國立臺灣大學公共衛生學院

流行病學與預防醫學研究所

生物醫學統計組

博士論文

Division of Biostatistics

Graduate Institute of Epidemiology and Preventive Medicine

College of Public Health

National Taiwan University

Doctoral Dissertation

隨機過程在肺結核及嚴重急性呼吸道症候群傳染病之運用

Stochastic Processes for SARS and TB Infectious Diseases

賴昭智

Chao-Chih Lai

指導教授：陳秀熙 博士

Supervisor：Hsiu-Hsi Chen, Ph.D.

中華民國 104 年 1 月

January, 2015

國立臺灣大學博士學位論文
口試委員會審定書

論文中文題目 **隨機過程在肺結核及嚴重急性呼吸道症候群傳染病之運用**

論文英文題目 **Stochastic Processes for SARS and TB Infectious Diseases**

本論文係 賴昭智 君（學號 D99849014 ）在國立臺灣大學流行病學與預防醫學研究所完成之碩士學位論文，於民國 103 年 1 月 29 日承下列考試委員審查通過及口試及格，特此證明。

口試委員：

__________________（簽名）
（指導教授）

__________________ __________________

__________________ __________________

__________________ __________________

__________________ __________________

誌謝

能順利完成博士班學業，首先要感謝指導教授陳秀熙博士，從我進入預防醫學研究所碩士班到流行病學與預防醫學研究所生統組博士班以來，對我一直耐心的指導與鼓勵，讓一位臨床醫師可以有機會從社區的角度來看問題，更進一步能踏入生物統計的範疇。另外，也要感謝葉彥伯局長，提供豐富及寶貴的資料，讓我有如此好的研究素材，可以順利完成此論文，並讓我有機會在彰化縣衛生局學習如何防治慢性病及傳染病。

還要感謝立昇、明芳、月暇、靜媛老師們，在過去學習過程中之指導，還有許辰陽博士適時的提點與協助，更要感謝口試委員葉彥伯局長、江大雄博士、王振泰醫師、潘信良醫師、于承平主任等，給了許多寶貴的意見與指導，才能順利完成此論文。

最後要感謝博士班的同學們，彼此的支持與鼓勵，讓此過程更加精采與值得懷念，當然也要感謝醫院同仁們與家人的體諒，讓我能順利完成學業。感謝大家，我才能有這些卑微的成長。

賴昭智謹誌于

國立台灣大學公共衛生學院
流行病學與預防醫學研究所
中華民國 104 年 1 月

中文摘要

研究背景

決定論模型(deterministic model)有助於推導非常重要的傳染病指標，用以評估傳染病的散播，如大規模的流行，地方性的流行等，此指標就是傳染病的閾值-基礎再生數(R_0)。然而，決定論模型在面對中小型人口及傳播傳染病滅絕(extinction)機率的估計是有疑問的，此時使用決定論模型是不適當。此外，當基礎再生數小於 1，卻仍發生小規模群突發時，使用決定論模型也是不適合的。此時，應用隨機模型便是其替代的方法。

在這些隨機模型中，分支過程(branching process) 是其中被考慮的隨機模型之一。因為它對於傳染病滅絕(extinction)的機率及基礎再生數(R_0)進行評估是較容易的。儘管有上述的優點，但因為我們常常得到的資料是在一段期間中，受到感染的人數及受感染的各不同子代間，時常相互重疊，以至於造成使用分支過程估計基礎再生數的困難。因此，連續時間馬可夫過程(continue time Markov process)中的生死過程(birth-death process)會是適合的方法。

研究目的

本論文的研究目標是在發展一系列隨機模型，利用嚴重急性呼吸道症候群及結核病兩個應用實例，估計基礎再生數及傳染病的滅絕機率。具體研究目的為：

(1) 應用決定性間隔模型，估計嚴重急性呼吸道症候群之基礎再生數。

(2) 以分支過程及生死過程，估計嚴重急性呼吸道症候群之基礎再生數及滅絕機率，並進一步擴展分支過程至 mortal 分支過程。

(3) 應用 Becker 的 SIR 模型，估計無法觀察之結核病從感染至發生症狀之潛伏期(incubation)及從感染至可傳染期之潛伏期(latent)。

(4) 發展三階段馬可夫模型結合生死過程，並以貝氏蒙地卡羅-馬可夫鏈(MCMC;

Markov chain Monte Carlo)方法，了解 IGRA(Interferon-γ release assays)於潛伏結核病感染率及結核病轉移率在不同疾病進程上所扮演的角色，且進一步運用生死過程計算潛伏結核病感染數及滅絕機率。

研究材料及方法

模擬資料

本論文以分支過程給定 2, 1.5, 1.1, 0.9 等不同基礎再生數下，模擬 9 代的資料且計算滅絕機率。給定出生率、死亡率及不同初始感染個案數以生死過程模擬資料，且計算從初始個案到達最後狀態的平均時間及其變異數。

臨床及社區資料

嚴重急性呼吸道症候群

本研究利用台灣地區 2002 年 11 月至 2003 年 7 月期間，罹患嚴重急性呼吸道症候群共 346 位確定個案，以及新加坡 2003 年 3 月 26 日至 4 月 15 日共 56 位在醫院感染嚴重急性呼吸道症候群之個案資料進行研究。

結核菌

長期照護機構結核病群突發流行資料

利用長期照護機構之結核病群突發資料，進行結核菌從受感染至發生症狀之潛伏期(incubation)及受感染至發生可傳染之潛伏期(latent)之估計。

結核病自然病史估計資料

利用彰化縣結核病 2009 至 2011 年監視系統共 2,420 位 30 歲以上結核病個案資料，與 2005 至 2011 年接觸者登記個案系統共 22,510 位 30 歲以上結核病接觸個案資料，以及 2012 至 2014 年結核病危險因子病例對照配對研究，加上 2011 至 2013 年一般族群 IGRA 調查資料進行分析。

模式發展及統計分析

本論文提出三種隨機過程，首先以分支過程及生死過程估計嚴重急性呼吸道症候群之基礎再生數、滅絕機率與擴散至最終感染數的時間。應用 Becker 的 SIR 模型估計無法觀察之結核菌從受感染至發生症狀之潛伏期(incubation)及受感染至發生可傳染之潛伏期(latent)，再以貝氏蒙地卡羅-馬可夫鏈方法，導入三階段馬可夫模型並結合生死過程，估計 IGRA 於潛伏結核病感染率及結核病轉移率在不同疾病進程之影響。

研究結果

1.模擬

分支過程

由假定布瓦松、二項及負二項等不同子代分佈，透過分支過程模擬 6 代的資料來估計基礎再生數。在不同分佈條件下，基礎再生數分別以無母數及母數方法進行估計，不同方法所估計之基礎再生數結果一致。然而不同方法下其變異數是具異質性。

純出生過程

在 λ=0.5 的假設下，經 1000 次的純出生過程(pure birth process)模擬並與精確公式所得之估計結果進行比較。經模擬的曲線與得自精確公式之結果曲線相異。然而當初始數越大，則經模擬的曲線就越接近精確公式所得結果曲線。而當 λ 大到 3 時，結果並没有改變太多。

2. 嚴重急性呼吸道症候群流行之基礎再生數估計

利用分支過程，在 16-22 代及 5-7 天的感染潛伏期假設下，基礎再生數為 0.9971 (0.5090~1.4852)，在布瓦松分佈的假設下，滅絕機率為 0.9912。在 Borel-Tanner 分佈且基礎再生數小於 1 的條件下，基礎再生數介於 0.9790 (0.8437 ~ 1.1143)與

1.0134 (0.8535 ~ 1.1733)之間，滅絕機率為 0.9709 ~ 0.9989。

由於線性生死過程無法最適配觀測資料，我們採以一般生死過程對觀察到的累積性個案資料進行配適。出生率觀察為 0.57 (於流行期小於 55 天)，11.45 (於流行期介於 55 天至 80 天)，以及 1.413(於流行期超過 80 天)。預期達到最終感染數 a 的時間(T_a)為:在 T_{32}、T_{300} 以及 T_{346} 分別為 54.97(10.09)天、80.00 (10.41)天以及 112.01 (11.47)天。

3. 結核病自然病史

結核病群突發

結核病受感染至發生可傳染之潛伏期(latent)經估計為 223.6 天[λ=0.0045 ($2.17*10^{-6}$)]，而症狀發生前之感染期經估計為 55.9 天[β=0.0179 ($3.47*10^{-5}$)]。因此從受感染至發生症狀之潛伏期(incubation)約為 279.6 天。而依據潛伏期的估計，感染至少 2 代至多 3 代。基礎再生數的範圍介於 0.9739 及 0.9796 間。

IGRA 對結核病發生的危險性

利用病例對照配對研究，在調整 TST 後，QFT-GIT 陽性對結核病發生的危險性為 2.47 (95% CI: 1.72-3.54)，若迴歸模式考量交互作用， TST 陽性者，QFT-GIT 陽性對結核病發生的危險性為 4.28 (95% CI: 1.16-15.76)， TST 陰性者，QFT-GIT 陽性對結核病發生的危險性為 1.15 (95% CI: 0.66-2.00)。

IGRA 在潛伏結核病感染率及結核病轉移率不同疾病進程之影響

整體結核病感染率(每人年)及轉移率(每人年自潛伏結核病感染轉移至結核病)經估計分別為 0.0168 (95% CI: 0.0157-0.0180)及 0.0113 (95% CI: 0.0098-0.0129)。感染率表現在年輕族群(30-44 歲)及男性都較高。那些陽性 IGRA 檢測值者，相較於陰性 IGRA 者，有 1.6 倍 (RR=1.59, 95% CI:1.39-1.84)的危險性，較易成為潛伏結核病感染個案。相對地，年老族群有較高轉移率，但男性轉移率仍較女性為高。陽性 IGRA 檢測者，相較於陰性 IGRA 檢測者有約 2 倍 (RR=2.12, 95% CI:1.57-2.85)

的危險性較易轉移至結核病。經年齡及性別的調整後，QFT-GIT 陽性者在結核病感染及轉移率之危險比分別為 1.71 (95% CI: 1.49-2.00) 及 1.58 (95% CI: 1.15-2.17)。

應用三階段馬可夫模式估計所得到參數於生死過程發現，在不考量共變數因子下，一個初始個案要擴散到 10 個個案約花 61 天，而要擴散到 30 個個案約花 87 天。年輕族群、男性及 QFT-GIT 陽性擴散愈迅速。年齡小於 45 歲且 QFT-GIT 陽性的男性，若在初始個案為 5 位的狀況下，擴散到最終為 10 位個案約需 1 週的時間。值得注意的是，若增加初始個案，則要達到預期的擴散個案數，所花費時間會越短。而初始個案若超過 5 位，則結核病滅絕的機率則幾乎不可能。

結論

本論文在結果的發現上可歸納出 5 個主要結論如下：

1. 當評估新加坡及台灣兩個地區嚴重急性呼吸道症候群的基礎再生數時，在新加坡發生的嚴重急性呼吸道症候群 3 至 8 代的感染資料，經評估發現基礎再生數介於 1 至 1.5 之間，利用分支過程可幾乎確定必定滅絕。而透過分支過程剖析台灣地區嚴重急性呼吸道症候群流行，其基礎再生數為 0.99，滅絕機率為 0.99。Borel-Tanner 分佈之分支過程也有相同的發現。
2. 估計結核菌從感染到發生症狀之潛伏期約 9 個月，其中從感染至可傳染之潛伏期約 7 個月，在症狀發生前之可傳染期約 2 個月。結核病群突發時，進行 TST 篩檢監視，追蹤 TST 陰性個案後來也發生結核病。所以針對長照機構的 TST 陰性年老族群，仍需進行監視，以期在結核病群突發時能獲得控制。
3. 本論文是在考量人口學特性及 TST 檢測結果下，針對 IGRA 對結核病發生之影響所進行的第一個病例對照研究。
4. 本論文也是第一個針對評估年齡、性別及 IGRA 在潛伏結核病感染(LTBI)及結核病轉移不同疾病進程影響之結核病自然病史研究。年輕族群有較高潛伏結核

病感染率，而年老族群有較高轉移率，男性在潛伏結核病感染率及轉移率均較女性高。考量年齡及性別因素後，IGRA 在潛伏結核病感染率及轉移率上均扮演重要角色。

5. 運用生死過程，在不同年齡、性別、IGRA 狀態下，可在給定潛伏結核病感染率(出生率)及轉移率(死亡率)下，計算到達潛伏結核病感染數目之預期時間及滅絕機率。

本論文在傳染病相關的方法學發展有三項貢獻綜述如下：

1. 提出多項統計模擬方法包含以分支過程模擬基礎再生數，或以生死過程估計滅絕機率及傳染病擴散時間。
2. 示範如何應用 Becker 的 SIR 模型結合分支過程估計潛伏期及感染潛伏期以監視結核病。
3. 以嶄新病例世代設計加上連續時間馬可夫模型，並結合生死過程以了解 IGRA 在潛伏結核病感染率及結核病轉移率之不同疾病進程上所扮演的角色。更進一步運用生死過程模擬 SARS，以計算其滅絕機率及到達最後感染人數之預期時間。兩種方法對於在傳染病之防治政策問題上，能提供相當助益。

關鍵字:隨機過程;分支過程;生死過程;肺結核;嚴重呼吸道症候群

Abstract

Background

Deterministic models are conducive to estimate a very important indicator for assessing the spread of infectious disease such as epidemic, endemic, and extinction, namely, the basic reproductive number (R_0). However, when small or moderate population size and the question of the probability of the extinction of infectious disease in question are involved the deterministic model is therefore not adequate. Furthermore, it may not be adequate when minor outbreak occurred if the R_0 is less than 1. The alternative is the application of stochastic model.

Of these stochastic models, the branching process is one of considerations because it can be easily applied to estimating both the extinction probability and R_0. In spite of these two advantages, because we often have the total number of infected individuals for a given period of time and generations usually overlap each other in reality that enables the branching processes difficult to estimate R_0. The continuous-time Markov process embodied with birth-death process may be more appropriate.

Objectives

The objectives of my thesis are to develop various types of stochastic models for estimating R_0 and the extinct probability of infectious disease by demonstrating the two examples of SARS and pulmonary TB. Specific aims are to

(1) apply the deterministic compartmental model to data on SARS poliomyelitis for estimating R_0;

(2) develop branching process and birth-death process to SARS dataset to estimate

both R_0 and the extinct probability and also extend the simple branching process to mortal branching process for measles;

(3) apply the Becker's SIR model to the data of TB for estimating latent period and incubation period;

(4) develop a novel three-state Markov model embodied with birth-death process to assess the effect of covariates (such as IGRA) on infection rate and conversion rate using Bayesian MCMC method and to further apply birth-death process to estimate extinct probability and the expected time to reach final size.

Materials and Methods

Generating Data by simulations

We simulated a branching process with 9 generations of data for a given offspring distribution under various values of R_0=2, 1.5, 1.1 and 0.9 for calculating the extinction probability. We simulated a birth-death process with given birth rate, death rate and different initial infected cases. We calculate the mean and variance of arrival time from the initial state.

Empirical Data

. SARS

The thesis used 346 confirmed cases with SRAS from November 2002 to July 2003 in Taiwan obtained from Taiwan CDC and also 22,520,776 population of Taiwan at the beginning of 2003**.** This thesis also made use of total 56 infected with SARS in a hospital in Singapore from Mar. 26 to Apr. 15, 2003. Only 3 generation of offspring was noted after outbreak investigation.

Mycobacterium **tuberculosis**

The outbreaks of TB in the Long-term Care Facility

The data on outbreak of TB in the LTCF provide empirical data for estimating the unobserved incubation period and latent period before onset of symptoms.

Data for estimating parameters of TB natural course

Various datasets were used including a total of 2,420 TB cases with age ≥ 30 enrolled in our cohort study from 2009 to 2011 (surveillance system for TB from 2009 to 2011 in Changhua County), a total of 22,510 TB contacts with age ≥ 30 enrolled in our cohort study from 2005 to 2011 (B contact registry database from 2005 to 2011 in Changhua County), a matched case-control study for risk factors of TB from 2012 to 2014 in Changhua County, and a IGRA survey for general population from 2011 to 2013 in Changhua County

Model Specification and Statistical Analysis

Three types of stochastic processes were applied and proposed. We first applied branching process and birth-death process to estimate R_0, extinct probability and the expected time to reach final size for SARS epidemics. We then applied the Becker's SIR model to estimate unobserved incubation period (including latent period) to the outbreak of TB to estimate its R_0 and extinct probability. The novel three-state Markov process embodied with birth-death process was develop to assess the effect of IGRA on the transition from susceptible to LTBI and the conversion from LTBI to TB with Bayesian MCMC method.

Results

Part I Simulation

Branching Process

The results of estimating R_0 on the generating data of a branching process with six generations for a given offspring distribution (such as Poisson, Binomial, and Negative Binomial distributions) are presented. The estimated R_0 were consistent with the nonparametric or parametric method with different distributions. However, the variances were heterogeneous by different methods.

Pure birth process

The simulated results of 1000 simulations for pure birth process assuming $\lambda=0.5$ compared with the true results estimated the exact equation for $E(T_a)$. It is very interesting to note that the simulated curve with mean value was still deviant from the curve obtained from the exact formula. However, when n_0 became larger, the simulated curve with mean value was close to the true curve obtained from the formula with larger n_0 but deviant from the formula with smaller n_0. When λ was enlarged to 3, the results were not changed at all.

Part II Estimation of R_0 for the outbreak of SARS in Taiwan

The estimated R_0 was 0.9971 (0.5090~1.4852) by using branching process given 16~22 generations assuming the incubation of 5 or 7 days. The estimated extinction probability was 0.9912 under the assumption of Poisson distribution.

Using Borel-Tanner distribution under the assumption of $R<1$, the stimulated R_0 was from 0.9790 (0.8437 ~ 1.1143) to 1.0134 (0.8535 ~ 1.1733). The estimated extinction probability was 0.9709 ~ 0.9989.

As linear birth-death process did not fit well with data apply instead general birth

death process to fit the observed cumulated SARS data. The estimated birth rates were 0.57 (< 55 day of outbreak), 11.45 (the 55^{th} ~ 80^{th} day of outbreak) and 1.413 (after the 80^{th} day of outbreak). The expected time to reach final size a (T_a) were 54.97(10.09), 80.00 (10.41) and 112.01 (11.47) days for T_{32} , T_{300} and T_{346}, respectively.

Part III Natural Course of TB

Outbreak of TB

The latent period was estimated about 223.6 days [λ=0.0045 ($2.17*10^{-6}$)] and the infectious period before symptoms onset was estimated about 55.9 days [β=0.0179 ($3.47*10^{-5}$)]. Hence, the incubation period was about 279.5 days. According to our estimation of latent period, there were at least two generations and at most 3 generations. R_0 was bounded between 0.9739 and 0.9796 in this cluster. The extinction probability was almost certain.

The effect of IGRA on the occurrence of TB with a case-control study

Using a match-case-control study, the estimated odds ratios in multivariable logistic regression mode for positive QFT-GIT after further adjustment for positive TST was 2.47 (95% CI: 1.72-3.54). After further considering the interaction term in the model, the odds ratio of QFT-GIT for subjects with positive TST was estimated as 4.28 (95% CI: 1.16-15.76) whereas the odds ratio of QFT-GIT for subjects with negative TST was estimated as 1.15 (95% CI: 0.66-2.00).

The effect of IGRA on the infection rate and conversion rate with multi-state Markov model

The overall estimated infection rate (per person-years) and conversion rate (per year) were 0.0168 (95% CI: 0.0157-0.0180) and 0.0113 (95% CI: 0.0098-0.0129). The

infection rate was higher for the young age group (30-44 years old) and male sex. Those with positive IGRA were 1.60 (RR=1.59, 95% CI:1.39-1.84) times likely to be susceptible to LTBI compared with negative IGRA. In contrast to the effect of age on infection rate, the older the subject was, the higher the conversion rate. Males still had higher conversion rate than females. Those with positive IGRA were two times (RR=2.12, 95% CI:1.57-2.85) likely to surface to TB compared with negative IGRA.

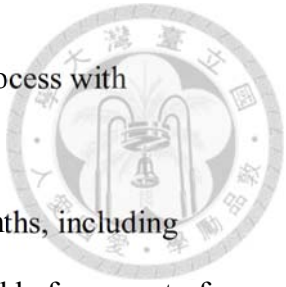
After taking the effect of age and sex on both infection rate and conversion rate into account, subjects with positive QFT-GIT still had higher risk of being infected and converting to tuberculosis with estimated RR being 1.71 (95% CI: 1.49-2.00) and 1.58 (95% CI: 1.15-2.17), respectively.

Application of birth-and-death process with the parameters obtained from three-state Markov model found one initial case may take about 61 days to have 10 of final size and 87 days to have 30 of final size without considering covariates. The young people, male and positive IGRA tended to spread quickly. The male aged less than 45 years with positive results of IGRA took only one week to reach final size given five initial cases. It should be noted that an increase in initial size reduced the time to reach the expected final size. When initial size was larger than five the extinct probability of TB was very unlikely.

Conclusion

There are five major conclusions on the practical findings reached as follows.

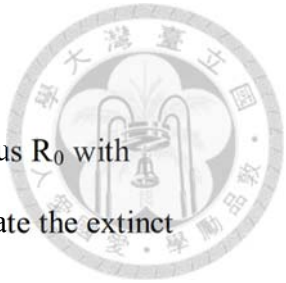
1. While evaluating SARS in the two regions, the estimation of R_0 given 3~8 generations was between 1 and 1.5, and the estimated extinct probability was almost certain using branching process in Singarepore. The SAS outbreak yielded 0.99 of R_0 using branching process in Taiwan. The estimated extinct probability was 0.99.



The similar findings were noted by using the mortal branching process with Borel-Tanner distribution.

2. Estimate unobserved incubation period with approximately 9 months, including seven months of latent period and two months of infectious period before onset of symptoms given data from an outbreak of TB occuring even among subjects with negative TST result after undergoing TB screening. Surveillance of the elderly people even with a negative TST after TB screening is still necessary given a long latent period if the outbreak of TB in a long-term care facility is to be controlled.
3. This is the first study to assess the effect of IGRA on the occurrence of TB by conducting a case-control study making allowance for demographic characteristics and induration size of TST.
4. This is the first study to assess the effects of age, gender, and IGRA on infection from susceptible to LTBI and also the conversion from LTBI to TB in the natural course of TB. The young age was at increased risk for being LTBI but the old age enhanced the risk of conversion from LTBI to TB. Male had higher risk for being infected as LTBI and also the conversion from LTBI to TB. The elevated IGRA plays a significant role not in the infection rate (from free of LTBI (susceptible) to LTBI) but also in the conversion rate after adjusting for age and gender.
5. The application of infection rate (birth rate) and conversion rate (death rate) gives the time expected to reach number of LTBI of final size and the extinct probability by various combinations of age, gender, and the results of IGRA. Subjects with positive IGRA results had shorter expected time to reach final size than those with negative result.

This thesis has also contributed to developing the methodological part related to



infectious disease consisting of three summary points:

1. Provide several statistical simulated methods for simulating various R_0 with branching process and also birth-and-death process so as to estimate the extinct probability and the expected time to reach final size.
2. Demonstrate how to apply the Becker's SIR model in conjunction with branching process to estimate incubation period and latent period for the surveillance of TB.
3. Develop a continuous-time Markov process embodied with birth-and-death process in conjunction with a novel case-cohort design data given the known sampling fraction to assess how covariates such as IGRA affect the infection rate and the conversion rate framed with a three-state Markov process. The further application of birth-and-death process used in the simulation of SARS process can compute the extinct probability and the expected time to reach final size, both of which provide a new insight into the golden period for the formulation of policy for the containment of infectious disease in question.

Keywords: stochastic process; branching process; birth-death process; TB; SARS

Catalog

Table

FIGURE

Stochastic Processes for SARS and TB Infectious Diseases

Chapter 1、 Introduction

1.1 Basic Reproductive number (R_0) and mathematical models

The application of the mathematical model to studying the effectiveness of vaccination against smallpox was pioneered by Daniel Bernoulli in 1760[1]. The literatures about mathematical epidemiology have increasingly gained attention after the threshold theorem was established by Kermack and McKendrick in 1927 [2]. Among these models, the deterministic model with ordinary differential equations (ODE) has been widely used for studying dynamics of infection disease, such as the SIR(Susceptible-Infectious-Recovered) model, the SIS (Susceptible-Infectious-Susceptible) model, and the SEIR (Susceptible-Exposed-Infectious- Recovered) model. Note that these models are population-based models and their underlying theories rely on the law of large numbers. While these models are applied during the epidemic process, they are conducive to yielding a very important indicator for assessing the spread of infectious disease such as epidemic outbreak, endemic, and extinction, with the threshold values of infection disease, namely, the basic reproductive number (R_0). These models were further applied to estimating effective reproduction number (R) after the containment of infectious disease with different strategies. Equation Section (Next)

To further elucidate the stage of infectious disease process, these models have been used to estimate latent or incubation period, the infectious period, the dynamic population (births, deaths or migrations) and the immunity status of the susceptible

(the exposure of infectious hosts or the vaccination). More importantly, the herd immunity can also be evaluated by these deterministic models.

