Modeling Infectious Diseases in the Big Data Era: Challenges and Promises

謝英恒 Ying-Hen Hsieh
中國醫藥大學公共衛生系
China Medical University
Taichung, Taiwan

1990's: HIV Data in Thailand

- HIV Serosurveillance of injecting drug users (IDU) and commercial sex workers (CSW)
- taken semi-annually in each of the 76 provinces during 1989 to 1995, with almost 20,000 serotests given every 6 months.

Map of Thailand

Thai HIV Serosurveillance data of IDUs and CSWs taken semi-annually during June 1993 to June 1995 (Hsieh et al. *Stat. Med.* 2000)

Table I. Thai sentinel data (Round 9–13) for intravenous drug users and commercial sex workers.

Date	IVDU			Direct CSW			Indirect CSW		
	HIV+	Total	%	HIV+	Total	%	HIV+	Total	%
06/93	1234	3515	35.11	2731	8979	30.42	608	7041	8.64
12/93	1276	3388	37.66	2412	8170	29.52	721	7793	9.25
06/94	1033	3234	31.94	2441	8653	28.21	703	8024	8.76
12/94	346	985	35.13	1313	4014	32.71	411	4186	9.82
06/95	1235	3585	34.45	—	—	—	—	—	—

— denotes not available.

1990's: HIV Data in Thailand

- The high mobility of these high-risk groups, especially of the CSWs (Brown & Sittitrai 1993), renders the provincial data highly volatile from survey in one year to another and difficult to use in our estimates.
- We therefore confine our study to using the estimates for national-wide totals.

TABLE 1. Estimates for the numbers of HIV-infected CSWs, HIV prevalence rates, and the estimated total CSW population sizes during 1993–1994. (Hsieh *JAIDS* 2002)

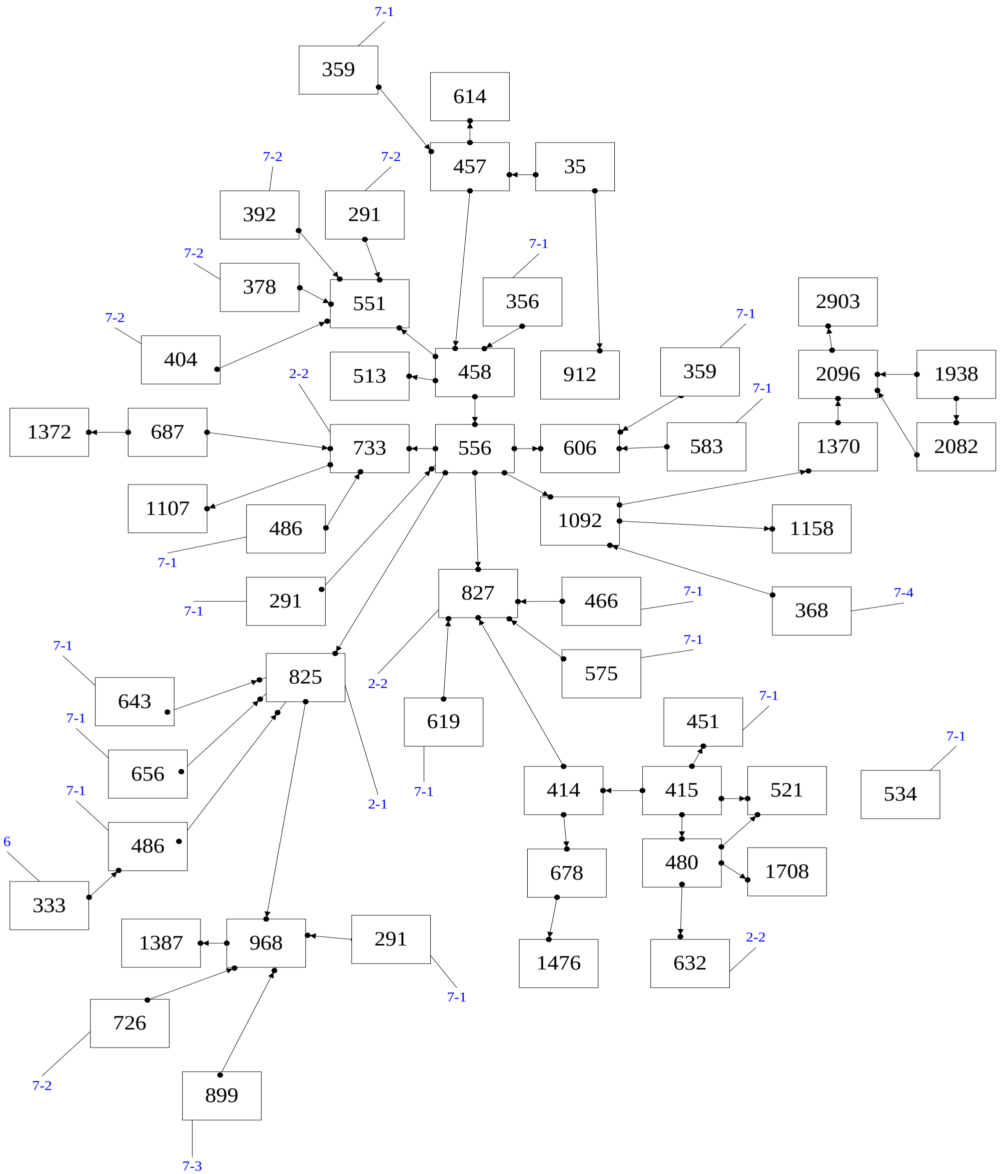
	Estimated number infected	HIV prevalence rate (%)[a]	Estimated total population size[b]
Direct CSW			
June 1993	54,595	30.42	179,471
December 1993	60,452	29.52	204,783
June 1994	64,157	28.21	227,426
December 1994	66,445	32.71	203,134
Indirect CSW			
June 1993	15,181	8.64	175,706
December 1993	16,275	9.25	175,946
June 1994	16,903	8.76	192,957
December 1994	17,171	9.82	174,857

(a) HIV prevalence rate computed from nationwide numbers of HIV-seropositive persons divided by number of tests in the HIV sentinel data.
(b) Median estimate divided by HIV prevalence rate.

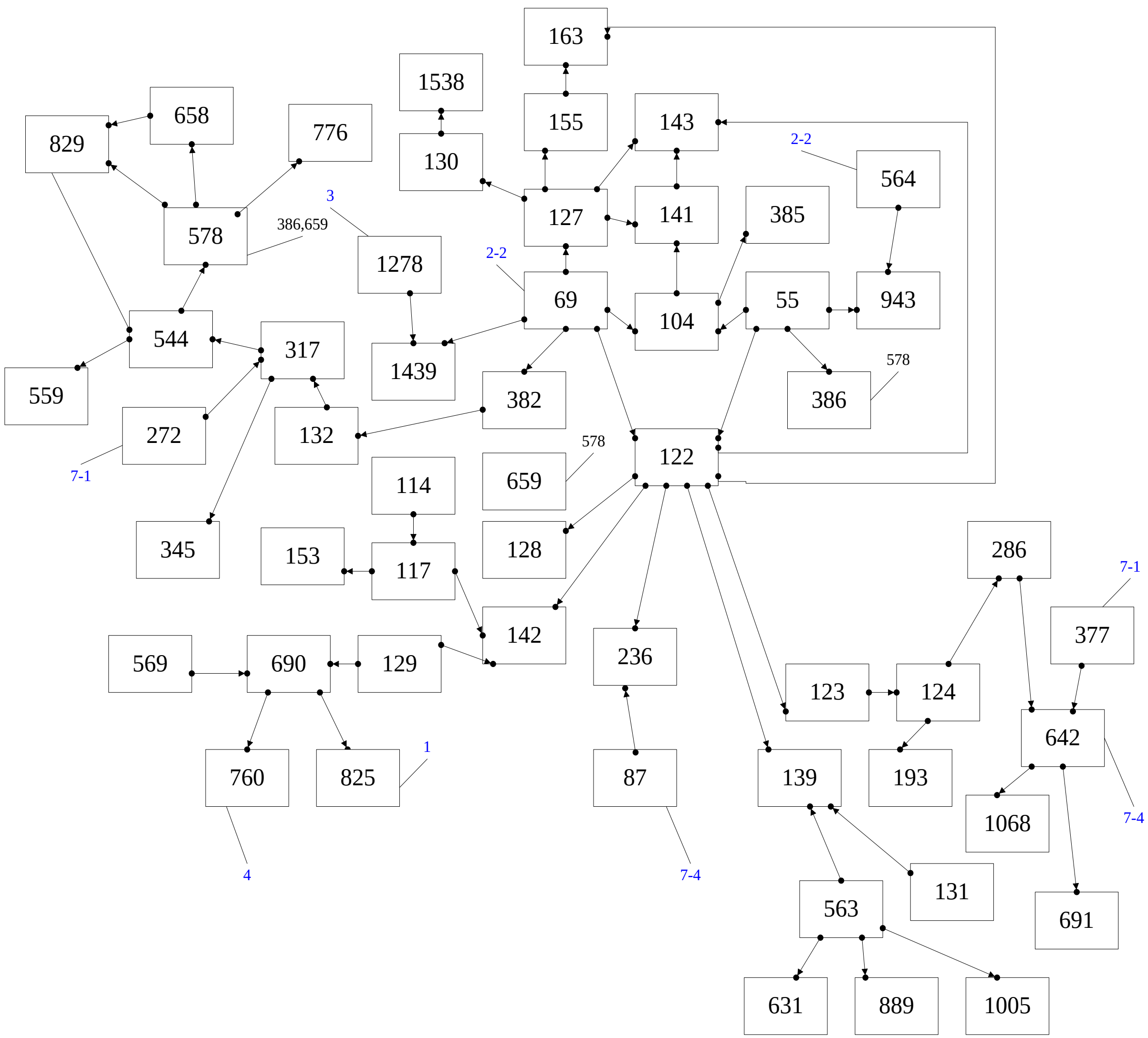
On Partner Notification Program in Cuba, 1991-2000 (Hsieh et al., *AIDS* 2000)

- Contact tracing of HIV-positive individuals to trace their sexual contacts.
- HIV tests were given to these contacts every 3 months for up to 1 year after the last sexual contact with an HIV-positive individual.
- Data from ~2500 cases in contact networks of up to 700 individuals, each case with 0~82 traced contacts within his/her network.

