Supplementary Information

A new approach to characterising infectious disease transmission dynamics from sentinel surveillance: application to the Italian 2009-2010 A/H1N1 influenza pandemic

I. Dorigatti*[1,2], S. Cauchemez[1], A. Pugliese[2] and N.M. Ferguson[1]

1 MRC Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, Imperial College London, Faculty of Medicine, Norfolk Place, London W2 1PG, UK

2 Department of Mathematics, University of Trento, via Sommarive 14, 38123 Povo (Trento), Italy

* Author for correspondence: i.dorigatti@imperial.ac.uk

phone: +44 (0)20 7594 3229 fax: +44 (0)20 7594 8321

Computation of R_0

For completeness, we define the infectivity function $A(\tau)$ at time τ after infection. Let $g(t)$ denote the probability density function of the variable T_E, the length of the latent period (*i.e.* the time spent in the classes E^1 and E^2). The probability of being in class I^1 at time τ (after infection) is given by

$$P(I^1, \tau) = \int_0^\tau g(t) e^{-\gamma(\tau - t)} dt \tag{26}$$

In a similar fashion, the probability of being in class I^2 at time τ (after infection) is given by

$$P(I^2, \tau) = \int_0^\tau g(t) \int_t^\tau \gamma e^{\gamma(u-t)} e^{-\gamma(\tau - u)} du dt \tag{27}$$

$A(\tau)$ is defined as follows

$$A(\tau) = h^1 P(I^1, \tau) + h^2 P(I^2, \tau) \tag{28}$$

Equation (4) needs the computation of

$$\int_0^{+\infty} A(\tau) d\tau = \frac{h_1}{\gamma} + \frac{h_2}{\gamma} \tag{29}$$

Computation of $P(ILI_t^i, P_t | T_t, \bar{Z}_t^i, \rho_i)$

This computation is based on the mathematical definition of the Beta function B for two variables $x, y > 0$

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}dt \tag{30}$$

From the definition given in (30) it follows that

$P(ILI_t^i, P_t | T_t, \bar{Z}_t^i, \rho_i) =$

$$\frac{\binom{T_t}{P_t}}{B(\alpha,\beta)}\Bigg((1-\rho_i)^{Z_t^i}\int_0^1 \pi_t^{P_t+\alpha-1}(1-\pi_t)^{ILI_t^i+T_t-P_t+\beta-1}d\pi_t +$$

$$+\sum_{F_t^i=1}^{\min(ILI_t^i,Z_t^i)}\binom{ILI_t^i-1}{F_t^i-1}\binom{Z_t^i}{F_t^i}\rho_i^{F_t^i}(1-\rho_i)^{Z_t^i-F_t^i}\int_0^1 \pi_t^{F_t^i+P_t+\alpha-1}(1-\pi_t)^{ILI_t^i-F_t^i+T_t-P_t+\beta-1}d\pi_t +$$

$$+\binom{Z_t^i}{F_t^i}F_t^i\rho_i^{F_t^i}(1-\rho_i)^{Z_t^i-F_t^i}\int_0^1 \pi_t^{P_t+\alpha-1}(1-\pi_t)^{ILI_t^i+T_t-P_t+\beta-1}d\pi_t\Bigg)$$

$$=\frac{\binom{T_t}{P_t}}{B(\alpha,\beta)}\Bigg((1-\rho_i)^{\bar{Z}_t^i}B(P_t+\alpha, ILI_t^i+T_t-P_t+\beta)+$$

$$+\sum_{F_t^i=1}^{\min(ILI_i^t,\bar{Z}_t^i)}\binom{ILI_t^i-1}{F_t^i-1}\binom{\bar{Z}_t^i}{F_t^i}\rho_i^{F_t^i}(1-\rho_i)^{\bar{Z}_t^i-F_t^i}B(F_t^i+P_t+\alpha, ILI_t^i-F_t^i+T_t-P_t+\beta)\Bigg)$$

Computation of $P(F_t^i|\bar{Z}_t^i, \rho_i, r)$

In this section we prove (22) given (21) and (12) using the probability generating function.

For simplicity of notation let's set

$$q_t^i = \frac{\bar{Z}_t^i}{\bar{Z}_t^i + r} \tag{31}$$

From (21) and (12) it follows that the probability generating function of Z_t^i is given by

$$G_{Z_t^i}(t) = \mathbb{E}\big[t^{Z_t^i}\big] = \frac{(1-q_t^i)^r}{(1-tq_t^i)^r} \tag{32}$$

and the probability generating function of F_t^i is given by

$$G_{F_t^i}(t) = \mathbb{E}\big[t^{F_t^i}\big] = (1-\rho_i + t\rho_i)^{Z_t^i} \tag{33}$$

where $\mathbb{E}$ denotes the expected value. Since

$$\mathbb{E}\big[t^{F_t^i}\big] = \mathbb{E}\big[\mathbb{E}[t^{F_t^i}|Z_t^i]\big] = \mathbb{E}\big[(1-\rho_i + t\rho_i)^{Z_t^i}\big]$$