1.2 Discrete-time stochastic model for estimating R_0

In a small or moderate population size, the deterministic model (SIR) may not be adequate. The expedient strategy with the deterministic model is to consider clinical symptoms while R is estimated given the herd immunity [3]. In my previous work, the 95% CI of R_0 through the simulations with assigning the proper distributions of these parameters was estimated by still using the deterministic model [4] (in submitting). However, the probability of the extinction of infectious disease in question cannot be evaluated by using the deterministic model. Furthermore, it may not be adequate when minor outbreak occurred if the R_0 is less than 1. The extinction of infectious disease was still noted even though the R_0 large than one for the spread of infectious disease at the beginning of stage. The alternative is the application of stochastic model.

The application of stochastic model in relation to the study of infectious disease was proposed by McKendrick in 1926 and continuous time version of the deterministic model in a stochastic manner was further provided by McKendrick and Kermack in 1927. The stochastic model for infectious disease has become popular since 1950s.[5-7] Of these stochastic models, the branching process is one of considerations. The merit of using branching process is two-fold. First, it can also be easily applied to the evaluation of disease extinction probability and R_0 or R. Second it enables one to start from individual level and sometimes is called an individual-based model when the focus is placed on the transmission probability that a contact between an infective and a susceptible. It can be extended from individual level to population level. The most difficulty with the application of branching

process is that mathematical computation is very intractable because of the complexity of mathematical property of branching process. To solve this issue, Cox and Miller proposed generating function for such a purpose but it has not been widely used in infectious diseases [8]. The alternative to generating function for overcoming the difficulty of the estimation in parameters of the complex stochastic model is the use of Bayesian Markov chain Monte Carlo (MCMC) method that is one of technical breakthrough in my thesis.

1.3 Continuous-time stochastic model for estimating R_0

Because we often have the total number of infected individuals for a given period of time and generations usually overlap each other in reality, it is difficult to estimate R_0 with branching processes. Hence, the birth-death process which is a continue time discrete state Markov process may be more appropriate. In addition, only the infected rate is considered and the recovery rate is not considered in the branching process. If we want to consider the both ones, birth death process may be used instead. However, this model has been barely addressed in infectious disease process. It is particularly useful for estimating the unobserved incubation period in chronic infectious disease such as TB. It is also very interesting to assess the effect of covariates on infectious process (from susceptible to infectives) and also conversion process (from infectives to disease) as the example of the effect IGRA on the natural course of TB.

1.4 Objectives

The objectives of my thesis are to develop various types of stochastic models for estimating R_0 and the extinct probability of infectious disease by demonstrating

the two examples of SARS and pulmonary TB. Specific aims are to

(1) apply the deterministic compartmental model to data on SARS for estimating R_0;

(2) develop branching process and birth-death process to SARS dataset to estimate both R_0 and the extinct probability and also extend the simple branching process to mortal branching process for measles;

(3) apply the Becker's SIR model to the data of TB for estimating latent period and incubation period;

(4) develop a novel three-state Markov model embodied with birth-death process to assess the effect of covariates (such as IGRA) on infection rate and conversion rate using Bayesian MCMC method and to further apply birth-death process to estimate extinct probability and the expected time to reach final size.

Chapter 2、Literature Reviews

The basic reproduction number is the key concept in infectious disease and is also used to assess the risk of an epidemic or pandemic in emerging infectious disease. Review literatures related to the important method used for the derivation of R_0 are described in the following section. Equation Section 2

2.1 Epidemic threshold method

The theoretical basis of herd immunity was derived from "the mass-action principle" which was introduced by Hamer [9]. "The ability to infection" was thought to be a function of the number of susceptibles in the underlying population, which can be expressed as

$$C_{t+1} = C_t S_t r \tag{2.1}$$

where S_t is the number of susceptibles at Time t, C_{t+1} and C_t are the number of cases at Time t+1, r is a function of the transmission parameter and contact rate. The number of susceptibles cycles reaching the equilibrium is described as "epidemic threshold" (Se). The equation (1.1) can be rewritten as

$$C_{t+1} / C_t = S_t r \tag{2.2}$$

,which reveals that the threshold is equal to 1/r. Only when $C_{t+1} / C_t > 1$ $(S_t > 1/r)$, the number of infectious disease cases increases. Therefore, if the proportion on immune is so high that result in the number of susceptibles below the epidemic threshold, then incidence will decrease. It can be easy to express as follows:

$$H = 1 - \frac{Se}{T} = 1 - \frac{1}{rT} \tag{2.3}$$

where H is the herd immunity threshold, T is the total population size, and Se is the

epidemic threshold.

The earliest reproduction number is described by Lotz's study the effectiveness of smallpox vaccination. Let a denote initial cases and q denote the infection ratio, the number of cases in the *n*th generation is aq^n. So, Lotz's 'infection ratio', q, is the earliest expression of the reproduction number.

The basic reproductive rate of infection R_0 is the average number of secondary cases produced by one primary infective in a wholly susceptible population. In other words, if the infectious disease can be spread and maintained in a susceptible population, the R_0 must be greater than one. Let S_t=T and C_t=1 in equation (2.1), we can easy get the equation R_0＝Tr. However, there are some persons with infected or immunity in the population. Hence, the average number of secondary cases by each infected individual is less than the basic reproduction number. It can be called as net reproduction rate or effective reproductive rate (R_n). So R_n=*f* R_0, where *f* denotes the percentage of the susceptibles in the population and it is equal to S_t/T. If R_n=*f* R_0 =1, R_0 can be estimated as 1/ *f* (T/Se). It can be inferred to the herd immunity threshold as follows:

$$H=1-\frac{1}{R_0}. \tag{2.4}$$

We can also estimate R_0 with epidemic threshold theory considering the heterogeneity of population. If the death rate is not associated with age in a steady population, R_0 can be expressed as follows:

$$R_0=\frac{T}{Se}=1+\frac{L}{A} \ \textit{defined as type II mortality in the population}, \tag{2.5}$$

and

$$R_0=\frac{L}{A} \ \textit{defined as type I mortality in the population} \tag{2.6}$$

, where L denotes the life span of this population and A denotes average age A at infection.[10]

If we consider the immunity of infants obtained from their mothers against from infectious diseases, R_0 can be expressed as follows:

$$R_0 = 1 + \frac{L-M}{A-M} \quad \textit{defined as type II mortality in the population}, \tag{2.7}$$

and

$$R_0 = \frac{L-M}{A-M} \quad \textit{defined as type I mortality in the population} \tag{2.8}$$

where M denotes the loss of immunity at average age M.[2, 11]

Population-based age-stratified serological surveys is suitable for using (2.5) –(2.8) estimators. Farrington studied the seroprevalence of hepatitis A in Bulgaria for estimating R [12]. The constant force of infection model is readily fitted by using a generalized linear model, giving an estimated value λ of 0.0494/year. Under this model,

$$R_0 \approx \frac{\lambda L}{1-\exp(-\lambda L)} \tag{2.9}$$

, where L denotes lifespan. R_0 was estimated as 3.6 (3.3-3.9) for hepatitis A in Bulgaria with assuming type 1 survival (L=70 years). He also estimated R_0 for rubella (1986-1987) and mumps (1991) in UK. So, R_0 was estimated as 2.62 (2.42-2.91) for rubella and 2.62 (2.42-2.91) for mumps with assuming type 1 survival (L=75 years).

If the age-specificity of infection is different and the "contact" means that the susceptible contact with not only the same age group but also the different age groups. Hence, the mass-action principle can be expanded as follows:

$$C_{a,t+1} = \sum_{i=1}^{n} C_{i,t} S_{a,t} R_{a \times i} \tag{2.10}$$

, where a denote age a group, i denotes the *i*-th age group (i=1,2,…n) and $R_{a \times i}$ denotes the contact parameters in different age groups. The parameter can be simplified as WAIFW (Who Acquires Infection From Whom) matrix [2, 13]. Farrington also studied rubella and mumps for estimating R_0 by using WAIFW (Who

Acquires Infection From Whom) matrix [12]. They use the age groups 0-3, 3-8, 8-15, 15-25 and 25-75 years and investigate three contrasting but plausible contact structures. These matrix are

$$\underset{matrix\ A}{\begin{pmatrix} 1 & 1 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \\ 3 & 3 & 3 & 4 & 5 \\ 4 & 4 & 4 & 5 & 5 \\ 5 & 5 & 5 & 5 & 5 \end{pmatrix}} \quad \underset{matrix\ B}{\begin{pmatrix} 1 & 1 & 4 & 4 & 5 \\ 1 & 2 & 4 & 4 & 5 \\ 4 & 4 & 3 & 4 & 5 \\ 4 & 4 & 4 & 5 & 5 \\ 5 & 5 & 5 & 5 & 5 \end{pmatrix}} \quad \underset{matrix\ C}{\begin{pmatrix} 1 & 5 & 5 & 5 & 5 \\ 5 & 2 & 5 & 5 & 5 \\ 5 & 5 & 3 & 5 & 5 \\ 5 & 5 & 5 & 4 & 5 \\ 5 & 5 & 5 & 5 & 5 \end{pmatrix}}$$

Each digit represents a distinct parameter value. R_0 was estimated for rubella as 3.3 (2.5-5.6) with matrix A, 3.6 (2.8-5.7) with matrix B, 4.2 (3.7-9.3) with matrix C. Estimated R_0 for mumps were 3.3 (3.1-3.5) with matrix A, 8.0 (4.6-11.5) with matrix B, 25.5(8.4-31.7) with matrix C.

Let S(a) be the probability that a newly infected individual remains infectious for at least time a and let b(a) denote the average number of newly infected individuals that an infectious individual will produce per unit time when infected for total time a. Then R_0 can be expressed as[14]

$$R_0 = \int_0^\infty b(a)S(a)da \tag{2.11}$$

The method can be extended to delineate models which a series of states are involved in the spread of infectious disease. S(a) can be defined as the probability that an individual in state 1 at time zero produces an individual who is in state 2 until at least time a. Hence, S(a) in Malaria can be expressed as

$$S(a)=\int_0^a \begin{array}{l} prob(\text{human infected at time } 0 \\ \text{exits at time t})\times \\ prob(\text{human infected } for \text{ tot.time } t \\ \text{infects mosquito})\times \\ prob(\text{nfect}ed \text{ mosquito } live\ to\ be \\ age \text{ a}-\text{t})\times \end{array} \quad (2.12)$$

Estimated values of the basic reproductive rate, R_0, and herd immunity threshold, H, for various infectious disease are listed in Table 2.1. It was found that there were different values of R_0 and H of the same disease in different areas or different epidemic times since R_0 depends on the duration of the infectious period, the contact rate and transmission probability of infectious disease.

The theory of epidemic threshold is the most widespread method. The key to using this method is to estimate the exact R_0.There are many methods to estimate the R_0, such as survival function, next generation method, derivations of threshold criteria and estimations from epidemiological data. [15]

2.2 Deterministic model

Deterministic models delineate the spread under the assumption of mass action, relying on the law of large numbers. It is the well-known SIR model introduced by Anderson[2]. It is also called the compartmental models.

There are many studies for estimating R_0, by using the SEIR model for many infectious diseases. One study built a deterministic model for estimating R_0 of 1918 influenza pandemic. [16] The differential equations are as follows:

$$\frac{ds}{dt} = -\beta si$$

$$\frac{di}{dt} = \beta si - ve$$

$$\frac{di}{dt} = ve - \gamma i \quad .$$

$$\frac{d\gamma}{dt} = \gamma i$$

, where β, ν and γ are rate constants for transformation of individuals from susceptible to exposed, from exposed to infectious, and from infectious to recovered and immune states, respectively. Hence, the estimator of R_0 is as follows:

$$R_0 = \frac{\beta}{\gamma}.$$

Basic reproduction number for 1918 influenza pandemic ranged between 1.3 and 3.1.

We developed a dynamic epidemic model to fit surveillance data to estimate the effective reproductive number (R) of HFMD by using the extended SIR (*Susceptible-Infected-Recovered*) differentiation method. The estimated effective reproductive number was about 1.22 in 2000, 1.21 in 2001, and 1.18 in 2008. As there were two waves of epidemic curves in 2005. The average effective R was 1.59 in 2005. We also carried out the sensitivity analysis on the basic reproductive number to assess the variability in R_0 that results from the uncertainty in the model parameters with next generation method. Monte Carlo procedure (simple random sampling) was used to measure the uncertainty of R_0 by given proper distributions of these parameters (β, α, τ_a, τ_s, and ρ) according to the findings from these outbreaks in 2000, 2001, 2005 and 2008. After 10000 times of simulation, the mean value of R_0 was estimated as 1.37 (95% CI: 0.23~5.71) by the next generation method with sensitivity analysis [17].

2.3 Markov model applied to infectious disease

(1) Binominal Chain Model

The above SIR model is not suitable for simulating the spread of disease in

small population, such as household. Binominal chain model is designed for this condition[18]. Let there be S_t (the number of susceptibles) of generation t exposed to the I_t (the number of infectives) infectives of generation t, the probability of x （x=0,1,2,…）infectives of generation t+1（I_{t+1}） is estimated as follows:

$$Pr(I_{t+1} = x \mid S_t = s, I_t = i) = \frac{s!}{s!(s-x)!} p_i^x q_i^{s-x} \quad x = 0,1,2,...,s,\ p_i = 1 - q_i \qquad (2.13)$$

，where q_i denotes the probability that a given susceptible escapes infection when exposed to the i infectives of a special generation. p_i is equal to 1-q_i, where i=1.2…. There are two famous models to simplify the binominal model. One is Reed-Frost model ($q_i = q_1^i$) and another is Greenwood model ($q_i = q$). The Reed-Frost model is suitable for the infection via close person-to-person contact. The Greenwood model is suitable for the infection under the environment “saturated” with infectious agent. The Greenwood model is suitable for the transmission mechanism of saturated infection such as the spread of common cold and measles.[18-20] However, the Reed-Frost model is suitable for the transmission mechanism of close contact such as sex transmitted disease.[18, 20, 21]

To modify the Reed-Frost model with the mass-action principle, we can get the formula as

$$C_{t+1} = S_t\{1-(1-p)^{C_t}\} \quad [21] \qquad (2.14)$$

, where p denotes the probability of effective contact, S_i, C_i denote the number of susceptible and case, given the time i. So, there are different probabilities of effective contact in different small groups, such as kindergarten, nursery school and elementary school.

The modified model is as follows:

$$I = I_0(1 - V)^C \quad [22] \qquad (2.15)$$

, where I denote the incidence, I_0 denote the incidence in unimmunized population, V denote the rate of vaccine coverage, and C denote the parameter of herd immunity effect. Hence, herd immunity effect is measured by the parameter C.

These models using the escaped probability to describe the herd immunity, therefore it is different from the concept of threshold. The main drawback is unsuited for large population.

Becker proposed the underlying models in term of absolute time, rather than generations. The logistic regression form is expressed as[18]

$$\log \frac{m_{tj} q_{tj}}{m_{tj}(1-q_{tj})} = \alpha_t + \beta_t \times j \qquad (2.16)$$

, where m_{tj} denotes the number of susceptibles exposed to j infectives in generation t-1. Effects of generation and number of infectives on escaped probability and relevant covariates can be included in the model.

Chain binomial models was utilized to elucidate the force of disease spread including tuberculosis [23, 24], human immunodeficiency virus [25], and influenza [26-28]. Chen et al. adopted the logistic form of (2.14) proposed in the Becker model and further extended the model with Bayesian approach to evaluate the effectiveness of vaccination at individual level and heterogeneity across household level [29]. They reported quantities of influenza transmission in the community, including the basic reproductive number of 2.56 (95% CI: 1.98-3.32), the transmission probability of 8.3% (95% CI: 7.0-9.6%) and the contact rate of 7.74 (95% CI: 7.70-7.79).

(2) Chain Binominal Chain and Markov Model[30]

The chain binomial model is a conditional model for the joint probability. Hence, the Markov assumption contained in the chain binomial model jointly describes the spread of infectious disease and facilitates the analysis of infectious disease data [31-33]. The matrix of transition probability is the result from the

construction of discrete state Markov model with discrete time. For the Greenwood model with the relation between escaped probability and number of infectives specified the transition probability is expressed as[30]

$$\begin{array}{c} \\ 0 \\ 1 \\ 2 \\ \vdots \\ X_0 \end{array} \begin{bmatrix} 1 & . & . & . & . \\ \binom{1}{0} q^0(1-q)^1 & \binom{1}{1} q^1(1-q)^0 & . & . & . \\ \binom{2}{0} q^0(1-q)^2 & \binom{2}{1} q^1(1-q)^1 & \binom{2}{2} q^2(1-q)^0 & \cdots & . \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \binom{X_t}{0} q^0(1-q)^{X_t} & \binom{X_t}{1} q^1(1-q)^{X_t-1} & \binom{X_t}{2} q^2(1-q)^{X_t-2} & \cdots & \binom{X_t}{X_t} q^{X_t}(1-q)^0 \end{bmatrix}$$

, (2.17)

which is an univariate Markov model with state space of the number of susceptible of the household [29, 33].

For the Greenwood model with the relation between escaped probability and number of infectives specified the transition probability is expressed as[30]

$$\mathbf{p}_{ij} = \begin{array}{c} X_t \backslash X_{t+1} \\ 0 \\ 1 \\ \vdots \\ j \\ j+1 \\ \vdots \\ x_0 \end{array} \begin{array}{c} \begin{array}{ccccccc} 0 & & 1 & & x_0 - j & & x_0 \end{array} \\ \begin{bmatrix} 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ (1-q^i)^j & 0 & \cdots & 0 & \cdots & 0 \\ 0 & (j+1)q^i(1-q^i)^j & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \binom{x_0}{x_0-j} q^{i(x_0-j)}(1-q^i)^j & \cdots & 0 \end{bmatrix} \end{array} . \quad (2.18)$$

Bailey proposed that the probability distribution of the process of disease transmission can be a geometric distribution [31]. In addition, Becker proposed chain binomial models with random effect to take the unobserved heterogeneity into account. [32]. Furthermore, O'Neill et al. extended this flexibility of the model by the

Markov Chain Monte Carlo approach [34]. However, an increase in the disease status of interest, the transition probability will become more complicated. Most of current application of chain binomial model assuming two states of disease transition. C.Y. Hsu et al. also analyzed the Flu data by using the continuous-time three-state Markov models with homogenous and non-homogenous transition rates specified by using a variety of distributions including exponential, Weibull, and lognormal distributions.[30] In addition, C.Y. Hsu et al. analyzed the Flu data with the multilevel structure by Bayesian generalized linear models to facilitate hypothesis testing.[35]

2.4 Branching processes

The literatures about branching process applying to the infection disease are limited. The earliest application of branching process to the study of smallpox from 1950 to 1970 in Europe was initiated by Becker [36]. There were 762 smallpox cases and 49 generations. He asserted the total number of offspring is a power series distribution on conditional extinction (Mortal branching process) and Poisson offspring distribution. Hence, let Y denote the total number of infected offspring by the end of the outbreak. The mortal branching process with *n* initial infected individuals as *n* independent mortal branching processes each with one initial infected individual.

If we can observe on $T_1, T_2, \ldots, T_y$, where T_i is the *i*-th interremoval time. For the offspring distribution it follows that

$$P\{X_i = x_i (i = 1, 2, ..., y)\} = \{\prod_{i=1}^{y} a(x_i)\}\theta^{\Sigma x_i} / A^y(\theta) \tag{2.19}$$

Hence, the distribution of Y has the following form

$$P(Y = y) = \frac{n}{y} P(X_1 + X_2 + \cdots + X_y = y - n) \tag{2.20}$$

, where y=n, n+1,n+2,… and $X_1, X_2, \ldots X_r$ are independent and identically distributed

random variables with the same distribution as X. We assume that the X_i are not observable. For the offspring distribution following the power series family, it follows that

$$P(Y = y) = b(y,n)\phi^y / B(\phi), \qquad y = n, n+1, n+2, \ldots \tag{2.21}$$

, where $\phi = \theta / A(\theta)$ *and* $B(\phi) = \theta^n$

The maximum likelihood estimate for θ is obtained by solving the following formula for θ.

$$\mu(\theta) = 1 - n / y \tag{2.22}$$

The maximum likelihood estimator of the mean of total numbers of removals can be written as $\hat{\nu} = y / n$.

The estimation of R was 0.936. The extinct probability was 0.9123.

Becker analyzed the data on smallpox in a closed community in Abakaliki in southeastern Nigeri. There were total 30 cases in a population of 120 susceptibles [37].

In fact, the number of generations and the generation sizes are usually not observed. Becker combined Reed-Frost model and branching process to estimate R as

$$\hat{\mu} = \frac{y_n}{\sum x_{j-1}(\mathrm{k} - \mathrm{y}_{j-1})} \quad , \tag{2.23}$$

where y_n denotes total number of *n* generation offspring, x_j denotes total number of the *j*-th generation offspring and k denote the initial size of the susceptible population.

Suppose it is determined that the number of generations lies strictly between μ and l. Then he derived the bound of R as follows:

$$\frac{ky}{k(y + x_0) - \mu\,\mathrm{y}^2 / 2(\mu - 1)} \le \hat{\mu} \le \frac{ky}{k(y + x_0) - (\mathrm{y} - l/2)^2 - l/2}, \tag{2.24}$$

where x_0 denotes initial size.

The reproductive number of small pox in one closed community of Nigeria was estimated as $1.10 \le R \le 1.26$. After changing the incubation period, the above data had 6 to 10 generations. Therefore, Becker obtained that $1.11 \le R \le 1.20$.

Farrington studied the surveillance of infectious diseases, measles outbreaks occurring between 1997 and 1999 in USA, after combining the branching process and Bayesian theory with application of Metropolis–Hastings sampling [38]. When the offspring distribution is Poisson the total outbreak size follows the Borel–Tanner distribution. In addition, when the offspring distribution is a geometric distribution, the total outbreak size follows the Haight distribution. The distribution of the outbreak duration *U*, given *s* introductory

cases, is then

$$f(\mathrm{u};\mathrm{s}) = \begin{cases} p_0(\lambda;\mathrm{s}) & u = 0 \\ \sum_{n=1}^{\infty} h_n(\mathrm{u})\, p_n(\lambda;\mathrm{s}) & 0 < u < \infty \\ 1 - q(\lambda)^s & u = \infty \end{cases} \qquad (2.25)$$

Here, he assumed serial intervals *z* are distributed as h (z). Let h_n *(z)* denote the distribution of the sum of *n* independent serial intervals and q (λ) denotes the extinction probability. The total number of cases in the 41 outbreaks was 207.

The probability of this event is $(1-\mathrm{e}^{-\lambda})^{41}$ for the Poisson offspring distribution and

$\lambda^{41}(1+\lambda)^{-41}$ for the geometric offspring distribution.

The estimated R was 0.66 (0.55, 0.78) assuming Poisson offspring and 0.60 (0.48, 0.75). He also combined simulations by Reed-Frost model to evaluate the bias of estimation.

Another study demonstrated that applying a Bayesian approach with mortal branching process to estimate the R of mumps in Bulgaria during the period 2005–

2008 [39]. Estimated R_0 was 0.83 ~0.91 for the region of Lovech and approximately 1 for Sofia city and the region of Kyustendil. [40]

One study on SARS outbreaks that occurred in Hong Kong, Vietnam, Singapore, and Canada in 2003 for estimating R_0 by likelihood-based procedure which is derived from the concept of branching process. The relative likelihood p_{ij} denotes case *i* has been infected by case *j*, given their difference in time of symptom onset $t_i - t_j$. This distribution for the generation interval is $w(\tau)$. So, the likelihood is expressed as

$$p_{ij} = \frac{w(t_i - t_j)}{\sum_{i \neq k} w(t_i - t_k)} \tag{2.26}$$

The effective reproduction number for case *j* is the sum over all cases *i*, weighted by the relative likelihood that case *i* has been infected by case *j*, it is expressed as follows.

$$R_j = \sum_i p_{ij} \tag{2.27}$$

The results are shown as next Table. The average effective reproduction numbers, before the first World Health Organization global alert was issued on March 12, 2003, were markedly similar (2.4-3.6) across the regions.

Another study also used likelihood based procedure to estimate the farm-to-farm reproduction number for outbreaks of highly pathogenic avian influenza in commercial poultry in Italy in 1999/2000, H7N7 in the Netherlands in 2003, and H7N3 in British Columbia, Canada in 2004. In these outbreaks the mean farm-to-farm reproductive number prior to controls ranged from 1.1 to 2.4, with the maximum farm-based reproductive number in the range 2.2 to 3.2.

2.5 Birth death process

The literatures about branching process applying to the infection disease are limited. A renewal equation with a birth-death process model is derived for the description of parasitic diseases on host populations with age structure.[40] The model has not been applied to empirical data on infectious disease.

Wood RM et al. proposed a model based on birth-death process by incorporating a fundamental mechanism undergone by intracellular bacteria, phagocytosis [41]. The model was applied to *Francisella tularensis*, for stochastic interaction between bacteria and cells of the immune system and heterogeneity in susceptibility to infection. The study reported a median infectious dose of about 23 organisms and an average incubation period of between 3 and 7 days depending on dose.