Network Diagram 1 (56 cases)



Network Diagram 2-1 (58 cases)



6 degrees of separation (small world, network model):

(Watts DJ, Strogatz SH. 1998. *Nature* 393 (6684): 440–442; cited >36500 time)

Collective dynamics of 'small-world' networks

Duncan J. Watts* & Steven H. Strogatz

Department of Theoretical and Applied Mechanics, Kimball Hall, Cornell University, Ithaca, New York 14853, USA

Networks of coupled dynamical systems have been used to model biological oscillators[1–4], Josephson junction arrays[5,6], excitable

• • • • • • • • • • • • •

synchronizability. In particular, infectious diseases spread more easily in small-world networks than in regular lattices.

But still unaccounted for…

- Timing of contacts
- Mode of contacts (homosexual? heterosexual?)
- Frequency of contacts (activity level)
- Partnership: steady? multiple? concurrency?

Geographical map of 8439 SARS cases as of 7/3, 2003 (# deaths later adjusted to 774) (Source: WHO)

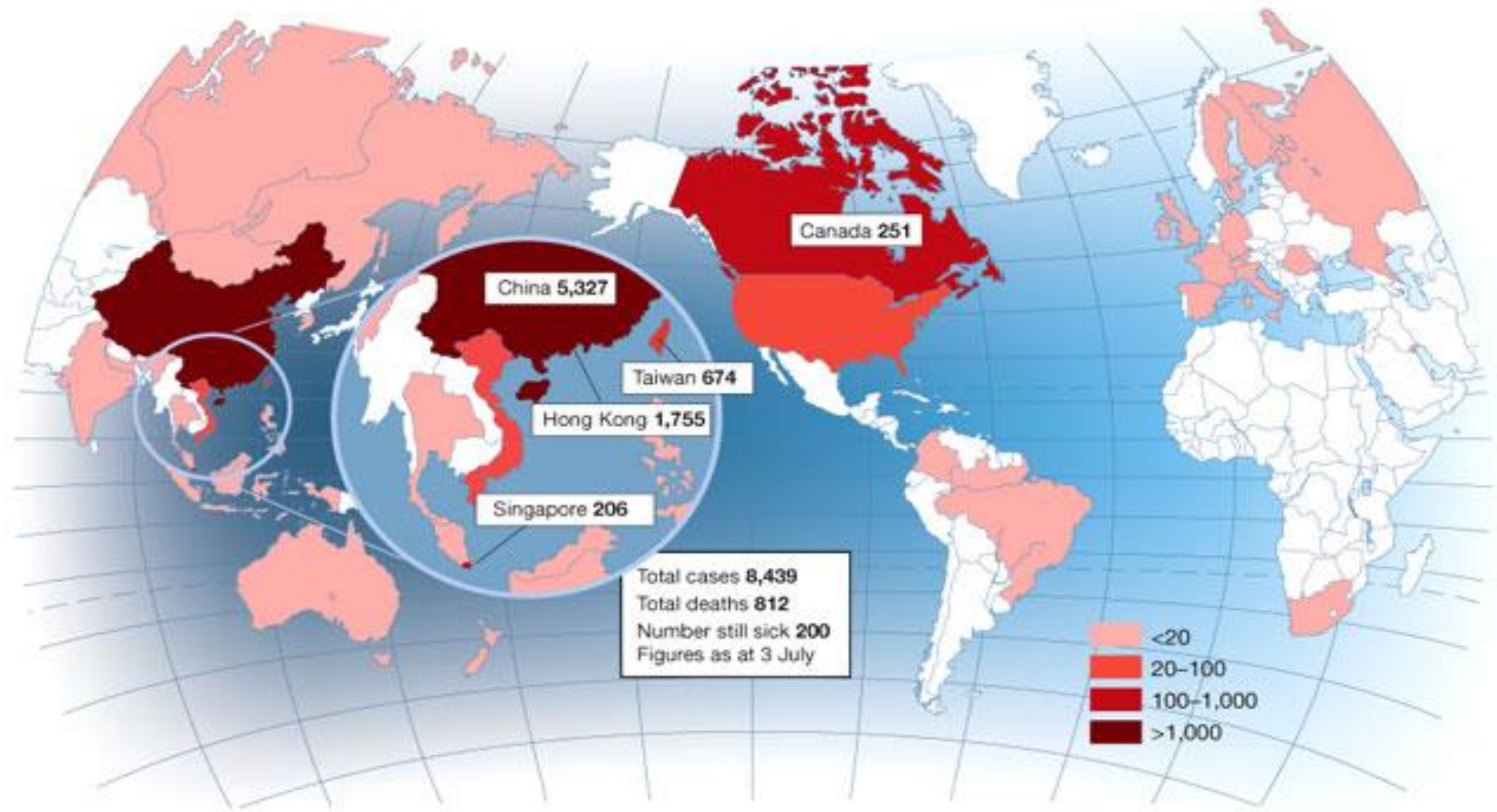


Table 4. Summary of quarantine in Taiwan during SARS outbreak, 2003. Total: 480 confirmed cases and 85 deaths, updated end of 2004. (reproduced from Hsieh et al. *EID*, 2005)

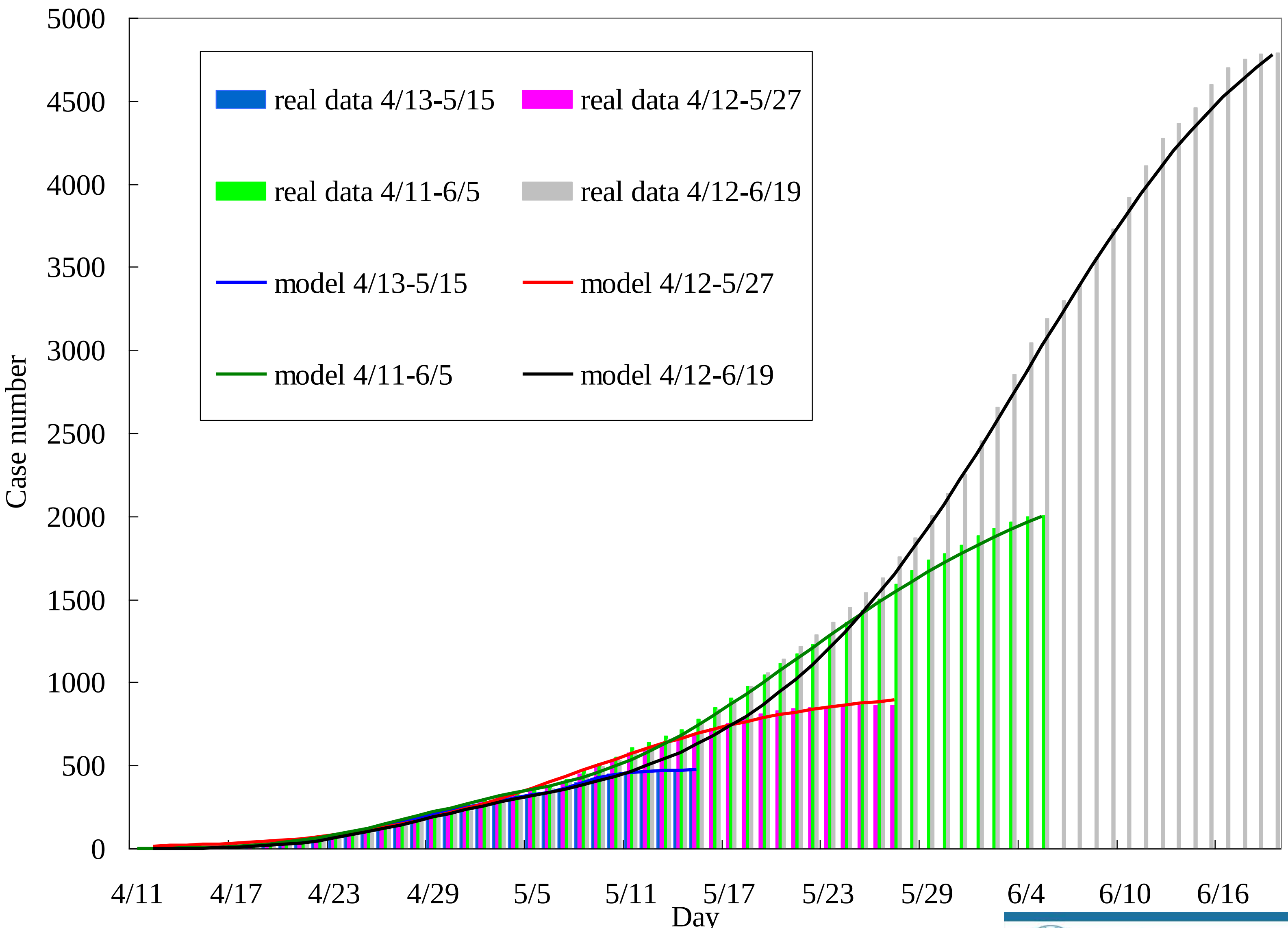
Level/Reason for quarantine	Number of persons quarantined	Officially confirmed cases with PCR(+) or antibody(+) [N=346]	Suspected or R/O cases with PCR(+) or antibody(+)[N=134]
Level A (from 3/17)			
Family members	7,921	8	2
Classmates and teachers	16,564	1	0
Health-care workers	2,409	0	2
Others[a]	19,224	6[d](1)	2
All others[b]	9,514	2	0
Subtotal	55,632	17	6
Level B (from 4/28)	95,828	0	(1)
Total	151,460	17(1)	7(1)