it follows that

$$\begin{aligned} G_{F_t^i}(t) &= \frac{\big(1-q_t^i\big)^r}{\big[1-(1-\rho_i+t\rho_i)q_t^i\big]^r} \\ &= \frac{\big(1-q_t^i\big)^r}{\big[1-q_t^i(1-\rho_i)-t\rho_i q_t^i\big]^r} \\ &= \frac{\Big[1-\dfrac{\rho_i q_t^i}{1-q_t^i(1-\rho_i)}\Big]^r}{\Big[1-t\dfrac{\rho_i q_t^i}{1-q_t^i(1-\rho_i)}\Big]^r} \end{aligned}$$

and hence by (31) we may conclude (22).

Tables

	0-4	5-14	15-24	25-64	65+
0-4	5.23	1.50	0.59	11.68	1.11
5-14	0.76	14.49	1.72	13.14	1.05
15-24	0.28	1.58	13.94	9.16	0.85
25-64	0.99	2.20	1.67	11.38	1.94
65+	0.26	0.49	0.43	5.38	2.93

Table SI-1: Symmetrised contact matrix of all reported contacts (physical and non-physical) in Italy, consisting of the average number of contact persons recorded per working day per survey participant (Polymod 2008). Row index represents the age class of the participant, column index represents the age class of the contact.

	0-4	5-14	15-24	25-64	65+
0-4	1.69	1.42	0.30	6.60	0.73
5-14	0.72	7.84	1.42	9.38	1.41
15-24	0.14	1.30	10.41	10.03	0.25
25-64	0.56	1.57	1.83	9.06	1.74
65+	0.17	0.65	0.13	4.81	0.57

Table SI-2: Symmetrised contact matrix of all reported contacts (physical and non-physical) in Italy, consisting of the average number of contact persons recorded per holiday day per survey participant (Polymod 2008). Row index represents the age class of the participant, column index represents the age class of the contact.

Basic no-overdispersion		
	Susceptibility model	Immunity model
DIC	9227.3	8822.9
log-likelihood	-4612.3	-4409.9
	(-4610.6, -4615.7)	(-4408.3, -4413.5)
R	1.354	1.357
	(1.348, 1.360)	(1.351, 1.363)
I_0	436	366
	(372, 504)	(312, 426)
$\rho_1 = \cdots = \rho_4$	0.090	0.099
	(0.088, 0.091)	(0.097, 0.101)

Table SI-3: Basic no-overdispersion models (susceptibility and immunity variants): mean and, in brackets, equal-tailed 95% credible interval of the marginal posterior distribution of the parameters for each specified model. Basic: reporting rates constant in time and across the age groups.

