Estimation of the reproduction number for the Pandemic Influenza A/H1N1 2009 in India

Sheikh Taslim Ali[1,2,†], A. S. Kadi[2] and Neil M. Ferguson[1,*]

[1] MRC centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, Imperial College London, UK
[2] Department of studies in Statistics, Karnatak University, Dharwad, INDIA

[†] Presenting Author, s.ali11@imperial.ac.uk; ali.sheikhtaslim@gmail.com

[*]Correspondence to : neil.ferguson@imperial.ac.uk

"Swiss Meeting for Infectious Disease Dynamics" @ FHS St. Gallen, Switzerland on August 30, 2012

Outline

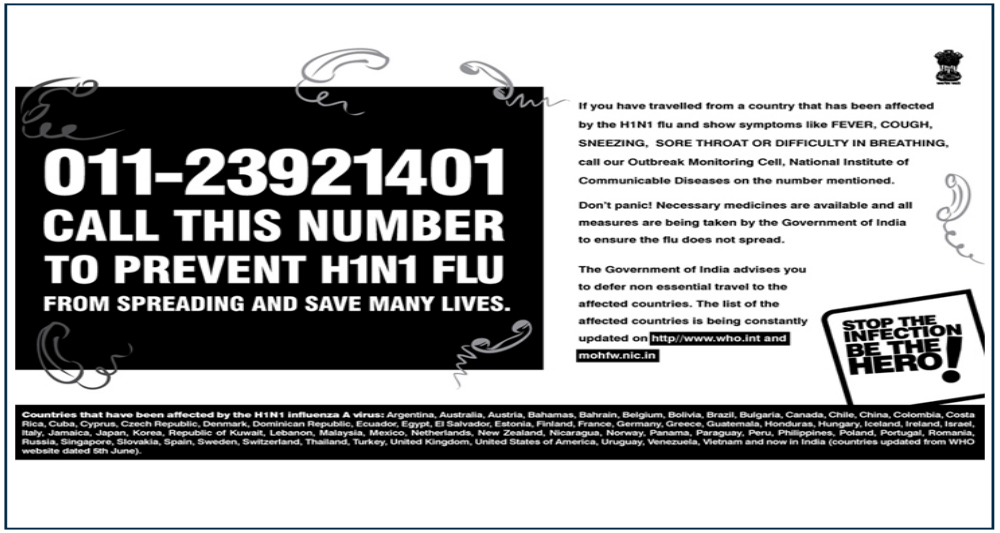
- Background
- Data Used
- Objectives/Aims
- Methodology

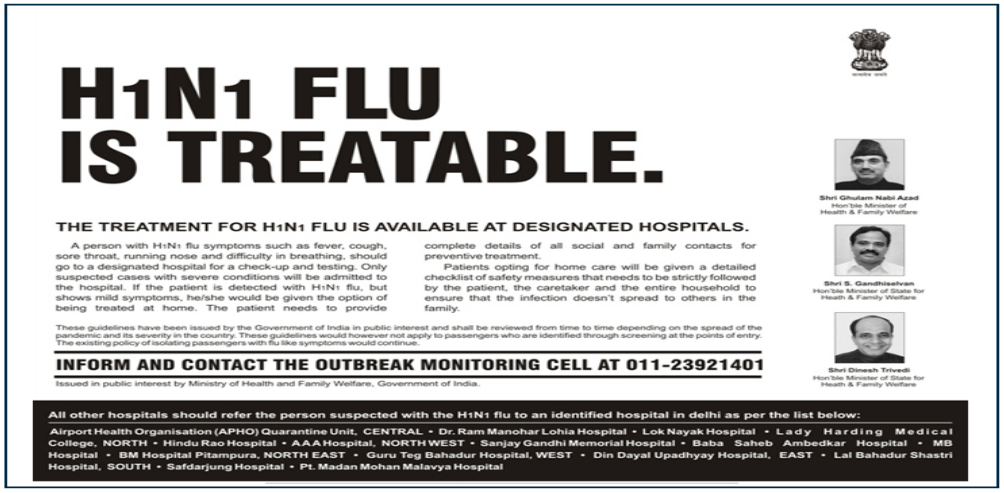
➢ **Instantaneous/ effective Reproduction Number, R_t using Bayesian Inference of Stochastic SIR/SEIR model**

➢ **R_t using Bayesian Choice of Branching Process (Time Since Infection Model)**

- Results & Discussion
- Conclusions

Fig: 1





Background

- After the detection of the ***first cases*** of H1NI Influenza in Mexico in April 2009, the virus spread rapidly around the world

- In India the ***first case (exogenous)*** of H1N1 2009-10 was identified on ***17th May, 2009*** at Hyderabad and then it was spread all over the country at varied intensities in almost all the states & Union Territories of India (***31 state/UTs***)

- The 2009 H1N1 influenza pandemic and recently published research on transmissible forms of highly pathogenic H5N1 has highlighted the need for continued ***public health preparedness*** against the threat of a pandemic.

- ***Mathematical models*** of disease transmission are useful tools for understanding epidemiological dynamics and their dependence on social mixing patterns.

Data Used

- Our analysis is based on ***Indian daily case reports*** of pandemic H1N1 2009. It was readily being available on the website of Ministry of Health & Family Welfare, India. http://pib.nic.in/h1n1/

- We used daily lab-confirmed case reports in ***a complete year*** of pandemic H1N1 from 2009, & ***stratified by region***. (namely: South, North-west, Mid-east and North-east).

- 1st May 2009 Ministry of Health & Family Welfare, India introduced the ***Screening Services*** for international travellers with ILI symptoms at different Airports and Railway Stations.