[a]Passengers and drivers of domestic public transportation traveling for 1 hour or more in the same bus or train cabin with a SARS case, persons who had contact with a person under quarantine for receiving care in a medical facility where cluster infection had occurred, and homeless persons.
[b]Co-workers and friends of SARS case, airplane passengers who sat within three rows from or stayed in the same room as a SARS case, and persons with missing information.
[c]There were two cases who were quarantined as classmates and teachers of SARS patients, but were also relatives of other cases.
[d]One case had onset of symptoms two days after the end of quarantine.
Note that the imported cases are in parenthesis ( ).

SARS 2003 in Taiwan

- In Taiwan, there were a total of 480 cases and 87 deaths (18.1%). (Hsieh et al. *EID* 2005; Hsieh et al. *BMB* 2007)
- 301 (77.3%) of 390 cases with a confirmed source of infection had been infected in a hospital, of which 67 (22.3%) had died. (Hsieh et al. *JTB* 2014)

“Real-time” epidemic modeling: Model fit of 2009 Canada pH1N1 data with the Richards model (Hsieh, Fisman, Wu, 2010).



中國醫藥大學
CHINA MEDICAL UNIVERSITY

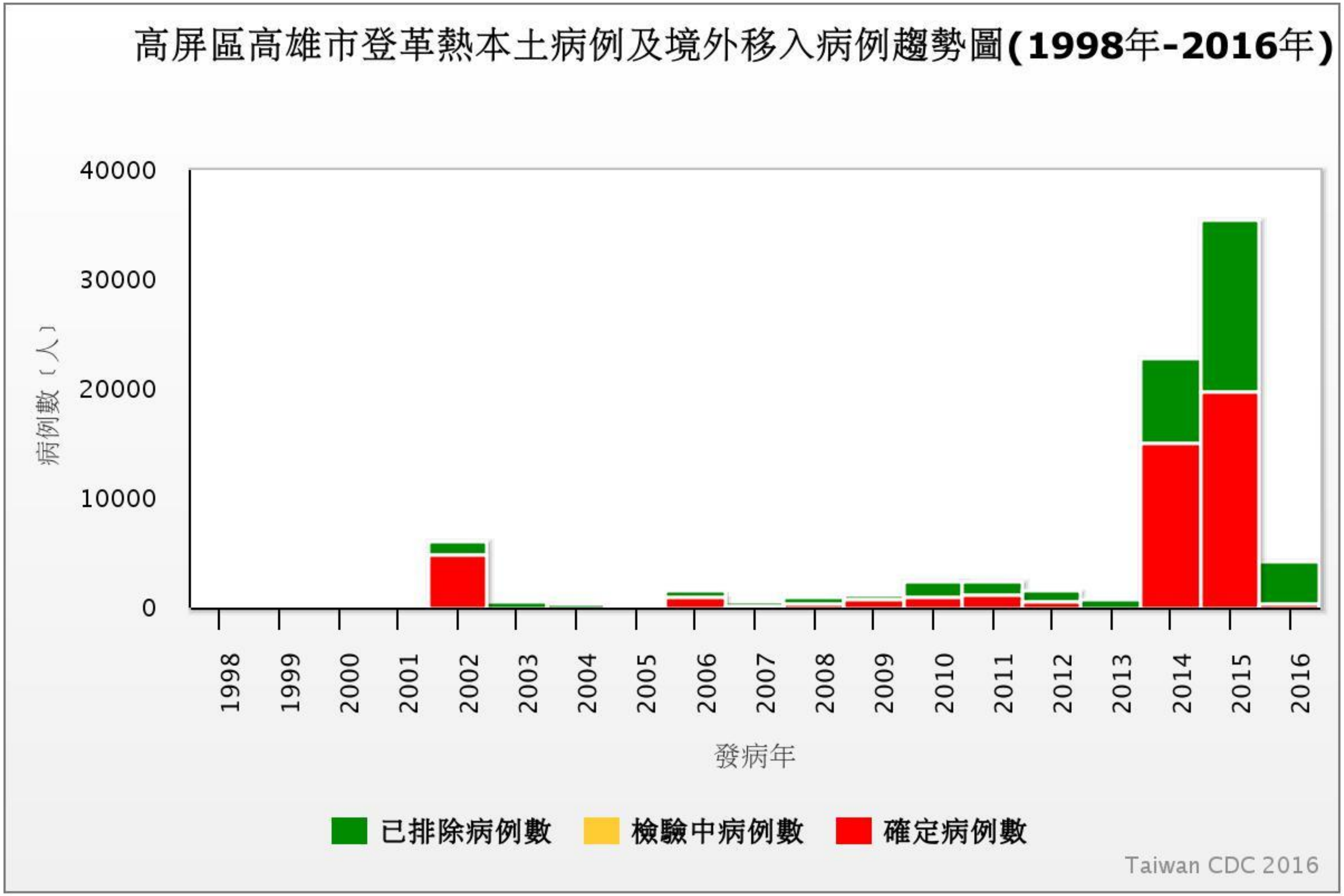
Y.H. Hsieh

About dengue…

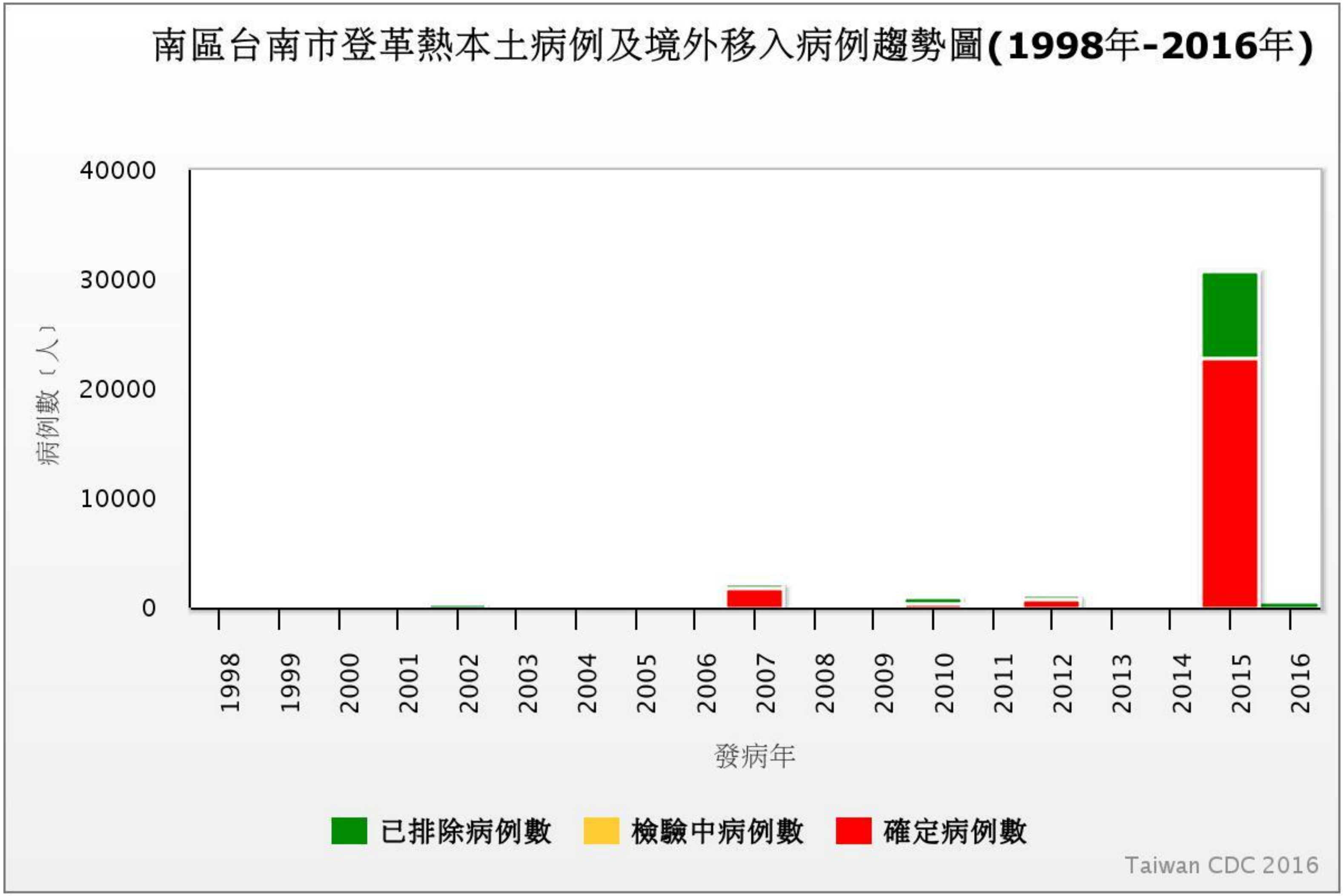
Question:

