those in the accompanying spreadsheet. Posterior growth rate and different MCMC chains are shown, with posterior mode (bold red line), posterior mean (lighter red line) and 95% credible intervals (dashed lines). Fitted functions are shown for each individual (black dashed line = overall mean). Finally, the posterior mean together with 100 randomly selected posterior fits within the 95% credible interval.

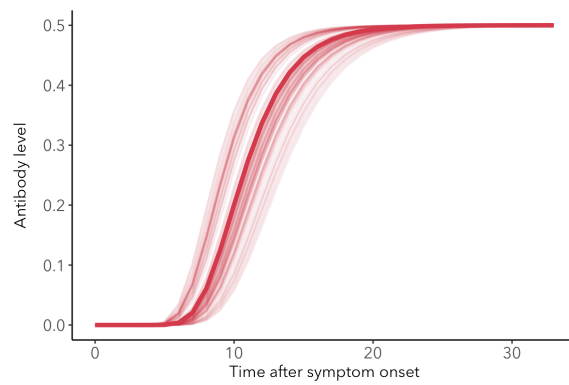
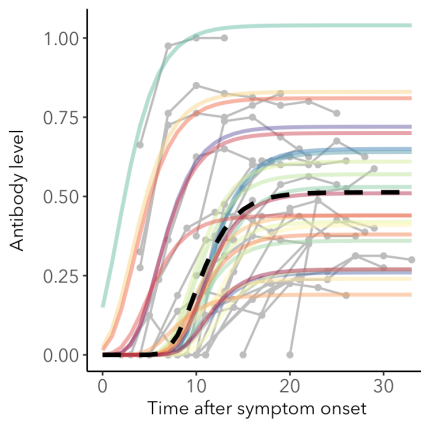
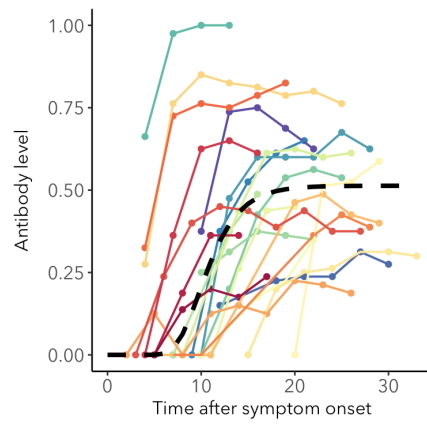