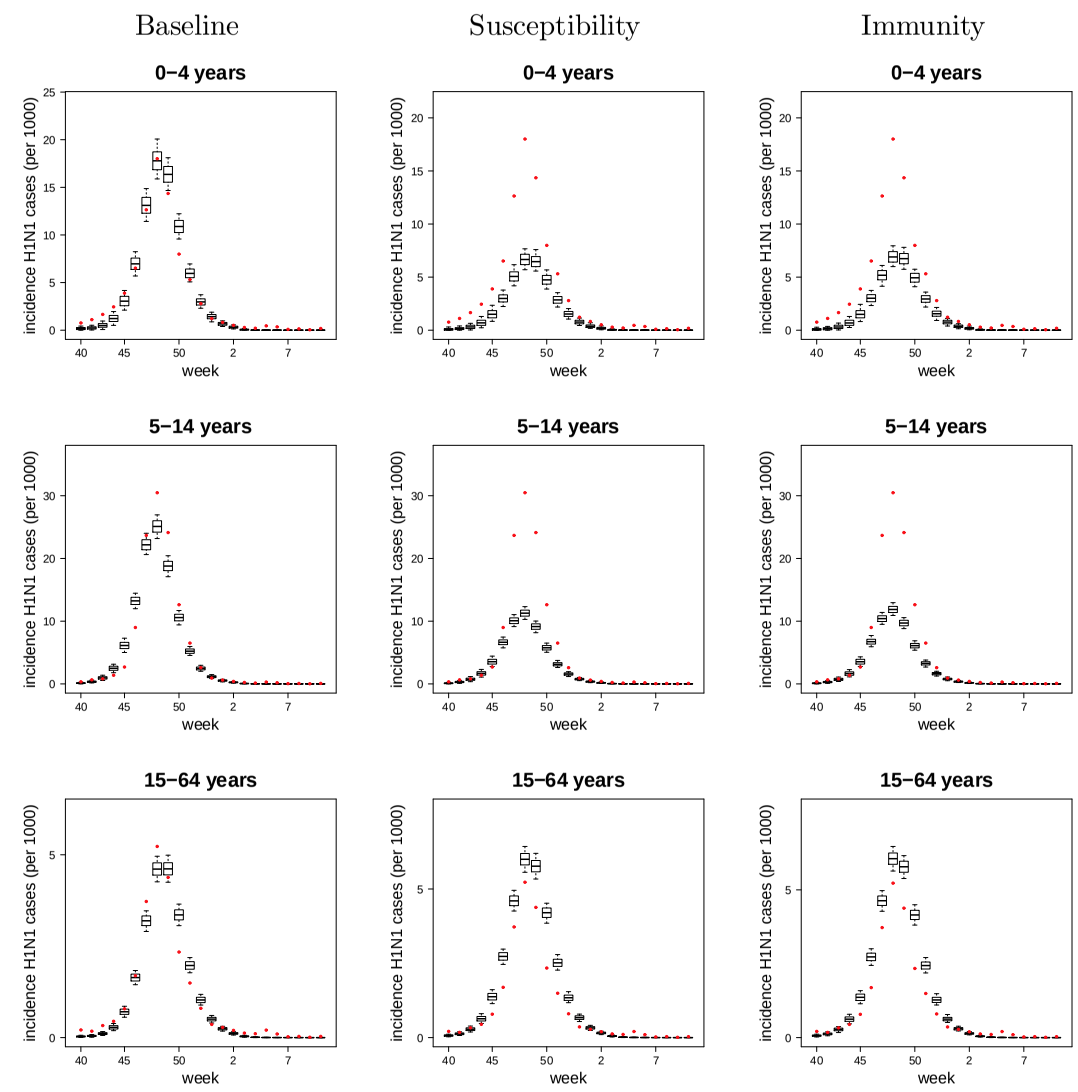
Basic with-overdispersion		
	Susceptibility model	Immunity model
DIC	1528.3	1525.3
log-likelihood	-762.3	-760.9
	(-760.4, -766.2)	(-764.8, -758.9)
R	1.293	1.293
	(1.261, 1.326)	(1.261, 1.325)
I_0	2842	2600
	(1172, 5804)	(1082, 5255)
$\rho_1 = \cdots = \rho_4$	0.139	0.152
	(0.117, 0.166)	(0.128, 0.180)
r	2.480	2.598
	(1.679, 3.468)	(1.751, 3.634)

Table SI-4: Basic with-overdispersion models (susceptibility and immunity variants): mean and, in brackets, equal-tailed 95% credible interval of the marginal posterior distribution of the parameters for each specified model. Basic: reporting rates constant in time and across the age groups.

No-overdispersion				
	Susceptibility model		Immunity model	
	ADR	TVR	ADR	TVR
DIC	2600.6	2531.4	2553.8	2502.7
log-likelihood	-1296.5	-1265.7	-1274.2	-1250.0
	(-1293.9, -1300.8)	(-1262.6, -1270.7)	(-1271.6, -1278.6)	(-1247.0, -1254.8)
R	1.405	1.368	1.402	1.369
	(1.398, 1.411)	(1.355, 1.382)	(1.396, 1.408)	(1.358, 1.380)
I_0	115	264	111	237
	(98, 133)	(178, 377)	(94, 130)	(170, 322)
ρ_1	0.182	0.174	0.197	0.188
	(0.176, 0.187)	(0.132, 0.217)	(0.191, 0.204)	(0.151, 0.234)
ρ_2	0.171	0.157	0.183	0.168
	(0.167, 0.175)	(0.120, 0.194)	(0.179, 0.188)	(0.136, 0.208)
ρ_3	0.059	0.056	0.066	0.063
	(0.057, 0.060)	(0.043, 0.070)	(0.064, 0.068)	(0.050, 0.078)
ρ_4	0.038	0.037	0.043	0.042
	(0.035, 0.040)	(0.028, 0.047)	(0.040, 0.046)	(0.033, 0.052)
a	-	1.309	-	1.291
	-	(0.981, 1.791)	-	(0.978, 1.653)
b	-	0.481	-	0.535
	-	(0.328, 0.692)	-	(0.371, 0.739)

Table SI-5: ADR and TVR no-overdispersion models (susceptibility and immunity variants): mean and, in brackets, equal-tailed 95% credible interval of the marginal posterior distribution of the parameters for each specified model. ADR: age-dependent reporting, *i.e.* reporting rates constant in time and age-specific. TVR: time-varying reporting, *i.e.* age-specific and time-dependent reporting rates as defined by the piecewise linear function given in Eq. (25).