Can we predict dengue outbreak?

Kaohsiung Annual Dengue DF/DHF case number (1998-2016)



Tainan Annual Dengue DF/DHF case number (1998-2015)



Can we predict future dengue outbreak?

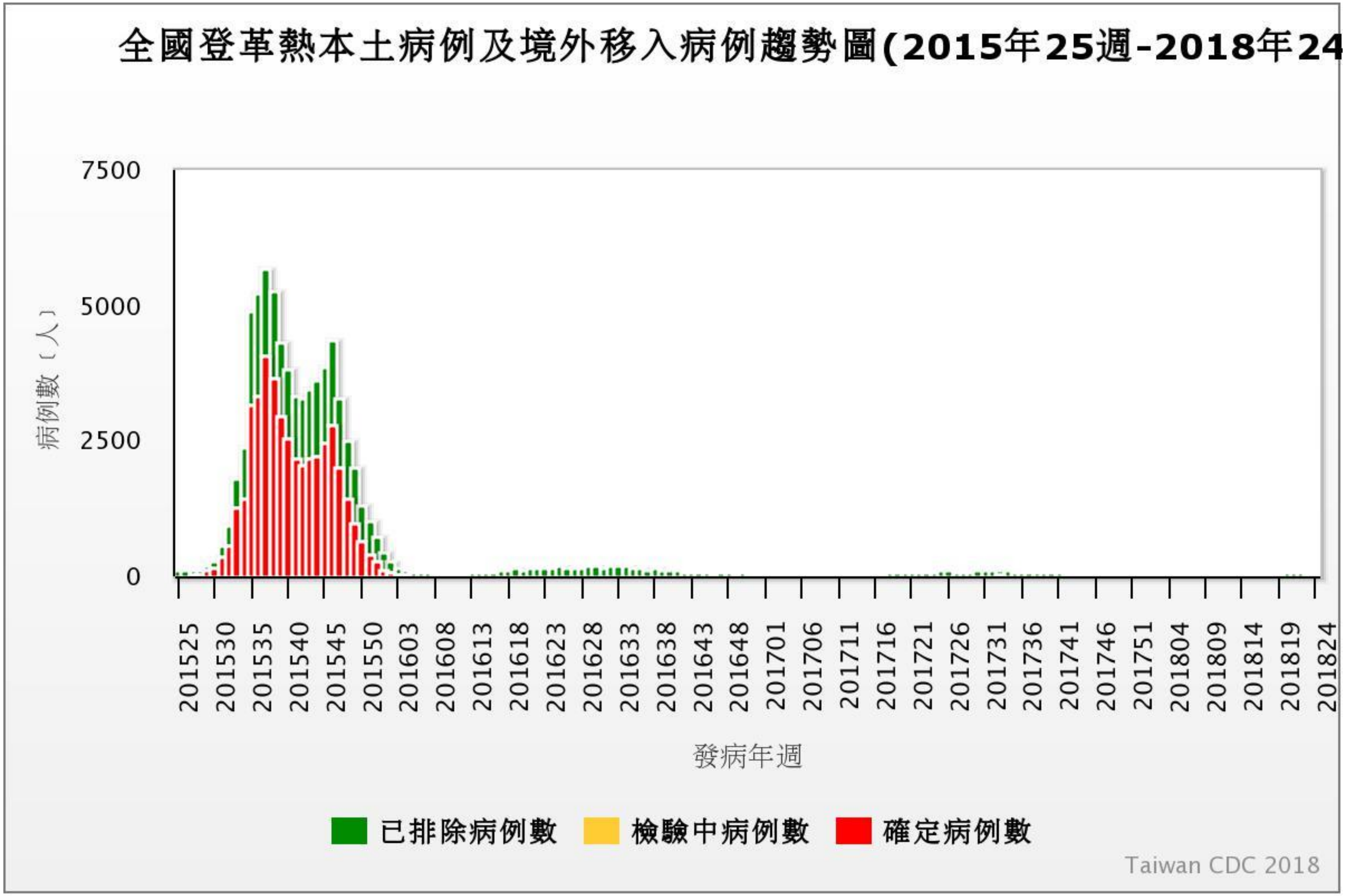
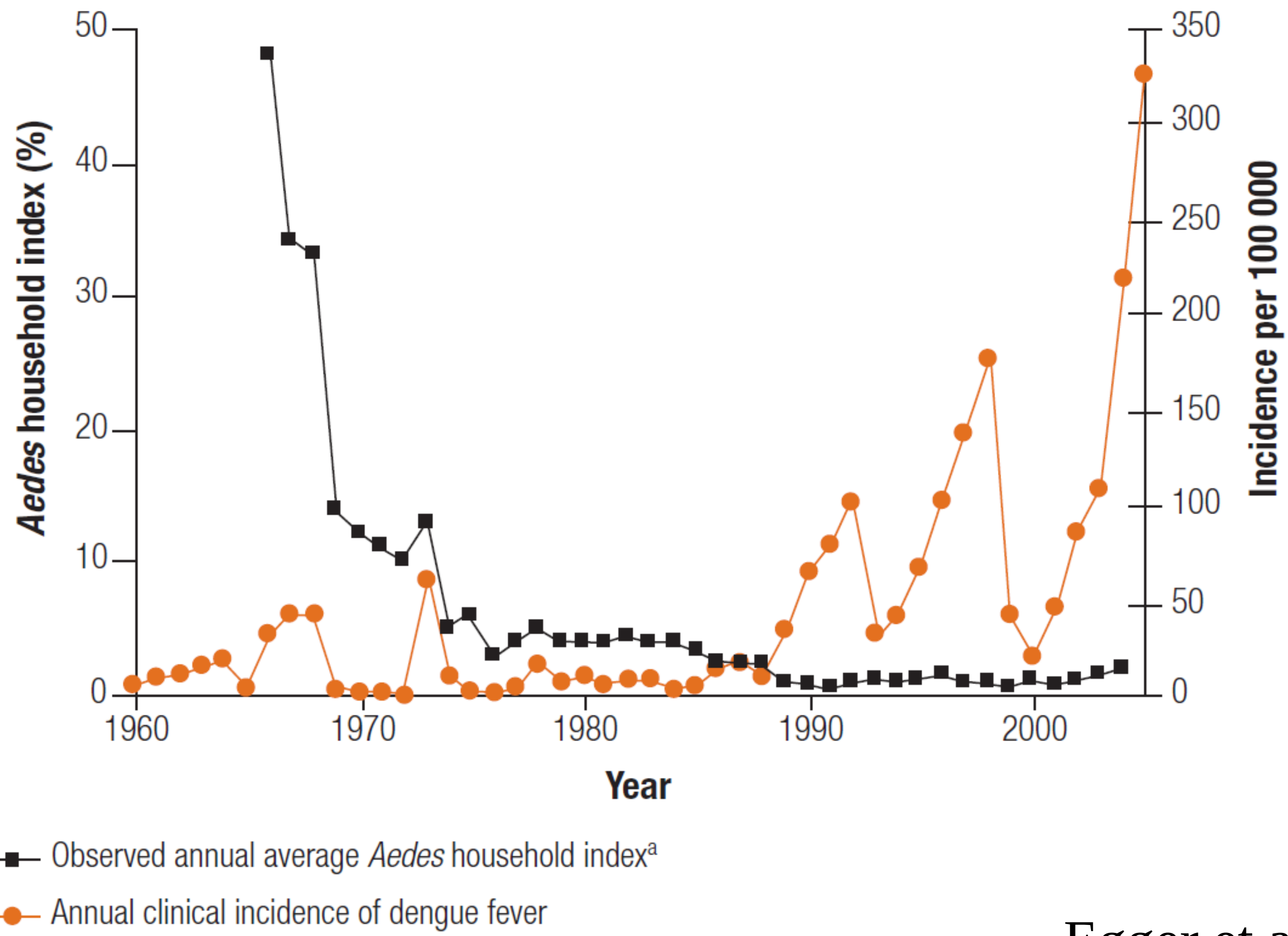
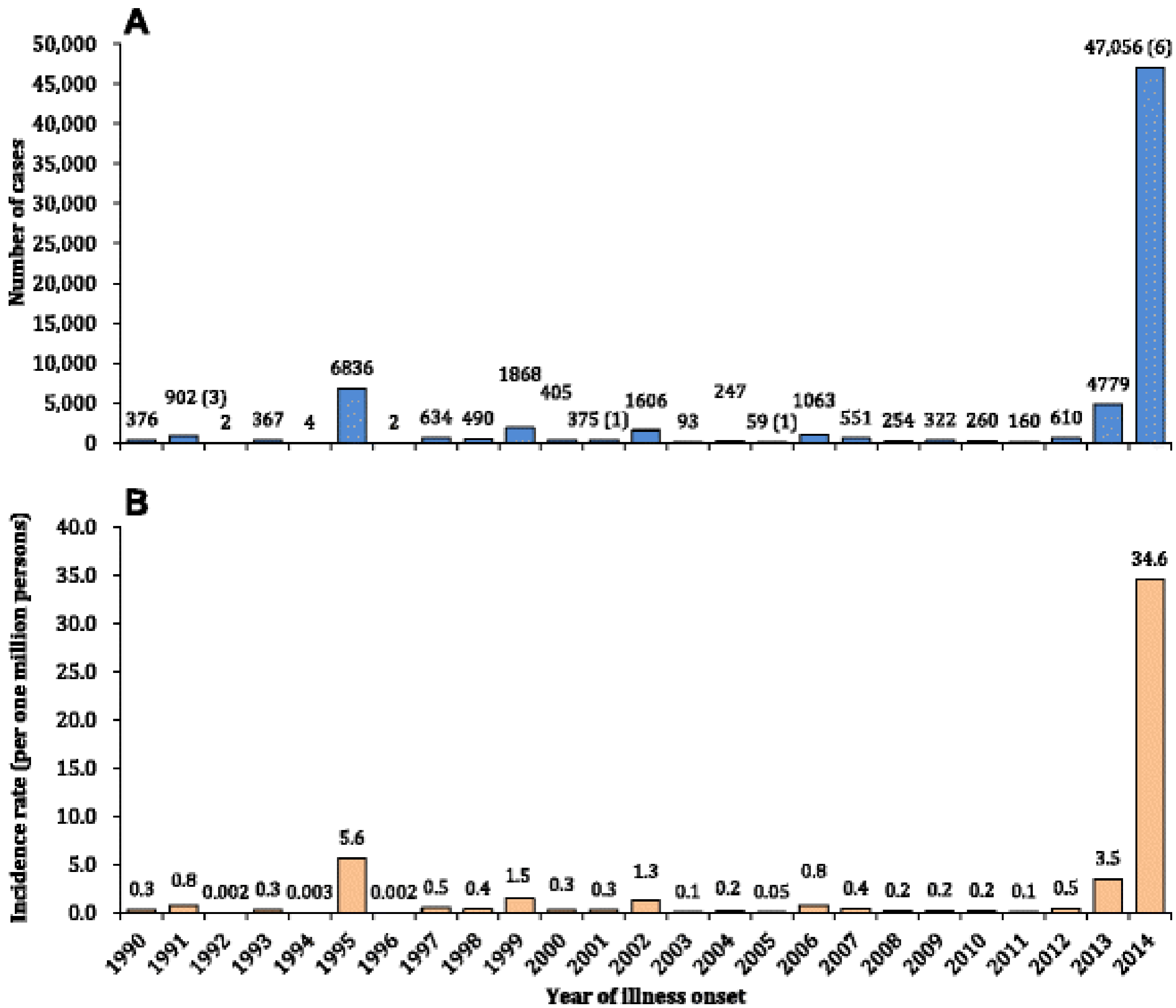