	01-May '09	14-Aug '09
No. of Centres to Screen	32	83
No. of Doctors in Service	96	225
No. of Paramedics in Service	0	172
No. of persons Screened on day	17949	39752
Cumulative No. of persons Screened	17949	4725725
Total no. of Airports under test	12	26

Table -1: Data Collection Pdm 2009 in India

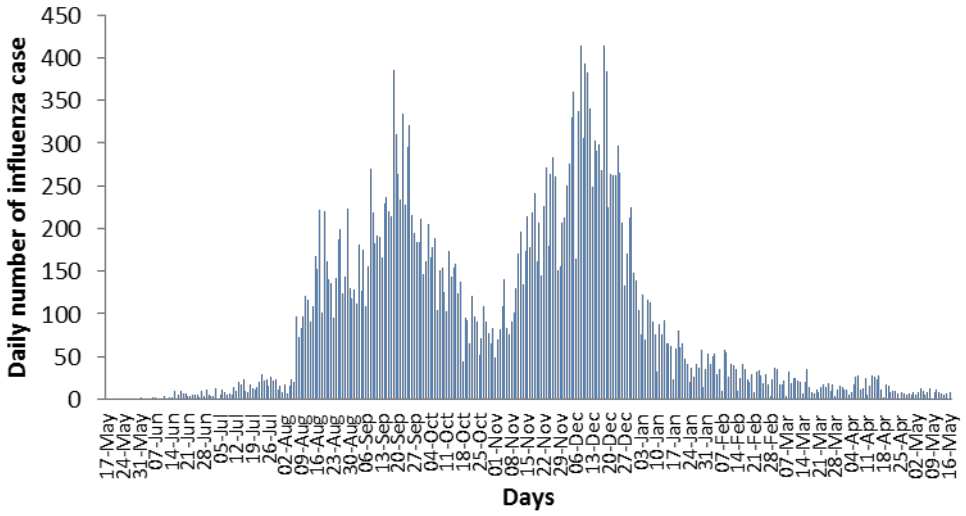


Fig-2 : Daily number of influenza A/H1N1 notifications

Objectives/Aims

- To quantify the ***transmission intensity*** of the pandemic through ***time varying estimation*** of the ***reproduction number***, a key epidemiological parameter which characterises the transmissibility of an emerging infectious disease.

- Here we ***compare different approaches to estimating the reproduction number*** of the 2009 H1N1 pandemic for different regions of India.

Definitions

- Reproduction number
- Basic reproduction number (R_o)
- Instantaneous reproduction number (R_t)
- Effective reproduction number (R_p)

$$R_t = [S(t)/N(t)]R_0 \qquad R_t \leq R_{0,\text{ for all } t}$$

$$R_p = (1-p)R_0$$

* p fraction of population is effectively protected from infection.

Methodology

I. Estimation of R_t using Bayesian Inference of Stochastic SIR/SEIR Model

- Following the method developed by ***Bettencourt & Ribeiro*** (2008).
- The sequential Bayesian estimation of effective reproduction number through a stochastic ***SIR model***

$$\frac{d}{dt}S(t) = -\frac{\beta S(t) I(t)}{N(t)}$$
$$\frac{d}{dt}I(t) = \frac{\beta S(t) I(t)}{N(t)} - \gamma I(t) \qquad \ldots\ldots (1)$$

- ***A stochastic version*** of this model can be formulated by taking the rates on the right-hand side of the population equations (1) to determine the ***mean changes*** (λ) over the time τ of the different compartments of population.

- This usually are evaluated from a ***probability distribution*** $P\{\lambda\}$, with average λ .

➔ $P(.)$ may be assumed as Poisson or **Negative Binomial**

- The ***number of new cases*** at time *t* is $\Delta C(t) = C(t) - C(t-\tau)$, *C(t)* commutative number of cases & $\tau = 1$ day. Then *C(t)* obey the equation

$$dC(t)/dt = \beta S(t) I(t)/N(t) \qquad \ldots.(2)$$

- To find the expression accounting for the evolution of new cases $\Delta C(t+\tau)$, integrate (1) for *I(t)* on to $(t,\ t+\tau)$ obtain

$$I(t+\tau) = I(t)\exp[\gamma(\int_{t}^{t+\tau}(R_0 \frac{S(t')}{N(t')} - 1)dt']$$

$$\approx I(t)\exp[\gamma((R_t - 1)]\tau \qquad \ldots\ldots(3)$$

- ➔ $I(t+\tau) = I(t)b(R_t)$, where

$$b(R_t) = \exp[\gamma\tau(R_t - 1)] = \exp[\tau\theta], \theta = \gamma(R_t - 1)$$

- ➔ $\Delta C(t+\tau) = \Delta C(t)b(R_t) \qquad \ldots\ldots(4)$

- ➔ Then $R_t = 1 + log(b(R_t))/\gamma\tau$

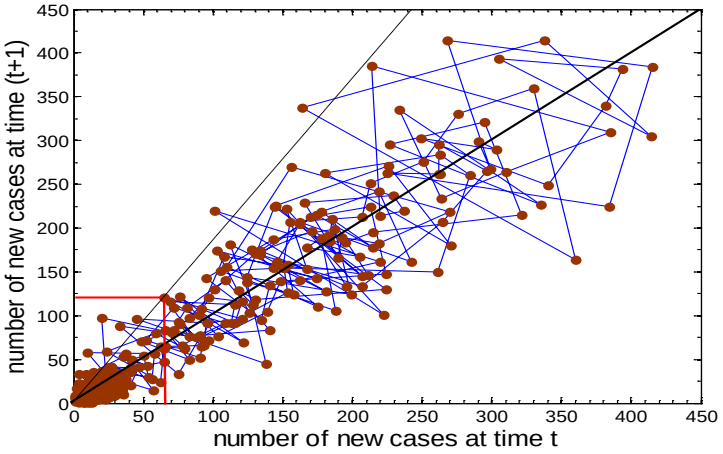


Fig-3: ***Time-delay trajectory diagram*** of $\Delta C(t)$ Vs $\Delta C(t+\tau)$ for Indian data

- But in practice, for emerging infectious diseases ***relative variations in case numbers are large*** (Figure 3, expressing a large fluctuation in new cases), therefore, this simple geometric approach becomes less realistic.