Figures


Basic no-overdispersion
Baseline
Susceptibility
Immunity
0–4 years
5–14 years
15–64 years
incidence H1N1 cases (per 1000)
week

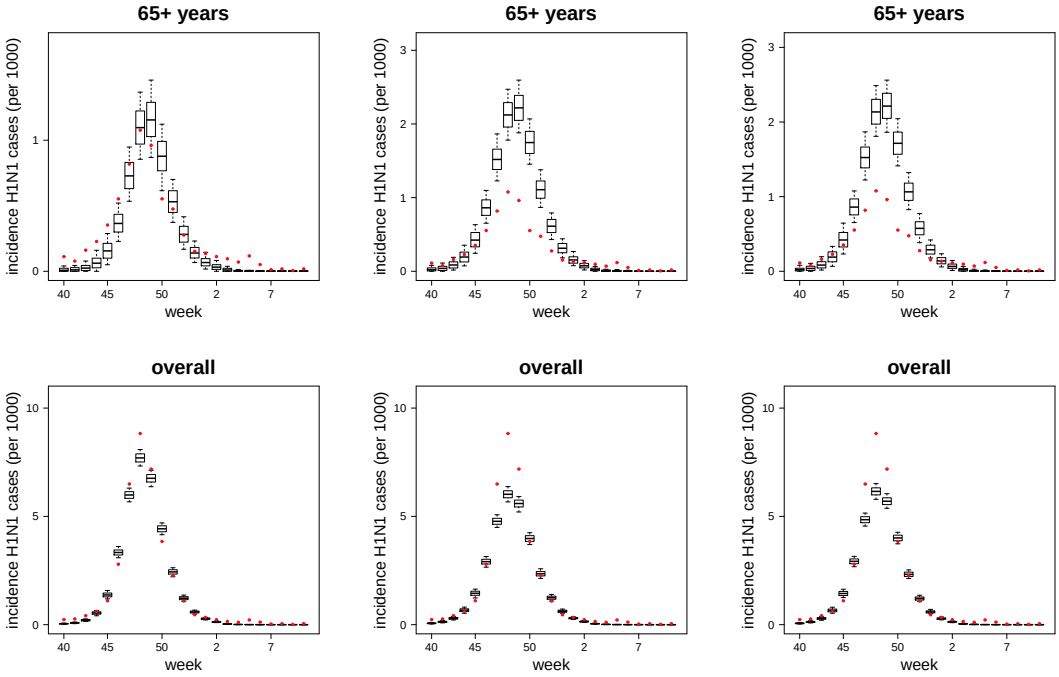
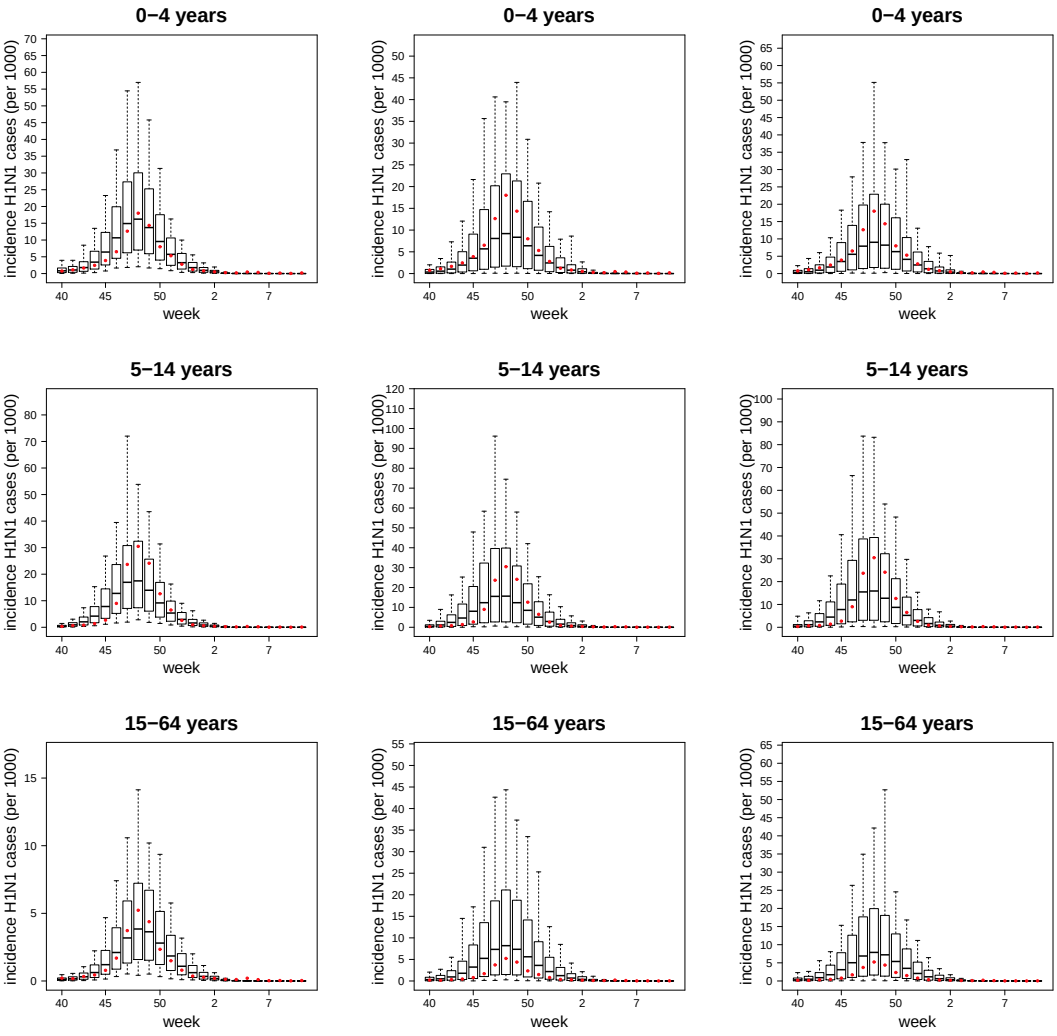