Fig. 1. **Observed annual average *Aedes* household index and annual clinical incidence of dengue fever**



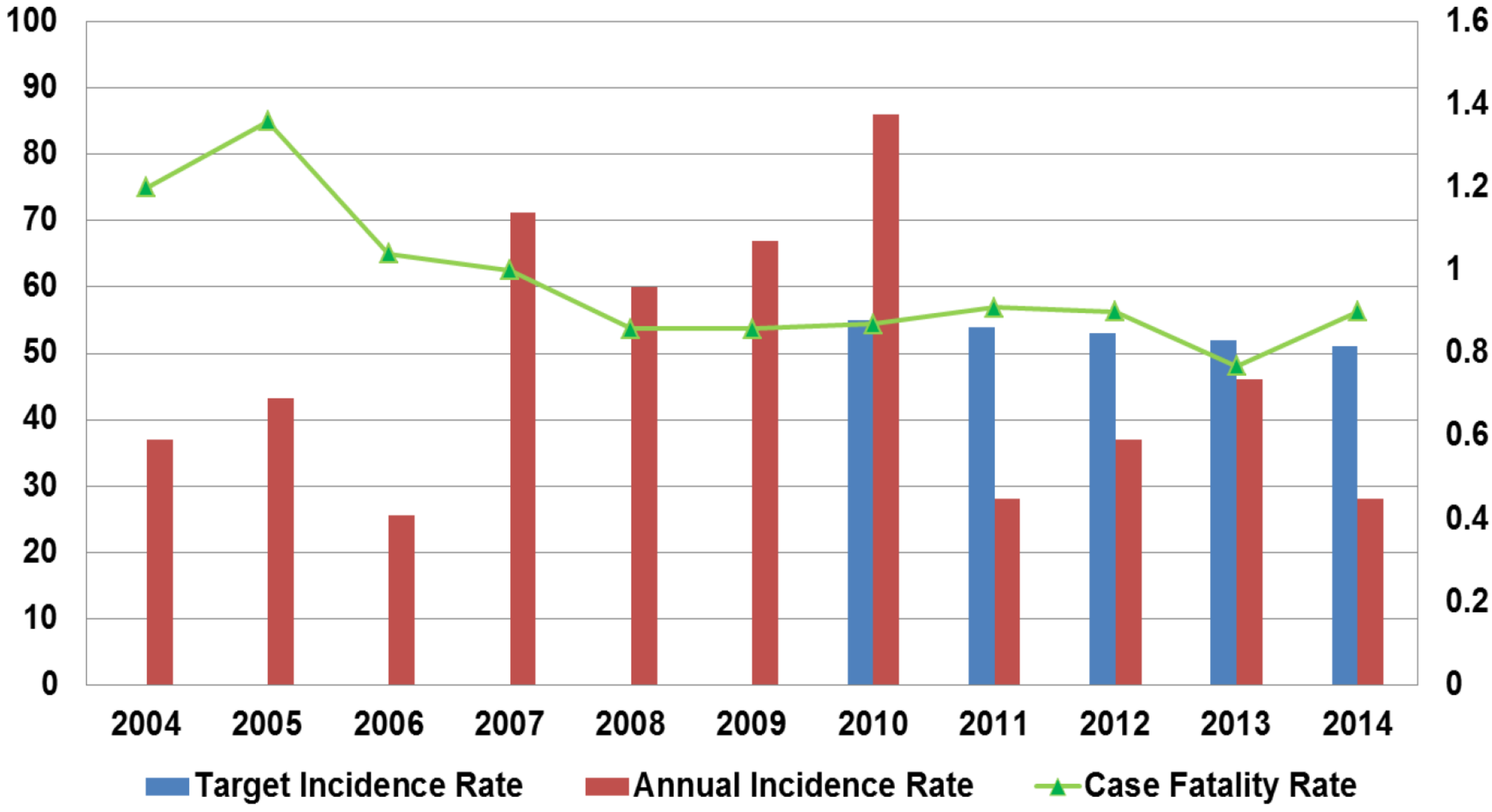
Egger et al. 2007

The incidence of dengue cases reported in China, 1990-2014 (N = 69,321). Lai, et al. 2015.



Dengue in Indonesia

Figure 6: Target and Actual Incidence and Case Fatality Rates, 2004-2014



Note: From 2014 to 2017, the target incidence rate is from 51/100,000 to 50/100,000.
Sources: Ministry of Health Indonesia. *Formulir 2, Rencana Kerja Kementerian/Lembaga (Renja-KL) Tahun Anggaran 2014*,

Table 1b. 2008-2016 Number of cases and serotyping data in Taiwan for indigenous and imported cases. (Source: Taiwan CDC)

Year	Case #		Serotests #	
	indigenous	imported	indigenous	imported
2008	**488**	**226**	205	120
2009	**848**	**204**	372	132
2010	**1592**	**304**	849	188
2011	**1545**	**157**	884	99
2012	**1271**	**207**	92	118
2013	**596**	**264**	38	117
2014	**15492**	**240**	70	123
2015	**43418**	**365**	200	135
2016	**380**	**363**	5	137
2017	**10**	**333**	4	134

The morbidity of imported (N = 2,061) and indigenous (N = 53,053) dengue cases by month per one million residents of affected provinces at the end of each year in **China**, 2005-2014. Lai et al. 2015.