- Which leads to find a ***stochastic estimation procedure*** evaluating the ***probability distribution*** of R_t instead.

- Realistic assumption: $\Delta C(t+\tau) \sim P\{\Delta C(t) b(R_t)\}$, $P\{\lambda\}$ is NB pmf with mean λ.

- In other words, for given R_t (and other parameters like, γ) and $\Delta C(t)$, one can predict the distribution of future case number as, $P[\Delta C(t+\tau) \leftarrow \Delta C(t)/R_t] = P\{\lambda\}$ $\lambda = \Delta C(t) b(R_t)$ for SIR model.

- With this uncertain measure, the parameter estimation can be achieved by using ***Bayesian approach***

$$P[R_t | \Delta C(t+\tau) \leftarrow \Delta C(t)] = \frac{P[\Delta C(t+\tau) \leftarrow \Delta C(t)/R_t] P[R_t]}{P[\Delta C(t+\tau) \leftarrow \Delta C(t)]} \quad (5)$$

- ***Prior*** $P[R_t]$ for the posterior at time *t+T*, *T=1,* We assumed initial prior **U (0, 3)**

The same can be derived for **SEIR** model. We have seen the effect of ***Fatal Risk*** to quantify the disease severity. ***Table: 2***

	SIR	SEIR
Excluding Fatal Risk $\delta = 0$	$\theta = \gamma(R_t - 1)$ *Equation (3)*	$\theta = \frac{k+\gamma}{2}\left[-1+\sqrt{1+\frac{4k\gamma(R_t-1)}{(k+\gamma)^2}}\right]$
Including Fatal Risk $\delta > 0$	$\theta = (\gamma+\delta)(R_t - 1)$	$\theta = \frac{k+\gamma+\delta}{2}\left[-1+\sqrt{1+\frac{4k(\gamma+\delta)(R_t-1)}{(k+\gamma+\delta)^2}}\right]$

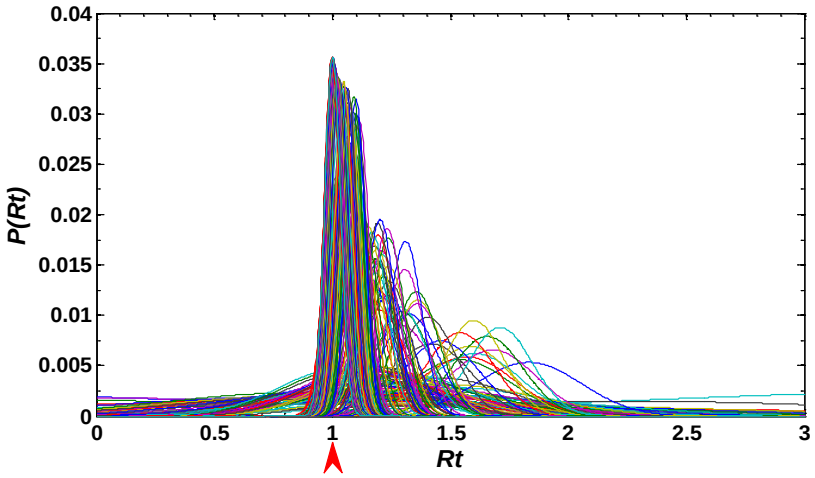


Figure 4: The complete density plot of Posterior Distribution of R_t through Sequential Bayesian Estimate from successive daily iteration.

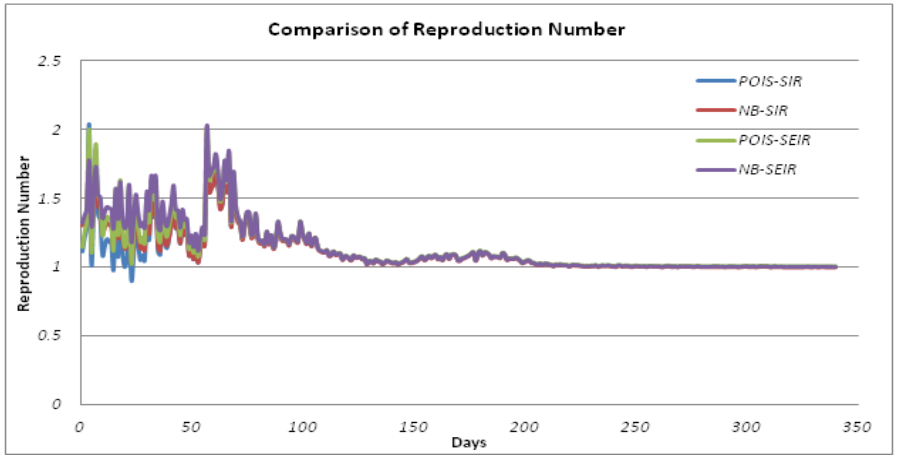
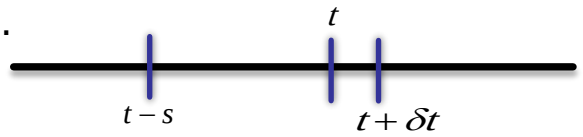


Figure 5: Estimation of R_t for different choice of uncertainty and Model choice.