Figure SI-1: Basic no-overdispersion models (baseline, susceptibility and immunity variants): plot of the mean, 95%CI, maximum and minimum simulated weekly reported incidences (per 1000) of symptomatic H1N1 cases (i.e. ILI & H1N1 cases) in the $0-4$, $5-14$, $15-64$, 65+ years age-classes and in the overall population. The dots represent the observed data (*i.e.* the H1N1-attributable ILI incidence curve). Basic: reporting rates constant in time and across the age groups.

Basic with-overdispersion



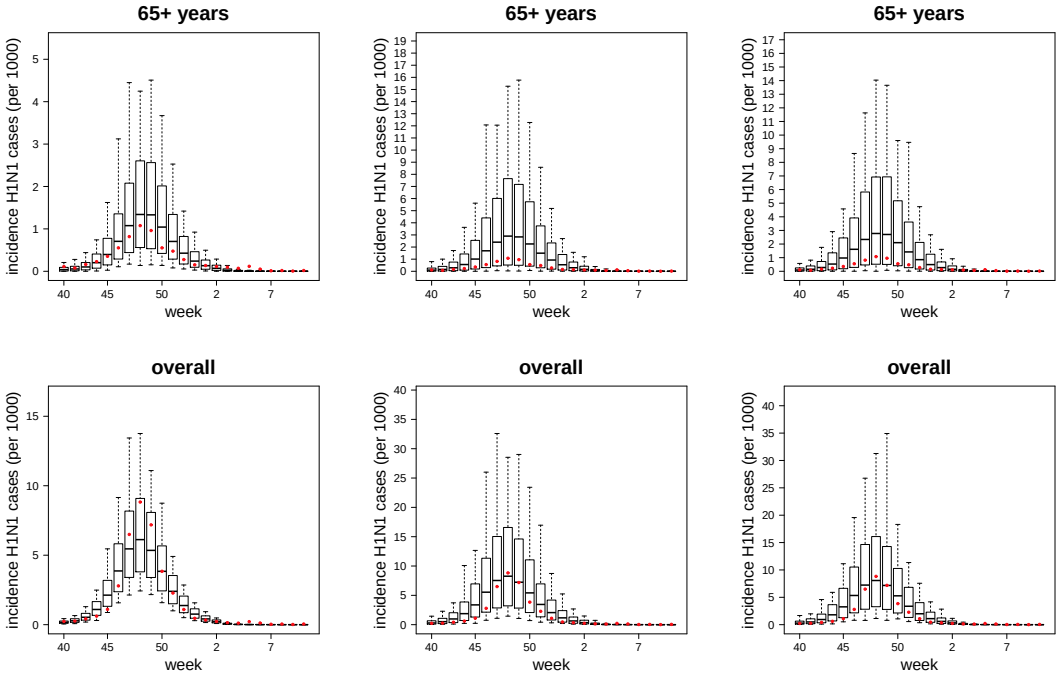
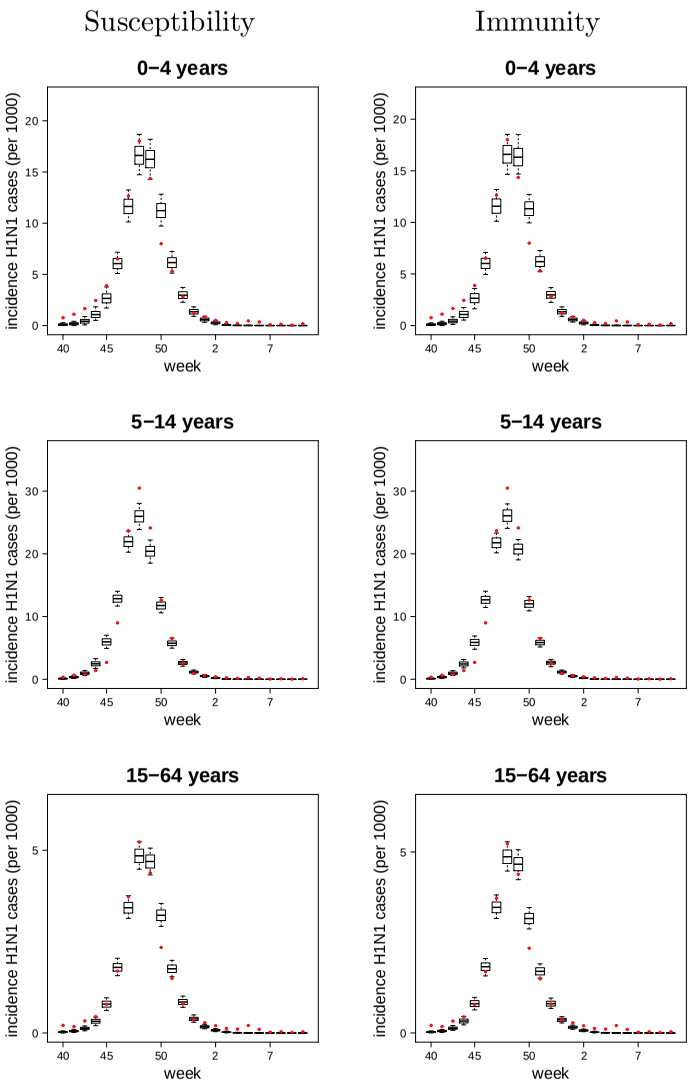