Chapter 3、Methodology

3.1 Becker's SIR model

The time between the (*i*-1)-th and *i*-th "infected event" was assumed as exponential distribution, we can estimated the latent period and incubation period of infectious disease by using Becker's method incorporating the information of contact history [18]. The conventional SIR model is the application of differential equation to capturing the infection rate (I), which is determined by the transmission probability, contact rate, and the recovery rate (R), the duration of disease process, during the susceptible (S) population. Here, we use an alternative way to model the two parameters by using Becker's concept that is widely applied to estimating the parameters related to the latent period and the infectious period following exponential distributions. My theses intend to apply the proposed method to estimated R_0 derived from the outbreak of TB (see data sources). Equation Section (Next)

If the beginning and end points of the infectious period for each patient with TB can be observed, let variable U denote the time to infectious period and variable W denote the time to removal. So Y=W-U denotes the duration of infectious period because we assumed that all TB cases are incommunicable after removal upon show of symptoms and treatment. Let X denote the duration of latent period. The method of deduction about characteristics of the latent period is less apparent because the precise time at which the infectious of TB contact occurs is unobservable. Hence, we assume that the infectiousness function is as follows:

$$f_y(t) = \begin{matrix} \beta e^{-\beta t}, & U \le t \le W \\ 0, & otherwise. \end{matrix} \tag{3.1}$$

We also assume that the latent period function is an exponential distribution.

$$f_x(x;\lambda,\gamma) = \lambda e^{-\lambda(x-\gamma)}, \qquad x \geq \gamma \tag{3.2}$$

To explain how inference can be made about the infectious rate β and characteristics of the latent period we consider for the moment affected LTCF of size two. We assumed that the observed TB outbreaks can be classified into the chains:

2 – two index TB cases;

1 - one index TB case, no secondary TB case;

$1 \rightarrow 1$- one index TB case, one secondary TB case.

Data from the outbreaks of TB in LTCF with two index TB cases contains no information about β, because β represents a rate of infection within LTCF. But it contains the available information about the latent period. Let U_1, W_1 and U_2, W_2 denote the infectious periods for the two index cases. We assume that outbreaks in LTCF develop independently of each other. It will also be reasonable to assume that the two index TB cases are infected in the meantime. Thus U_1-U_2=X_1-X_2, where X_1 and X_2 are the durations of the two latent periods. As E(U_1-U_2)=0 and Var(U_1-U_2)=2Var(X) it is clear that U_1-U_2 is informative primarily about the dispersion of the duration of the latent period. From the latent period function density (3.7), U_1-U_2 has a distribution depending on parameter λ only. So the likelihood function is as follows:

$$f_y(y_1) f_y(y_2) \int_0^\infty f_x(x+u_1-u_2) f_x(x)\,dx \tag{3.3}$$

Consider affected LTCF of size two with one index TB case and no secondary TB case. These data contain no information about latent period, but they do contain information about the infection parameter β. The likelihood is as follows:

$$f_Y(w-u) e^{-\beta(w-u)} \tag{3.4}$$

The exponential term represents the probability of escaping infection in the

susceptible.

Finally, consider affected LTCF of size two with one index TB case and one secondary TB case. Now U_2-U_1=X_2+Z, where X_2 is the duration of the latent period for the secondary case and Z represents the time from U1 until the infection of the secondary case. The conditional density of U_2-U_1, given that the index TB case has an infectious period of duration y and chain $1 \to 1$ is observed, is expressed as

$$f_{U_2-U_1}(\mathrm{a} \mid \mathrm{Y}_1 = \mathrm{y}_1, 1 \to 1) = \int_0^{y_1} \beta e^{-\beta z} f_X(a-z)\,\mathrm{dz} / (1-\mathrm{e}^{-\beta y_1}) \tag{3.5}$$

So, because infectious make a contribution, the likelihood is as follows:

$$f_Y(\mathrm{y}_1) f_Y(\mathrm{y}_2) \int_0^{y_1} \beta e^{-\beta z} f_X(\mathrm{u}_1 - \mathrm{u}_2 - \mathrm{z})\,\mathrm{dz} \tag{3.6}$$

But the data now only consist of the end points of the infectious period of all TB cases. Hence, the two time points W_1 and W_2 are observed in the cases which do contain some information about characteristics of the distribution of X+Y. Specifically,

$$W_2 - W_1 = (\mathrm{X}_2 + \mathrm{Y}_2) - (\mathrm{X}_1 + \mathrm{Y}_1) \tag{3.7}$$

and the observation w_2-w_1 contributes the likelihood function of observed chain 2 (two index TB cases) .

$$\int_0^{\infty} f_{X+Y}(\mathrm{x} + \mathrm{w}_2 - \mathrm{w}_1) f_{X+Y}(\mathrm{x})\,\mathrm{dx} \tag{3.8}$$

The likelihood of an affected LTCF of size with observed chain is as follows:

$$E(\mathrm{e}^{-\beta Y}) = M(-\beta) \tag{3.9}$$

Here, M is the moment generating function of the duration of the infectious period Y. Now consider affected LTCF of size two with one index TB case and one secondary TB case. The likelihood is expressed as

$$\int_0^{\infty} f_{X+Y}(\mathrm{a}+\mathrm{w}_2-\mathrm{w}_1)\beta \mathrm{e}^{\beta a}\int_a^{\infty} e^{-\beta\mu} f_Y(u)\,\mathrm{d}u da \tag{3.10}$$

To simplify problems, Y, the duration of infectious period, is constant (μ_y). The likelihood contributions for outbreaks 2,1, and 1 → 1 in LTCF of two are expressed by formula (3.13), (3.14) and (3.15) respectively. Now they simplify to give

$$\int_0^{\infty} f_X(\mathrm{x}+\mathrm{w}_2-\mathrm{w}_1) f_X(x)dx, \tag{3.11}$$

$$\exp(-\beta\mu_y), \tag{3.12}$$

and

$$\beta\int_0^{\mu_y} e^{-\beta x} f_X(\mathrm{w}_2-\mathrm{w}_1-\mathrm{x})dx \tag{3.13}$$

Therefore, the likelihood function of data from outbreaks of LTCF is as follows:

$$L(\beta,\lambda \mid w) = \prod_{i=2}^{n}\int_0^{\mu} \beta e^{-\beta x} f_x(w_i - w_1 - x)dx$$

where n denotes the number of the infectives (n=1,2,…).

3.2 Various Types of Stochastic Processes for Estimating R_0

3.2.1 Deterministic model

Deterministic models delineate the spread under the assumption of mass action, relying on the law of large numbers. It is the well-known SIR model introduced by Anderson[2]. There are also called compartmental models.

We applied the deterministic model to data on the outbreak of SARS in Taiwan. Let S, I and R denote the size of subpopulation of susceptibles, infectives, removed individuals (meaning immune or dead or in quarantine, but neither susceptible nor infective). We assume that the force of infection is proportional to I and an infected individual becomes immediately infectious. Let β denotes transmission rate from S to I and α denotes remove rate from I to R. The ordinary differential equations are as follows:

$$\frac{dS}{dt} = -\beta SI \tag{3.14}$$

$$\frac{dI}{dt} = \beta SI - \alpha I \tag{3.15}$$

$$\frac{dR}{dt} = \alpha I \quad . \tag{3.16}$$

In addition, S+I+R=N, where N denotes the total population size. Therefore, the estimator of R_0 is as follows:

$$R_0 = \frac{\beta}{\alpha} \tag{3.17}$$

If the parameters of models are dependent on time, models become dynamic models; otherwise models are static model. Next generation method [42] is a general method of deriving R_0 in models with several disjoint compartments. R_0 is defined as

$$\mathrm{R}_0 = \rho(\mathrm{FV}^{-1}), \tag{3.18}$$

where $\rho(\cdot)$ is the matrix dominant eigenvalue (spectral radius), F is the newly

infection rate matrix, and V is the transition matrix between compartments.

It is suitable for large population, but not for the small population, such as household. Note that only the dynamic model can predict the herd effect of intervention (vaccination).

3.2.2 Markov Renewal process

Consider a state space Ω. Let X_t represent the number of infected cases in the interval (0, t] and it is associated states. Let T_1 denote the time of first "infected event", and T_i denote the time between the (*i*-1)-th and *i*-th "infected event" for $i \in \{2,3,...\}$. $T = (T_1, T_2, ...)$ is assumed as an independent, identically distributed (*F*) sequence of random variables. The stochastic process is defined as renewal process.

Let S_n denote the time of the *n*-th renewal

$$S_n = \sum_{i=1}^{n} T_i \tag{3.19}$$

In renewal process, we have

$$P\{X_t \geq n\} = P\{S_n \leq t\} = F_n(t) \tag{3.20}$$

$$\begin{aligned} P\{X_t = n\} &= P\{X_t \geq n\} - P\{X_t \geq n+1\} \\ &= P\{S_n \leq t\} - P\{S_{n+1} \leq t\} = F_n(t) - F_{n+1}(t) \end{aligned} \tag{3.21}$$

Hence, the sequence (X_n, T_n) is called Markov renewal process if

$$\begin{aligned} &P(T_{n+1} \leq t, X_{n+1} = j \,|\, (X_0, T_0), (X_1, T_1), ...(X_n = i, T_n)) \\ &= P(T_{n+1} \leq t, X_{n+1} = j \,|\, X_n = i,) \quad \forall n \geq 1, t \geq 0, i, j \in \Omega \end{aligned}$$

3.2.3 Birth-death process

If the common distribution *F* is an identically exponential distribution in a Markov renewal process, the renewal process is simplified as a linear birth-death process.

(1) Linear birth-death process [8, 43]

Suppose that the infective population size changes by infected and recovery. If there is one group of infectious people, the model is to assume that in (t, t+$\triangle t$) each infective individual present at t has a chance $\lambda\triangle t + o(\triangle t)$ of giving birth to a new infective individual and a chance $\mu\triangle t + o(\triangle t)$ of recovery. We also suppose that the probability of newly infected and recovery are independent. The intervals between events are exponential distribution with parameters $\lambda + \mu$ in this model. The probability of infected is $\frac{\lambda}{\lambda+\mu}$ and the probability of recovery is $\frac{\mu}{\lambda+\mu}$.

Let N(t) denote the number of infected individual at time t ($0 \le t < \infty$). If N(t)=i, the conditional probability of newly infected and recovery are $i\lambda\triangle t + o(\triangle t)$, $i\mu\triangle t + o(\triangle t)$, respectively.

In the deterministic theory, the infected population size is $n_0 e^{(\lambda-\mu)t}$ if the initial infectious individuals have n_0. For stochastic theory, let

$P_i(t) = \Pr\{N(t) = i\} \quad i = 0,1,\ldots,$

and then

$$P_i(t+\triangle t) = [1 - i(\lambda+\mu)]\triangle t P_i(t) + (i-1)\lambda\triangle t P_{i-1}(t) + (i+1)\mu\triangle t P_{i+1}(t). \qquad (3.22)$$

So

$$\frac{P_i(t+\triangle t) - P_i(t)}{\triangle t} = P_i'(t) = -i(\lambda+\mu)P_i(t) + (i-1)\lambda P_{i-1}(t) + (i+1)\mu P_{i+1}(t) \qquad (3.23)$$

$P_i(0) = \delta_{in_0}$.

In formula (3.23), with i=0, $P_{-1}(t) = 0$.

If the p.g.f. of N(t) is $G(z,t) = \sum_{i=0}^{\infty} P_i(t) Z^i$, from (3.23) we can obtain that

$$P_i'(t)Z^i = -i(\lambda+\mu)P_i(t)Z^i + (i-1)\lambda P_{i-1}(t)Z^i + (i+1)\mu P_{i+1}(t)Z^i \qquad (3.24)$$

since

$$G(z,t)=\sum_{i=0}^{\infty}P_i(t)Z^i \qquad \frac{\partial G(z,t)}{\partial t}=\sum_{i=0}^{\infty}P_i'(t)Z^i \qquad \frac{\partial G(z,t)}{\partial z}=\sum_{i=0}^{\infty}iP_i\ (t)Z^{i-1}\ .$$

Hence,

$$\frac{\partial G(z,t)}{\partial t}=-(\lambda+\mu)Z\frac{\partial G(z,t)}{\partial z}+\lambda Z^2\frac{\partial G(z,t)}{\partial z}+\mu\frac{\partial G(z,t)}{\partial z} \qquad (3.25)$$

Then, we have that

$$\frac{\partial G(z,t)}{\partial t}=(\lambda Z-\mu)(Z-1)\frac{\partial G(z,t)}{\partial z} \qquad (3.26)$$

The auxiliary equation is

$$\frac{dt}{1}=\frac{-dz}{(\lambda z-\mu)(z-1)} \quad dG(z,t)=0 \qquad (3.27)$$

The procedure of solving the problem is as follows:

$$\frac{dt}{1}=\frac{-dz}{(\lambda z-\mu)(z-1)}=\left(\frac{1}{z-1}-\frac{\lambda}{\lambda z-\mu}\right)\frac{1}{\lambda-\mu}dz \qquad (3.28)$$

$$\therefore -t=\frac{1}{\lambda-\mu}(\log|z-1|-\log|\lambda z-\mu|)+c_1 \qquad (3.29)$$

$$G=c_2(\because dG=0) \qquad (\lambda-\mu)t+\log|z-1|-\log|\lambda z-\mu|=c$$

$$G(z,t)=a((\lambda-\mu)t+\log|z-1|-\log|\lambda z-\mu|)+b=F((\lambda-\mu)t+\log|z-1|-\log|\lambda z-\mu|)$$

$$G(z,t)=F(\log(e^{(\lambda-\mu)t}\frac{z-1}{\lambda z-\mu}))=f(w)$$

$$w=e^{(\lambda-\mu)t}\frac{z-1}{\lambda z-\mu} \qquad (3.30)$$

Therefore, the solution satisfying the initial condition is that

$$G(z,0) = f(w) = Z^{n_0} \quad \text{when t=0,}$$

$$w = \frac{z-1}{\lambda z - \mu} \quad \rightarrow \quad Z = \frac{\mu w - 1}{\lambda w - 1}, \text{ and}$$

$$\therefore G(z,0) = (\frac{\mu w - 1}{\lambda w - 1})^{n_0} \tag{3.31}$$

When Z=1, $w = 0$ $G(1,t) = f(0) = \sum_{i=0}^{\infty} P_i(t) = 1$.

$$G(z,t) = (\frac{\mu w - 1}{\lambda w - 1})^{n_0} = (\frac{ue^{(\lambda-\mu)t}\frac{z-1}{\lambda z-\mu} - 1}{\lambda e^{(\lambda-\mu)t}\frac{z-1}{\lambda z-\mu} - 1})^{n_0} = (\frac{(1-z)\mu - (\mu-\lambda z)e^{-(\lambda-\mu)t}}{\lambda(1-z) - (\mu-\lambda z)e^{-(\lambda-\mu)t}})^{n_0}$$

(3.32)

Let $\alpha(t) = \mu\frac{1-e^{(\lambda-\mu)t}}{\mu - \lambda e^{(\lambda-\mu)t}}$ $\beta(t) = \frac{\lambda}{\mu}\alpha(t)$, then

$$G(z,t) = \{\frac{\alpha(t) + [1-\alpha(t)-\beta(t)]z}{1-\beta(t)z}\}^{n_0} \tag{3.33}$$

The mean of the size of infective population is

$$E(N(t)) = G'(1,t) = n_0(\frac{1-\beta(t)}{1-\beta(t)})^{n_0-1}\frac{[1-\alpha(t)-\beta(t)](1-\beta(t)) + [\alpha(t) + (1-\alpha(t)-\beta(t))]}{(1-\beta(t))^2}$$

$$E(N(t)) = n_0\frac{1-\alpha(t)}{1-\beta(t)} = n_0 e^{(\lambda-\mu)t} \tag{3.34}$$

At finally time t, the size of infective population is $n_0 e^{(\lambda-\mu)t}$. The estimation is the same as previous result in deterministic model.

The variance of the size of infective population is

$$\mathrm{Var}\left(\mathrm{N}\left(\mathrm{t}\right)\right) = G''(1,t) + G'(1,t) - [G'(1,t)]^2$$

$$\mathrm{Var}\left(\mathrm{N}\left(\mathrm{t}\right)\right) = n_0\frac{[1-\alpha(t)][\alpha(t)+\beta(t)]}{[1-\beta(t)]^2} \tag{3.35}$$

$$\mathrm{Var}\left(\mathrm{N}\left(\mathrm{t}\right)\right) = n_0(\frac{\lambda+\mu}{\lambda-\mu})e^{(\lambda-\mu)t}[e^{(\lambda-\mu)t} - 1] \tag{3.36}$$

The zero state, corresponding to extinction, is an absorbing state. The probability of infection extinction has occurred at or before t is the coefficient of z^0 in (3.33).

When $\lambda \neq \mu$, the probability of extinction is

$$P_0(t) = G(0,t) = (\frac{\mu - \mu e^{-(\lambda-\mu)t}}{\lambda - \mu e^{-(\lambda-\mu)t}})^{n_0} \qquad (3.37)$$

If $\lambda > \mu$, then

$$P_0(t \to \infty) = (\frac{\mu}{\lambda})^{n_0} \qquad (3.38)$$

If $\lambda \leq \mu$, then $P_0(t \to \infty) = 1$.

The basic reproductive number R_0=pcd, where p denotes the probability of infected, c denotes the contact rate between the infective and the susceptible. For simplification of problems, we assume c=1. Then the basic reproductive number is expressed as

$$R_0 = \frac{\lambda}{\mu} \qquad (3.39)$$

In summary, if the rate of recovery is not lower than the rate of infected, it is that $R_0 = \frac{\lambda}{\mu} \leq 1$, the probability of extinction at or before time t is one. On the contrary, if the rate of recovery is lower than the rate of infected, $R_0 = \frac{\lambda}{\mu} > 1$, the probability of extinction at or before time t is $(\frac{\mu}{\lambda})^{n_0}$ when the time is long enough (t→∞).

(2) Generalized birth-death process [43]

If the rate of infected and recovery is not constant and it is dependent on time, the simple linear birth death process is not suitable for use. For example, the latent tuberculosis has higher disease occurrence rate in the initial two year. In addition, the disease incidence rate drops with time. Hence, the newly infected rate is varied by

time.

Similarly, let N(t) denote the number of infected individual at time t ($0 \le t < \infty$). If N(t)=i, the conditional probability of newly infected and recovery are $i\lambda(t)\triangle t + o(\triangle t)$, $i\mu(t)\triangle t + o(\triangle t)$, respectively.

Therefore, we can change the λ, μ to λ(t), μ(t) in formula (3.23)

$$P_i'(t) = -i\{\lambda(t)+\mu(t)\}P_i(t) + (i-1)\lambda(t)P_{i-1}(t) + (i+1)\mu(t)P_{i+1}(t) \quad (3.40)$$

Similarly, we can obtain that

$$\frac{\partial G(z,t)}{\partial t} = (\lambda(t)Z - \mu(t))(Z-1)\frac{\partial G(z,t)}{\partial z} \quad (3.41)$$

The auxiliary equation is

$$\frac{dt}{1} = \frac{-dz}{(\lambda(t)z - \mu(t))(z-1)} \qquad dG(z,t) = 0 \quad (3.42)$$

Let $s = (1-z)^{-1}$, then

$ds = (1-z)^{-2}dz$ $\quad dz = (1-z)^2 ds$ $\quad z = \frac{s-1}{s}$.

We can obtain that

$$[\lambda(t)z - \mu(t)](1-z)dt = dz = (1-z)^2 ds \quad (3.43)$$

$$\lambda(t)z - \mu(t) = (1-z)\frac{ds}{dt}$$

$$\frac{ds}{dt} - (\frac{s-1}{s})\frac{ds}{dt} - \lambda(t)(\frac{s-1}{s}) + \mu(t) = 0$$

$$\frac{ds}{dt} - [\lambda(t) - \mu(t)]s + \lambda(t) = 0 \quad (3.44)$$

Let $r(t) = -\int_0^t [\lambda(\tau) - \mu(\tau)]d\tau$, then the formula (3.100) multiply $e^{r(t)}$

$$e^{r(t)}\frac{ds}{dt} - [\lambda(t) - \mu(t)]se^{r(t)} + \lambda(t)e^{r(t)} = 0 \quad (3.45)$$

$$\frac{d}{dt}[se^{r(t)}] + \lambda(t)e^{r(t)} = 0 \quad (3.46)$$

Integrating the above formula and replacing s with $(1-z)^{-1}$, we can obtain that

$$\frac{1}{1-z}e^{r(t)}+\int_0^t \lambda(\tau)e^{r(\tau)}d\tau = c \tag{3.47}$$

, where c is constant.

From formula (3.47) and the auxiliary equation, we can obtain that

$$G(z.t) = F\{\frac{1}{1-z}e^{r(t)}+\int_0^t \lambda(\tau)e^{r(\tau)}d\tau\} \tag{3.48}$$

Suppose $\mathrm{N}(0)=\mathrm{n}_0$, $\therefore G(z,0)=F\{\frac{1}{1-z}\}=z^{n_0}$ $\quad |z|<1$

Let $\theta=\frac{1}{1-z}$ $\quad z=1-\frac{1}{\theta}$ $\quad \therefore F(\theta)=(1-\frac{1}{\theta})^{n_0}$ $\quad \left|1-\frac{1}{\theta}\right|<1$, hence

$G(z,t)=\{1-\dfrac{1}{\frac{1}{1-z}e^{r(t)}+\int_0^t \lambda(\tau)e^{r(\tau)}d\tau}\}^{n_0}$, then

$$G(z,t)=\{1-\frac{1-z}{e^{r(t)}+(1-z)\int_0^t \lambda(\tau)e^{r(\tau)}d\tau}\}^{n_0} \tag{3.49}$$

Let $\alpha(t)=1-\dfrac{1}{\frac{1}{1-z}e^{r(t)}+\int_0^t \lambda(\tau)e^{r(\tau)}d\tau}$ $\quad \beta(t)=1-e^{r(t)}[1-\alpha(t)]$, then

$$G(z,t)=\{\frac{\alpha(t)+[1-\alpha(t)-\beta(t)]z}{1-\beta(t)z}\}^{n_0} \tag{3.50}$$

From the definition of r(t) and $-\int_0^t[\lambda(\tau)-\mu(\tau)]e^{r(\tau)}d\tau = e^{r(\tau)}\Big|_0^t = e^{r(t)}-1$, we obtain

that

$$e^{r(t)}+\int_0^t \lambda(\tau)e^{r(\tau)}d\tau = 1+\int_0^t \mu(\tau)e^{r(\tau)}d\tau \tag{3.51}$$

Hence, the probability of extinction is

$P_{n_0,0}(t) = G(0,t) = \{\alpha(t)\}^{n_0} = \{1 - \frac{1}{e^{r(t)} + \int_0^t \lambda(\tau) e^{r(\tau)} d\tau}\}^{n_0}$, then

$$P_{n_0,0}(t) = \{\frac{\int_0^t \mu(\tau) e^{r(\tau)} d\tau}{1 + \int_0^t \mu(\tau) e^{r(\tau)} d\tau}\}^{n_0} \quad (3.52)$$

.

If $\int_0^t \mu(\tau) e^{r(\tau)} d\tau$ is diverge, that is $\lambda(\tau) < \mu(\tau)$, then

$P_{n_0,0}(t \to \infty) = 1 \qquad \tau \geq 0$.

If $\int_0^t \mu(\tau) e^{r(\tau)} d\tau$ is converge, that is $\lambda(\tau) \geq \mu(\tau)$, then

$$P_{n_0,0}(t \to \infty) = \{\frac{\int_0^t \mu(\tau) e^{r(\tau)} d\tau}{1 + \int_0^t \mu(\tau) e^{r(\tau)} d\tau}\}^{n_0} \quad (3.53)$$

To sum up, if the rate of recovery is not lower than the rate of infected, it is that $R_0(t) = \frac{\lambda(t)}{\mu(t)} \leq 1$, the probability of extinction at or before time t is one. On the contrary, if the rate of recovery is lower than the rate of infected, $R_0(t) = \frac{\lambda(t)}{\mu(t)} > 1$, the probability of extinction at or before time t is $\{\frac{\int_0^t \mu(\tau) e^{r(\tau)} d\tau}{1 + \int_0^t \mu(\tau) e^{r(\tau)} d\tau}\}^{n_0}$ when the time is long enough ($t \to \infty$).