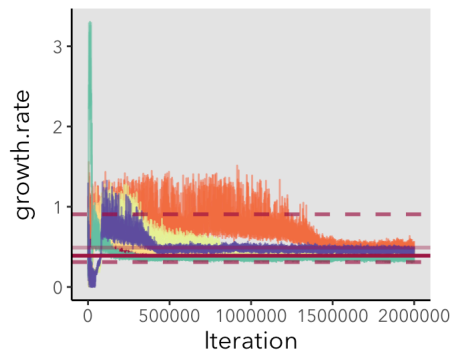
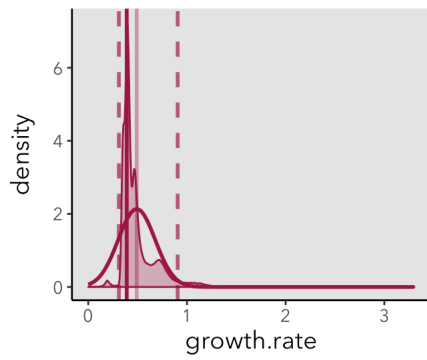
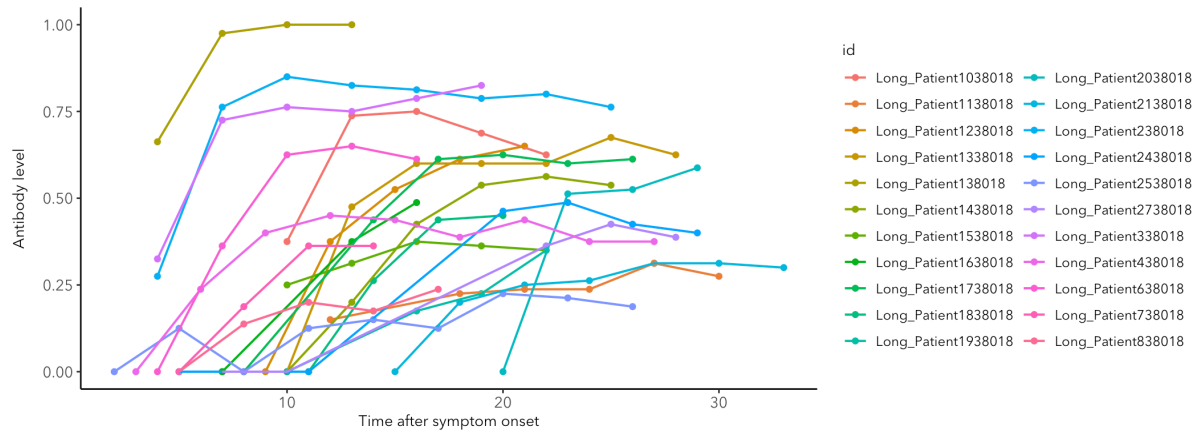
IgG ELISA-Spike



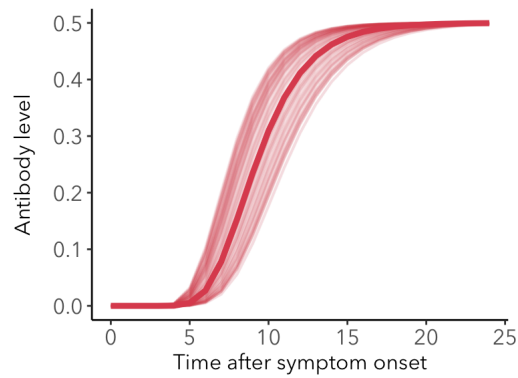
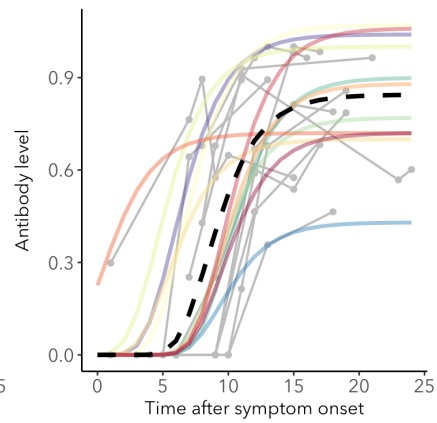
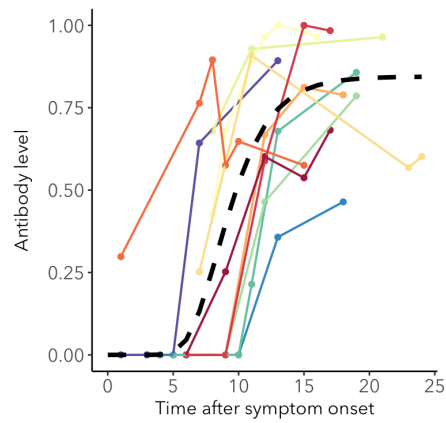