Figure SI-2: Basic with-overdispersion models (baseline, susceptibility and immunity variants): plot of the mean, 95%CI, maximum and minimum simulated weekly reported incidences (per 1000) of symptomatic H1N1 cases (i.e. ILI & H1N1 cases) in the 0-4, 5-14, 15-64, 65+ years age-classes and in the overall population. The dots represent the observed data (*i.e.* the H1N1-attributable ILI incidence curve). Basic: reporting rates constant in time and across the age groups.

ADR no-overdispersion



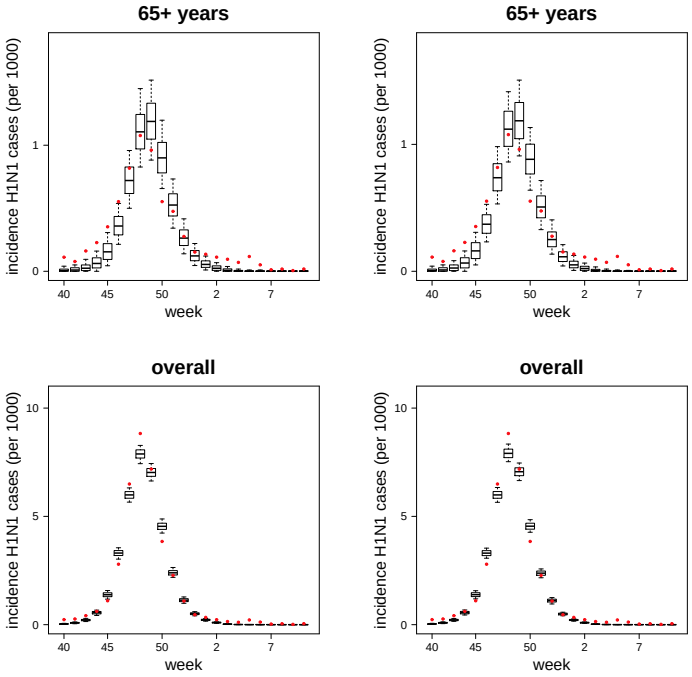
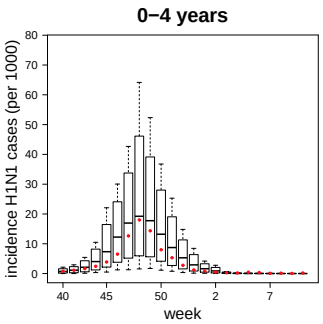
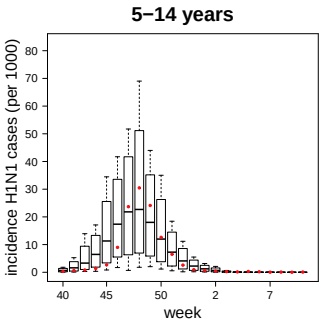


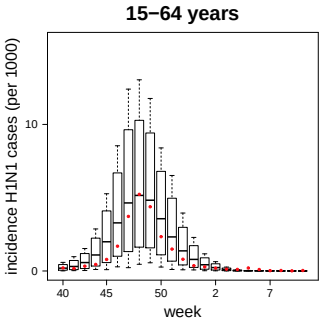
Figure SI-3: ADR no-overdispersion models (baseline, susceptibility and immunity variants): plot of the mean, 95%CI, maximum and minimum simulated weekly reported incidences (per 1000) of symptomatic H1N1 cases (i.e. ILI & H1N1 cases) in the 0-4, 5-14, 15-64, 65+ years age-classes and in the overall population. The dots represent the observed data (*i.e.* the H1N1-attributable ILI incidence curve). ADR: age-dependent reporting, *i.e.* reporting rates constant in time and age-specific.