3.2.4 Multi-state continue time Markov process

The progression of tuberculosis can be categorized into three stages: free of tuberculosis infection (state 0), latent tuberculosis infection (LTBI, state 1), and development of tuberculosis (state 2).[44] The national registry of tuberculosis and contact tracing provide the information on both the time frame and the proportion of subjects belonging to above three states among population. It is thus suitable for applying multi-state model to elucidate the natural history of disease progression and the effect of relevant covariates on the rates of disease progression, namely infection rate and conversion rate in infectious disease.[45] Based on the context of the progression of tuberculosis, the random variable of stochastic process is denoted by $X(t)$ with the state space defined by $\Omega = \{0,1,2\}$ corresponding to free of TB infection, LTBI, and tuberculosis. For a subject detected as LTBI during survey performed at age t_1 then progress to TB at time t_2 can be expressed as $\{X(0)=0, X(t_1)=1, X(t_2)=2\}$. The joint probability of the sequence representing TB progression can be decomposed into the product of a series of conditional probabilities as the follows

$$\begin{aligned}&\Pr\left\{X\left(0\right)=0, X\left(t_1\right)=1, X\left(t_2\right)=2\right\}\\&=\Pr\left\{X\left(0\right)=0\right\}\times\Pr\left\{X\left(t_1\right)=1\mid X\left(0\right)=0\right\}\times\Pr\left\{X\left(t_2\right)=2\mid X\left(t_1\right)=1, X\left(0\right)=0\right\}\end{aligned}$$

which can be further reduced to

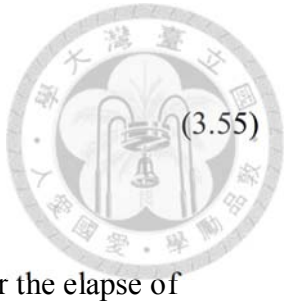
$$\begin{aligned}&\Pr\left\{X\left(t_1\right)=1\mid X\left(0\right)=0\right\}\times\Pr\left\{X\left(t_2\right)=2\mid X\left(t_1\right)=1\right\}\\&=P_{01}(0,t_1)P_{12}(t_1,t_2)\end{aligned}$$

based on Markov property and the fact that Pr(X(0)=0)=1. This can be simplified as

$$P_{01}(t_1)P_{12}(t_2-t_1) \tag{3.54}$$

assuming a time homogenous process. The transition rate dominate the generation of observed sequence of disease progression can thus be written as

$$Q = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} -\lambda & \lambda & 0 \\ 0 & -\mu & \mu \\ 0 & 0 & 0 \end{bmatrix} \end{array} \tag{3.55}$$



The probability of having observation of transition of state *i* to *j* after the elapse of period t can be derived as the following [46]

$$P(t) = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} P_{00}(t) & P_{01}(t) & P_{02}(t) \\ 0 & P_{11}(t) & P_{12}(t) \\ 0 & 0 & 1 \end{bmatrix} \end{array} \tag{3.56}$$

$$P_{00}(t) = e^{-\lambda t} \tag{3.57}$$

$$P_{01}(t) = \frac{\lambda_1 (e^{-\lambda t} - e^{-\mu t})}{\lambda_1 - \lambda_2} \tag{3.58}$$

$$P_{02}(t) = 1 - \frac{\lambda_2 e^{-\lambda t}}{\lambda - \mu} + \frac{\lambda_1 e^{-\mu t}}{\lambda - \mu} \tag{3.59}$$

$$P_{11}(t) = e^{-\mu t} \tag{3.60}$$

$$P_{12}(t) = \mu e^{-\mu t} \tag{3.61}$$

Bayesian revision for data with case-cohort design

The framework of our case-cohort study design is shown in Figure 3.1. In addition to the information on disease progression, observed data on QFT-GIT survey and matched case-control study were sampled from the cohort of general population, TB contacts, and TB cases. Such a data is collected based on the case-cohort design with the underpinning of multistate disease progression. Bayesian revision can be

utilized to derive the joint probability of having observed data [47, 48]. Denote the sampling fraction for state j ($j \in \Omega$) at time t_i for individual i by

$$\Pr(S=1 \mid 0 \to j; t_i) = f_j^{t_i} \tag{3.62}$$

,which can be simplified as f_j given a time-homogenous sampling proccss.

The probability of observing a subject with such an transition can be derived as the following

$$\begin{aligned}\Pr(0 \to j, t_i \mid S=1) &= \frac{\Pr\left(S=1 \mid 0 \to j, t_i\right)\Pr(0 \to j, t_i)}{\sum_{j=0}^{2}\Pr\left(S=1 \mid 0 \to j, t_i\right)\Pr(0 \to j, t_i)} \\ &= \frac{f_j \Pr(0 \to j, t_i)}{\sum_{j=0}^{2} f_j \Pr(0 \to j, t_i)} \\ &= \frac{f_j P_{0j}(t_i)\Pr(0 \to j, t_i)}{\sum_{j=0}^{2} f_j P_{0j}(t_i)}\end{aligned} \tag{3.63}$$

Likelihood function for observed data

QFT-GIT survey for general population using sampled data

The samples of QFT-GIT survey study for general population were derived from four sources of cohorts: general population without contact history, contact cohort remaining in the state of free of TB infection, contact cohort with LTBI, and the cohort of TB cases. Denote sampling for these cohorts are denoted as follows

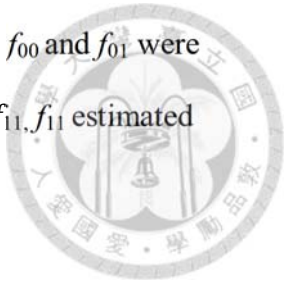
f_0 : sampling fraction of subjects belong to general cohort without contact history

f_{00} : sampling fraction of subjects belong to the contact cohort free of TB infection

f_{01} : sampling fraction of subjects belong to the contact cohort with LTBI

f_{02} : sampling fraction of subjects belong to the cohort of TB cases

Hence, f_0 was estimated from the information of IGRA survey, f_{00} and f_{01} were estimated from the information of TB contact registry database; f_{10} ,f_{11}, f_{11} estimated from the information of matched case-control study for TB.



The likelihood for subjects with contact history detected as free of TB infection and LTBI are thus

$$l_{00} = \frac{P_{00}(t) \times f_{00}}{\Delta_0} \tag{3.64}$$

and

$$l_{01} = \frac{P_{01}(t) \times f_{01}}{\Delta_0} \tag{3.65}$$

The likelihood for subjects detecting as TB cases is

$$l_{02} = \frac{P_{02}(t) \times f_{02}}{\Delta_0} \tag{3.66}$$

and those without contact history and detected as free of TB infection and LTBI are

$$l_{03} = \frac{P_{00}(t) \times f_0}{\Delta_0}, \tag{3.67}$$

and

$$l_{04} = \frac{P_{01}(t) \times f_0}{\Delta_0} \tag{3.68}$$

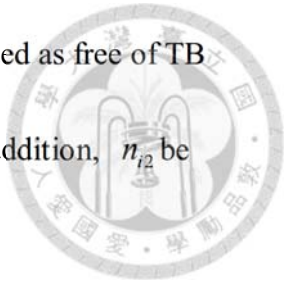
,where

$$\Delta_0 = l_{00} + l_{01} + l_{02} + l_{03} + l_{04}.$$

The likelihood for observed data during survey is hence

$$L_{S0} = \prod_i \left(\frac{P_{00}(t_i) \times f_{00}}{\Delta_0} \right)^{n_{i0}^c} \left(\frac{P_{01}(t_i) \times f_{01}}{\Delta_0} \right)^{n_{i1}^c} \left(\frac{P_{02}(t_i) \times f_{02}}{\Delta_0} \right)^{n_{i2}} \left(\frac{P_{00}(t_i) \times f_0}{\Delta_0} \right)^{n_{i0}} \left(\frac{P_{01}(t_i) \times f_0}{\Delta_0} \right)^{n_{i1}}$$

, where $n_{i0}, n_{i1}, n_{i0}^{c}$, and n_{i1}^{c} represent the count of subjects *i* detected as free of TB infection and LTBI with or without contact history, respectively. In addition, n_{i2} be the number of subjects with TB during survey at age t_i.



After initial survey, subjects were free of TB infection (state 0) or LTBI (state 1). However, they may be transited into TB after the follow-up period (t_f). The likelihood are $P_{02}(t_f)$ and $P_{12}(t_f)$ for those ,who were free of tuberculosis or LTBI at initial survey, had tuberculosis, respectively. If subjects are free of TB after following the period t_f till the end of study, those belong to censored data. The likelihood for these subjects ,who identified as free of TB infection or LTBI, can be written as

$$P_{00}(t_f) + P_{01}(t_f) \tag{3.69}$$

and

$$P_{11}(t_f). \tag{3.70}$$

Likelihood function for follow-up data after initial survey is as follows,

$$L_{F0} = \prod_i \left(P_{02}\left(t_{fi}\right)\right)^{n_{i02}^f} \left(P_{12}\left(t_{fi}\right)\right)^{n_{i12}^f} \left(P_{00}(t_{fi}) + P_{01}(t_{fi})\right)^{n_{i0}^f} \left(P_{11}(t_{fi})\right)^{n_{i1}^f} \tag{3.71}$$

, where $n_{i02}^f, n_{i12}^f, n_{i0}^f$, and n_{i1}^f represents the count of subjects turning into TB cases, censored initially identified as free of TB infection, and LTBI after the follow-up time of t_{fi}. The likelihood function for observed data on QFT-GIT survey using sampled data from general population is thus

$$L_{S0} \times L_{F0}. \tag{3.72}$$

Case-control study for TB cases and contacts

The likelihood for TB cases and contacts can be derived similarly using the transition probability of tuberculosis based on continuous time Markov model and the sampling fraction for each cohort the observed data was derived in Case-control study for TB. The observed data were derived by sampling from three types of cohorts: the cohort of TB cases, that of contact and infected by TB (LTBI), and that of contact by remain free of TB infection. Symbols denote the sampling fraction from these three cohorts are as follows

f_{10} : sampling fraction of subjects belongs to the contact cohort free of TB infection
f_{11} : sampling fraction of subjects belongs to the contact cohort with LTBI
f_{12} : sampling fraction of subjects belongs to the cohort of TB cases.

During initial survey using QFT-GIT performed at time *t*, observed subjects belongs to one of the three categories during their progression of tuberculosis: free of TB infection, LTBI and TB cases. The likelihood can be written as follows:

likelihood for TB case identified during survey

$$l_{12} = \frac{P_{02}(t) \times f_{12}}{\Delta_1} \quad , \tag{3.73}$$

likelihood for subjects with LTBI identified during survey

$$l_{11} = \frac{P_{01}(t) \times f_{11}}{\Delta_1} , \tag{3.74}$$

and likelihood for subjects free of LTBI identified during survey

$$l_{10} = \frac{P_{00}(t) \times f_{10}}{\Delta_1} \quad , \tag{3.75}$$

where $\Delta_1 = l_{10} + l_{11} + l_{12}$.

Thus the likelihood of observed data of case-control study during initial survey is

$$L_{S1} = \prod_i \left(\frac{P_{00}(t_i) \times f_{10}}{\Delta_1} \right)^{n_{i0}} \left(\frac{P_{01}(t_i) \times f_{11}}{\Delta_1} \right)^{n_{i1}} \left(\frac{P_{02}(t_i) \times f_{12}}{\Delta_1} \right)^{n_{i2}} \quad (3.76)$$

,where n_{i0}, n_{i1} and n_{i2} represent the count of subjects at age *i* identified as free of TB infection among contacts, LTBI among contacts and TB cases. After initial survey, the control group, namely TB contacts may develop TB or being followed for time t_f till the end of the study. The likelihood for subjects, who identified as free of TB infection or LTBI, developing TB after the follow-up period t_f are written as $P_{02}(t_f)$ and $P_{12}(t_f)$, respectively. The likelihood for those avoid from turning into TB is written as

$$P_{00}(t_f) + P_{01}(t_f) \quad (3.77)$$

and

$$P_{11}(t_f) \quad (3.78)$$

for subjects initially identified as free of TB and LTBI, respectively.

Total likelihood for follow-up data after initial survey is thus

$$L_{F1} = \prod_i \left(P_{02}(t_{di}) \right)^{n_{i02}^f} \left(P_{12}(t_{di}) \right)^{n_{i12}^f} \left(P_{00}(t_{ei}) + P_{01}(t_{ei}) \right)^{n_{i0}^f} \left(P_{11}(t_e) \right)^{n_{i1}^f} \quad (3.79)$$

The likelihood function for the observations in data of case-control study is thus

$$L_{S1} \times L_{F1} . \quad (3.80)$$

Combining the likelihood function for QFT-GIT survey for general population using sampled data and that for case-control study, total likelihood function is thus

$$L_{S0} \times L_{F0} \times L_{S1} \times L_{F1} \quad (3.81)$$

Incorporation the effect of covariates

The effect of covariates such as age group, sex, and the result of QFT-GIT on the rates of infection and conversion can be assessed based on the concept of generalized linear stochastic process [30]. For infection rate, the effect of covariate can be incorporated into the model as the follows

$$\eta_{1i} = \beta_{11} \times age\ group1_i + \beta_{12} \times age\ group2_i + \beta_{13} \times sex_i + \beta_{14} \times IGRA_i \qquad (3.82)$$

$$\lambda_i = \lambda_0 e^{\eta_{1i}},$$

$$age\ group1_i = \begin{cases} 1, if\ 45 \le \text{age}_i < 65 \\ 0, \quad \text{otherwise} \end{cases},$$

$$age\ group2_i = \begin{cases} 1, if\ 65 \le \text{age}_i \\ 0, \quad \text{otherwise} \end{cases}.$$

Similarly, the effect of covariates can be assessed by using

$$\eta_{2i} = \beta_{21} \times age\ group1_i + \beta_{22} \times age\ group2_i + \beta_{23} \times sex_i + \beta_{24} \times IGRA_i \qquad (3.83)$$

$\mu_i = \mu_0 e^{\eta_{2i}}$.

3.2.5 Branching process

If T denotes the generations of infectious disease in the Markov renewal process, the renewal process is simplified as branching process. The classical branching process, arising from Francis Galton's study of the extinction of family names in 1873, is a discrete time Markov chain. It is also called Galton-Watson branching process. The earliest application of branching process to build the infectious disease model for the outbreak of smallpox in Europe from 1950 to 1976 is initiated by Becker [36, 37, 49]. There were some similar studies on infectious disease based on branching process and these stochastic individual based models were applied to the spread of infectious disease in the population [38, 50].

A simple branching process model with different distributions is applied to the spread of disease under the assumption of homogeneity. However, it cannot be applied

to more complicated situations. Multi-type branching processes was derived for the heterogeneity and the continue time branching process was developed from the discrete time branching process. Therefore, we can flexibly build the model with branching process for SARS outbreak in hospital of Singapore, TB outbreak in LTCF and measles in Taiwan.

(a) Simple discrete branching process [8, 43, 51]

We assumed that the susceptible can be infected independently and all of them are infectious. Let random variable X_n denote the size of an infected population in generation n, n=1,2,…. The initial infected population size X_0 is taken as a known constant. We suppose that each infected individual produces offspring independently with the same offspring distribution. Each infected individual produces k infected persons (offspring) with the same probability P_k; therefore the new infected population size by each infected individual is the same distribution $\{P_k\}_{k=0}^{\infty}$. We can obtain $\sum_{k=1}^{\infty} P_k = 1$ from the probability axioms. Let $G_n(z)$ is the probability generating function (p.g.f.) of X_n and $G(z) = \sum_{k=0}^{\infty} P_k z^k \quad 0 \le z \le 1$, then we can obtain

$$G'(1) = \sum_{k=0}^{\infty} kP_k Z^{k-1} = \sum_{k=1}^{\infty} kP_k = R_0 . \tag{3.84}$$

If $X_n=j$, then X_{n+1} denotes the sum of j independent random variables $(Z_1,Z_2,\ldots\ldots,Z_j)$ and each random variable has the distribution $\{P_k\}$. Hence,

$$P_{jk} = \Pr(X_{n+1} = k \mid X_n = j) = \Pr(Z_1 + Z_2 + \cdots + Z_j = k) \tag{3.85}$$

That is, P_{jk} is equal to the coefficient of Z^k in $\{G(z)\}^j$.

If there was only one infectious person in initial status, then $X_0=1$, let us define

$$P_k^{(n)} = \Pr(X_n = k) \tag{3.86}$$

thus $\{P_k^{(n)}\}$ is the distribution of the number of nth generation infective of a single person. Since

$$P_k^{(n)} = \sum_{j=0}^{\infty} p_j^{n-1} p_{jk} \tag{3.87}$$

$P_k^{(n)} = coefficient\ of\ z^k\ in\ \sum_{j=1}^{\infty} p_j^{n-1}\{G(z)\}^j$.

We define the p.g.f. of the size of the nth generation as

$$F_n(z) = \sum_{k=0}^{\infty} p_k^{(n)} z^k \tag{3.88}$$

,we have

$$F_n(z) = \sum_{j=0}^{\infty} p_j^{(n-1)}\{G(z)\}^j = F_{n-1}(G(z)). \tag{3.89}$$

Since $F_1(z) = G(z),\ F_2(z) = F_1(G(z)),\ \cdots, F_n(z) = G(G(\cdots G(G(z))\cdots)$,

we can also write

$$F_n(z) = G(F_{n-1}(z)) \tag{3.90}$$

But $F_n(z)$ will be difficult to get from the above recurrence relation (3.90). However, the moments of the size of nth generation can be obtained. Let $z=e^{-\theta}$ and define the cumulant generating functions

$$K(\theta) = log\ G\left(e^{-\theta}\right) \tag{3.91}$$

$$K_n(\theta) = log\ F_n\left(e^{-\theta}\right) \tag{3.92}$$

Then it becomes

$$\log F_n(e^{-\theta}) = \log G(F_{n-1}(e^{-\theta})) = \log G(e^{K_{n-1}(\theta)}) \tag{3.93}$$

$$\therefore \quad K_n(\theta) = \log F_n(e^{-\theta}) = \log G(e^{K_{n-1}(\theta)}) = K(-K_{n-1}(\theta)) \tag{3.94}$$

Let μ, σ^2 denote the mean and variance of the number of infected offspring per infectious individual, respectively. If n is finite, the average number (μ) and variance (σ^2) of persons by one infectious person is as follows:

$$\mu = -K'(0) \qquad \sigma^2 = K''(0) \tag{3.95}$$

We can obtain the mean and variance of the size of the *n*-th generation from (3.94) as follows:

$$\mu_n = -K_n'(0), v^n = K_n''(0)$$

$$\mu_n = \mu\mu_{n-1} \qquad \therefore \mu_n = \mu^n \tag{3.96}$$

$$v^n = K_n''(0) = \{K'(-K_{n-1}'(0)) \times [-K'(0)]\}' = K''(-K_{n-1}'(0)) \times [-K_{n-1}'(0)]^2 + K'(-K_{n-1}'(0)) \times [-K_{n-1}''(0)]$$

$$\because -K_{n-1}'(0) = 0$$

$$\therefore v_n = \sigma^2\mu^{2(n-1)} + \mu v_{n-1} = \sigma^2(\mu^{n-1} + \cdots + \mu^{2(n-1)}) = \sigma^2\mu^{n-1}\frac{1-\mu^n}{1-\mu} \tag{3.97}$$

From the definition of R_0, we know $R_0 = \mu$.

When $R_0 < 1$, both μ_n and v_n approach 0 as $n \to \infty$. It means that the infection will be extinct.

When $R_0 = 1$, $\mu_n = 1$ and $v_n = n\sigma^2$ as $n \to \infty$. It means that the infection will be continued but it will not be enlarged.

When $R_0 > 1$, both μ_n and v_n approach ∞ as $n \to \infty$. It means that the infection will be continued and enlarged.

The infection will be extinct when $R_0 < 1$, but how about the probability of extinction is? Let q_n denote the probability of no one be infected in the n-th generation, then

$q_n = P_0^{(n)} = F_n(0)$. Now $P_0^{(n)} = \Pr(X_n = 0 \mid X_1 = 1)$ and 0 is an aperiodic recurrent state. In fact, it is an absorbing state. From the limit theorem for Markov chain, $\lim_{n\to\infty} p_0^{(n)} = F_n(0) = \xi$ .From (2) we have $F_n(0) = G(F_{n-1}(0))$, and we can obtain the relation

$$\xi = G(\xi) \ as \ n \to \infty \tag{3.98}$$

Therefore, ξ is the roots of equation x=G(x), and we must seek roots satisfying $0 \le \xi \le 1$ because ξ is to be a probability. The roots of (3.98) are the values of x at which the curve y=G(x) and the line y=x intersect graphically.

G(x) is a convex monotonically increasing function if $G'(1) > 0$ and $G''(x) > 0$ for x>0. Thus the curve y=G(x) can intersects the line y=x in at most two points. It is clear that x=1 is one point of intersect. Let $\xi_1 \le \xi_2$ are the two positive roots of (3.98). Since G(x) is increasing, we have

$$F_1(0) = G(0) < G(\xi_i) = \xi_i \quad \text{for any positive roots} \quad \xi_i$$

$$F_2(0) = G(G(0)) < G(\xi_i) = \xi_i .$$

By induction

$$F_n(0) < \xi_i \qquad (n = 1, 2, \cdots)$$

Since $\xi = \lim_{n\to\infty} F_n(0)$, ξ must be the smallest positive root of (3.98).

Then we can see the following results graphically,

if $R_0 = \mu = G'(1) > 1, \qquad \xi_1 < \xi_2 = 1$ (Figure 3.2(a));

if $R_0 = \mu = G'(1) = 1, \qquad \xi_1 = \xi_2 = 1$ (Figure 3.2(b))

if $R_0 = \mu = G'(1) < 1, \quad \xi_1 = 1 < \xi_2$ (Figure 3.2(b))

Conclusion: For one initial infective individual and $R_0 \le 1$, the probability of ultimate extinction ξ=1 and extinction is certain. When $R_0 > 1$, then the probability ξ<1 and the probability of enlarged infected population is 1-ξ.

Similarly, for x_0 initial infective individuals and $R_0 \le 1$, the probability of ultimate extinction ξ^{k_0}=1 and extinction is certain since the offspring of any one initial individual are independent of those of any other. When $R_0 > 1$, then the probability ξ^{k_0}<1 and the probability of enlarged infected population is 1-ξ^{k_0}.

Now we introduce some nonparametric estimators for $\hat{\mu}_n$ [49]:

$$(1)\ \hat{\mu}_n = \frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_{i-1}} \quad (if \quad X_{i-1} > 0); \hat{\mu}_n = 1 \quad (if \quad X_{i-1} = 0) \qquad (3.99)$$

$$(2)\ \bar{\mu}_n = \frac{X_n}{X_{n-1}} \quad (if \quad X_{n-1} > 0); \bar{\mu}_n = 1 \quad (if \quad X_{n-1} = 0) \qquad (3.100)$$

$$(3)\ \tilde{\mu}_n = \left(\frac{X_n}{X_0}\right)^{1/n} \quad (if \quad X_n > 0); \tilde{\mu}_n = 1 \quad (if \quad X_n = 0) \qquad (3.101)$$

All the above estimators are consistent estimators for μ. The $\hat{\mu}_n$ is the Maximum likelihood estimator when the offspring distribution belongs to a certain exponential family distributions. The estimator of standard error for $\hat{\mu}_n$ is

$$\hat{\sigma} / \left(\sum_{i=1}^{n} X_{i-1}\right)^{1/2} \qquad (3.102)$$

The consistence estimator of variance (σ^2) is

$$\hat{\sigma}^2 = \sum_{i=1}^{n} X_{i-1}(X_i / X_{i-1} - \hat{\mu}_n)^2 / n \qquad (3.103)$$

,where *n* denotes the generation.

It is convenient for analysis to let the infected offspring distribution $P\equiv(P_0,P_1,P_2,\ldots,P_k)$ in particular be the power series distribution (psd) $P_{n,\theta}\equiv(P_{0;\theta},P_{1;\theta},P_{2;\theta},\ldots,P_{k;\theta})$. We assume that the offspring distribution of X belongs to the power series family

$$\Pr(X = k) = P_{k,\theta} = \frac{a_k \theta^k}{A(\theta)} \qquad k = 0,1,\ldots, \qquad 0 < \theta \le \infty \qquad (3.104)$$

Here $(a_0, a_1, \cdots)\equiv$a are non-negative constants with >0 and >0 for at least one k $\geqq$ 2, θ>0 is the unknown parameter, $A(\theta) = \sum a_k \theta^k$.

The p.g.f. of $P_{n,\theta}$ is

$G_k(s) = \frac{A(\theta s)}{A(\theta)}$ $\quad 0 \le s \le 1$. Hence,

$$\mathrm{E}_\theta(\mathrm{X}) = \mu_\theta = \frac{\theta A'(\theta)}{A(\theta)} \qquad (3.105)$$

$$Var_\theta(X) = \theta \frac{d\mu}{d\theta} \qquad [52]\ (3.106)$$

In the power series distribution, the parameters of geometric distribution, binomial distribution, negative binomial distribution and Poisson distribution are shown as Table 3.1. The probability of no one be infected is q and q is the smallest roots of equation $G(s) = s$ .Since

$$\frac{G(qs)}{q} = \frac{A(qs\theta)}{q\{A(\theta)\}} = \frac{A(qs\theta)}{A(q\theta)\}} \qquad (3.107)$$

q is the smallest roots of equation:

$$qA(\theta) = A(q\theta) \qquad (3.108)$$

If we want to know the conditional distribution of the total size of the n-th

infected generation for given the extinction ($R \le 1$) or explosive ($R > 1$) occurred, we can infer it by Bayes Theorem [53, 54]. Therefore, the conditional probability is as follows:

$$\Pr(X_n = k \mid extinction) = \frac{\Pr(extinction \mid X_n = k)\Pr(X_n = k)}{\Pr(extinction \mid X_0 = n_0)} = \frac{\mu^k p_k^{(n)}}{\mu^{n_0}} = \mu^{k-n_0} p_k^{(n)}$$

(3.109)

$$\Pr(X_n = k \mid extinction) = \frac{\Pr(extinction \mid X_n = k)\Pr(X_n = k)}{\Pr(extinction \mid X_0 = n_0)} = \frac{\mu^k p_k^{(n)}}{\mu^{n_0}} = \mu^{k-n_0} p_k^{(n)} \quad (3.110)$$

(b)Mortal branching process

Usually, we do not obtain complete information about the number of infected individuals by each infectious individual. But we can have the total number of infected individuals for a given period of time. If the infected offspring distribution in a branching process is a power series distribution, the total number of offspring also has a power series distribution conditional on extinction [36].