II. R_t using Bayesian Choice of Branching Process (Time Since Infection Model)

- Following Fraser (2007), we assume that the ***distribution of infectiousness through time after infection is independent of calendar time***.



- Transmission can be modelled as a ***Poisson process.***

i.e. the probability that someone with symptoms onset at time $t-s$ infects someone else who will show symptoms in a short time period $[t;t+\delta t]$ is $R_t w_s \delta t$, where R_t is the instantaneous reproduction number at time *t* and w_s is the discrete SI distribution.

- Therefore the incidence at time t, is ***Poisson distributed*** with ***mean*** $R_t \sum_{s=1}^{t} I_{t-s} w_s$.

- Assume, transmissibility is ***constant over a time period*** $[t-\tau;t]$, measured by a reproduction number $R_{[t-\tau;t]}$, the ***likelihood*** of the incidence during this time period, $I_{t-\tau},\ldots,I_t$, given the reproduction number $R_{[t-\tau;t]}$, conditional on the previous incidences $I_0,\ldots I_{t-\tau-1}$, is:

$$P\left(I_{t-\tau},\ldots,I_t \mid I_0,\ldots,I_{t-\tau-1}, w, R_{[t-\tau;t]}\right) = \prod_{s=t-\tau}^{t} \frac{\left(R_{[t-\tau;t]}\Lambda_s\right)^{I_s} e^{-R_{[t-\tau;t]}\Lambda_s}}{I_s!} \qquad (5)$$

- Where $\Lambda_t = \sum_{s=1}^{t} I_{t-s} w_s$.

- Using a ***Bayesian framework*** with a ***Gamma distributed prior*** with parameters (a,b) for $R_{[t-\tau;t]}$, the ***posterior joint distribution*** of $R_{[t-\tau;t]}$ is proportional to:

$$R_{[t-\tau;t]}^{\,a+\sum_{s=t-\tau}^{t} I_s - 1} e^{-R_{[t-\tau;t]}\left(\sum_{s=t-\tau}^{t}\Lambda_s + \frac{1}{b}\right)} \prod_{s=t-\tau}^{t} \frac{\Lambda_s^{I_s}}{I_s!} \qquad (6)$$

- Therefore, the ***posterior distribution*** of $R_{[t-\tau;t]}$ is a ***Gamma distribution*** with parameters $\left(a+\sum_{s=t-\tau}^{t} I_s, \frac{1}{\frac{1}{b}+\sum_{s=t-\tau}^{t}\Lambda_s}\right)$.
- In particular, the ***posterior mean*** of $R_{[t-\tau;t]}$ is $\frac{a+\sum_{s=t-\tau}^{t} I_s}{\frac{1}{b}+\sum_{s=t-\tau}^{t}\Lambda_s}$, and the ***posterior coefficient of variation*** of $R_{[t-\tau;t]}$ is $\frac{1}{\sqrt{a+\sum_{s=t-\tau}^{t} I_s}}$.

- Choice of the time period $[t-\tau;t]$:Imposing a ***posterior CV smaller than a predetermined threshold value*** $CV_{threshold}$. This gives a ***minimum bound*** to the number of incident cases in each time period as $\sum_{s=t-\tau}^{t} I_s \geq \frac{1}{CV_{threshold}^{\,2}} - a$, which is independent of serial interval distribution . (7 days)

- When can we start estimating *R* ? : i. Estimation of $R_{[t-\tau;t]}$ depends on all observations in $[t-\tau;t]$.

ii. The SI distribution also provides the guideline on the **START**: indeed, estimation before at least one ***generation of cases*** has been observed is difficult.

iii. Advisable START:Estimating $R_{[t-\tau;t]}$only after 12 cases have been observed at total (1/0.3)^2=11.11.

Results & Discussion

- Overall Two Waves
- Varied Dynamics Patterns for Different States/regions
- CFP is not uniform across the country

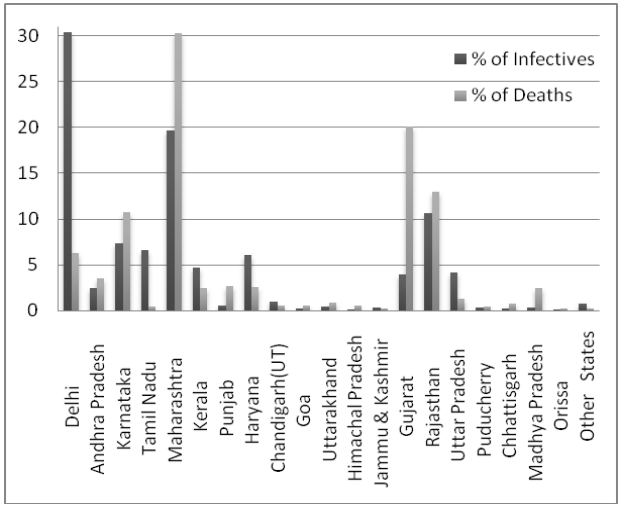


Figure 6: Percentage of infective, and percentage of deaths for different states of INDIA