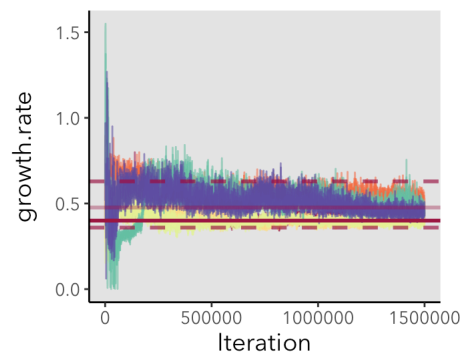
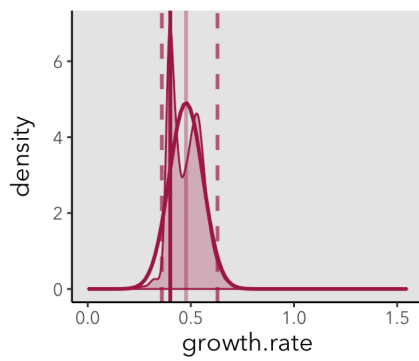
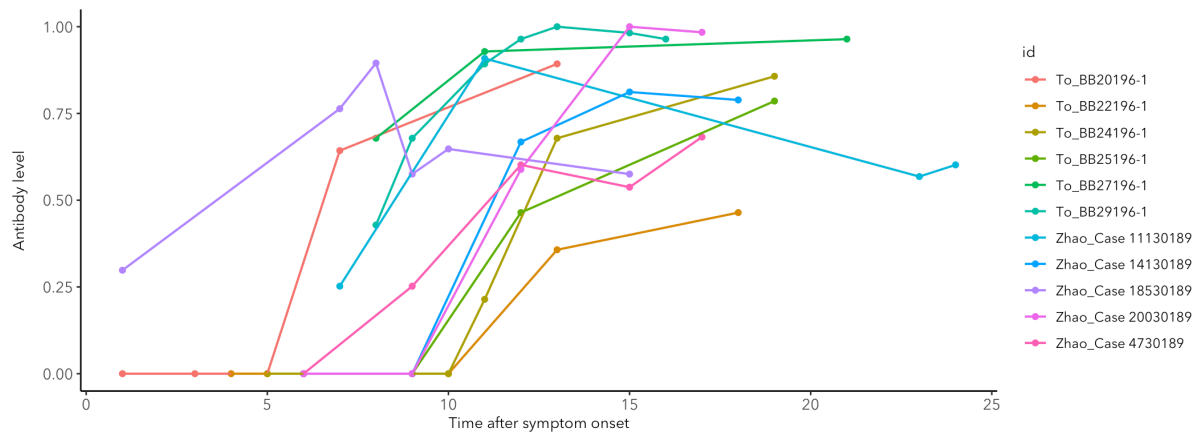
IgG ELISA-NP



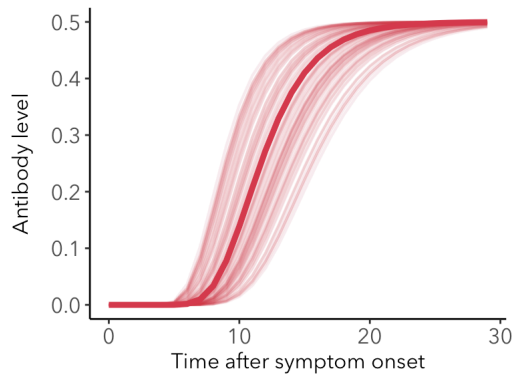
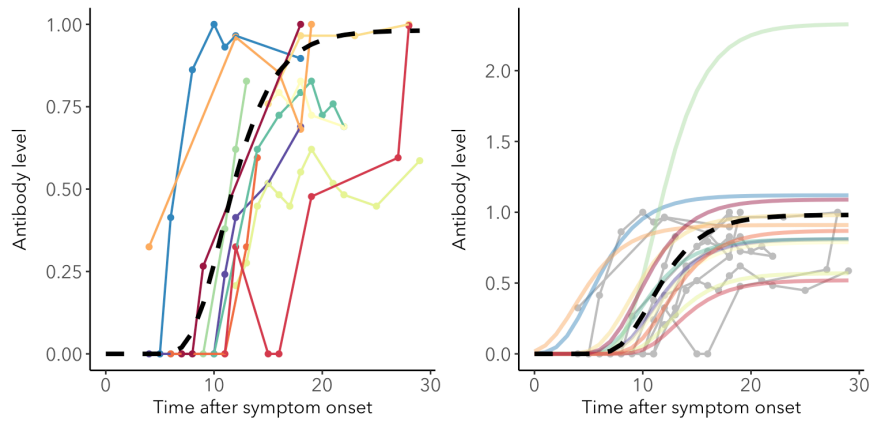
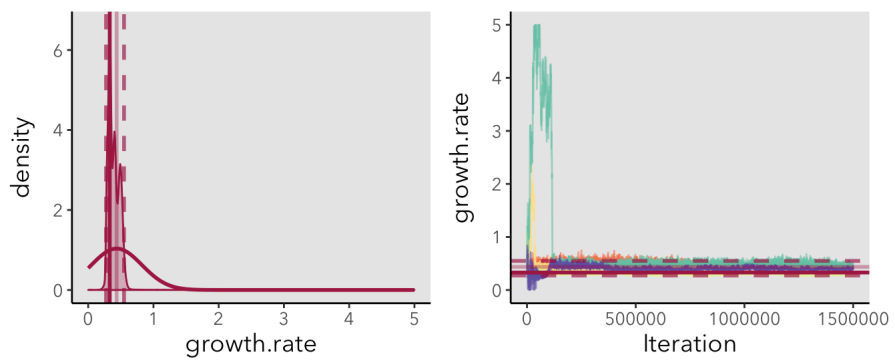