Let Y denote the total number of infected offspring by the end of the outbreak. It is defined as follows:

$$Y = \sum_{n=0}^{\infty} X_n \qquad (3.111)$$

We can view the mortal branching process with n initial infected individuals as n independent mortal branching processes each with one initial infected individual. If we can observe on $T_1, T_2, \ldots, T_y$, where T_i is the *i*-th inter-removal time. For the offspring distribution it follows that

$$P\{X_i = x_i (i = 1, 2, ..., y)\} = \{\prod_{i=1}^{y} a(x_i)\}\theta^{\Sigma x_i} / A^y(\theta) \qquad (3.112)$$

Hence, the distribution of Y has the form

$$P(Y = y) = \frac{n}{y} P(X_1 + X_2 + \cdots + X_y = y - n) \qquad (3.113)$$

, where y=n, n+1,n+2,… and $X_1, X_2, \ldots X_r$ are independent and identically distributed random variables with the same distribution as X. We assume that the X_i are not observable.

For the offspring distribution following the power series family, it follows that

$$P(Y = y) = b(y, n)\phi^y / B(\phi), \qquad y = n, n+1, n+2, \ldots \qquad (3.114)$$

, where $\phi = \theta / A(\theta)$ *and* $B(\phi) = \theta^n$

Note that

$$\frac{d\phi}{d\theta} = \frac{1-\mu}{A(\theta)} \tag{3.115}$$

The mean and variance of the Y are nν and $n\nu_2$, where ν and ν_2 are the mean and variance of the total number Y_i of removals for a mortal process with one initial infected individual. We can also note that

$$\nu = 1/(1-\mu),\ \nu_2 = \mu_2/(1-\mu)^3 \qquad where\ \mu_2 = \theta\frac{d\mu}{d\theta}$$

The maximum likelihood estimate for θ is obtained by solving the following formula for θ.

$$\mu(\theta) = 1 - n/y \tag{3.116}$$

The maximum likelihood estimator of the mean of total numbers of removals can be written $\hat{\nu} = y/n$. Fisher's information measure on the mean is

$$I(\nu) = n/\nu_2 = n(1-\mu)^3/\mu_2 \tag{3.117}$$

Borel-Tanner distribution

Let the infected offspring distribution be Poisson,

$P(X=k) = \frac{e^{-\lambda}(\lambda)^k}{k!}$, k=0,1,2,….

Thus the distribution of the total offspring is:

$$P(Y=y) = \frac{n}{y}P(X_1 + X_2 + \cdots + X_y = y-n) = \frac{n}{y}\frac{e^{-\lambda y}(\lambda y)^{y-n}}{(y-n)!} \tag{3.118}$$

, y=*n*, *n*+1, *n*+2,….

Therefore, Y has a Borel-Tanner distribution [55]. The minimum variance unbiased estimator for θ=μ is

$$\left(1+1/\mathrm{n}\right)\left(1-\mathrm{n}/\mathrm{Y}\right) \tag{3.119}$$

The minimum variance unbiased estimator for the variance of the above estimator is

$$(Y-n)(Y+n+n^2)/n^2Y^2 \tag{3.120}$$

3.3 Parameter Estimation

3.3.1 Deterministic model

Inputs of clinical Parameters

The parameters except the transmission coefficient (β) were estimated by published literatures. The incubation period of SARS is about 5 days (1~10days), and the infectious period occurred after the onset of symptoms. the infectious period is not yet understood. The maximum period of communicability in less than 21 days. [56] Therefore, we let α=0.2~0.4. We applied the deterministic model to fit SARS data to estimate the transmission coefficient (β) and effective reproductive number in Taiwan.

Goodness of fit

The total number of simulated cases to fit the total number of observed SARS cases with the Pearson's chi-squared test was also performed.

3.3.2 Estimation of latent period for TB

In September 2011, a resident of the small-scale LTC facility was referred to a hospital and then reported as TB. Following diagnosis of the index case, contact investigation was performed by the local health authority according to the guideline of Centers for Disease Control, Taiwan.[57] All facility staff and residents were considered close household contacts, and visitors and family members were considered close non-household contacts if the shared airspace > 8 hours/day or > 40 hours exposure with cases and were offered TB screening, including a review of symptoms, chest radiography, sputum smear and culture for *Mycobacterium tuberculosis* and the tuberculin skin test (TST). TST which contained 2 tuberculin unit

of purified protein derivative (PPD) of the RT23 strain (1 TU PPD RT23) was performed immediately, and an induration ≧ 10 mm was considered positive.

The local health authority of Changhua applied genotyping for early detection of the TB outbreak when two or more cases of TB occurred within 2-year period in the same congregate setting. Isolates of *M. tuberculosis* were sent to the National Reference Laboratory of Mycobacteriology for genotyping using IS*6110* restriction fragment length polymorphism (RFLP) [58, 59] and spacer oligonucleotide genotyping (spoligotyping).[60] The RFLP and spoligotype analyzed using Bionumerics® software, version 6.6 (Applied Maths, Kortijk, Belgium). Clustered cases, defined as isolates with matching strains, were considered to reflect recent transmission events.[61] An intensified contact investigation targeted at all the facility staffs and residents should be initiated when the first clustered TB case occurred.

An outbreak is defined as at least two epidemiologic-linked cases infected with identical genotypes of *Mycobacterium tuberculosis* isolates. Suspected cases defined as clinical TB cases, on the basis of symptoms, physical findings, and radiologic evidence, had epidemiologic links without laboratory confirmation. LTBI patients were not compelled to receive treatment in Taiwan. However, we still provided the LTBI treatment if they consent.

We estimated the latent and infectious period before symptoms onset following

the exponential distributions with parameters λ and β respectively by maximum likelihood estimation (MLE) method assuming the fixed duration of infectious period [18]. The likelihood function is obtained from this outbreak (appendix). Basic reproductive number (R) was also computed by branching process.[37, 49] The study was approved by the Taipei City Hospital Institutional Review Board.

The likelihood function is as follows:

$$L(\beta,\lambda \mid w)=\prod_{i=2}^{5}\int_{0}^{\mu}\beta e^{-\beta x} f_x(w_i-w_1-x)dx \qquad (3.121),$$

then

$$\int_{0}^{\mu}\beta e^{-\beta x}\lambda^4 e^{-\lambda[(220-x)+(260-x)+(306-x)+(332-x)]}dx$$

$$=\beta\lambda^4\int_{0}^{\mu}e^{-1118\lambda+(4\lambda-\beta)x}dx$$

If $4\lambda-\beta=0$,

$$L=\beta\lambda^4\int_{0}^{\mu}e^{-1118\lambda}dx=\beta\lambda^4 e^{-1118\lambda}\mu=4\lambda^5 e^{-1118\lambda}\mu \qquad (\because \beta=4\lambda)$$

$$\frac{dL}{d\lambda}=4\lambda^4\mu[5e^{-1118\lambda}-1118\lambda e^{-1118\lambda}]=0$$

$$5e^{-1118\lambda}-1118\lambda e^{-1118\lambda}=0 \qquad (\because \lambda,\mu\neq 0)$$

$$\lambda=\frac{5}{1118}=0.0045 \qquad \beta=\frac{20}{1118}=0.0179$$

Therefore, E(X)=1118/5=223.6 days, E(Y)=1118/20=55.9 days and E(X+Y)=279.5 days.

$$I(\lambda)=\frac{d^2L}{d\lambda^2}=80\lambda^3\mu e^{-1118\lambda}-24*1118\lambda^4\mu e^{-1118\lambda}+4*1118^2\lambda^4\mu e^{-1118\lambda}$$

We can get the variance of λ and β from inverse expected information matrix:

$\mathrm{Var}(\lambda)=2.41*10^{-8}\mu=2.17*10^{-6}$ $(if\ \ \mu=90)$. Similarly, we have,

$$\mathrm{Var}(\beta) = Var(4\lambda) = 16 * 2.17 * 10^{-6} = 3.47 * 10^{-5}$$

If $4\lambda - \beta \neq 0$

$$L = \beta\lambda^4 \int_0^{\mu} e^{-1118\lambda + (4\lambda-\beta)x} dx = \frac{\beta\lambda^4}{4\lambda-\beta} \int_0^{\mu} (4\lambda-\beta) e^{-1118\lambda + (4\lambda-\beta)x} dx$$

$$= \frac{\beta\lambda^4}{4\lambda-\beta} e^{-1118\lambda} [e^{(4\lambda-\beta)\mu} - 1]$$

$$\frac{dL}{d\beta} = 4\beta\lambda^3(3\lambda-\beta)e^{(4\lambda-\beta)\mu} - 4\beta\lambda^3(3\lambda-\beta) + [\beta\lambda^4(4\lambda-\beta)(4\mu-1118)]e^{(4\lambda-\beta)\mu-1118\lambda}$$

$$\frac{dL}{d\lambda} = \lambda^4(1+\beta+4\lambda\beta-\beta^2)[e^{4(\mu-1118)\lambda-\beta\mu} - 1]$$

Let $\frac{dL}{d\beta} = \frac{dL}{d\lambda} = 0$, we can get (a) If $\lambda = 0, \beta$ is arbitrary. (b) If $1+\beta+4\lambda\beta-\beta^2 = 0$, then $\beta = \frac{1+\sqrt{5}}{2} = 1.62, \lambda = 0$;(c) if $4(\mu-1118)\lambda - \beta\mu = 1$, then it is insoluble.

All parameters (β,μ,λ) must be positive numbers. Therefore, we obtained the solution, $\lambda = \frac{5}{1118}, \beta = \frac{20}{1118}$.

(3) A matched case-control study for risk factors of TB in Changhua County, Taiwan

We estimated univariate odds ratios (OR) and 95% confidence intervals (CI) using conditional logistic regression. Conditional logistic regression model with a stepwise selection procedure (*P* to enter < 0.1; *P* to remove > 0.05) was used to identify the most important determining factors for active TB.

3.3.3 Three state Continue time Markov process

In the three-state Markov process for depicting the natural course of TB we used Bayesian MCMC to estimate the parameters including infection rate (birther rate), and conversion rate (death rate), and the corresponding regression coefficients of each covariates (age, gender, and IGRA) as mentioned above.

Bayesian method and MCMC

We use Markov chain Monte Carlo (MCMC) method to evaluate the posterior distribution. Assume that the target distribution is known up to a normalizing constant. We would like to construct a chain with π as its stationary distribution.

We take a proposal distribution $q(x, y) = q(y \mid x)$, where the proposal for a new value of a chain is *y*, given that the chain is at value *x*.

Thus *q* defines transition kernel $Q(A, x) = \int_A q(y \mid x)dx$, which is the probability of transition to some $y \in A$.

A Markov Chain with transition density $q(x, y) = q(y \mid x)$ satisfies detailed balance equation if there exists a distribution *f* such that

$$q(y \mid x) f(x) = q(x \mid y) f(y) \tag{3.122}$$

The distribution *f* is stationary and the chain is reversible.

If (3.122) holds,

$$\int q(x \mid y) f(y)dy = \int q(y \mid x) f(x)dy = f(x)\int q(y \mid x)dy = f(x),$$

which is the definition of invariant distribution.

Metropolis-Hastings Algorithm is universal. One can select an arbitrary proposal distribution that is admissible. Of course such arbitrary distribution/kernel cannot be expected to satisfy the detailed balance equation (3.122) for the target distribution π

i.e, $q(x \mid y)\pi(y) \neq q(y \mid x)\pi(x)$

$q(y \mid x)\pi(\mathrm{x}) \neq q(x \mid y)\pi(y)$ Then there is a factor $\rho(x, y) \leq 1$ such that the above inequality is balanced,

$$q(x \mid y)\pi(y)\rho(x, y) = q(y \mid x)\pi(x) \times 1,$$

suppose

$$q(x \mid y)\pi(y) > q(y \mid x)\pi(x).$$

By solving with respect to $\rho(x, y)$, we get

$$\rho(x, y) = \frac{q(x \mid y)\pi(y)}{q(y \mid \mathrm{x})\pi(\mathrm{x})} \wedge 1$$

where $a \wedge b$ denotes $\min\{\mathrm{a,b}\}$.

Metropolis-Hastings Algorithm

(1) Set any $\mathrm{X}_0 = \mathrm{a}$

(2) Generating Y_{k+1} from the density function, $q(\mathrm{Y} \mid \mathrm{X}_k)$; here $q(\mathrm{Y} \mid \mathrm{X})$ is a proposal function.

(3) Generating random variable u from uniform distribution U(0,1)

(4) If $u \leq \rho$ then let $X_{k+1} = Y_{k+1}$; otherwise, let $X_{k+1} = X_k$, where

$$\rho = \min(1, \frac{\pi(Y)q(\mathrm{X} \mid \mathrm{Y})}{\pi(\mathrm{X})q(Y \mid X)})$$

(5) Repeat Step 2 to Step 4.

Elicitation of prior distribution

Non-informative priors using $\mathrm{N}(0,10^6)$ and Uniform (-5,5) were assigned for the exponent of baseline rates of infection and conversion and the regression coefficient of factors.

Parameter	Prior distribution
e^{λ_0}	$N(0,10^6)$
e^{μ_0}	$N(0,10^6)$
$\beta_{11},\beta_{12},\beta_{13},\beta_{14}$	*Uniform* (-5,5)
$\beta_{21},\beta_{22},\beta_{23},\beta_{24}$	*Uniform* (-5,5)

Gibbs sampling and model comparison

A Gibbs sampler was used to derive samples of a stationary distribution by which inferences on posterior distributions were made. The initial values for the regression coefficients, including the parameters of covariates ($\beta_{11},\beta_{12},\beta_{13},\beta_{14},\beta_{21},\beta_{22},\beta_{23},\beta_{24}$) were set to 0.5 and the initial value of 0.001 and 0.01 were assigned for parameter of the exponent of baseline infection rate and conversion rate (e^{λ_0},e^{μ_0}), respectively. Full conditional distributions of models were used to update the process of sampling. All of the estimates of parameters to derive summary statistics of posterior distributions were computed by 50,000 iterations carried out with a thinning interval of 10 after a burn-in period of 50,000 iterations. Such a MCMC simulation yielded a total of 5,000 updated posterior samples. Estimates based on the posterior distributions of parameters were derived using the MCMC method, which was carried out using SAS 9.4. The comparison between models was guided by the deviance information criterion (DIC) values of models.

3.3.4 Discrete time Markov process

(a) Complete information data

Models of branching process and application of the nonparametric method allows us to estimate the R_0 using simulated and real data. We also apply the parametric methods to the simulated and real data for estimating the R_0. It was assumed that the offspring distribution of branching process belongs to the family of generalized power series distribution.

The extinction probability can be computed with a generating-function approach. After having defined and constructed a branching process with offspring distribution, the extinction probability (q) is the smallest non-negative solution of the equation (6).

(b) Incomplete information data

Usually we do not have complete information about the number of infected individuals by each infectious individual. Therefore, we can use "approximations to branching process by means of diffusion processes" or "mortal branching process" with Bayesian method, or linear birth death process to estimate the R_0.

(c)Nonparametric method to estimate the R_0 and the extinction probability

We applied the nonparametric estimators for $\hat{\mu}_n$. It is expressed by the formula

$$\hat{\mu}_n = \frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_{i-1}} \quad (if \quad X_{i-1} > 0); \hat{\mu}_n = 1 \quad (if \quad X_{i-1} = 0)$$

and the estimator of standard error for $\hat{\mu}_n$ is as formula

$$\hat{\sigma} / (\sum_{i=1}^{n} X_{i-1})^{1/2}$$

The consistence estimator of variance (σ^2) is applied as formula

$$\hat{\sigma}^2 = \sum_{i=1}^{n} X_{i-1}(X_i / X_{i-1} - \hat{\mu}_n)^2 / n,$$

(d)Parametric method to estimate the R_0 with Maximum Likelihood Estimation

Let $f(\mathrm{x} \mid \theta)$ denote the density function for the offspring, $X = (x_1, x_2, \ldots, x_n)$. The likelihood function is the density function regarded as a function of θ,

$$L(\theta \mid x) = f(x \mid \theta) \qquad \theta \in \Theta \quad (3.123)$$

The maximum likelihood Estimator is

$$\hat{\theta}(x) = \arg\max_{\theta} L(\theta \mid \mathrm{x}) \qquad (3.124)$$

(e)Nonparametric method to estimate the extinction probability

If k denotes the number of infected offspring by one infectious case, the probability of k offspring was observed as P_k. Hence, PGF $G(s) = \sum_{k=0}^{n} p_k s^k$. We solved the equation x=G(x), and obtained the extinction probability.

(f) Parametric method to estimate the extinction probability

If the distribution of infected offspring was assumes as Poisson, Geometric, Binomial and negative binomial distribution, we can obtain the extinction probability after solving the equation x=G(x). G(x) was the probability generation function of Poisson, Geometric, Binomial or negative binomial distribution.

(g) Mortal branching process

While the infectious disease was extinct, the total number of infected cases (Y) may be a Borel-Tanner distribution under the assumption of R<1. Therefore, the minimum variance unbiased estimator for R_0 is

$$(1+1/\mathrm{n})(1-\mathrm{n}/\mathrm{Y}),$$

where n is the generation.

The minimum variance unbiased estimator for the variance of the R_0 is

$$(Y-n)(Y+n+n^2)/n^2Y^2 .$$

3.4. Simulation

3.4.1 Simulation of birth-death process

(a) Generating data and estimated by approximation equations

We simulated the birth-death process with Matlab 2014a and the algorithm is as follows:

(1) Set the initial number of case (n).

(2) Set the initial time (t).

(3) Set the birth rate λ

(4) Set the death rate μ

(5) Set the final state a

(6) Let U ~ uniform distribution (0,1)

(7) Let the time_λ = -log(U)/n λ

(8) Let the time_μ = -log(U)/ nμ

(9) Let time=min (time_λ, time_μ)

(10) If time_λ < time_μ, then n=n+1

(11) If time_λ < time_μ, then n=n-1

(12) Let state=n.

(13) Let interval time (t_jump) = time.

(14) If n=0 then stop the process

(15) If n=a then stop the process

(16) Go to step 5 till satisfy the condition of stopping.

To simulate the pure birth process, we started with different initial cases after given fixed parameter λ. We used the approximation equation as follows:

$$\sim \frac{1}{\lambda}\log(\frac{a}{n_0}), \tag{3.125}$$

$$\sim \frac{1}{\lambda}\log(a-1), \tag{3.126}$$

and

$$E(T_a)=\frac{1}{\lambda}\sum_{j=n_0}^{a-1}\frac{1}{j} \tag{3.127}$$

Where 'a' denotes the final state, n_0 denotes the number of initial case, and λ denotes the parameter of pure birth process. We also calculate the variance of arrival time with following equation:

$$V(T_a)=\frac{1}{\lambda}\sum_{j=n_0}^{a-1}\frac{1}{j^2} \tag{3.128}$$

(b) Simulated model with birth-death process for SARS

We simulated the birth-death process with Matlab 2014a and the algorithm is as follows:

(1) Set the initial number of case (n).

(2) Set the initial time (t).

(3) Set the birth rate λ

(4) Set the death rate μ

(5) Set the final state a

(6) Let U ~ uniform distribution (0,1)

(7) Let the time_λ = -log(U)/n λ or [the time_λ = -log(U)/ nλ]

(8) Let the time_μ = -log(U)/ nμ or [the time_μ = -log(U)/ μ]

(9) Let time=min (time_λ, time_μ)

(10) If time_λ < time_μ, then n=n+1

(11) If time_λ < time_μ, then n=n-1

(12) Let state=n.

(13) Let interval time (t_jump) = time.

(14) If n=0 then stop the process

(15) If n=a then stop the process

(16) Go to step 5 till satisfy the condition of stopping.

The total number of simulated cases to fit the total number of observed SARS cases and mean time of state in simulation data fit the observed data. Calculating the extinct probability was also performed via simulation.

3.4.2 Simulation of a branching process

We simulated the branching process with Matlab 2014a and the algorithm is as follows:

(1) Set the initial number of case (x).

(2) Set parameters of determined distribution

(3) Set the generation g

(4) Set i=1

(5) Let j=x

(6) Let offspring y ~ given distributions

(7) Let Y[i,j]=y

(8) Let s=s+y

(9) Let j=j-1

(10) If j>0 then go to step (5) else j=s, i=i+1

(11) If i >g then stop

We simulated the branching process for a given offspring distribution. We assumed all the offspring distribution is the independent and identically distribution and it belongs to the family of generalized power series distribution. We carried out the simulations using MATLAB version 8.3.0.532 to generate the data of a branching process.

We also calculated the extinction probability and the average number of offspring in each generation under 10000 times simulations for a given distributions with different parameters (R=2, 1.5, 1.1 and 0.9).

Chapter 4、Data sources

4.1 Generating Data by simulations

We simulated a branching process with 9 generations of data for a given offspring distribution which is the power series distribution under various values of R_0=2, 1.5, 1.1 and 0.9. We simulated a branching process with 10 generations of data for calculating the extinction probability.

We simulated a birth-death process with given birth rate, death rate and different initial infected cases. We calculate the mean and variance of arrival time from the initial state.

4.2 Empirical Data

4.2.1 SARS

SARS is a viral respiratory disease caused by the SARS coronavirus (SARS-CoV). The incubation period for SARS is 3 to 10 days (median 5 days) [56, 62]. Few cases with SARS were noted even though there were many susceptibles in the population and it cannot be protected by vaccination till now. 8098 cases were reported in the worldwide SRAS outbreak from November 2002 to July 2003 [62].

(a) The outbreak of SARS in Taiwan

The cases with SARS of Taiwan in 2003 were obtained from Taiwan CDC [Figure 4.1(a)]. Three hundred and forty-six patients with SARS were reported and case fatality was 10.7% (37 patients died) from Feb. 25 to Jun. 15, 2003. In addition, there were 22,520,776 population of Taiwan at the beginning of 2003.

(b) SARS in a tertiary hospital in Singapore

Two hundred and thirty-eight patients with SARS were reported and case fatality was 13.9% (33 patients died) from Mar. 6 to May 22, 2003. There were total 56 infected with SARS in a hospital in Singapore from Mar. 26 to Apr. 15. Only 3 generation of offspring was noted after outbreak investigation. [Figure 4.3(b)].

4.2.2 *Mycobacterium* tuberculosis

Tuberculosis (TB) is caused by *Mycobacterium* tuberculosis. Although most people received BCG (Bacillus Calmette-Guerin) vaccine against TB in Taiwan, there are still many susceptibles for TB. There are many cases with TB because long incubation period and latent TB present. The incubation period for TB is unknown. The period from infection to primary lesion or significant tuberculin reaction is about 2 to 10 weeks [56]. Using genotyping data may identify small clusters that are likely to become outbreaks and define areas for location-based TB screenings in previous studies [63-66]. Hence, using genotyping can enhance TB outbreak monitoring and the targeted interventions.

Changhua County is located in the middle Taiwan, with a population of around 1,300,000, 12.5% of which were older than 65 years. Nearly 0.7% of the Changhua people resided in LTC facilities. Changhua County is divided into 1 city, 7 urban townships and 18 rural townships. There are around 669,000 people with age ≥ 30.

(a) The outbreaks of TB in the Long-term Care Facility

The outbreak of TB in the LTCF is shown in [Figure 4.2(a)]. The Data was

obtained from Changhua County Public Health Bureau in Taiwan.

(b)The cohort study for TB in Changhua County

We built up a continue-time Markov process to estimate the parameters for the evaluation of the role of QFT-GIT in this cohort study. In addition, we also calculated the extinct probability by the simulation with birth-death process. We built the continue time Markov process with four classes of data, which are explained as follows:

(1)Data from the surveillance system for TB from 2009 to 2011 in Changhua County

The incidence of TB in Changhua County is shown in Figure 4.2(b). A total of 2,420 TB cases with age ≥ 30 enrolled in our cohort study from 2009 to 2011.

(2)Data from TB contact registry database from 2005 to 2011 in Changhua County

A total of 22,510 TB contacts with age ≥ 30 enrolled in our cohort study from 2005 to 2011.

(3)A matched case-control study for risk factors of TB from 2012 to 2014 in Changhua County

The matched case-control study was conducted in Changhua County from March 1, 2012 to December 31, 2013. All participants were continued follow-up till the end of 2014. The study was approved by the Changhua Christain hospital Institutional Review Board. TB cases were confirmed by *Mycobacterium tuberculosis* isolates or clinical TB cases, based on symptoms, physical findings, and radiologic evidence without laboratory confirmation.

All participants received questionnaires with interview, the tuberculin skin test (TST) and QuantiFERON-TB Gold In-Tube test (QFT-GIT) after gave consent to

participate in the study. The interpretation criteria approved by FDA in 2007 for QFT-GIT were followed.[67] One TB case was matched to four TB contacts with living the same TB incidence area. According to TB incidence rate in 26 sub-areas in Changhua County, we divided into three areas; The TB incidence rate > 79.7/100,000 per year was defined as high incidence area, 62.6 – 79/100,000 per year was medium incidence area and <62.6/100,000 per year was low incidence area.

All participants received questionnaires with interview, the tuberculin skin test (TST) and QuantiFERON-TB Gold In-Tube test (QFT-GIT) after they had given informed consent to participate in the study. The questionnaires were designed to collect socio-demographic information, clinical history of TB and factors considered relevant for the disease, such as living location, travel history, contact with ill persons, Cigarettes, alcohol, betel nuts, medications taken, and previous morbidity.