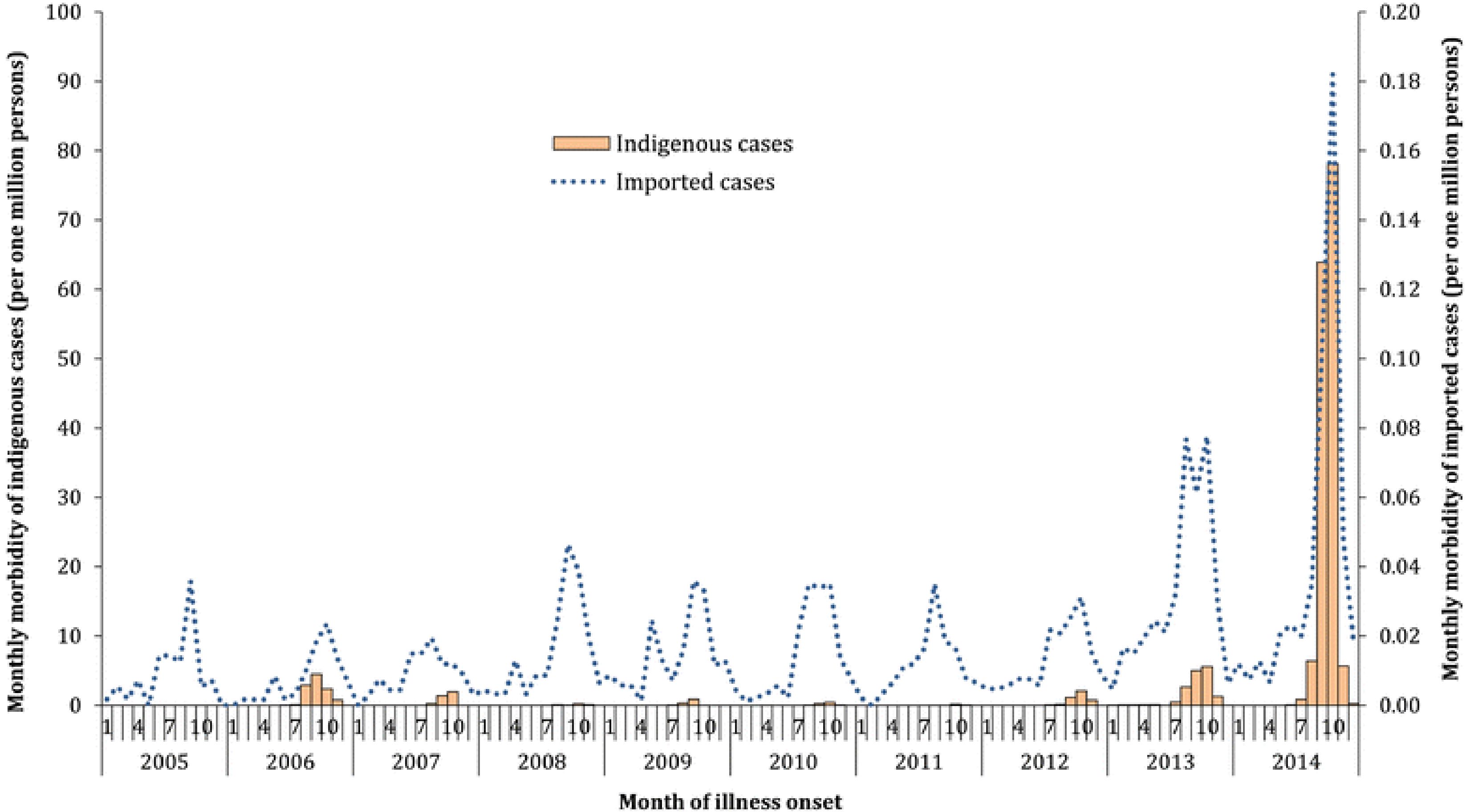


Figure 1a. **Percentages of 4 serotypes** among indigenous dengue cases with test results in Taiwan during 1998-2016.

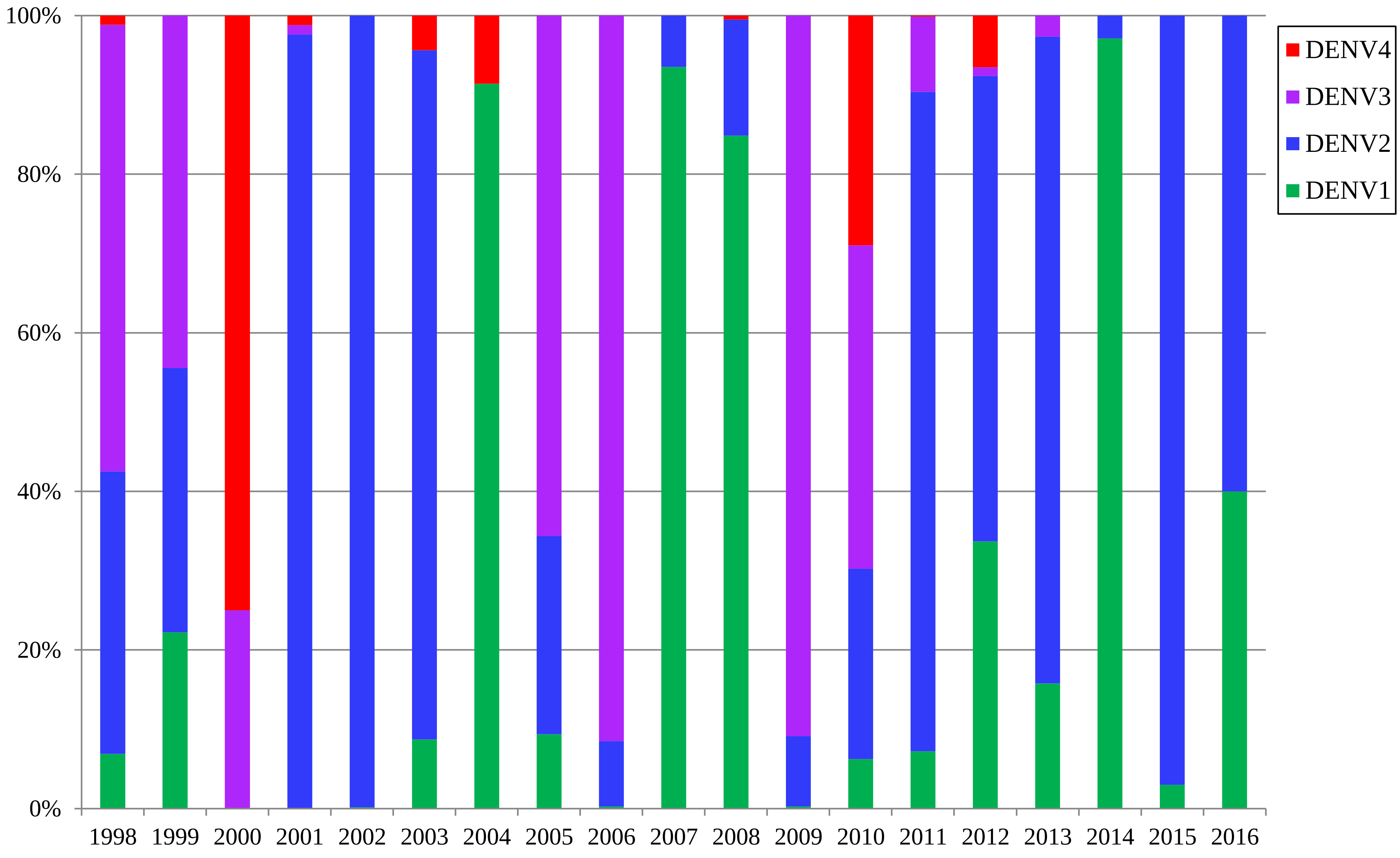
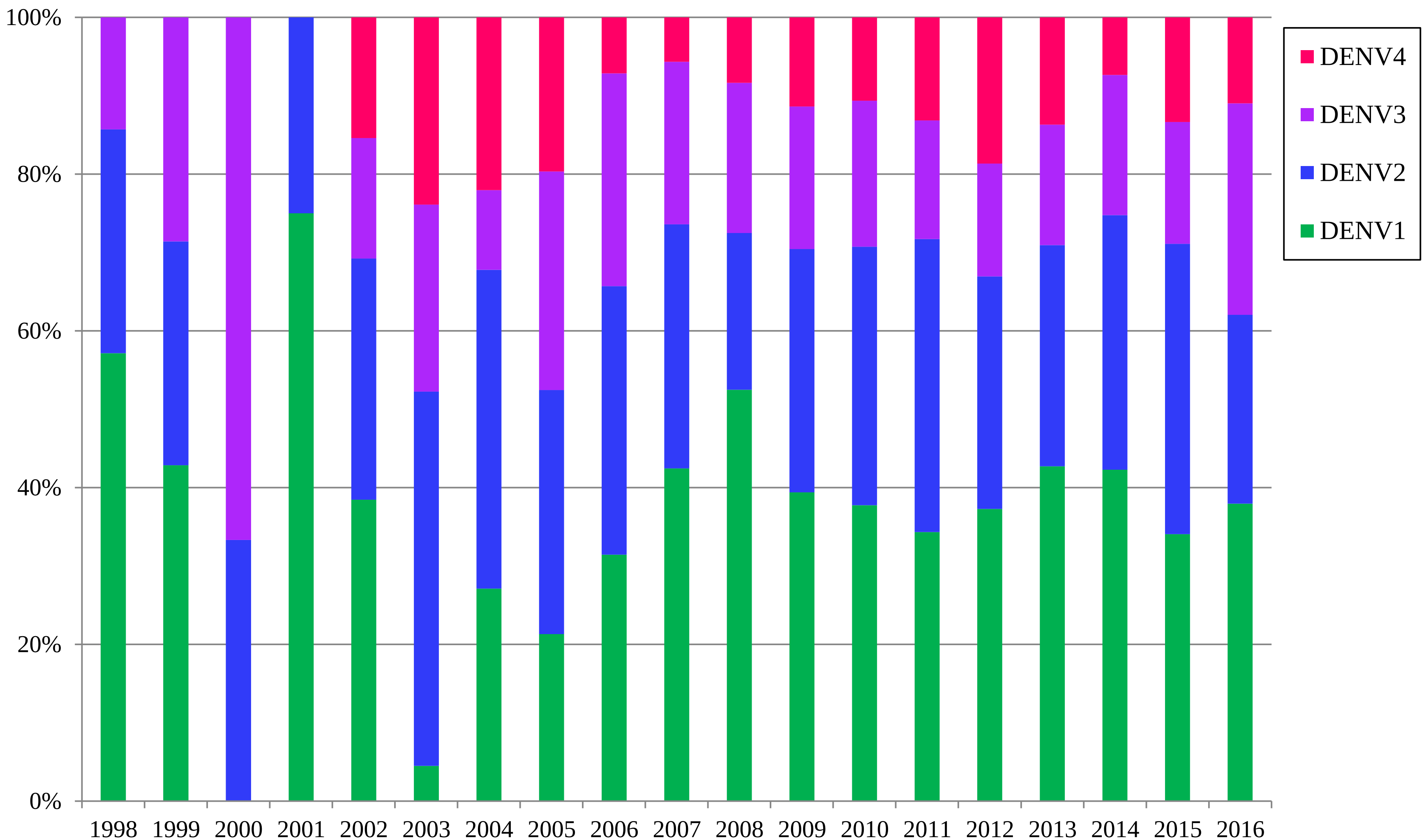