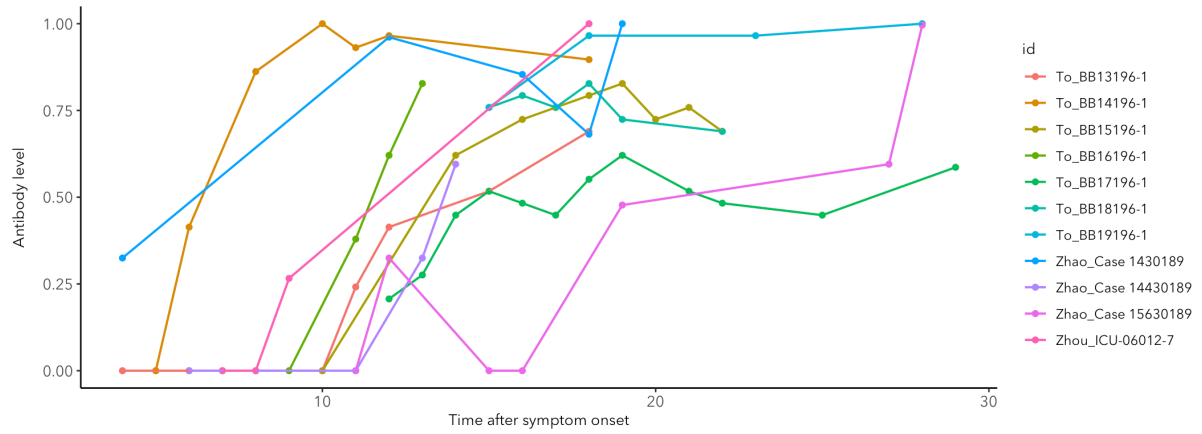
IgG MCLIA



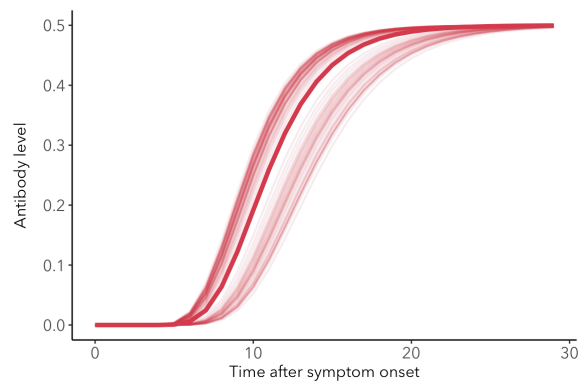
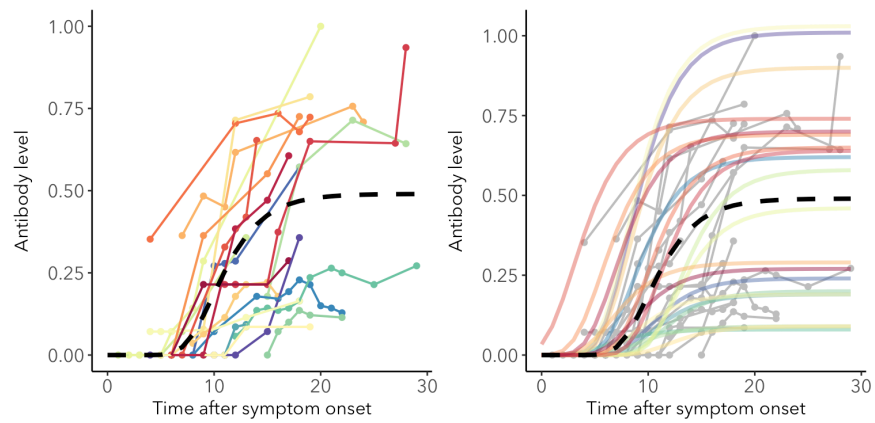
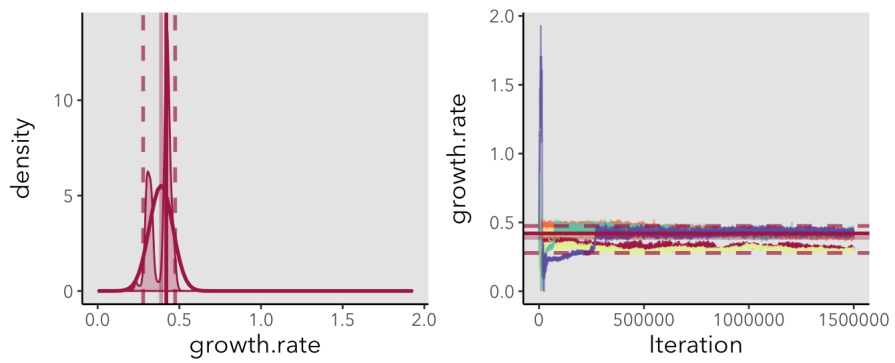
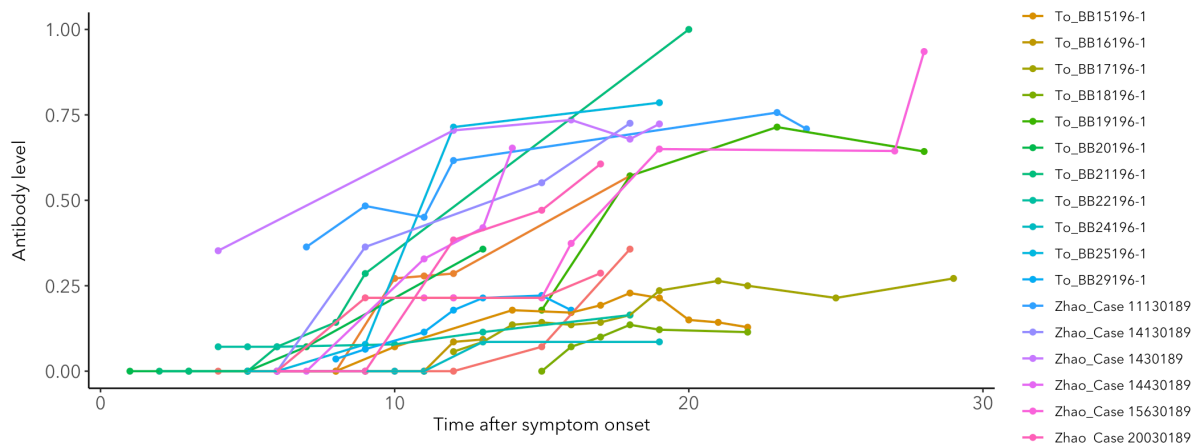
IgG Mild/Moderate cases