Of the 213 confirmed TB patients and 954 TB contacts attending this study, 212 TB cases and 948 TB contacts met the study criteria. The questionnaires were designed to collect socio-demographic information, clinical history of TB and factors considered relevant for the disease, such as living location, travel history, contact with ill persons, Cigarettes, alcohol, betel nuts, medications taken, and previous morbidity.

Specimen collection and Laboratory Methods. All TB cases were confirmed by *Mycobacterium tuberculosis* isolates or clinical TB cases, on the basis of symptoms, physical findings, and radiologic evidence without laboratory confirmation. TST which contained 2 tuberculin unit of purified protein derivative (PPD) of the RT23 strain (1 TU PPD RT23) was performed immediately, and an induration $\geqq$ 10 mm was considered positive. The IFN-γ assay was performed in 2 stages according to the manufacturer's instructions. The cutoff value of 0.35 IU/mL was defined as a positive

response.[68, 69]

(4)IGRA survey for general population from 2011 to 2013 in Changhua County

A community-based study was conducted in Changhua city from 2011 to 2013. [70] The attendee of Changhua community-based integrated screen (CHCIS) were enrolled after gave consent to participate in the study. All participants underwent both QFT-GIT and TST. A total of 492 subjects who met the inclusion criteria underwent QFT-GIT and TST in the study**. (**Figure 4.3**)**

Chapter 5、Results

5.1 Simulation of R_0 with different methods

5.1.1. Simulation of R_0 using branching process

The results of estimating R_0 on the generating data of a branching process with six generations for a given offspring distribution, such as Poisson (Table 5.1-5.2), Binomial (Table 5.3-5.6), Geometric (Table 5.7-5.8) and Negative Binomial distributions (Table 5.9-5.16), are shown in Table 5.1-5.16. The estimated R_0 were consistent with the nonparametric or parametric method with different distributions. However, the variances were heterogeneous by different methods. The results show that the extinction probability was similar across different methods except the parametric method with geometric assumption. Estimated extinction probability under geometric assumption was higher than that using other methods. Apparently, the assumption of negative binomial offspring distribution was not adequate to estimate the R_0 even though the offspring, in fact, was a negative binomial distribution.

The results of the extinction probability and the average number of offspring in each generation under 10000 times of simulations are listed as Table 5.17-5.22. The calculated extinction probability was a bit higher than previous estimated under the same condition (means the same initial cases size and the same R_0). Calculated extinction probability for a given geometric offspring distribution was higher than those for others. When R_0 increases, the probability of extinction of an infectious disease decreased apparently. An increase in initial infected population size, the size of total infected population increased and the probability of extinction of an infectious disease decreased. When the offspring distribution follows the negative binomial distribution [NB(r,p)], the probability of extinction of an infectious disease decreased with the increase of the value of parameter ‘r’, number of the failure before success in

the same conditions. Final size of total population was affected not only initial infected individuals but also the underlying offspring distribution. (Figure 5.1-5.2).

The results of estimated R_0 with non-parametric method after given generations are showed in Figure 5.3.

2. Estimated time to outbreak size and extinct probability from simulation data with birth death process

Generating data with simulation with birth-death process after given λ= 0.5, 3 and different initial infected cases was performed by Matlab 2014a. We estimated the time to infected size a (T_a) with formula 3.125-3.127. Results of simulated data and estimation with these formulas are showed in Figure 5.4-5.5.

5.1.2 Simulation of R_0 using birth-death process

Figure 5.4 (Appendix Table 1) showed the simulative results of 1000 simulations for pure birth process assuming λ=0.5 compared with the true results estimated the exact equation for E(T_a) which has been derived in the methodological section. The different results given different birth rate (λ=3) are given in Figure 5.5 (Appendix Table 2).

It is very interesting to note that the simulated curve with mean value was still deviant from the curve obtained from the formula (1) $\frac{1}{\lambda}\log(\frac{a}{n_0})$ (2) $\frac{1}{\lambda}\log(a-1)$. However, when n_0 and $\frac{a}{n_0}$ became larger [8], the simulated curve with mean value was close to the true curve obtained from the formula (1) but deviant from formula (2). When λ was enlarged to 3, the results were not changed at all.

5.2 SARS (Severe Acute Respiratory Syndrome)

5.2.1 Outbreak of SARS in Singapore

Simple branching process

According to the outbreak investigation, there were only 3 generations among 55 patients with SARS in the hospital. Estimated R_0 was 2.38 (-3.52~8.11) with non-parametric method, but the extinct probability was one estimated by solving generating function (see the equation 3.67) with empirical observed data. In addition, R_0 was 1 (0.7533~1.3016) for Poisson offspring distribution, 1 (0.6262~1.3738) for geometric offspring distribution. The estimated extinct probability was one by parametric method.

Branching process with generation size determined by incubation period of infectious disease

As we had only the information on the date of onset of SARS onset date (Figure 4.2.b.ii) without the detailed information on contact investigation we assumed one generation had taken 5-7 days. The estimated R_0 was 1 (0.57182~2.5718) with non-parametric method by median days of incubation. Hence, there were 3~8 generations and the estimated R_0 was from 1 (0.1080~ 1.8920) to 1.5 (-3.4609 ~ 6.4609)

5.2.2. SARS in Taiwan

(a) Deterministic model

We simulated the SARS cases with SIR model to fit the cumulated cases in the period of outbreak. The size of total population in Taiwan was 22,520,776 in 2003. The transmission rate (β) of SARS infection was estimated as $7.3*10^{-9}$ in the first 8 weeks, $1.96*10^{-8}$ from the 8^{th} to 11^{th} weeks and $5.4*10^{-9}$ after 11 weeks of outbreak period assuming 2.5 days of infectious period (α=0.4) defined in the methodological section. The R_0 was estimated as 0.411, 1.1035 and 0.30 in the three different periods, respectively. The average R_0 was 0.49 estimated from data on three periods.

A total of observed cases were 346 during the period of 112 days (16 weeks) and 345.1 cases were simulated by SIR model (Figure 5.6). The Goodness of fit test was chi-square =0.0021 with one degrees of freedom (p=0.9634) for SARS cases in 2003. Similarly, the transmission rate (β) was estimated as $3.5*10^{-9}$ in the first 8 weeks, $8*10^{-8}$ from the 8^{th} to 11^{th} weeks and $8.4*10^{-9}$ after 11 weeks of outbreak under the assumption of 5 days of infectious period (α=0.4). The Goodness of fit test was chi-square =0.0021 with one degrees of freedom (p=0.9634). The R was estimated as 0.394, 9.01 and 0.946 in the three different periods, respectively. The average R_0 was 2.52 on three periods. Hence, the estimated R_0 ranged from 0.49 to 2.52.

(b)Branching process

There were 16~22 generations by the incubation of 5 or 7 days. Estimated R_0 was from 0.9971 (0.3308~0.6663) to 0.9971 (0.5090~1.4852) with non-parametric method. Estimated extinction probability was 0.9912 under the assumption of Poisson distribution.

While the SARS went to extinction, the total number of SARS (Y) may be

describe as Borel-Tanner distribution under the assumption of R<1. Therefore, Estimated R_0 was from 0.9790 (0.8437 ~ 1.1143) to 1.0134 (0.8535 ~ 1.1733) according to the information of 16 ~ 22 generations. Estimated extinction probability was 0.9709 ~ 0.9989.

(c) **Birth-death process**

We simulated the SARS cases with pure birth process to fit the cumulated cases in the period of outbreak. Estimated birth rate was 0.0577 per day given the mean time to final outbreak size of 346 with 112.29 (23.72) days. The result of observed data is showed as Figure 5.7(a). The birth rate of the *n* infective cases was $n\lambda$.

The simulated data was not similar to the observed data. We simulated the SARS cases with general birth death process to fit the observed cumulated SARS data. Finally, the estimated birth rates were 0.57 (< 55 day of outbreak), 11.45 (the 55^{th} ~ 80^{th} day of outbreak) and 1.413 (after the 80^{th} day of outbreak). It was as Figure 5.7(b). The birth rate of the *n* infective cases was λ. The result was obtained after fitting the means of T_{32} , T_{300} and T_{346}, where Ta denotes the time to outbreak size of a. The observed T_{32} , T_{300} and T_{346} were 55, 80 and 112 days, respectively. In this model, T_{32} , T_{300} and T_{346} were 54.97(10.09), 80.00 (10.41) and 112.01 (11.47) days.

5.3 TB

5.3.1 **Application to TB in the Long-term Care Facility (LTCF)**

Over a period of 13 months, 7 confirmed cases and 2 suspected cases were associated with a pulmonary TB outbreak which began in September 2011 and continued following until September 2013 (Figure 5.8). The index case of active pulmonary TB with cavitation was diagnosed in Sep. 2011. The contact investigation did not found any new TB cases. Eight months later, two TB cases were reported in May and June 2012, separately, whose genotyping matched that of the index case. Therefore, contact investigations were conducted by the local health authority in July 2012. Totally, there were five TB cases found by the intensified contact investigation. Even through the LTC is a three-floor building housing 63 beds in 13 rooms, there are only 40 beds registered and approved for use by the local health authority. The fresh air exchange rate was not enough in this building, because there was no exhaust unit in the central air conditioning system. There were 62 residents and 18 staff members in this LTC during the investigation period (Figure 5.9). In addition, another 4 case residents had been reported. Four cases located at the third floor and two cases located at the second floor during the investigation period (Figure 5.10). Tracing back the history, these cases and the index case were contacted with each other in the same rooms at the third floor. In addition, the resident case located at the first floor had been cared by the suspected staff case who had the contact history with these cases located at the third floor. Characters of residents and staff are listed in Table 5.23. A new TST conversion rate was 25.0% among residents and 0% among staff. All these resident cases except for the first-floor case had been contacted with each other in the same room during the period of communicability. Among these TB cases, first 5 resident cases presented fever and the others had no symptoms. Sputum smears were

positive for TB but the sputum culture showed no growth in one resident-case and one staff case. In addition, results of RFLP and spoligotyping showed that these active cases were infected by M. tuberculosis with an identical genotype (Figure 5.11). The attack rate was 12.1% (8/66) for residents and 5.6% (1/18) for staff. All of the case-residents were bedridden because of stroke (5; 62.5%), chronic obstructive pulmonary disease with hypoxic encephalopathy (3; 37.5%) and bladder cancer with distal metastasis (1; 12.5%). In addition, half of them needed tracheostomy suction, and the others needed oral and nasal tracheal suction. The strain was susceptible to all first-line anti-TB drugs. However, 5 resident cases were dead during treatment. Of total 46 suspected LTBI persons, only 26 LTBI persons agreed with receiving treatment. Two cases died from other causes during the treatment and one case stopped the treatment due to side effects.

The latent period was estimated about 223.6 days [λ=0.0045 ($2.17*10^{-6}$)] and the infectious period before symptoms onset was estimated about 55.9 days [β=0.0179 ($3.47*10^{-5}$)]. Hence, the incubation period was about 279.5 days. According to our estimation of latent period, there were at least two generations and at most 3 generations. R_0 was bounded between 0.9739 and 0.9796 in this cluster. Control measures including contact tracing and cases follow-up were performed by the facility with assistance of the local public health authority. Closing the facility to admissions, increasing ventilation rates in the building and decreasing the numbers of residents in one room were implemented. All cases of suspected or confirmed TB disease should be placed in an isolated room or transferred to hospitals for treatment. All contact residents and staff members have to receive chest x-ray examination every 6 months till 2 years after the last one confirmed TB case. Sputum culture should be done if any symptom or sign was noted in contacts for suspected TB disease.
The probability of extinction was one under assumption of binomial distribution with

total 80 persons. Of course, the probability of extinction was one under Poisson or geometric offspring distribution's assumption.

5.3.2 A matched case control study for TB

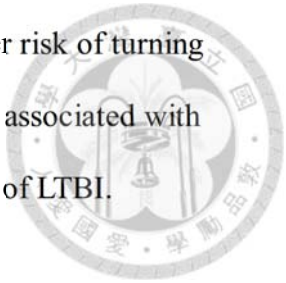
Of the 213 confirmed TB patients and 954 TB contacts attending this study, 212 TB cases and 948 TB contacts met the study criteria. One TB case and six TB contacts were excluded because results of TST and QFT-GIT were all missing. Characteristics of TB cases and TB contacts are listed in Table 5.24. Positive TST and positive QFT-GIT results had the same distribution (68.4% vs 69.4%) among active TB cases. However, the percentage of positive TST (52%) was higher than those of positive QFT-GIT (39.3%).

Applying conditional logistic regression model to data on TB cases and contacts with matched case-control study design shows that both TST and QFT-GIT were independent predictor for the development of tuberculosis adjusted for age group and sex. Results of univariate analysis showing factors significantly associated with TB are presented in Table 5.25. The estimated odds ratios in multivariable logistic regression mode for positive QFT-GIT after further adjusted for positive TST was 2.47 (95% CI: 1.72-3.54, Table 5.26).

This result showed that the effect of QFT-GIT on turning into TB cases depended on whether TST is positive. After considering the interaction term in the model, the odds ratio of QFT-GIT for subjects with positive TST was estimated as 4.28 (95% CI: 1.16-15.76), on the other hand, the odds ratio of QFT-GIT for subjects with negative TST was estimated as 1.15 (95% CI: 0.66-2.00)(Table 5.27).

After the stepwise selection procedure and comparing AIC values, the model considering the main effect of age, sex, TST, and QFT-GIT as well as the interaction term between TST and QFT-GIT is the most parsimonious model (Table 5.28).This

result showed that subjects with LTBI, positive QFT-GIT have higher risk of turning into TB cases. However, persons with positive QFT-GIT may not be associated with higher risk of turning into TB cases for those who were not the state of LTBI.



5.3.3 A case-cohort study for nature history of TB in Changhua County, Taiwan

The application of the case-cohort design in conjunction with the sampling fractions for each state of diseases using a three-state Markov model embodied with birth-and-death process yielded the univariate results as shown in Table 5.29. The overall estimated infection rate (per person-years) and conversion rate (per year) were 0.0168 (95% CI: 0.0157-0.0180) and 0.0113 (95% CI: 0.0098-0.0129). The effect of each covariate on the infection rate is shown in the upper middle panel. Only one parameter (λ) was incorporated with each of three covariates and μ was estimated without covariates. Three models are listed as follows:

(1) $\lambda = \lambda_0 \exp(-0.5316 \times Age(45-64) - 0.6016 \times Age(\geq 65))$

(2) $\lambda = \lambda_0 \exp(0.403 \times Sex)$

(3) $\lambda = \lambda_0 \exp(0.4662 \times IGRA)$

The infection rate was higher for the young age group (30-44 years old) and male sex. Those with higher IGRA were 1.60 (RR=1.59 (95% CI:1.39-1.85) times likely to be susceptible to LTBI compared with low IGRA.

In a similar vein, the effect of each covariate on the conversion rate is shown in the lower middle panel. Only one parameter (μ) was incorporated with each of three covariates and μ was estimated without covariates. Three models are listed as follows:

(4) $\mu = \mu_0 \exp(0.7315 \times Age(45-64) + 1.8319 \times Age(\geq 65))$,

(5) $\mu = \mu_0 \exp(0.403 \times Sex)$,

(6) $\mu = \mu_0 \exp(0.4662 \times IGRA)$.

In contrast to the effect of age on infection rate, the older the subject was, the

higher the conversion rate. Males still had higher conversion rate than females. Those with higher IGRA were two times (RR=2.20 (95% CI:1.60-1.85) likely to surface to TB compared with low IGRA.

The effects of one covariate affecting both of two parameters (λ and μ) are shown in Table 5.30. The three three-state Markov models are listed as follows:

(1) $\lambda_1=\lambda_0 \exp(-0.5348\times Age(45-64)-0.6302\times Age(\geq 65))$
$\mu_1=\mu_0 \exp(0.7074\times Age(45-64)+1.846\times Age(\geq 65))$

,

(2) $\lambda_2=\lambda_0 \exp(0.4007\times Sex)$
$\mu_2=\mu_0 \exp(0.5277\times Sex)$,

(3) $\lambda_3=\lambda_0 \exp(0.4671\times IGRA)$
$\mu_3=\mu_0 \exp(0.7495\times IGRA)$.

The joint effect of each covariate (age, gender, and IGRA) on infection rate and conversion rate were similar to the marginal effect of each covariate.
Results of multivariate analysis with the incorporation of three covariates into infection rate or conversion rate are presented in the middle panel of Table 5.31. A total of three models are listed as follows:

(1)

$\lambda_1=\lambda_0 \exp(-0.5825\times Age(45-64)-0.7585\times Age(\geq 65)+0.3576\times Sex+0.5307\times IGRA)$

μ_1

(2)

λ_2

$\mu_2=\mu_0 \exp(0.6126\times Age(45-64)+1.6308\times Age(\geq 65)+0.4233\times Sex+0.4934\times IGRA)$

After adjustment for age and gender, the effect of IGRA on infection rate (RR= 1.70; 95%CI:1.46-1.95) remained the same as that in univariate analysis but the effect on conversion rate were slightly reduced but were statistically significant (RR=1.64;

95%CI:1.19-2.25).

Results of multivariate analysis with the incorporation of three covariates into infection rate and conversion rate simultaneously are presented in the right panel of Table 5.31. The model form is listed as follows.

(3)
$$\lambda_3=\lambda_0 \exp(-0.5799\times Age(45-64)-0.7896\times Age(\geq 65)+0.3583\times Sex+0.5371\times IGRA)$$
$$\mu_3=\mu_0 \exp(0.4256\times Age(45-64)+1.4886\times Age(\geq 65)+0.4104\times Sex+0.4578\times IGRA)$$

(5.1)

After taking the effect of age and sex on both infection rate and conversion rate into account, subjects with positive QFT-GIT still had higher risk of being infected and converting to tuberculosis with estimated RR being 1.71 (95% CI: 1.49-2.00) and 1.58 (95% CI: 1.15-2.17), respectively (Table 5.31). Table 5.32 show the model selection with deviance information criterion (DIC). The trace plus of the parameters in equation 5.1 with Bayesian MCMC method are presented in Figure 5.12.

Given the parameters of λ and μ estimated for subjects with various combinations of three covariates (Formula 5.1), we simulated the time to final size (10) and final size (30) of LTBI. The results are listed in Table 5.33 and Table 5.34. The results suggest one initial case may take 60.94 days to have 10 of final size and 87.36 days to have 30 of final size without considering covariates. The young people, male and positive IGRA tended to spread quickly. The male aged less than 45 years with positive results of IGRA took only one week to reach final size given five initial size. It should be noted that an increase in initial size reduced the time to reach the expected final size. The corresponding results of estimated extinct probability with various combinations of covariates and different initial cases of LTBI are shown in

Table 5.35. It is clearly seen that when initial size was larger than five the extinct probability of TB was very unlikely.

Chapter 6 Discussion

6.1 Summary of findings

The main contribution made from this thesis embrace two aspects, clinical and epidemiological epidemic investigation of SARS and TB and the methodological development including the application of various stochastic processes to throwing light on the infectious process (including latency period and infectious period) in the axis of infection and incubation period in the axis of disease process by estimating R_0, the expected time to reach final size, and the extinct probability and the development of a novel infectious model by combining the flexible multi-state Markov process with birth-and-death process.

6.1.1 Clinical and epidemiological findings on outbreak of SARS and TB

The main contributions of clinical and epidemiological epidemic profiles of SARS and TB include the following points.

1. Applications to SARS in the two regions: In Singapore, the estimation of R_0 given 3~8 generations was between 1 and 1.5. The estimated extinct probability was almost certain using branching process. The three period of SAS outbreak yielded 0.99 of R0 using branching process in Taiwan. The estimated extinct probability was 0.99. The similar findings were noted by using the mortal branching process with Borel-Tanner distribution.
2. Estimate unobserved incubation period with approximately 9 months, including seven months of latent period and two months of infectious period before onset of symptoms given data from an outbreak of TB occuring even among subjects with negative TST result after undergoing TB screening. Surveillance of the elderly

people even with a negative TST after TB screening is still necessary given a long latent period if the outbreak of TB in a long-term care facility is to be controlled.

3. This is the first study to assess the effect of IGRA on the occurrence of TB by conducting a case-control study making allowance for demographic characteristics and induration size of TST.
4. This is the first study to assess the effects of age, gender, and IGRA on infection from susceptible to LTBI and also the conversion from LTBI to TB in the natural course of TB. The young age was at increased risk for being LTBI but the old age enhanced the risk of conversion from LTBI to TB. Male had higher risk for being infected as LTBI and also the conversion from LTBI to TB. The elevated IGRA plays a significant role not in the infection rate (from free of LTBI (susceptible) to LTBI) but also in the conversion rate after adjusting for age and gender.
5. The application of infection rate (birth rate) and conversion rate (death rate) gives the time expected to reach number of LTBI of final size and the extinct probability by various combinations of age, gender, and the results of IGRA. Subjects with positive IGRA results had shorter expected time to reach final size than those with negative result.

6.1.2 Methodological development

This thesis has contributed to developing the methodological part related to infectious disease as follows.

1. Provide several statistical simulated methods for simulating various R_0 with branching process and also birth-and-death process so as to estimate the extinct probability and the expected time to reach final size.
2. Demonstrate how to apply the Becker's SIR model in conjunction with branching process to estimate incubation period and latent period for the surveillance of TB.
3. Develop a continuous-time Markov process embodied with birth-and-death

process in conjunction with a novel case-cohort design data given the known sampling fraction to assess how covariates such as IGRA affect the infection rate and the conversion rate framed with a three-state Markov process. The further application of birth-and-death process used in the simulation of SARD process can compute the extinct probability and the expected time to reach final size, both of which provide a new insight into the golden period for the formulation of policy for the containment of infectious disease in question.

6.2 Clinical Usefulness

In spite of numerous studies that have already estimated R_0 in the previous studies, it is lacking of a systematic approaches to estimate R_0 by proposing various stochastic processes to accommodate various types of infectious diseases. The proposed model can be flexibly adapted to large and small population of outbreak, time to each period, latent period, incubation period exactly known or unknown form empirical data such as TB, the probability of extinction, and the expected time to reach final size.

6.2.1 SARS

The estimated R_0 depends on the determinants of infectious disease. Variant values of R_0 for SARS were obtained using the SIR model even though all simulations have not rejected the Goodness of fit test. The consistent results were obtained from simple branching process, mortal branching process and birth-death process. The effective R for SARS was also heterogeneous in the different stage which was before or after global alert in previous reports.[71, 72] In addition, results of these studies

were similar to ours.

To compare previous studies, R_0 was estimated 1.1 for the nationwide of Singapore [73]. Even though our data was partial information of outbreak in Singapore and especially it was derived from hospital, our findings were similar to the previous report. By using branching process, we can provide not only the information about 95% CI of R_0 but also the expected time to reach final size and the extinct probability of SARS. Such information is very useful for how to devise the policy of the containment in order to forestall the outbreak before the expected time reaching to final size and also the likelihood of extinction.

From the results of simulations with birth-death process, we found that the parameter λ more tended to fit the observed data than the parameter $n\lambda$. The infection rate was not associated with cumulated cases because it may be resulted from isolation policy.

6.2.2 TB

6.2.2.1 TB outbreak in LTCF

In a LTC facility, 7 confirmed cases and 2 suspected cases were associated with a pulmonary TB outbreak which began in September 2011 and continued following until September 2013. The TST positive rate of contacts was 50.0 % (29/58) and 76.5% (13/17) among residents and staff, respectively. By comparison, TB contacts had 69% positive TST response in Taiwan community-based study.[74] However, a new TST conversion rate was 25.0% for residents. LTBI treatment should be

considered even though in BCG-vaccinated populations, which was suggested in previous studies.[75, 76] Furthermore, monitoring and documented TST results for residents regularly should be considered in the LTC facility and other health care facility.

Using genotyping data may identify small clusters that are likely to become outbreaks and define areas for location-based TB screenings in previous studies.[63-66] Hence, using genotyping can enhance TB outbreak monitoring and the targeted interventions (e.g., intensified contact investigation). However, LTC facilities were never reported as high risk for TB outbreaks by using genotyping monitor in previous studies. This strategy also enables us to identify which LTC facilities are in high-risk status in our study.

Two sputum culture-positive TB cases with normal CXR was noted, which was not uncommon [77]. Screening by sputum culture should be suggested for all contacts. In addition, early diagnosis of active TB occurs in individuals who initially have negative TST results is important to reduce the transmission in a LTC facility. Two active TB cases that tested TST-negative were later diagnosed by sputum examinations in this outbreak. Sputum examination and chest radiograph should be performed regardless of a negative TST result. In addition, providing LTBI treatment to elderly contacts with comorbidities regardless of LTBI test results should be

considered. The above findings were due to the lower sensitivity of the TST among the elderly. [78-82] However, we are also concern about the risk of adverse reactions to the medication for LTBI among the elderly population. Therefore, further study about the tolerance of LTBI treatment among the elderly with high comorbidity in LTC facilities or other health care settings should be considered.