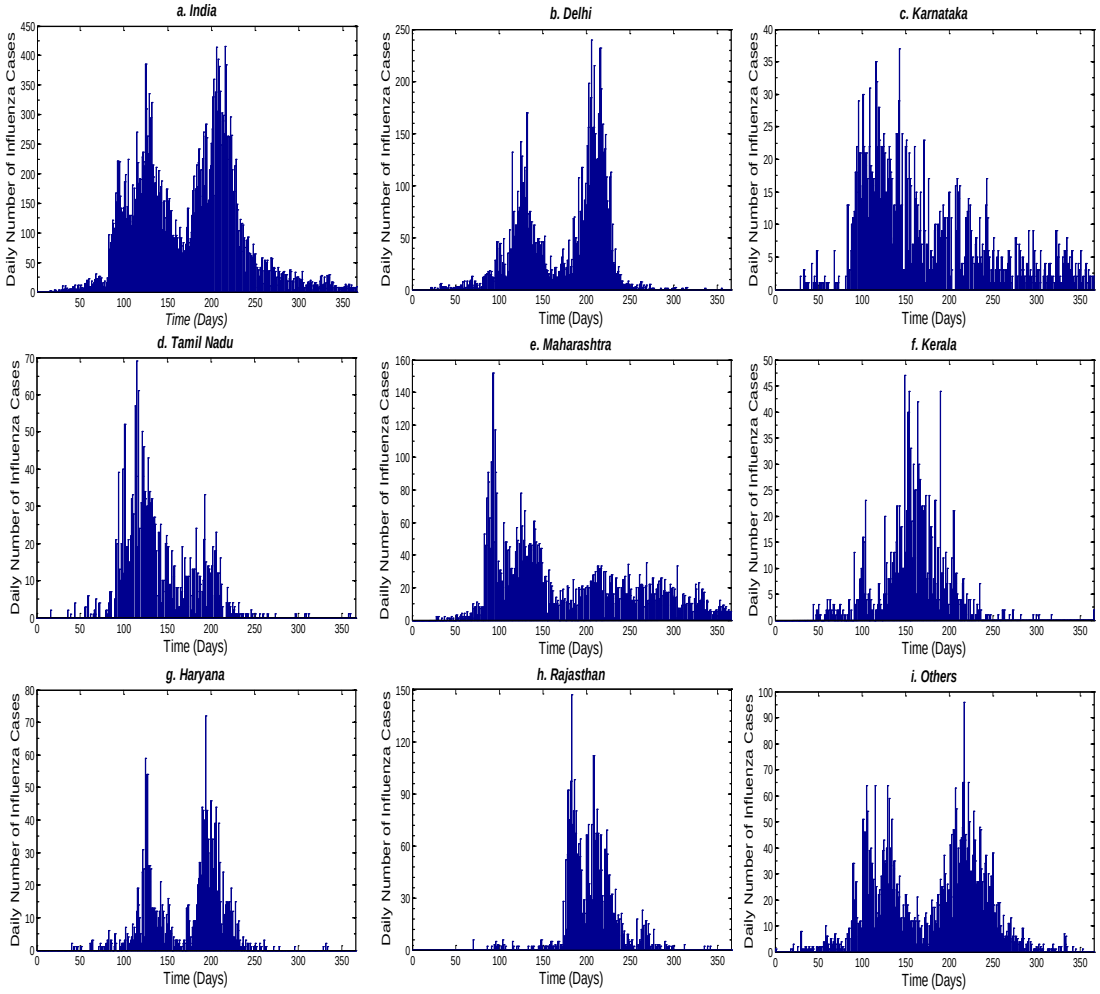


Figure 7: Daily number of influenza A/H1N1 notifications in different segments (fairly affected states) of India during pandemic 2009-10,

- The availability of ***stratified data***
- Provides a unique opportunity to ***compare the spread*** of a single virus in different region of the country
- Also encourages to ***gain insight*** into the ***dynamics of spread*** and the ***factors modifying transmission intensity***.

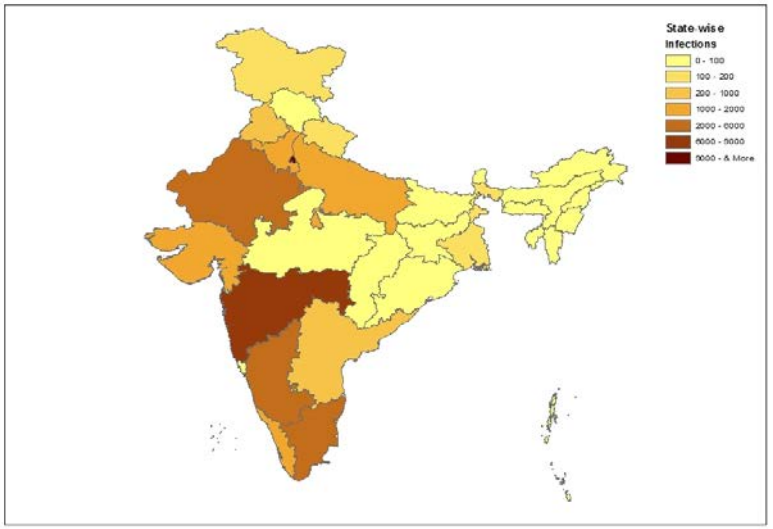


Fig 8: Spread of H1N1 virus (as on 17th May 2010)