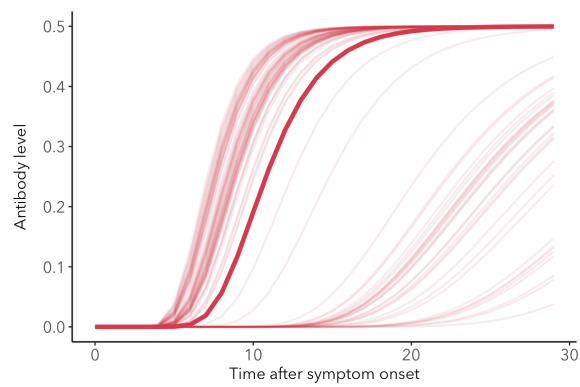
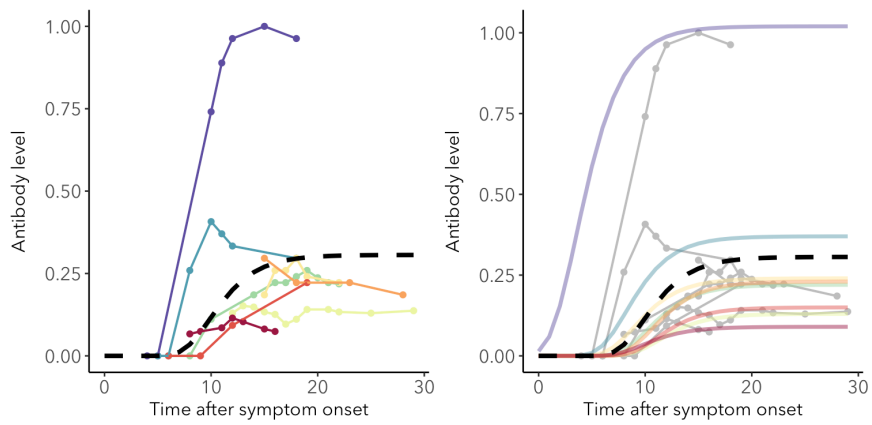
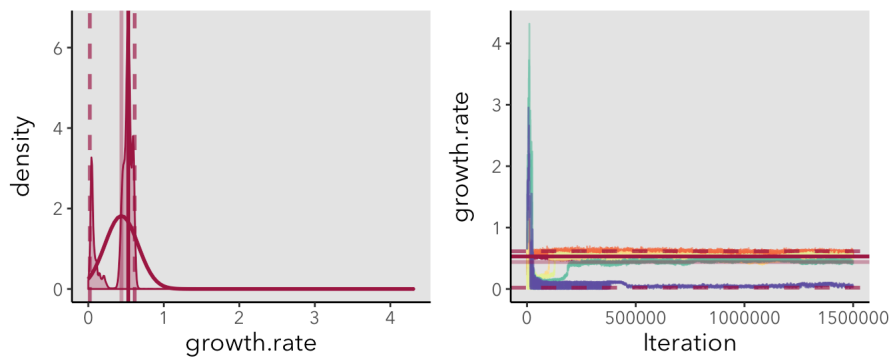
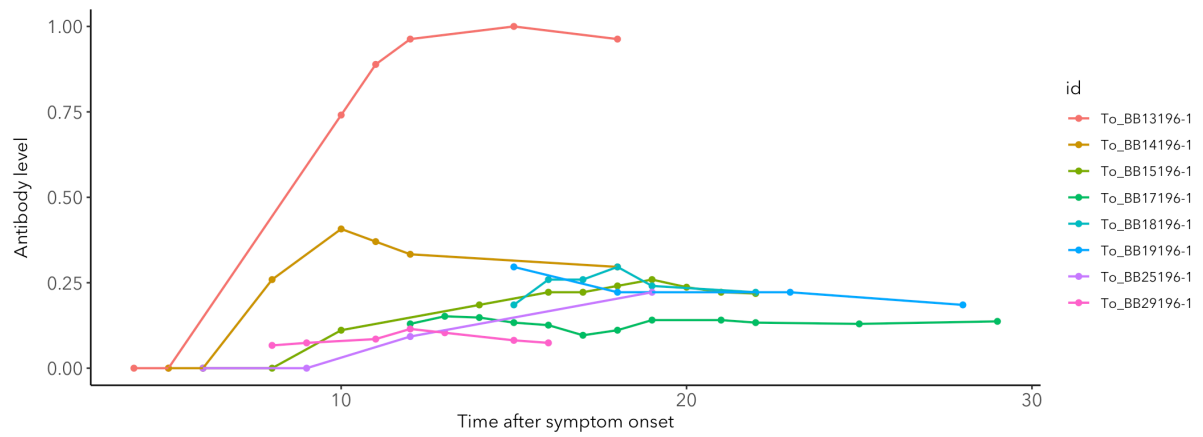
IgG Severe/Critical cases



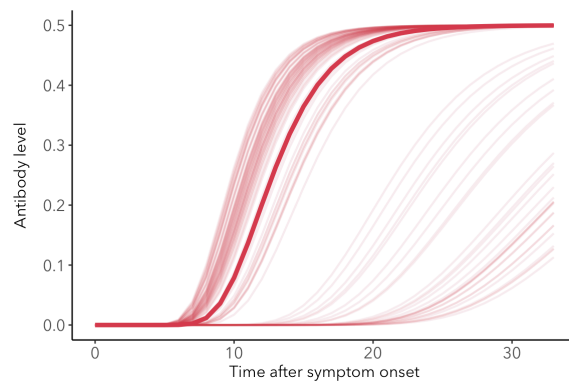
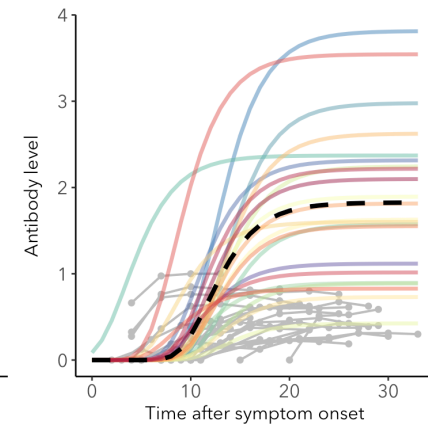