Studies on estimation of the latent and infectious period of TB are rare. The infectious period before symptoms onset was estimated about 2 months in this outbreak. Based on expert opinions, the infectious period defined as 3 months before symptom onset is recommended[83]. The latent period was estimated about 32 weeks. Contact investigations and LTBI treatment become important strategies in TB control and elimination within the latent period. The shorter incubation periods in our study than previous study (45% within one year; median 1.26 years)[84] may be attributed in part to the outbreak study, and perhaps also in part to some bias resulting from only 2-year follow-up.

Previous studies on R_0 for TB are various and no available data from Taiwan TB for comparison. Our estimation of R_0 excluding reactivation of TB and very slowly progressing TB (longer than 2 years) may be underestimated [85].

There were some limitations in this study. Without two-step testing to measure tuberculin reactivity in this investigation and no previous documented baseline TST

results within the prior 12 months, false-negative TST results are a greater concern. Because boosting is more common in the elderly, occurring in 15% of elderly subjects screened.[86] False-positive TST results are not common. The effect of remote vaccination with BCG on positive tuberculin response in adults aged >30 years is probably negligible. [74] LTBI or *M. tuberculosis* infection should be assumed in an older patient with a positive TST result. The air exchange rate was low in field investigation, but the number of air changes per hour was not measured. Aerosolization of the TB patient's secretions from repeated suctioning and inadequate ventilation in his room were routes for spreading in this outbreak. Prior investigations of TB outbreaks in health care settings have found the same transmission routes.[78] We inferred that sputum suction and low air ventilation rate were the risk factors for TB transmission among residents in this outbreak.

This outbreak reinforced the importance of considering comprehensive TB screening included sputum examinations and chest radiograph for all residents and staff regardless of a negative TST result given the estimated long latency period. LTBI treatment to elderly contacts should be considered if the tolerance of treatment for the elderly with high comorbidity can be accepted.

6.2.2.2 IGRA & LTBI/TB

The tuberculin skin test (TST) for detecting latent TB infection (LTBI) was

performed for decades. However, the lower sensitivity of the TST among the elderly was reported in previous studies because of booster phenomenon, [78-82] and the results of TST were affected by BCG (bacillus Calmette-Guérin) vaccine. [87] Though the effect of remote vaccination with BCG on positive tuberculin response in adults aged >30 years is probably negligible,[74] Interferon-gamma release assays (IGRAs) have emerged as attractive alternatives because it was unaffected by BCG and most NTM exposure and it has no booster phenomenon.

Interferon-gamma release assays (IGRAs) appears to demonstrate better specificity than the TST and good correlation to the risk of exposure to TB. [88-90] IGRAs may be better at detecting recent rather than remote infection.[90] However, the performance of IGRA was different between high and low TB incidence settings in previous reports, and relatively lower sensitivity in high-incidence countries was also reported.[91, 92] The IGRAs have dynamic characteristics over time.[93] Therefore, the difference between IGRAs and TST in the course of TB infection is still unclear.

Diel et al. evaluated progression from latent tuberculosis infection to active TB in close contacts of TB patients. Of 954 subjects, 20.8% were QFT positive. 12.9% of a TB progression rate among subjects with QFT positive was found over the observation period.[94] Corresponding progression rate of 4.8% for the TST positive (>10mm) was significant lower. The progression rate of 28.6% for QFT-positive children was significantly higher than 10.3% for adults in Diel's study. Another prospective cohort study for QuantiFERON screening of adult contacts reported 13.4% (2-year rate) in QFT positive adults for developing TB [95]. Similar results

were found in our study. The infection rate was estimated as 1.68% in our study. The high IGRA is more likely to be susceptible to LTBI compared with low IGRA. The infection rate was higher in the young age group. Haldar et al., also reported the positive predictive value of QFT loses significance for older contacts (≥36 years). However, our study presents the gender difference for developing TB. The infection rate was higher in male population in our study but there was lacking of statistical significant in sex for progression of TB in Diel's study.[94]

The heterogeneity of tuberculosis incidence was observed among different age groups and sex in previous studies[96]. This observation may due to the differences in rate of conversion between age groups and sex as well as that in rate of infection. The effect of QFT-GIT on the progression of tuberculosis is faced with similar argument.

Applying conditional logistic regression to the collected data on case control study of TB cases and their contacts, we evaluated the effects QFT-GIT on the probability of turning into TB cases adjusted for age, sex, and TST results. Although this approach shows significant correlation between positive QFT-GIT and the probability of turning into TB cases and hence provides the information of the role of QFT-GIT on TB progression. However, this conventional method is not capable of telling the mechanism of these effects.

In addition to prove the correlation between positive QFT-GIT and the development of TB cases, the role of QFT-GIT on different stages of disease progression is also of great interest. Although previous studies showing the correlation between QFT-GIT the intensity of TB exposure implying the correlation between TB infection and positive QFT-GIT result [97], our results further demonstrated that positive QFT-GIT was also an independent predictor for the progression to tuberculosis.

The gold standard for diagnosis of active TB is bacteriological diagnostic tests.

However, they have limitations because smear microscopy test is less sensitive and culture test is time-consuming. Though QFT-GIT was developed for diagnosis of LTBI, IGRAs can be used as diagnosis of active TB in previous studies.[98, 99] But IGRAs cannot distinguish active TB from latent infection because its utility for active TB diagnosis is low in TB-endemic countries, where most of the people are believed to be latently infected with M. tuberculosis.[100, 101] QFT-GIT reduced utility for active TB diagnosis because it is able to diagnose both active TB disease and LTBI, [102] which leads to the different sensitivity of QFT-GIT for active TB diagnosis in high or low TB endemic areas.

IGRA was an independent risk factor associated with active TB and positive IGRA with positive TST had more risk to developed active TB in our study. Dosanjh DP et al. reported that IGRA in combination with TST can be used to rule out the suspicion of active TB disease.[103] The role of QFT-GIT on different stages of disease progression is still unclear.

Applying generalized linear stochastic model enable us to assess the effect of positive QFT-GIT on the force of disease progression, namely infection rate and conversion rate, while taking relevant factors such as age and sex into account. By using multistate method, we are able to quantify these effects on the rate of infection and conversion. However, it is often not admissible to have information on the results of biomarker such as QFT-GIT at population level. This is especially true for the study on diseases with relatively low incidence such as tuberculosis. To tackle this difficulty, we proposed case-cohort study method based on multistate disease progression and derive the probability of having observed sample given sampling fractions for each underlying cohort in terms of the state of disease progression using Bayesian revision. In addition to elucidate the mechanism on disease progression, the strength of the proposed method make it possible to evaluated the effect of candidate

biomarkers using sampled data, which is more feasible and cost-saving (see below).

6.3 Strength and concerns of methodology

6.3.1 The novelty of multi-state Markov model in conjunction with birth-death process

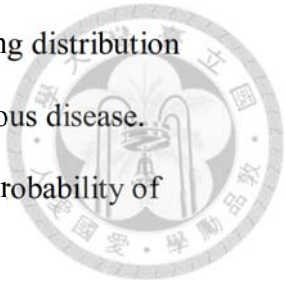
The most breakthrough point in the methodological development of this thesis is that we made use of a continuous-time Markov process embodied with birth-and-death process to capture the disease natural history of infectious disease consisting of infectious process and also the conversion of infected to disease process. The parameters on birth rate and death rate were further applied to calculate tie expected time to reach final size and also the extinct probability. Doing so has numerous advantages for modelling the natural course of susceptible-infection-disease like TB. First, as mentioned earlier, covariate-specific such as IGRA-specific infection rate and conversion rate provide a new insight into personalized infection control when screening for TB with IGRA is considered. LTBI with positive IGRA may consider the use of prophylactic treatment and require intensive surveillance. Second, the identification of high-risk subjects in the LTBI may provide an evidence on personalized isolated and quarantine suggestion in order to reduce the likelihood of spread of TB. Third, quantifying the disease natural history of chronic infectious disease would provide a pseudo control group for evaluation of any intervention related to TB control when a randomized controlled trial design cannot be used. Fourth, the parameters also provide best-case estimates for future cost-effectiveness analysis.

6.3.2 Expedient use of non-standard case-cohort design

Another novelty of this thesis is to make use of multi-state Markov process in conjunction with a non-standard case-cohort design. The transition parameters underpinning the disease natural history using case-cohort design in Markov model by Chen et al. enable one to evaluate the effectiveness of intervention using surrogate endpoint for dispensing with long-term follow-up.[48] This approach is applied to modelling of progression of adenoma to colorectal cancer as an example.[47] A case-cohort study was conducted to review and collect a set of random samples to estimate the disease natural history. A total of 305 normal cases, 300 polyp patients, and 116 CRCs were extracted from a total of 13908 subjects in the routine medical care setting and collected with the full information on pathological findings. It is less costly and efficient to review all medical charts. A Markov process was further applied to this case-cohort study to elucidate the disease progress from adenoma to carcinoma with and without consideration of covariates, e.g., adenoma size. This thesis applied this novel and efficient design together with a novel statistical method for estimating multi-state TB natural history that would render the study very efficient.

6.3.3 Simulations

According to our results of simulations with branching process, we can estimate the R_0 with parametric or non-parametric method. It depends on the structure of information about infection disease. If the sample size (the number of generations) is large enough, we can obtain unbiased estimation of R_0 using the method of simulation with branching process. However, if small generations of infection disease, we can't obtain the unbiased R_0.



Combination of non-parametric method with Poisson offspring distribution assumption, we can easily calculate the extinct probability of infectious disease. Though deterministic model can be used to estimate R_0, the extinct probability of infectious disease is unable to be known.

Similarly, the precise of estimated R using the method of simulation with birth-death process depends on the initial infected cases and final infection size.]

From results of our simulations, R_0 can be consistently estimated by using any method. However, with the decrease of generation size, estimation of R_0 by using nonparametric method may tend to increase variance than the parametric method if the offspring size is larger than the generation size. It depends on the sample size. Hence, the results of estimation by parametric method are unbiased than by nonparametric method in our illustration of SARS in Singapore. The extinct probability was consistent by both methods because sample size is the same.

In addition, the method of branching process and the SIR model all depend on the incubation period seemed to have a good estimation.

Data of TB outbreak was used for the validation of time prediction from birth-death process model. There were at least 9 newly onset of LTBI during 2 months periods in the outbreak of LTCF (Table 5.23). If these night LTBI cases were infected by one initial active TB case, it was fitted for our estimation (about 60 days). However, it should be noted that the expected time to reach final size is very easily affected by small initial size and final size when birth-death process is considered. This is worthy of being investigated in the future.

6.5 Limitation

In fact, human contact patterns are more heterogeneous than assumed by homogeneous-mixing models, such as SIR model and simple birth-death process. This

simplifying assumption makes the analysis easy to control but may not adequately reflect reality. Without two-step testing to measure tuberculin reactivity in the cohort study, lower TST positive rate was estimated because of booster phenomenon. Hence, LTBI cases were lower estimated. Furthermore, reactivation or reinfection of TB was not considered in the three-state Markov model because those cases were few in TB.

6.6 Conclusion

Several conclusions could be reached in this thesis, including

1. the simulation of various R_0 with branching process and also birth-and-death revealed a full knowledge of the extinct probability and the expected time to reach final size given initial size;
2. the demonstration of how to apply the Becker's SIR model in conjunction with branching process to estimate incubation period and latent period for the surveillance of TB;
3. the development of a novel continuous-time Markov process embodied with birth-and-death process in conjunction with a novel case-cohort design data given the known sampling fraction to assess how covariates such as IGRA affect the infection rate and the conversion rate framed with a three-state Markov process. The further application of birth-and-death process used in the simulation of SARS process can compute the extinct probability and the expected time to reach final size, both of which provide a new insight into the golden period for the formulation of policy for the containment of infectious disease in question.

References

1. Bacaër, N., *Daniel Bernoulli, d'Alembert and the inoculation of smallpox (1760)*, in *A Short History of Mathematical Population Dynamics*. 2011, Springer. p. 21-30.
2. Anderson, R. and R. May, *Infectious disease of humans: dynamics and control*. 1991: Oxford University Press.
3. Lai, C.-C. and H.-H. Chen, *Dynamic Infection model Associated with Herd Immunity - Illustrations of Poliomyelitis and Hand-Foot-Mouth Disease* in *Graduate Institute of Preventive Medicine,College of Public Health*. 2007, National Taiwan University.
4. Lai, C.-C., et al., *A Dynamic Model of Enterovirus Outbreaks: Hand Foot and Mouth Disease in Taiwan.* BMC infectious disease, 2014.
5. Bartlett, M.S., *Deterministic and stochastic models for recurrent epidemics.* Proc. of the Third Berkeley Symp. on Mathematical Stats. and Probability, 1956. **4**: p. 81-108.
6. Bailey, N.J.T., *On estimating the latent and infectious periods of measles. 1. Families with two susceptibles* Biometrika, 1956. **43**: p. 15-22.
7. Bartlett, M.S., *Measles periodicity and community size.* J. R. Statist. Soc., 1957. **A120**: p. 48-70.
8. Cox, D. and H. Miller, *The theory of stochastic processes*. 1965: Chapman and Hall/CRC.
9. WH, H., *Epidemic disease in England - the evidence of variability and of persistency of type.* Lancet, 1906. **1**: p. 733-9.
10. K., D., *Transmission and control of arbovirus disease.* Society for Industrial and Applied Mathematics, 1975: p. 104-21.
11. RM, A. and M. RM., *Immunisation and herd immunity.* Lancet, 1990. **335**(8690): p. 641-5.
12. Farrington, C.P., M.N. Kanaan, and N.J. Gay, *Estimation of the basic reproduction number for infectious diseases from age-stratified serological survey data.* Journal of the Royal Statistical Society Series C-Applied Statistics, 2001. **50**: p. 251-283.
13. RM, A. and M. RM., *Age-related changes in the rate of disease transmission: implications for the design of vaccination programmes.* J Hyg. , 1985. **94**: p. 365-436.
14. Heffernan, J., R. Smith, and L. Wahl, *Perspectives on the basic reproductive ratio.* J R Soc Interface, 2005. **2**(4): p. 281-93.

15. JM, H., S. RJ, and W. LM, *Perspectives on the basic reproductive ratio.* J R Soc Interface, 2005. **2**(4): p. 281-93.
16. G, S., et al., *Key transmission parameters of an institutional outbreak during the 1918 influenza pandemic estimated by mathematical modelling.* Theor Biol Med Model, 2006. **3**(38): p. 1-7.
17. Lai, C.-C., et al., *A Dynamic Model of Hand Foot and Mouth Disease Outbreaks: Hand Foot and Mouth Disease in Taiwan.* under submitted, 2015.
18. Becker, N.G., *Analysis of Infectious Disease Data*. 1989: Chapman & Hall/CRC
19. Greenwood, M., *On the Statistical Measure of Infectiousness.* J Hyg (Lond), 1931. **31**(3): p. 336-51.
20. Thomas, J.C. and D.J. Weber, *Epidemiologic methods for the study of infectious disease*. the first ed. 2001, United States: Oxford University Press.
21. WH, F., *Some conceptions on epidemics in general.* Am J Epidemiol, 1976. **103**: p. 141-51.
22. T, S., B. BP, and A. GL, *Quantifying the impact of hepatitis A immunization in the United States, 1995-2001.* Vaccine, 2004. **22**(31-32): p. 4342-50.
23. Akhtar, S., T. Carpenter, and S. Rathi, *A chain-binomial model for intra-household spread of Mycobacterium tuberculosis in a low socio-economic setting in Pakistan.* Epidemiology and infection, 2007. **135**(01): p. 27-33.
24. Brooks-Pollock, E., et al., *Epidemiologic inference from the distribution of tuberculosis cases in households in Lima, Peru.* Journal of Infectious Diseases, 2011. **203**(11): p. 1582-1589.
25. Ng, J. and E.J. Orav, *A generalized chain binomial model with application to HIV infection.* Mathematical biosciences, 1990. **101**(1): p. 99-119.
26. Hsu, C.-Y., et al., *Surveillance of influenza from household to community in Taiwan.* Transactions of The Royal Society of Tropical Medicine and Hygiene, 2014. **108**(4): p. 213-220.
27. Nishiura, H. and G. Chowell, *Household and community transmission of the Asian influenza A (H2N2) and influenza B viruses in 1957 and 1961.* 2007.
28. Azman, A.S., et al., *Household transmission of influenza A and B in a school-based study of non-pharmaceutical interventions.* Epidemics, 2013. **5**(4): p. 181-186.
29. 許辰陽, *應用感染鍊二項模式及隨機概念分析流行性感冒家戶資料 - 運用貝式統計*, in *Application of Chain Binomial Model with Stochastic Concept to Influenza Household Data - A Bayesian Approach*. 2009, 臺灣大學.
30. Hsu, C.-Y., *Generalized Linear Stochastic Process with Applications to*

Infectious Diseases, in *Graduate Institute of Epidemiology and Preventive Medicine,*
College of Public Health. 2014, National Taiwan University: Taiwan.
31. Bailey, N.T., *The mathematical theory of infectious diseases and its applications*. 1975: Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE.
32. Becker, N.G., *Analysis of Infectious Disease Data*. 1989, London, New York: Chapman and Hall.
33. Daley, D.J., J. Gani, and J.M. Gani, *Epidemic modelling: an introduction*. Vol. 15. 2001: Cambridge University Press.
34. O'Neill, P.D., et al., *Analyses of infectious disease data from household outbreaks by Markov chain Monte Carlo methods.* Journal of the Royal Statistical Society: Series C (Applied Statistics), 2000. **49**(4): p. 517-542.
35. Hsu, C., et al., *Analysis of household data on influenza epidemic with Bayesian hierarchical model.* Mathematical biosciences, 2015. **261**: p. 13-26.
36. Becker, N., *On Parametric Estimation for Mortal Branching Processes.* Biometrika, 1974. **61**(3): p. 393-399.
37. Becker, N., *Estimation for an epidemic model.* Biometrics, 1976. **32**(4): p. 769-77.
38. Farrington, C.P., M.N. Kanaan, and N.J. Gay, *Branching process models for surveillance of infectious diseases controlled by mass vaccination.* Biostatistics, 2003. **4**(2): p. 279-295.
39. Angelov, A.G. and M. Slavtchova-Bojkova, *Bayesian estimation of the offspring mean in branching processes: Application to infectious disease data.* Computers and Mathematics with Applications, 2012. **64**: p. 229-235.
40. Kretzschmar, M., *A renewal equation with a birth-death process as a model for parasitic infections.* J Math Biol, 1989. **27**(2): p. 191-221.
41. Wood, R.M., J.R. Egan, and I.M. Hall, *A dose and time response Markov model for the in-host dynamics of infection with intracellular bacteria following inhalation: with application to Francisella tularensis.* J R Soc Interface, 2014. **11**(95): p. 20140119.
42. Diekmann, O. and J.A.P. Heesterbeek, *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation.* 2000, New York: Wiley.
43. Chiang, C.L., *An Introduction to Strochastic Processes and Their Application*. 1980: Robert E. Krieger Publishing Company.
44. Blower, S.M., et al., *The intrinsic transmission dynamics of tuberculosis epidemics.* Nat Med, 1995. **1**(8): p. 815-21.

45. Hsu, C.-Y., *Generalized linear stochastic process with applications to infectious diseases* in *Graduate Institute of Epidemiology and Preventive Medicine, College of Public Health*. 2014, National Taiwan University: Taiwan.
46. Chen, T.H., et al., *Estimation of sojourn time in chronic disease screening without data on interval cases.* Biometrics, 2000. **56**(1): p. 167-72.
47. Chen, C.D., et al., *A case-cohort study for the disease natural history of adenoma-carcinoma and de novo carcinoma and surveillance of colon and rectum after polypectomy: implication for efficacy of colonoscopy.* Br J Cancer, 2003. **88**(12): p. 1866-73.
48. Hsiu-Hsi Chen, T., et al., *Stochastic model for non-standard case-cohort design.* Stat Med, 2004. **23**(4): p. 633-47.
49. Becker, N., *Estimation for discrete time branching processes with application to epidemics.* Biometrics, 1977. **33**(3): p. 515-22.
50. Farrington, C.P. and A.D. Grant, *The distribution of time to extinction in subcritical branching processes: Applications to outbreaks of infectious disease.* Journal of Applied Probability, 1999. **36**(3): p. 771-779.
51. Diekmann, O. and J. Heesterbeek, *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation.* 2000, New York: Wiley.
52. Noack, A., *On a class of discrete random variables.* Annals of Mathematical Statistics, 1950. **21**: p. 127-132.
53. Perlman, M.D. *Stochastic comparisons among conditional processes derived from a semicritical Galton-Watson branching process*. Technical Report No. 594 2012; Available from: http://www.stat.washington.edu/research/reports.
54. Guttorp, P. and M.D. Perlman. *Extinction or Explosion in a Galton-Watson Branching Process: Testing or Prediction?* Technical Report no. 602 2012; Available from: http://www.stat.washington.edu/research/reports.
55. Haight, F.A. and M.A. Breuer, *The Borel-Tanner Distribution.* Biometrika, 1960. **47**(1/2): p. 143-150.
56. David L. Heymann, M., *Control of communicable disease manual*. 18th ed. 2004: American Public Health Association.
57. Centers for Disease Control, D.o.H., R.O.C.(Taiwan), *TB Control Manual*. 2nd ed. 2009: Centers for Disease Control, Department of Health, R.O.C.(Taiwan).
58. van Embden, J.D., et al., *Strain identification of Mycobacterium tuberculosis by DNA fingerprinting: recommendations for a standardized methodology.* J Clin Microbiol, 1993. **31**(2): p. 406-9.
59. Cave, M.D., et al., *Stability of DNA fingerprint pattern produced with IS6110 in strains of Mycobacterium tuberculosis.* J Clin Microbiol, 1994. **32**(1): p.

262-6.
60. Kamerbeek, J., et al., *Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology.* J Clin Microbiol, 1997. **35**(4): p. 907-14.
61. Driver, C.R., et al., *Which patients' factors predict the rate of growth of Mycobacterium tuberculosis clusters in an urban community?* Am J Epidemiol, 2006. **164**(1): p. 21-31.
62. Chan-Yeung, M. and R.H. Xu, *SARS: epidemiology.* Respirology, 2003. **8 Suppl**: p. S9-14.
63. Althomsons, S.P., et al., *Using routinely reported tuberculosis genotyping and surveillance data to predict tuberculosis outbreaks.* PLoS One, 2012. **7**(11): p. e48754.
64. de Vries, G., et al., *A Mycobacterium tuberculosis cluster demonstrating the use of genotyping in urban tuberculosis control.* BMC Infect Dis, 2009. **9**: p. 151.
65. Moonan, P.K., et al., *What is the outcome of targeted tuberculosis screening based on universal genotyping and location?* Am J Respir Crit Care Med, 2006. **174**(5): p. 599-604.
66. Kik, S.V., et al., *Tuberculosis outbreaks predicted by characteristics of first patients in a DNA fingerprint cluster.* Am J Respir Crit Care Med, 2008. **178**(1): p. 96-104.
67. Mazurek, G.H., et al., *Updated guidelines for using Interferon Gamma Release Assays to detect Mycobacterium tuberculosis infection - United States, 2010.* MMWR Recomm Rep, 2010. **59**(RR-5): p. 1-25.
68. Lalvani, A., et al., *Rapid detection of Mycobacterium tuberculosis infection by enumeration of antigen-specific T cells.* Am J Respir Crit Care Med, 2001. **163**(4): p. 824-8.
69. Brock, I., et al., *Comparison of tuberculin skin test and new specific blood test in tuberculosis contacts.* Am J Respir Crit Care Med, 2004. **170**(1): p. 65-9.
70. Hsieh, M.-C., *A community-based study for latent TB infection in a country with an intermediate TB burden, using QuantiFERON-TB Gold In-Tube and tuberculin skin test*, in *Graduate Institute of Epidemiology and Preventative Medicine College of Public Health*. 2013, National Taiwan University: Taiwan.
71. Wallinga, J. and P. Teunis, *Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures.* Am J Epidemiol, 2004. **160**(6): p. 509-16.
72. Choi BC, P.A., *A simple approximate mathematical model to predict the*

number of severe acute respiratory syndrome cases and deaths. J Epidemiol Community Health, 2003. **57**(10): p. 831-5.