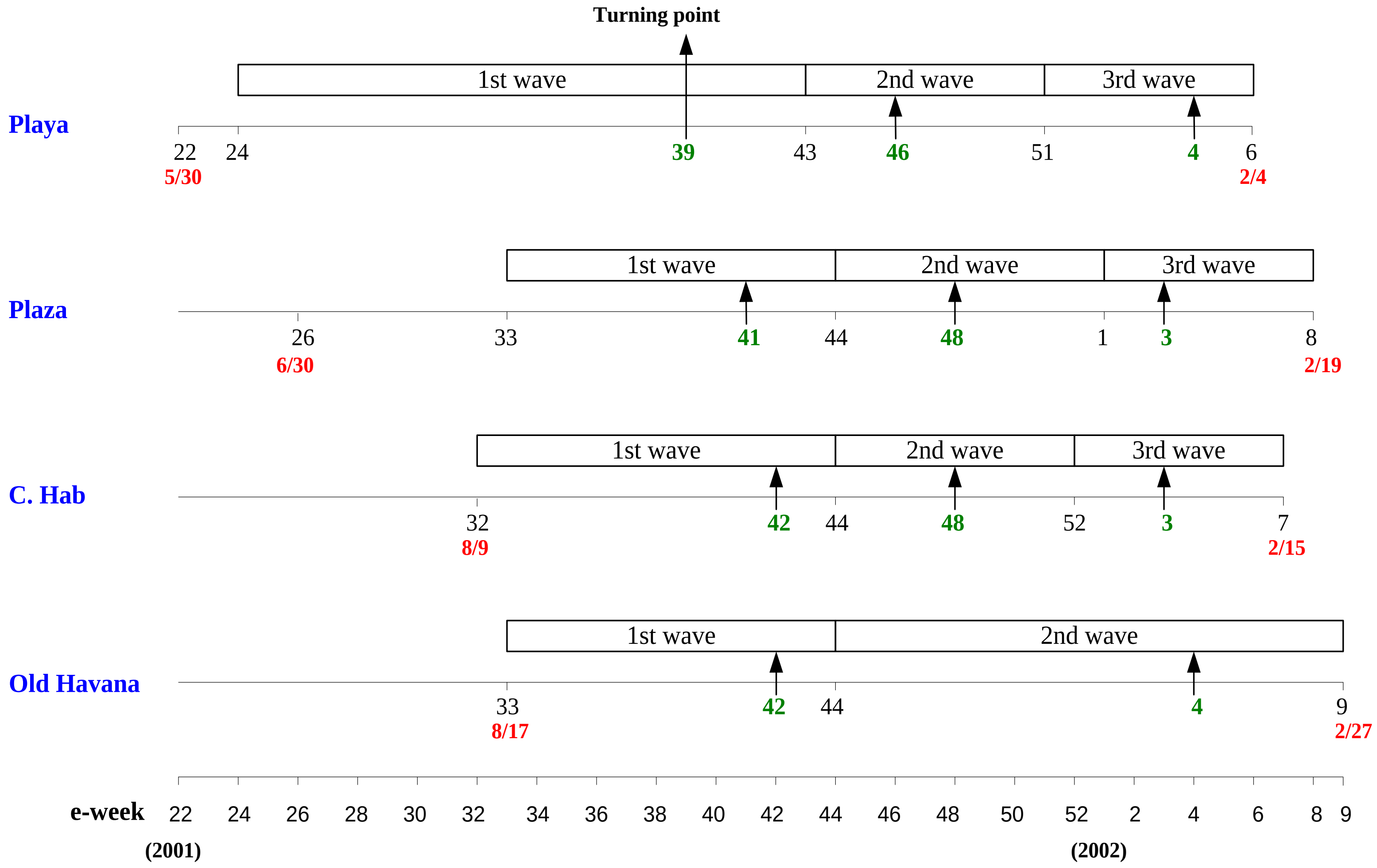
Figure 1b. **Percentages of 4 serotypes** among imported dengue cases with test results in Taiwan during 1998-2016



Impact of climate (Hsieh & Chen, *TMIH* 2009)

- The first two turning points of the two-wave Tainan dengue outbreak in 2007 were partially attributable to 2 typhoons around early August that brought sharp drop in temperature and substantial rainfall.
- This highlights the possible impact of climate change on spread of infectious diseases, **if the timing of climatological events is right**.