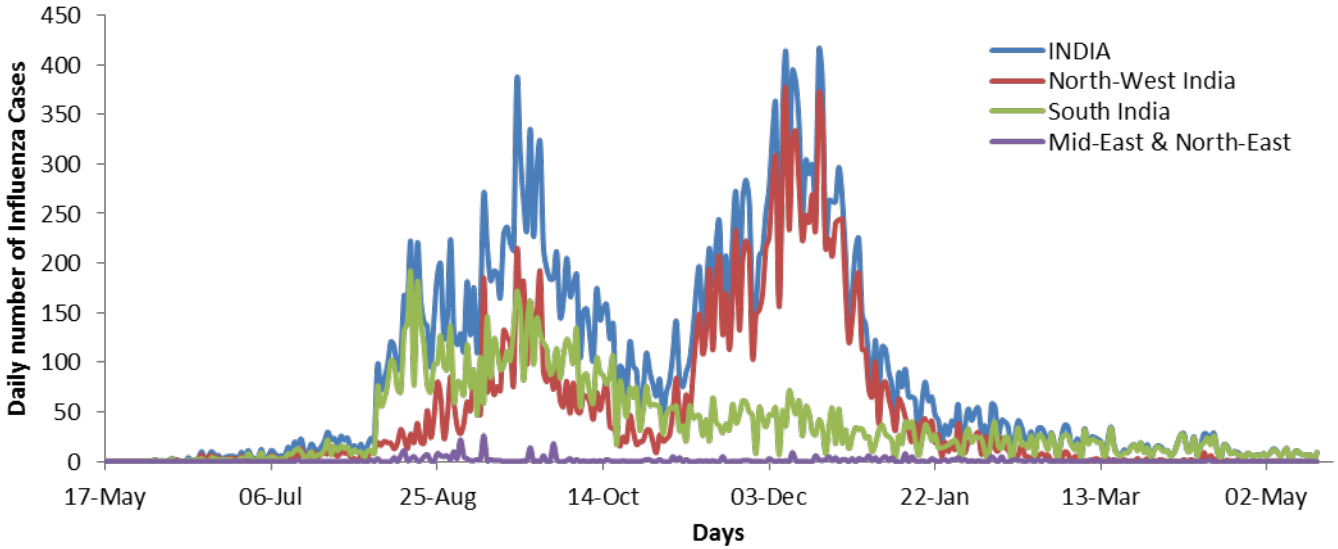
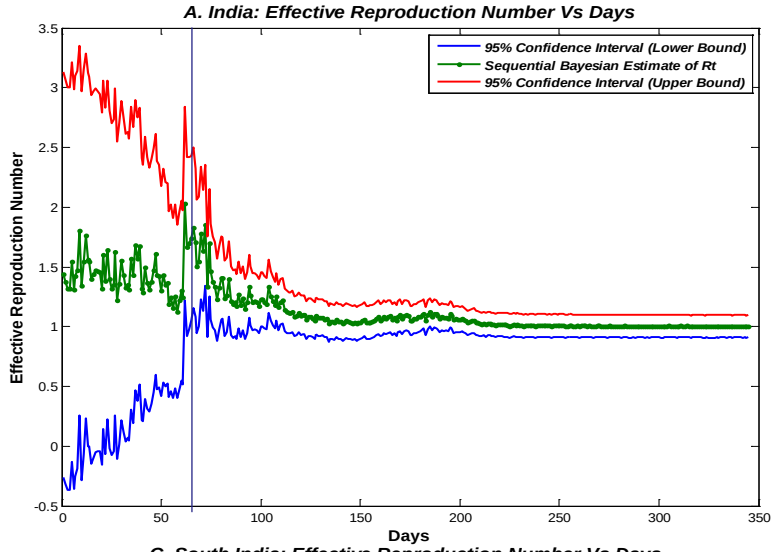
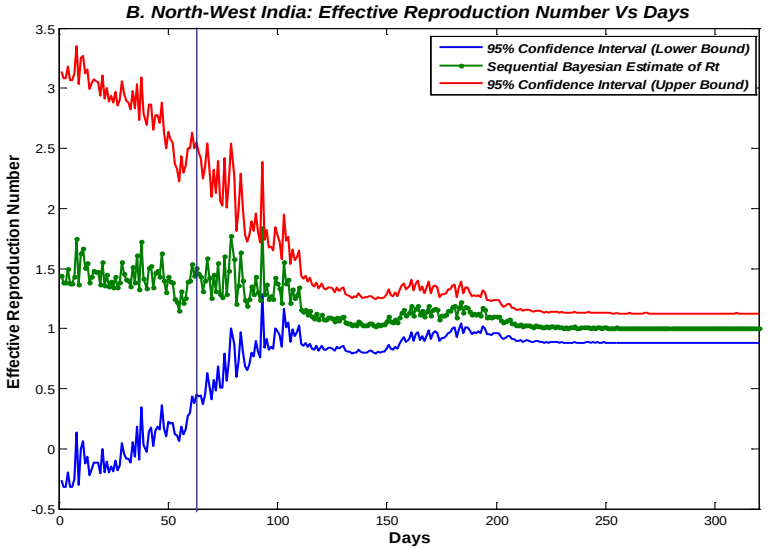