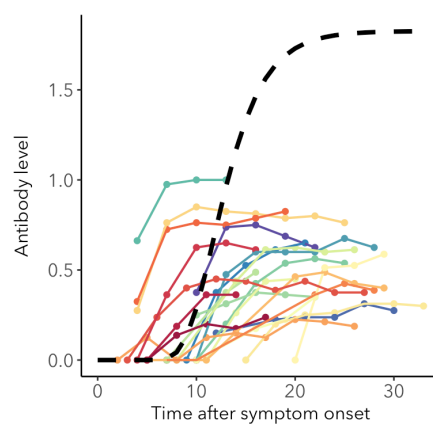
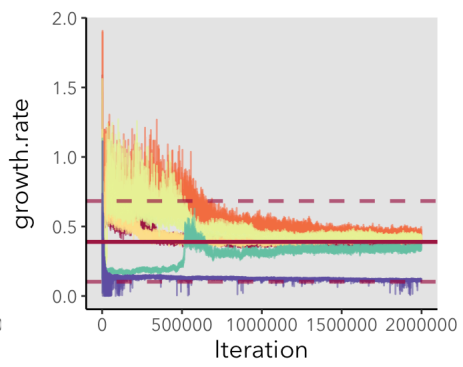
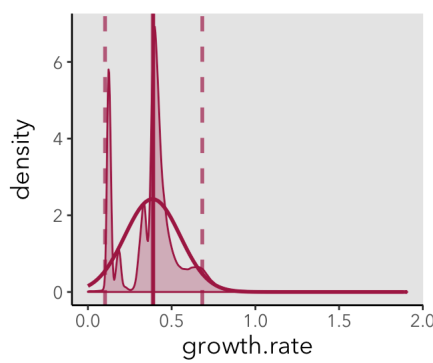
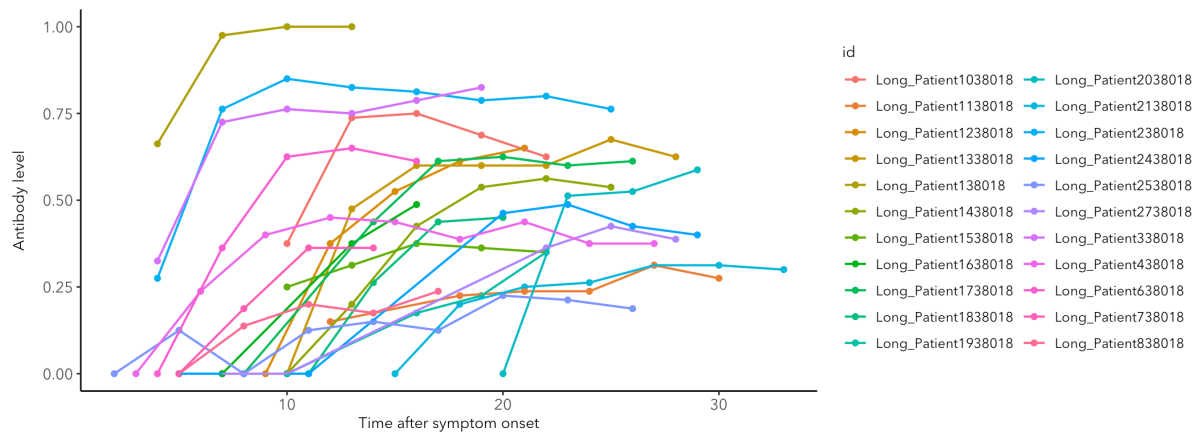
IgM ELISA-Spike



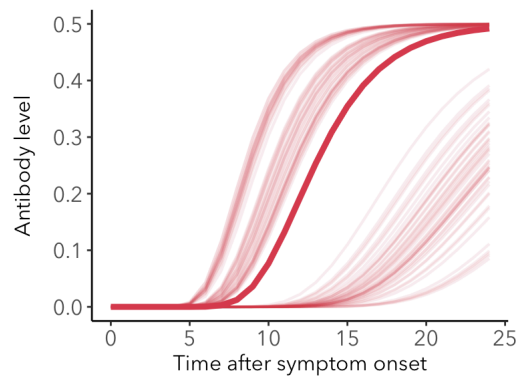
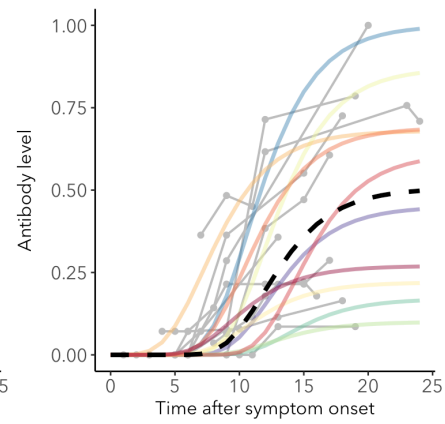