ADR with-overdispersion

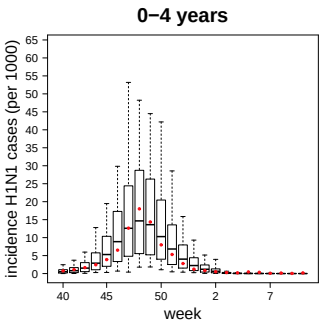
Susceptibility

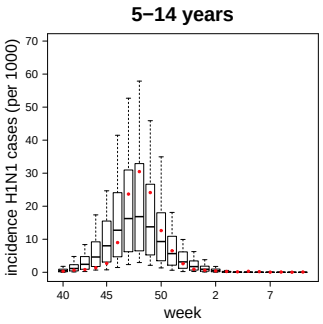


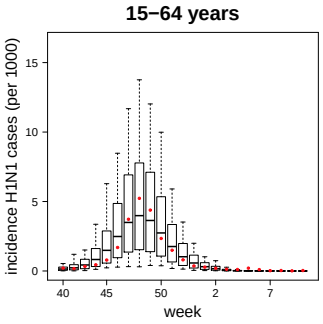




Immunity







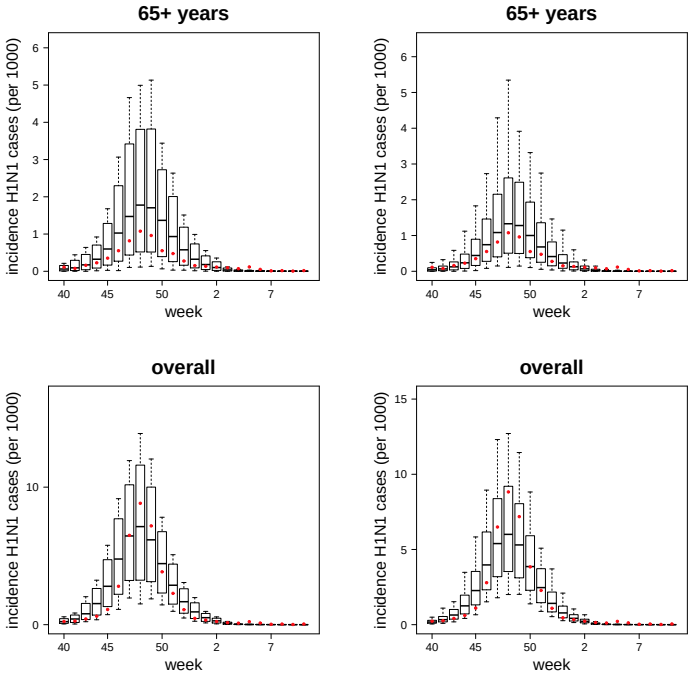
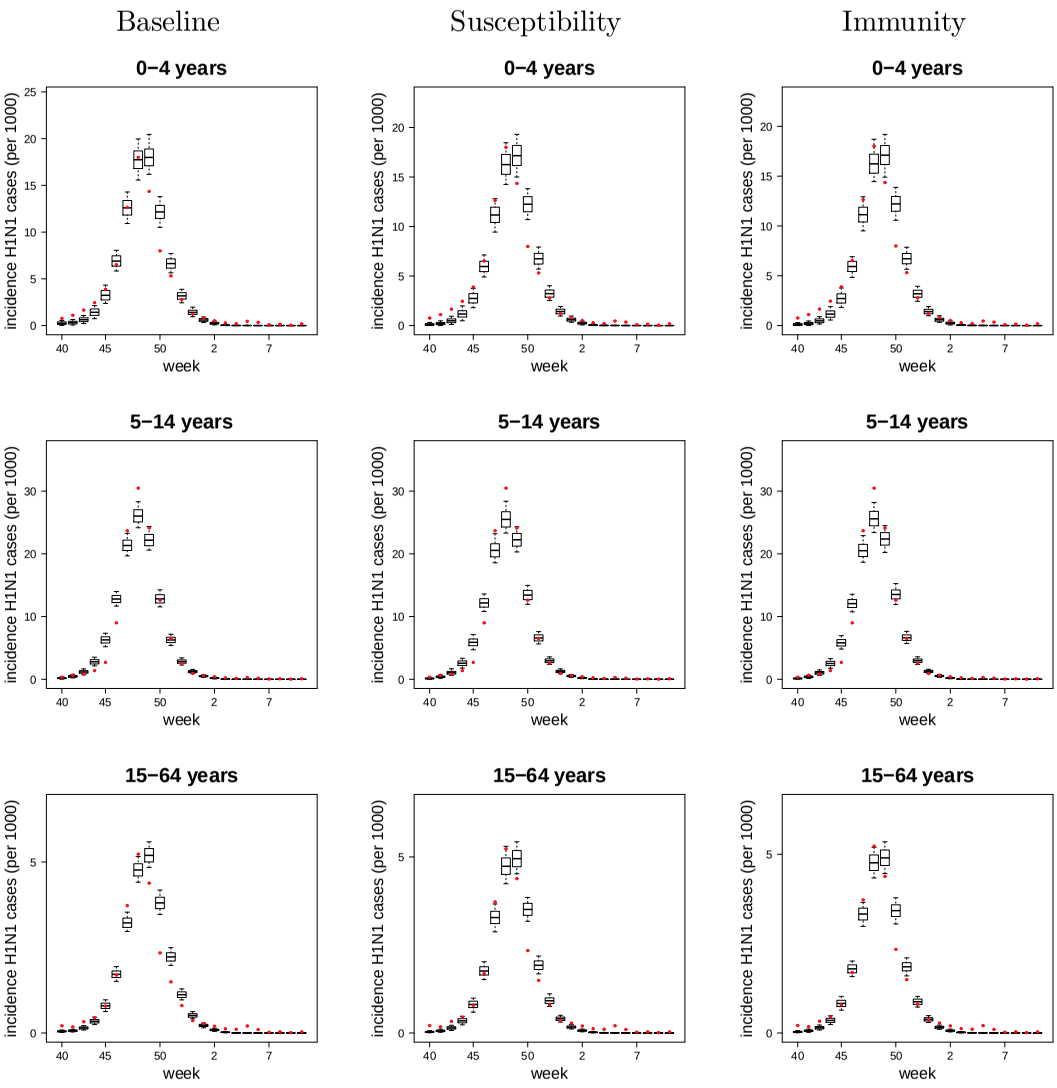