Fig 9: Disease dynamics for the three stratified regions of India

Method-I





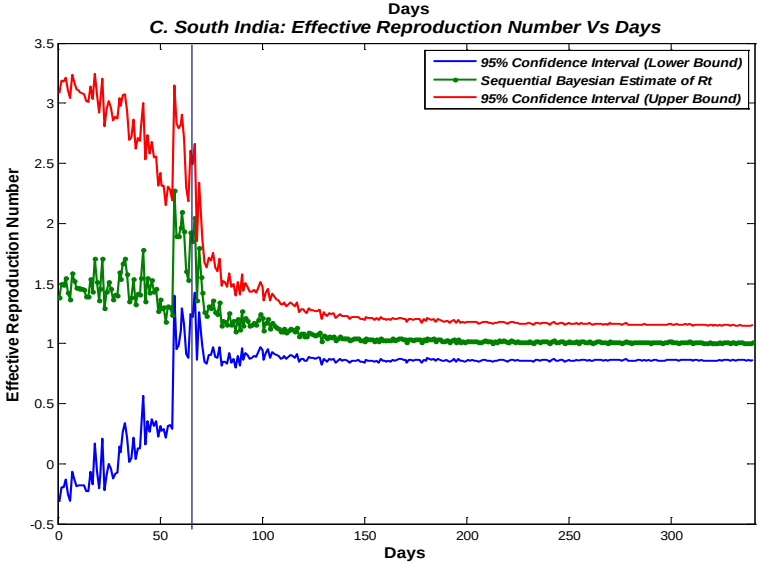
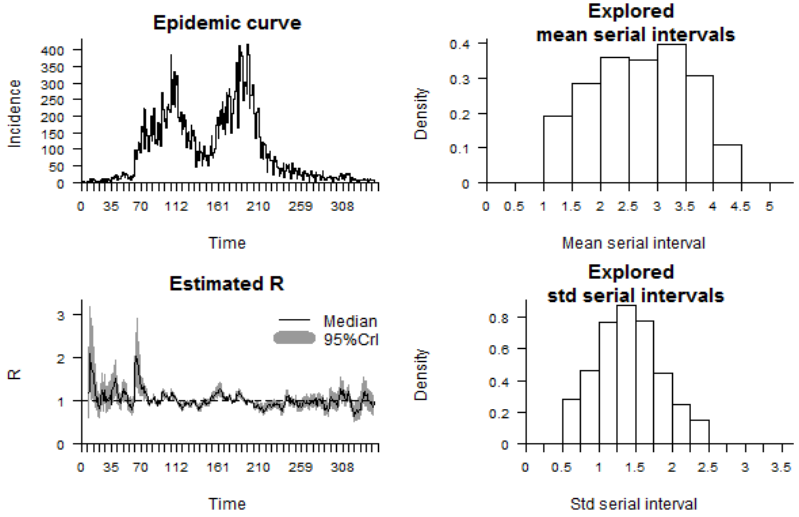


Fig-10: Sequential Bayesian Estimate of Rt with 95% CI

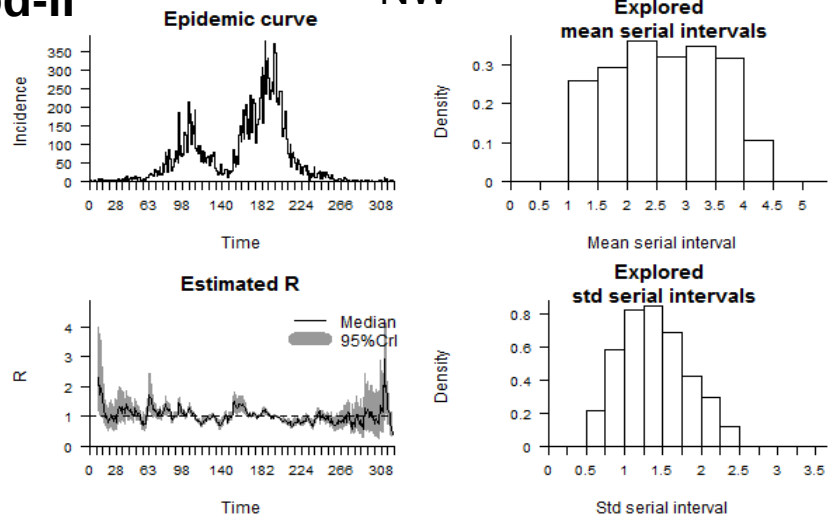
- Vertical line of Cut-off been derived from the Exponential Growth rate (64 days)
- At the end of the outbreak R_t is tending to the value 1.
- NW population has a higher effect of second wave.

Method-II

INDIA



NW



South

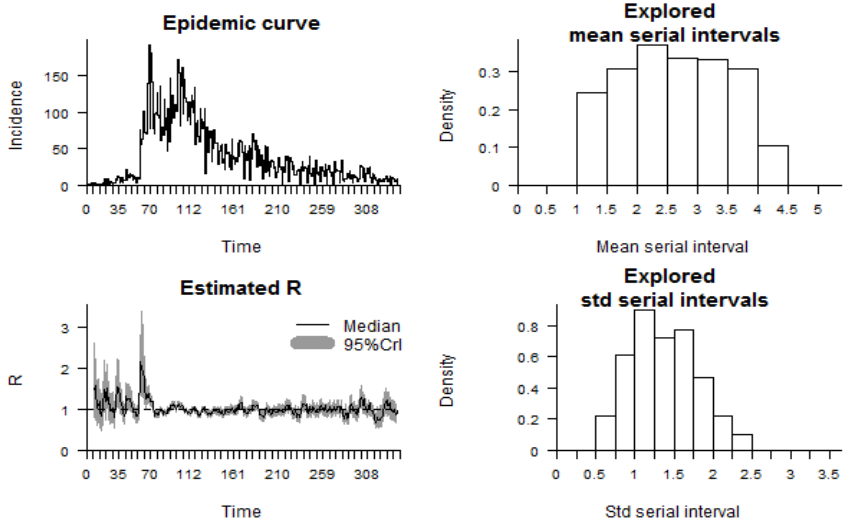


Table-3: Parameter Choice

Mean Prior	5
Std Prior	5
SI Uncertainty	Y
Parametric SI	Y
Sample Size R	500
Sample Size SI	500
Length	7
Estimation Time Step	1

Fig 11. Estimated reproduction number *R(t)* (posterior median in black line) with 95% CrI (grey zone) for different regions

Estimation of reproduction number under different model frameworks

Table-4

Model Choice	INDIA	North West	South
R_0 : (Exponential Growth rate of SEIR)	1.46 (1.11, 1.99)	1.42 (1.05, 2.07)	1.48 (1.01, 2.09)
R_t : (Bayesian choice of Stochastic SIR)	1.41 (0.22, 2.74)	1.41 (0.01, 2.87)	1.46 (0.13, 2.91)
R_t : (Bayesian choice of Stochastic SEIR)	1.30 **(0.14, 2.48)**	1.31 (0.06, 2.70)	1.37 (0.14, 2.71)
R_t : (Bayesian Choice of Branching Process (Time Since Infection Model)	1.24 **(0.90, 1.68)**	1.21 (0.80, 1.79)	1.37 (0.87, 1.73)

- In the beginning of the outbreak SOUTH region had a ***higher intensity*** compare to the North-West region of India.
- Estimates through Bayesian choice of Stochastic models are ***less confident*** than that of trough Time Since Infection Model.
- And off-course $R_t < R_0$

Table 5: Estimates of basic reproduction numbers and effective reproduction numbers with 95% confidence intervals of the influenza pandemic 2009-10 for India and its different segments (fairly affected states).