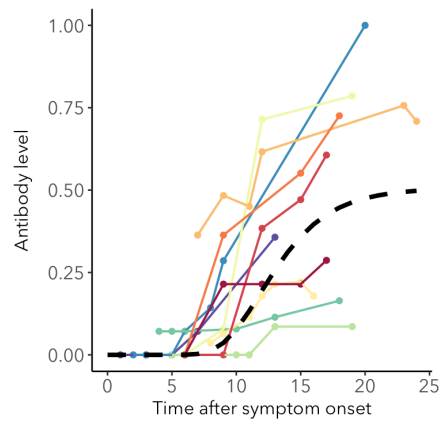
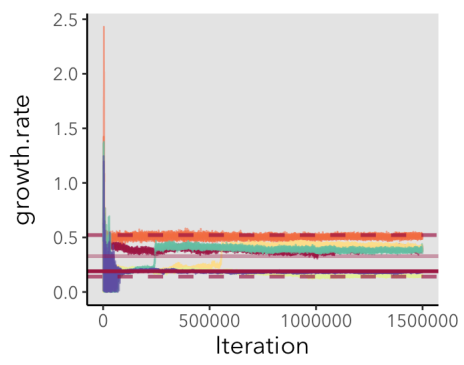
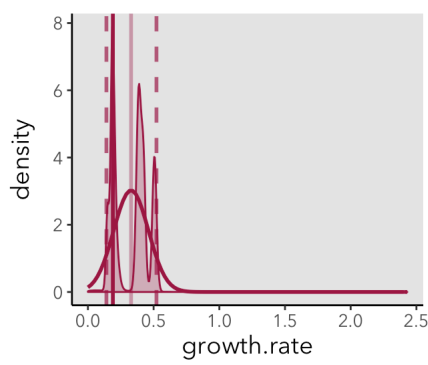
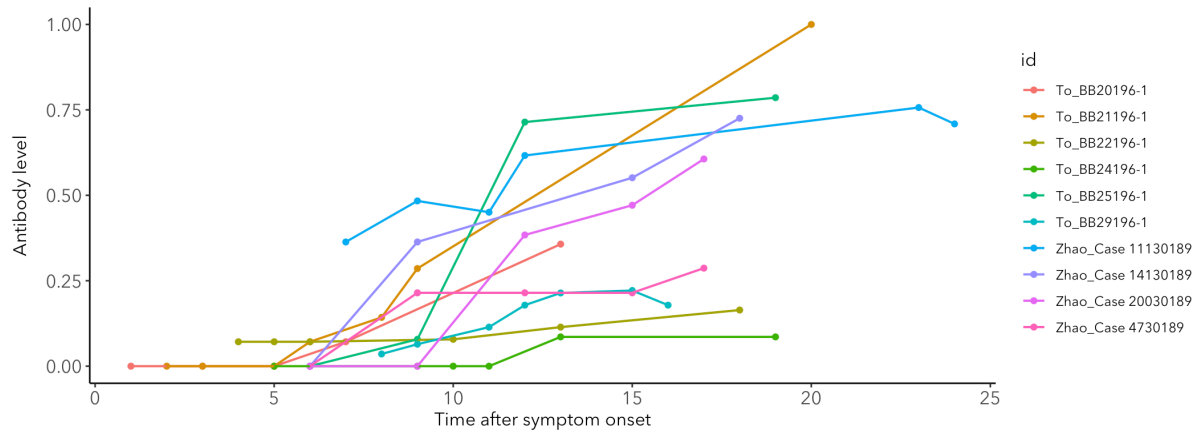
IgM ELISA-NP



IgM MCLIA



IgM Mild/Moderate cases



IgM Severe/Critical cases