Figure SI-4: ADR with-overdispersion models (baseline, susceptibility and immunity variants): plot of the mean, 95%CI, maximum and minimum simulated weekly reported incidences (per 1000) of symptomatic H1N1 cases (i.e. ILI & H1N1 cases) in the 0-4, 5-14, 15-64, 65+ years age-classes and in the overall population. The dots represent the observed data (*i.e.* the H1N1-attributable ILI incidence curve). ADR: age-dependent reporting, *i.e.* reporting rates constant in time and age-specific.

TVR no-overdispersion



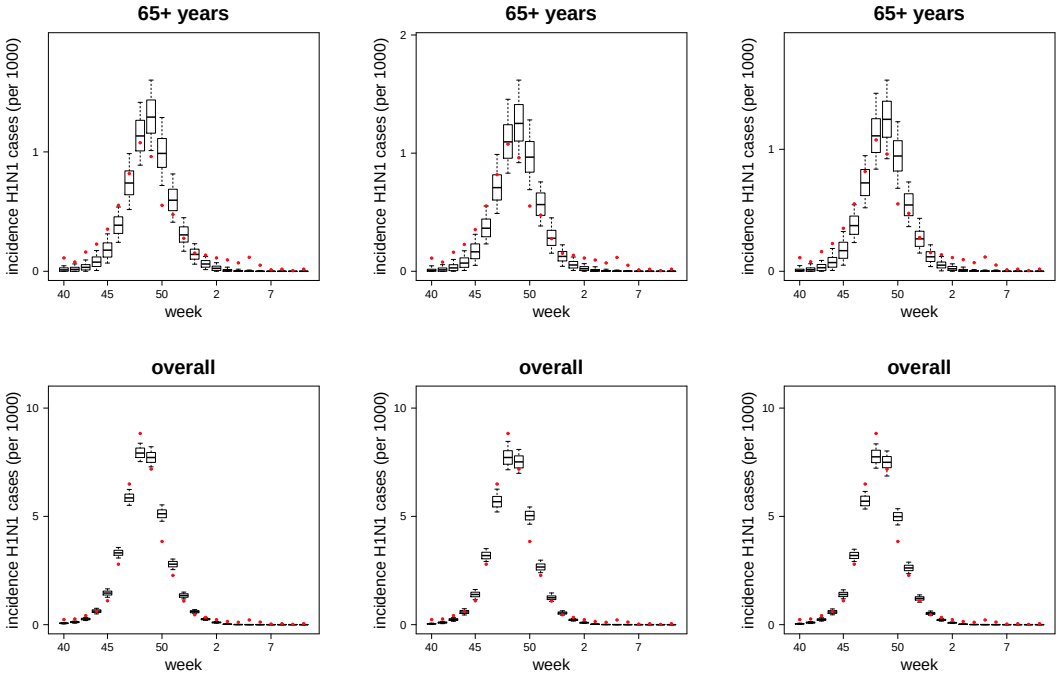


Figure SI-5: TVR no-overdispersion models (baseline, susceptibility and immunity variants): plot of the mean, 95%CI, maximum and minimum simulated weekly reported incidences (per 1000) of symptomatic H1N1 cases (i.e. ILI & H1N1 cases) in the 0-4, 5-14, 15-64, 65+ years age-classes and in the overall population. The dots represent the observed data (*i.e.* the H1N1-attributable ILI incidence curve). TVR: time-varying reporting, *i.e.* age-specific and time-dependent reporting rates as defined by the piecewise linear function given in Eq. (25).