* Effective Reproduction Number (R_t) has been calculated for each segment at the time t = cut-off point, which is derived for the estimation of Basic Reproduction Number (R_0).

Sl. No.	States	Total Number of Infections	Percentage of Infections	Total Number of Deaths	Percentage of Deaths	Case Fatality Proportion (CFP) (per 100)	Basic Reproduction Number-BRN (R_0)	Effective Reproduction Number -ERN (R_t) * (with 95% CI)	Doubling Time
1	Delhi	9697	30.38	95	6.23	0.98	**1.52**	1.24 (0.23, 2.25)	8.89
2	Karnataka	2350	7.36	164	10.75	6.98	1.32	1.50 (0.60, 2.41)	13.86
3	Tamil Nadu	2090	6.55	7	0.46	0.34	1.50	1.68 (0.79, 2.57)	9.12
4	Maharashtra	6283	19.68	461	30.23	7.34	1.49	1.35 (0.25, 2.44)	9.24
5	Kerala	1482	4.64	38	2.49	2.56	1.35	1.03 (0.21, 1.85)	12.84
6	Haryana	1948	6.10	39	2.56	2.00	1.33	1.31 (0.00, 2.74)	13.33
7	Rajasthan	3380	10.59	198	12.98	5.86	**1.17**	1.75 (0.92, 2.58)	25.67
8	Others	4694	14.70	523	34.30	11.14	1.45	1.29 (1.07, 1.51)	10.05
9	INDIA	31924	100.00	1525	100.00	4.78	**1.46**	**1.46 (1.15, 1.77)**	9.90

Gani, S. R.; Ali, Sk. Taslim; Kadi, A. S.; The transmission dynamics of pandemic influenza A/H1N1 2009-2010 in India; Current Science (00113891);10/25/2011, Vol. 101 Issue 8, p1065

Conclusions

- A ***variety of different model frameworks*** have been utilised successfully to characterise the transmission dynamics of the 2009 H1N1 pandemic in India.
- The estimates of reproduction number ***similar*** to those seen in other countries. (European)
- The similarity of the estimates obtained with different methods demonstrates a degree of ***robustness*** to the values obtained.
- More work is required to understand the ***causal factors*** underlying the variation in the temporal dynamics of the pandemic seen in different regions of India.

Future Work

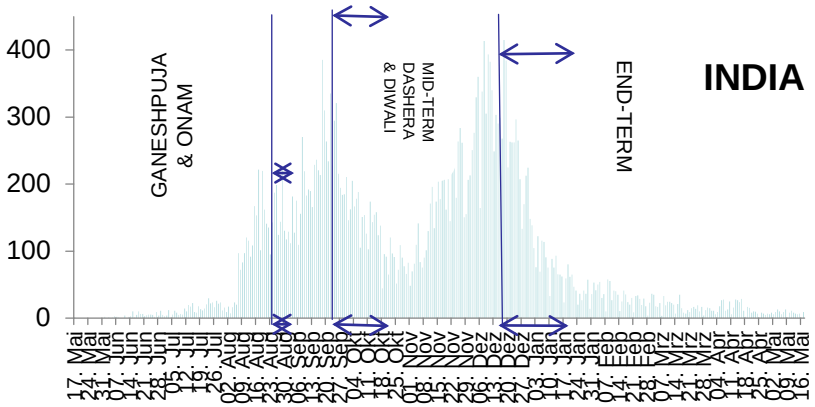


Fig-12: Role of School closure/ public holidays in dynamics

References

[1] C. Fraser, Estimating individual and household reproduction numbers in an emerging epidemic, PLoS ONE 2 (2007) e758.
[2] Cori A, Fraser C, Ferguson N, Cauchemez S, A new tool to estimate time-varying reproduction numbers: supplementary material (under publication)
[3] Ministry of Health and Family Welfare, Government of India, 20 May, 2010. http://www.mohfw.nic.in/ and http://pib.nic.in/h1n1/h1n1.asp.
[4] Bettencourt, L. M. A. & Ribeiro, R. M., Real time Bayesian estimation of the epidemic potential of emerging infectious diseases, PLoS one, 2008 May, vol. 3, issue 5, e2185.
[5] Lipsitch, M. et. al. Transmission dynamics and control of severe acute respiratory syndrome. Science 2003, 300, 1966–1970.
[6] Gani, S. R.; Ali, Sk. Taslim; Kadi, A. S.; The transmission dynamics of pandemic influenza A/H1N1 2009-2010 in India; Current Science (00113891);10/25/2011, Vol. 101 Issue 8, p1065

Acknowledgement:

Supervisors:

- **Prof. Neil M Ferguson**, DIDE, Imperial College London, UK
- **Prof. A S Kadi**, Department of Statistics, Karnatak University, India

Institutional/Funding

- MRC Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, Imperial College London, UK.
- Commonwealth Scholarship Commission, UK
- University Grant Commission (UGC-MANF), New Delhi, India.

Thank You!