2001 DENV-3 outbreak in Cuba
(Hsieh et al. *TMIH* 2013)

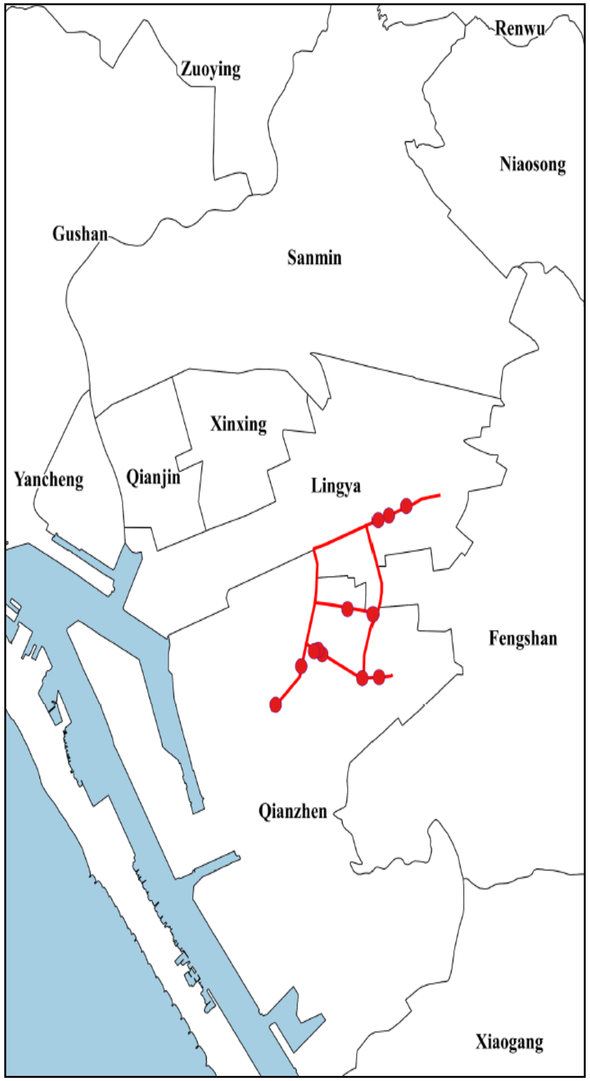


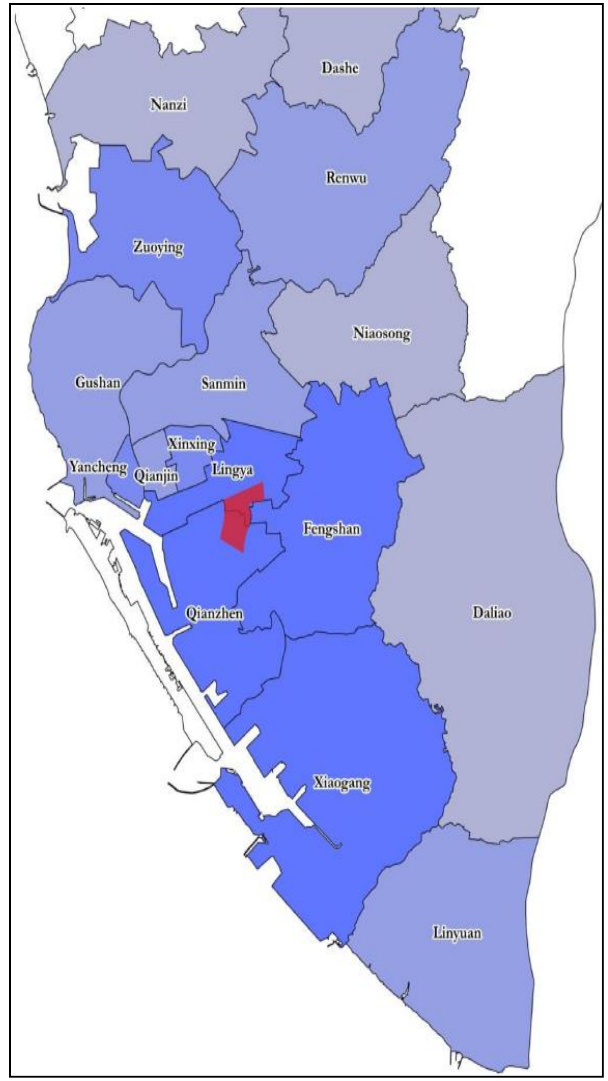
Interpreting the Turning Point (peak)

Hurricane Michelle, the most destructive hurricane in the history of Cuba, struck Cuba on November 4, 2001, in e-week 46, the down turning point of the second and largest wave.

2014年7/31高雄氣爆事件是否引發氣爆地區數日後開始出現的第一波大型登革熱疫情?

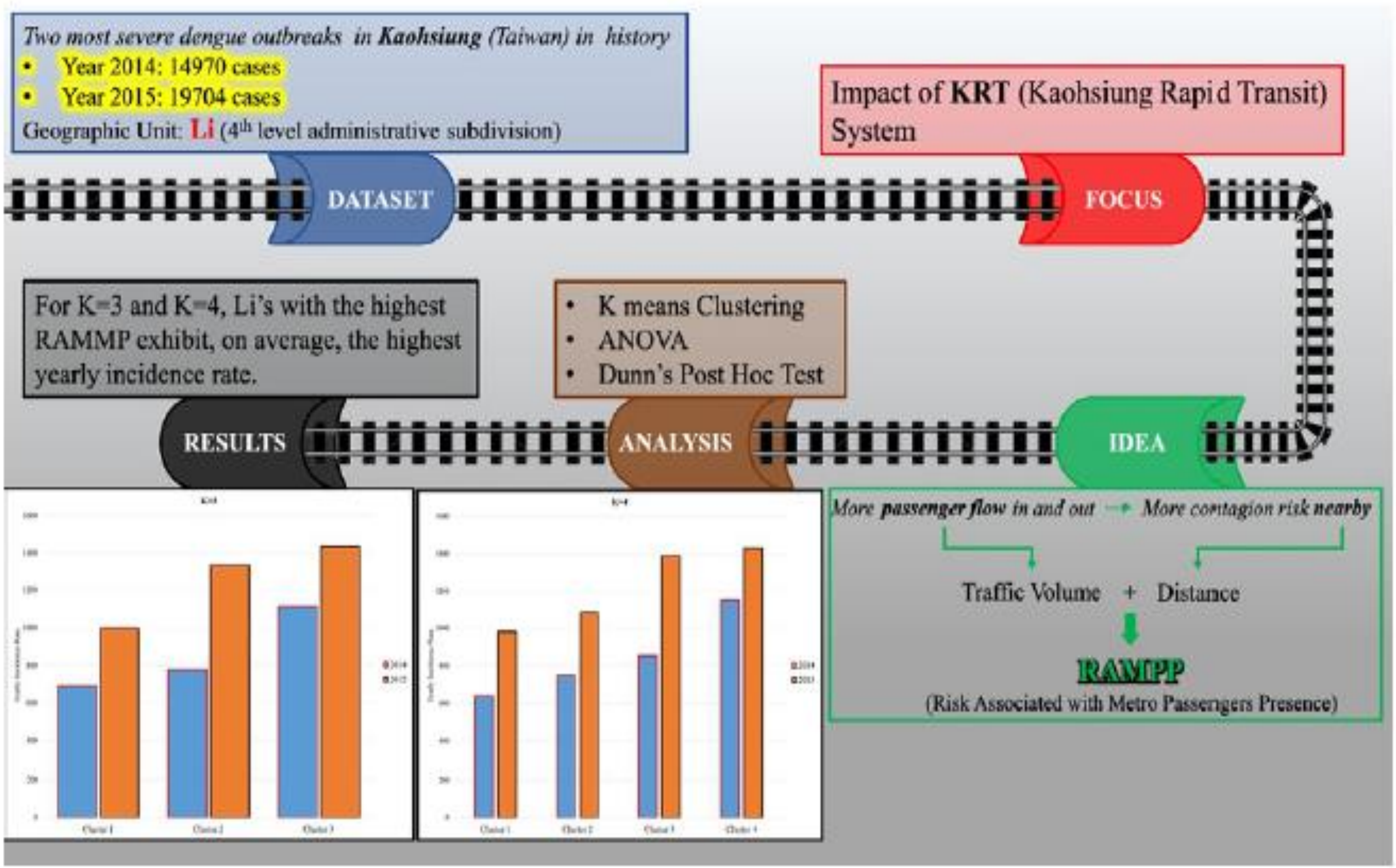
Hsieh YH. Ascertaining the Impact of Catastrophic Events on Dengue Outbreak: The 2014 Gas Explosions in Kaohsiung, Taiwan. *PLoS ONE*, 2017, 12(5): e0177422.





高雄捷運有扮演助長高雄登革熱疫情散播之角色?

Sanna M, Hsieh YH*. Ascertaining the impact of public rapid transit system on spread of dengue in urban settings. *Science of Total Environment*, 2017, 598: 1151-1159.



Conditions for Predicting Dengue Outbreak (天時-地利-人和)

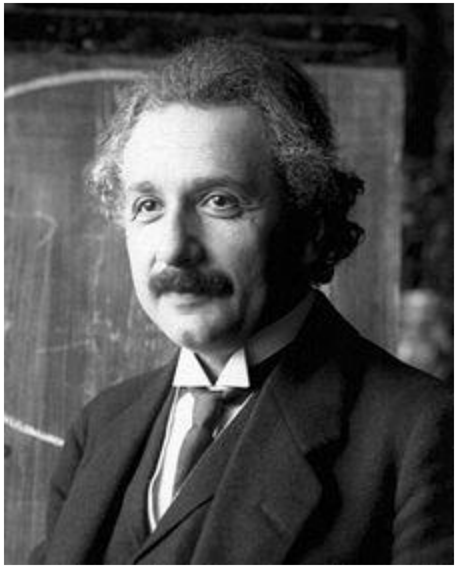
- 天時: Timing – weather conditions, timing of importation
- 地利: Spatial – location of cluster infections, spatial spread of vectors
- 人和: human movement and mobility – imported cases, spatial spread by humans

Challenge and Opportunity

- Information on infectious diseases pertaining to its epidemiology, etiology, immunology, and related data on climatology/geography/sociology.
- These multi-facet information must be consolidated into **one single multi-layered model** incorporating different types of data and **stochastic variations**, in order to truly ascertain the threat to humans.

Promises for Future

- Individual-based network model based in detailed human data holds great promise, as one of the layers of the model that combined with:
- traditional compartmental modeling depicting disease transmission, spatial disease spread and climate/environmental data as other layers
- If the ethical and legal ramifications of modeling with individual data can be satisfactorily resolved.



In 1921

Albert Einstein (1879–1955):

- Models should be as simple as possible,

but not more so.

Thank You for Your Attention

謝謝聆聽 敬請指教