73. Chowell, G., et al., *SARS outbreaks in Ontario, Hong Kong and Singapore: the role of diagnosis and isolation as a control mechanism.* J Theor Biol, 2003. **224**(1): p. 1-8.
74. Yeh, Y.P., et al., *Tuberculin reactivity in adults after 50 years of universal bacille Calmette-Guerin vaccination in Taiwan.* Trans R Soc Trop Med Hyg, 2005. **99**(7): p. 509-16.
75. Chan, P.C., L.M. Huang, and J. Suo, *It is time to deal with latent tuberculosis infection in Taiwan.* J Formos Med Assoc, 2009. **108**(12): p. 901-3.
76. Bowerman, R.J., *Community-wide INH treatment of latent TB infection in a BCG-vaccinated population: experience in rural Taiwan.* Int J Tuberc Lung Dis, 2007. **11**(4): p. 470-2.
77. Marciniuk, D.D., et al., *Detection of pulmonary tuberculosis in patients with a normal chest radiograph.* Chest, 1999. **115**(2): p. 445-52.
78. Harris, T.G., J. Sullivan Meissner, and D. Proops, *Delay in diagnosis leading to nosocomial transmission of tuberculosis at a New York City health care facility.* Am J Infect Control, 2013. **41**(2): p. 155-60.
79. Stead, W.W., *Tuberculosis among elderly persons: an outbreak in a nursing home.* Ann Intern Med, 1981. **94**(5): p. 606-10.
80. Centers for Disease, C., *Tuberculosis in a nursing care facility--Washington.* MMWR Morb Mortal Wkly Rep, 1983. **32**(9): p. 121-2, 128.
81. Ijaz, K., et al., *Unrecognized tuberculosis in a nursing home causing death with spread of tuberculosis to the community.* J Am Geriatr Soc, 2002. **50**(7): p. 1213-8.
82. Malone, J.L., et al., *Investigation of healthcare-associated transmission of Mycobacterium tuberculosis among patients with malignancies at three hospitals and at a residential facility.* Cancer, 2004. **101**(12): p. 2713-21.
83. National Tuberculosis Controllers, A., C. Centers for Disease, and Prevention, *Guidelines for the investigation of contacts of persons with infectious tuberculosis. Recommendations from the National Tuberculosis Controllers Association and CDC.* MMWR Recomm Rep, 2005. **54**(RR-15): p. 1-47.
84. Borgdorff, M.W., et al., *The incubation period distribution of tuberculosis estimated with a molecular epidemiological approach.* Int J Epidemiol, 2011. **40**(4): p. 964-70.
85. MA, S. and B. SM, *Uncertainty and sensitivity analysis of the basic reproductive rate. Tuberculosis as an example.* Am J Epidemiol, 1997. **145**(12): p. 1127-37.

86. Thrupp, L., et al., *Tuberculosis prevention and control in long-term-care facilities for older adults.* Infect Control Hosp Epidemiol, 2004. **25**(12): p. 1097-108.
87. *The role of BCG vaccine in the prevention and control of tuberculosis in the United States. A joint statement by the Advisory Council for the Elimination of Tuberculosis and the Advisory Committee on Immunization Practices.* MMWR Recomm Rep, 1996. **45**(RR-4): p. 1-18.
88. Campbell, J.R., et al., *A Systematic Review on TST and IGRA Tests Used for Diagnosis of LTBI in Immigrants.* Mol Diagn Ther, 2015.
89. Zwerling, A., et al., *Interferon-gamma release assays for tuberculosis screening of healthcare workers: a systematic review.* Thorax, 2012. **67**(1): p. 62-70.
90. Pai, M., et al., *T-cell assays for the diagnosis of latent tuberculosis infection: moving the research agenda forward.* Lancet Infect Dis, 2007. **7**(6): p. 428-38.
91. Dheda, K., et al., *T-cell interferon-gamma release assays for the rapid immunodiagnosis of tuberculosis: clinical utility in high-burden vs. low-burden settings.* Curr Opin Pulm Med, 2009. **15**(3): p. 188-200.
92. Diel, R., R. Loddenkemper, and A. Nienhaus, *Evidence-based comparison of commercial interferon-gamma release assays for detecting active TB: a metaanalysis.* Chest, 2010. **137**(4): p. 952-68.
93. Pai, M. and R. O'Brien, *Serial testing for tuberculosis: can we make sense of T cell assay conversions and reversions?* PLoS Med, 2007. **4**(6): p. e208.
94. Diel, R., et al., *Negative and positive predictive value of a whole-blood interferon-gamma release assay for developing active tuberculosis: an update.* Am J Respir Crit Care Med, 2011. **183**(1): p. 88-95.
95. Haldar, P., et al., *Single-step QuantiFERON screening of adult contacts: a prospective cohort study of tuberculosis risk.* Thorax, 2013. **68**(3): p. 240-6.
96. Raviglione, M.C., D.E. Snider, Jr., and A. Kochi, *Global epidemiology of tuberculosis. Morbidity and mortality of a worldwide epidemic.* JAMA, 1995. **273**(3): p. 220-6.
97. Kang, Y.A., et al., *Discrepancy between the tuberculin skin test and the whole-blood interferon gamma assay for the diagnosis of latent tuberculosis infection in an intermediate tuberculosis-burden country.* JAMA, 2005. **293**(22): p. 2756-61.
98. Bartu, V., M. Havelkova, and E. Kopecka, *QuantiFERON-TB Gold in the diagnosis of active tuberculosis.* J Int Med Res, 2008. **36**(3): p. 434-7.
99. Chee, C.B., et al., *Comparison of sensitivities of two commercial gamma interferon release assays for pulmonary tuberculosis.* J Clin Microbiol, 2008.

46(6): p. 1935-40.

100. Metcalfe, J.Z., et al., *Interferon-gamma release assays for active pulmonary tuberculosis diagnosis in adults in low- and middle-income countries: systematic review and meta-analysis.* J Infect Dis, 2011. **204 Suppl 4**: p. S1120-9.
101. Sester, M., et al., *Interferon-gamma release assays for the diagnosis of active tuberculosis: a systematic review and meta-analysis.* Eur Respir J, 2011. **37**(1): p. 100-11.
102. Prabhavathi, M., et al., *Role of QuantiFERON-TB Gold antigen-specific IL-1beta in diagnosis of active tuberculosis.* Med Microbiol Immunol, 2014.
103. Dosanjh, D.P., et al., *Improved diagnostic evaluation of suspected tuberculosis.* Ann Intern Med, 2008. **148**(5): p. 325-36.
104. Bauch, C.T., et al., *Dynamically modeling SARS and other newly emerging respiratory illnesses: past, present, and future.* Epidemiology, 2005. **16**(6): p. 791-801.
105. Choi, B.C. and A.W. Pak, *A simple approximate mathematical model to predict the number of severe acute respiratory syndrome cases and deaths.* J Epidemiol Community Health, 2003. **57**(10): p. 831-5.
106. Chowell, G., et al., *Model parameters and outbreak control for SARS.* Emerg Infect Dis, 2004. **10**(7): p. 1258-63.
107. Breban, R., J. Riou, and A. Fontanet, *Interhuman transmissibility of Middle East respiratory syndrome coronavirus: estimation of pandemic risk.* Lancet, 2013. **382**(9893): p. 694-9.
108. Reichler, M.R., et al., *Outbreak of paralytic poliomyelitis in a highly immunized population in Jordan.* J Infect Dis, 1997. **175 Suppl 1**: p. S62-70.
109. Sanchez, M.A. and S.M. Blower, *Uncertainty and sensitivity analysis of the basic reproductive rate. Tuberculosis as an example.* Am J Epidemiol, 1997. **145**(12): p. 1127-37.
110. WJ, E., et al., *The pre-vaccination epidemiology of measles, mumps and rubella in Europe: implications for modelling studies.* Epidemiol Infect, 2000. **125**(3): p. 635-50.
111. P., M., et al., *The pre-vaccination regional epidemiological landscape of measles in Italy: contact patterns, effort needed for eradication, and comparison with other regions of Europe.* Popul Health Metr., 2005. **3**(1): p. 1.
112. Ma, E., et al., *Estimation of the basic reproduction number of enterovirus 71 and coxsackievirus A16 in hand, foot, and mouth disease outbreaks.* Pediatr Infect Dis J, 2011. **30**(8): p. 675-9.
113. G, C., et al., *Estimation of the reproductive number of the Spanish flu epidemic*

in Geneva, Switzerland. Vaccine, 2006. **24**: p. 6747-50.
114. Mills, C.E., J.M. Robins, and M. Lipsitch, *Transmissibility of 1918 pandemic influenza.* Nature, 2004. **432**: p. 904–906.
115. E, M., et al., *The 1918 influenza A epidemic in the city of Sao Paulo, Brazil.* Med Hypotheses, 2007. **68**(2): p. 442-5.
116. Hsu, C.Y., et al., *Surveillance of influenza from household to community in Taiwan.* Trans R Soc Trop Med Hyg, 2014. **108**(4): p. 213-20.
117. PD, G., M. A, and E. VC, *Encouraging prospects for immunization against primary cytomegalovirus infection.* Vaccine, 2001. **19**(11-12): p. 1356-62.
118. PG, C., et al., *Mathematical models of Haemophilus influenzae type b.* Epidemiol Infect., 1998. **120**(3): p. 281-95.
119. MN., K. and F. CP., *Matrix models for childhood infections: a Bayesian approach with applications to rubella and mumps.* Epidemiol Infect., 2005. **133**(6): p. 1009-21.
120. E, M., et al., *Dengue and the risk of urban yellow fever reintroduction in Sao Paulo State, Brazil.* Rev Saude Publica, 2003. **37**(4): p. 477-84.
121. G, C., et al., *The basic reproductive number of Ebola and the effects of public health measures: the cases of Congo and Uganda.* J Theor Biol., 2004. **229**(1): p. 119-26.
122. DL, S., et al., *Revisiting the Basic Reproductive Number for Malaria and Its Implications for Malaria Control.* PLoS Biol, 2007. **5**(3): p. e42.
123. G, D., *Further studies of the basic factors concerned in the transmission of malaria.* Trans R Soc Trop Med Hyg., 1955. **49**(4): p. 339-50.
124. G, D. and D. CC, *Field studies on some of the basic factors concerned in the transmission of malaria.* Trans R Soc Trop Med Hyg., 1953. **47**: p. 522–535.
125. TR, B., et al., *Human malaria transmission studies in the Anopheles punctulatus complex in Papua New Guinea: sporozoite rates, inoculation rates, and sporozoite densities.* Am J Trop Med Hyg, 1988. **39**(2): p. 135-44.
126. R, H., et al., *Malaria and its possible control on the island of Principe.* Malar J, 2003. **2**: p. 15.
127. ME, W., et al., *Heterogeneities in the transmission of infectious agents: implications for the design of control programs.* Proc Natl Acad Sci U S A, 1997. **94**(1): p. 338-42.
128. JJ, P., et al., *Chlamydia transmission: concurrency, reproduction number, and the epidemic trajectory.* Am J Epidemiol, 1999. **150**(12): p. 1331-9.
129. RH, G., et al., *Stochastic simulation of the impact of antiretroviral therapy and HIV vaccines on HIV transmission; Rakai, Uganda.* AIDS, 2003. **17**(13): p. 1941–1951.

130. N, B., A. X, and Y. J, *Modeling the HIV/AIDS epidemic among injecting drug users and sex workers in Kunming, China.* Bull Math Biol, 2006. **68**(3): p. 525-50.
131. F, F., *Nature, nurture and my experience with smallpox eradication.* The Medical Journal of Australia, 1999. **171**: p. 638-41.

Table2.1Reproduction number for infectious disease

Infectious disease	Area	Epidemic period	R_0	References
SARS	Toronto	2003	0.86,1.2,2,2.5	[73, 104-106]
	Hong Kong	2003	1.2,1.7,2.1,2.7,3.4,3.6	
	Singapore	2003	1.1	
	Beijing	2003	1.1, 3.3	
	Taiwan	2003	4.2	
	All locations	2003	3	
MERS	Saudi Arabia	2013	0·60 (0·42–0·80)	[107]
Poliomyelitis	USA	1955	5-6	[2]
	Netherlands	1960	6-7	
	Senegal	1981	18	[108]
	Dominican	1963	19	
	French Morocco	1953	25	
	Burma	1979	30	
	Taiwan	1982	2.19-3.15	[3]
Mycobacterium tuberculosis	Worldwide	1990s	1.1-31.26 < 8.93（90%）； >1（>99%）	[109]
Measles	England	1947-50	13-14	[2]
	England and Wales	1950-68	16-18	
	USA	1918-21	5-6	
	Italy	1964-79	6.1	[110]
	Denmark	1983	9.7	
	England and Wales	1956-65	10.2	
	Canada	1912-13	11-12	[2]
	England	1912-13	11-12	
	Ghana	1960-8	14-15	
	Eastern Nigeria	1960-8	16-17	
	Italy：Homogeneous mixing	1949-1976	10-14.5	[111]
	South		13-20	
	Centre		8.5-12	
HFMD	Taiwan	2000-2008	1.37 (0.23~5.71)	[3]
	Hong Kong	2004-2009	5.48(4.20-5.61)EV71 2.6 (1.963.67) Cox A16	[112]

SARS: Severe Acute Respiratory Syndrome; MERS: Middle East Respiratory Syndrome; HFMD: Hand Foot Mouth Disease

Table 2.1. Reproduction number for infectious disease (continue)

Infectious disease	Area	Epidemic period	R_0	References
Influenza	Switzerland	1918（1st wave）	1.45-1.53	[113]
		（2nd wave）	3.57-3.93	
	USA	1918	2.9	[114]
	Brazil	1918	2.68	[115]
	New Zealand	1918	1.3-3.1	[16]
	Taiwan	2001-2001	2.56	[116]
CMV	London	1975-82	2.4	[117]
		1983-85	2.7	
Haemophilus influenzae type b	England and Wales	1993	3.3	[118]
Diphtheria	USA	1918-19	4-5	
	USA	1908-17	4-5	
Rubella	Finland	1979	3.4	[110]
	England and Wales	1986-7	3.7	
	Denmark	1983	4.2	
	Italy	1970-81	4.2	
	UK	1986	6.1(4.3-9.2）	[119]
	Netherlands	1958-74	6.4	[110]
	East Germany	1978-89	7.8	
	England and Wales	1960-70	6-7	[2]
	West Germany	1970-7	6-7	
	Czechoslovakia	1970-7	8-9	
	Poland	1970-7	11-12	
	Gambia	1976	15-16	
Varicella	USA	1913-17	7-8	[2]
	USA	1912-21	7-8	
	USA	1943	10-11	
	England and Wales	1944-68	10-12	

Table 2.1. Reproduction number for infectious disease (continue)

Infectious disease	Area	Epidemic period	R_0	References
Dengue fever	Brazil	2001	2.74-11.57	[120]
Yellow fever	Brazil	2001	1.57-6.61	[120]
Scarlet fever	USA	1908-17	7-8	[2]
	USA	1908-17	5-6	
	USA	1918-19	6-7	
Ebola	Congo	1995	1.83	[121]
	Uganda	2000	1.34	
Malaria	African	2000-3	1-3000	[122-127]
	Northwest Tanzania	1950s	1600	
	Central Uganda	1950s	2000-5000	
	Papua New Guinea	1980s	>500	
			7 （age seroprevalence）	
	the island of Principe, Gulf of Guinea	1999	1.6	
	Papua New Guinea	-	3.89	
	Tanzania	1990	3.7	
Schistosoma haematobium	Zimbabwe	-	3.02	[127]
Schistosoma spp.	Mali	-	2.9	[127]
Bacterial STDs	France	-	12.01	[127]
Chlamydia	Colorado	1996-7	0.55	[128]
			46.5% R=0;	
			21.3% R<1;	
			5.4% R≧2.0	
Pertussis	England and Wales	1944-78	16-18	[2]
	USA	1943	16-17	
	Canada	1912-13	10-11	

Table 2.1. Reproduction number for infectious disease (continue)

Infectious disease	Area	Epidemic period	R_0	References
HIV type I	Uganda（ART）	1994-98	1.44	[129]
	England and Wales（male homosexuals）	1981-5	2-5	[2]
	Kenya （female prostitutes）	1981-5	11-12	
	Uganda （heterosexuals）	1985-7	10-11	
	UK		13.82	[127]
	Kunming,China（IDUs）	1994-2003	32	[130]
Smallpox	Boston, USA	1721	4.3	[131]
	Burford, England	1758	3.4	
	Chester, England	1774	5.8	
	Warrington, England	1773	4.7（4.0-5.3）	
	Paris, France	1766	> 4-5	
	London	1836-70	> 5	
	Kosovo	1972	10.8	
	Europe	1958-73	10-12	
Mumps	Denmark	1983	3.6	[110]
	East Germany	1968-72 1983-90	4.0	
	Italy	1964-81	4.2	
	Netherlands	1970s	4.3	
	England and Wales	1986-87	4.5	
	USA	1943	7-8	[2]
	England and Wales	1960-80	11-14	
	Netherlands	1970-80	11-14	
	UK	1986	19.3（4-31.5）	[119]

HIV: Human Immunodeficiency Virus
ART：antiretroviral therapy
IDUs：injecting drug users

Table3.1The parameters of power series distribution.

	θ	a(k)	$A(\theta)$	$A'(\theta)$	pmf	pgf
Geometric distribution G(π)	$1-\pi$	1	$\frac{1}{1-\theta}$	$(1-\theta)^{-2}$	$\pi(1-\pi)^k$	$\frac{\pi}{1-(1-\pi)s}$
Binomial distribution Bin(n,π)	$\frac{\pi}{1-\pi}$	$\binom{n}{k}$	$(1+\theta)^n$	$n(1+\theta)^{n-1}$	$\binom{n}{k}\pi^k(1-\pi)^{n-k}$	$(1-\pi+\pi s)^n$
negative binomial distribution NB(r,π)	$1-\pi$	$\binom{r+k-1}{k}$	$(1-\theta)^{-r}$	$r(1-\theta)^{-r-1}$	$\binom{r+k-1}{k}\pi^r(1-\pi)^k$	$(\frac{1}{1-(1-\pi)s})^r$
poisson distribution P(λ)	λ	$\frac{1}{k!}$	e^{θ}	e^{θ}	$\frac{e^{-\lambda}\lambda^k}{k!}$	$e^{\lambda(s-1)}$

pmf: probability mass function; pgf: probability generating function

Table 5.1Data generated with Poisson distribution

R_0	X_0	X_1	X_2	X_3	X_4	X_5	X_6	Total
2	1	2 (2)	3 (1,2)	3 (1,2,0)	3 (1,1,1)	5 (1,1,3)	19 (5,3,6,3,2)	36
1.5	1	2 (2)	3 (1,2)	5 (1,1,3)	8 (2,1,2,0,3)	8 (0,2,1,3,0,2,0,0)	13 (3,0,1,2,2,3,1,1)	40
1.1	1	2 (2)	2 (0,2)	2 (1,1)	2 (2)	1 (0,1)	0 (0)	10
0.9	1	2 (2)	1 (1,0)	1 (1)	0 (0)	-	-	5
2	5	10 (0,3,3,2,2)	14 (0,0,2, 5,1,2,0, 2,1,1)	20 (3,1,1,1, 2,2,0,2,1, 0,1,3,3,0)	37 (2,1,3,4, 0,0,1,2, 3,2,1,4, 1,3,1,1, 3,1,1,3)	90 (3,0,3,3,2,1,0 1,1,3,8,0,1,3,2, 1,3,4,1,0,4,6,4, 1,2,1,3,1,1,3,0, 4,7,4,1,5,3)	196 (2,1,2,3,1,3,2,3,2,1, 1,1,3,2,1,2,4,0,5,3, 2,4,2,2,4,0,2,1,1,3, 5,2,2,1,0,2,1,0,2,3, 2,1,2,1,0,2,0,4,2,2, 0,1,2,2,3,2,0,1,2,3, 0,3,1,1,4,2,4,1,1,3, 0,2,4,2,4,3,5,5,2,4, 4,3,1,4,2,2,7,3,3,3,)	372
1.5	5	11 (3,2,1,3,2)	13 (1,0,2,1,2, 2,1,0,1,1, 2)	19 (0,2,1,0,1, 2,1,2,2,0, 4,2,2)	28 (2,0,2,1,2, 1,2,2,1,2, 1,3,0,1,1, 2,2,1,2)	41 (2,0,2,1,0,0,1,3,3,0, 2,2,1,2,0,2,2,2,2,0, 0,2,4,0,4,0,2,2)	62 (1,4,2,1,1,2,0,2,3,2, 2,0,3,1,0,2,4,2,1,2, 0,2,2,1,1,2,0,2,1,1, 3,2,0,0,3,0,2,1,2,1,1)	179
1.1	5	4 (1,1,2,0,0)	3 (3,0,0,0)	5 (1,4,0)	7 (2,1,1,2,1)	6 (0,2,1,0,0,3,0)	9 (0,4,0,1,0,4)	39
0.9	5	6 (2,0,3,0,1)	7 (2,2,1,0,0, 2)	5 (0,0,0,0,3, 2,0,)	3 (2,0,0,1,0)	3 (0,2,1)	3 (0,2,1)	32

R_0 :the basic reproductive number; X_i : the total number of the *i*-th generation offspring.

Table 5.2 Estimation for simulated data with Poisson distribution

R_0	X_0	Method / Estimate	Non-parametric method	Nbin	Poisson	Bin (n=1000)	Bin (n=100)	Geo
2	1	$\hat{R}$	2.0588 (1.1286~2.9890)	2.0588 (1.3461~2.7716)	2.0588 (1.4340~2.8633)	2.0588 (1.9700~2.1477)	2.0588 (1.7805~2.3371)	2.0588 0.8659~3.2518)
		$\hat{q}$	0.104	0.2069	0.1879	0.1874	0.1836	0.4857
1.5	1	$\hat{R}$	1.4444 (1.2046~1.6842)	-*	1.4444 (1.0271~1.9746)	1.4444 (1.3700~1.5189)	1.4444 (1.2106~1.6783)	1.4444 (0.7357~2.1532)
		$\hat{q}$	0.4000		0.4553	0.4552	0.4512	0.6923
1.1	1	$\hat{R}$	1.0000 (0.5999~1.4000)	-*	1.0000 (0.4573~1.8983)	1.0000 (0.9381~1.0619)	1.0000 (0.8050~1.1950)	1.0000 (0.0760~1.9240)
		$\hat{q}$	1		1	1	1	1
0.9	1	$\hat{R}$	0.8000 (0.1353~1.4647)	-*	0.8000 (0.2180~2.0483)	0.8000 (0.7446~0.8554)	0.8000 (0.6254~0.9746)	0.8000 (-0.2518~1.8518)
		$\hat{q}$	1		1	1	1	1
2	5	$\hat{R}$	2.0852 (1.8358~2.3347)	2.0852 (1.8618~2.3087)	2.0852 (1.8719~2.2986)	2.0852 (1.9958~2.1746)	2.0852 (1.8052~2.3653)	2.0852 (1.7105~2.4600)
		$\hat{q}$	0.2003	0.201	0.1814	0.1810	0.1772	0.4795
1.5	5	$\hat{R}$	1.4872 (1.3465~1.6279)	-*	1.4872 (1.2662~1.7082)	1.4872 (1.4117~1.5627)	1.4872 (1.2499~1.7244)	1.4872 (1.1387~1.8357)
		$\hat{q}$	0.3550		0.4256	0.4252	0.4213	0.6725
1.1	5	$\hat{R}$	1.1333 (0.8526~1.4141)	1.1333 (0.6413~1.6253)	1.1333 (0.7849~1.5837)	1.1333 (1.0674~1.1993)	1.1333 (0.9259~1.3408)	1.1333 (0.5769~1.6898)
		$\hat{q}$	0.8456	0.8538	0.7744	0.7742	0.7723	0.8825
0.9	5	$\hat{R}$	0.9310 (0.7431~1.1190)	0.9310 (0.5488~1.3133)	0.9310 (0.6136~1.3546)	0.9310 (0.8713~0.9908)	0.9310 (0.7428~1.1193)	0.9310 (0.4430~1.4191)
		$\hat{q}$	1	1	1	1	1	1

Nbin: Negative Binomial distribution; Bin: Binomial distribution; Geo: Geometric distribution; $\hat{R}$: estimated reproductive number; $\hat{q}$: estimated extinction probability.

*: No results because it cannot be estimated

Table 5.3Data generated with Binomial distribution (n=1000)

R_0	X_0	X_1	X_2	X_3	X_4	X_5	X_6	Total
2	1	5 (5)	6 (0,1,2,1,2)	11 (1,2,2,2,1 3)	23 (2,1,3,1,1, 2,3,4,2,0, 4)	44 (0,1,6,2,0,4,0,2,1,0, 0,1,2,0,2,4,5,2,2,2, 2,3,3)	93 (3,1,5,1,2,1,1,3,7,1, 3,3,3,1,1,2,1,1,3,1, 2,3,0,3,1,3,1,3,0,6, 2,3,2,4,1,3,2,1,1,3, 2,2,1,0)	183
1.5	1	2 (2)	5 (3,2)	7 (1,2,0,4,0)	14 (5,3,2,1,2,0,1)	18 (1,1,2,0,3,1,1,2,2,0, 1,1,1,2)	24 (1,1,2,1,3,1,2,2,2,2, 0,1,0,1,0,0,2,3,)	71
1.1	1	1 (1)	2 (2)	3 (1,2)	2 (1,0,1)	1 (0,1)	3 (3)	13
0.9	1	1 (1)	2 (2)	1 (0,1)	1 (1)	2 (2)	0 (0,0)	8
2	5	8 (2,3,1,2,0)	17 (3,1,1,2,3, 2,1,4)	32 (1,5,1,0,5, 2,1,0,3,1, 3,3,1,3,1, 1,1)	71 (2,4,2,3,2, 6,4,2,2,2, 5,1,2,2,3, 2,2,1,1,1, 0,1,2,3,1, 1,0,6,3,1, 1,3)	139 (0,4,4,2,4,3,0,1,1,1, 4,3,0,1,2,2,3,2,3,3, 3,1,2,3,2,3,2,3,3,1, 2,3,4,2,3,3,0,1,2,0, 0,1,4,2,1,1,2,4,3,1, 1,0,1,2,2,2,1,1,0,2, 1,2,3,2,1,4,0,2,4,2,1)	286 (4,2,2,2,1,1,2,4,4,3, 1,0,1,2,1,1,2,1,3,1, 4,1,1,0,2,0,3,4,1,0, 2,4,3,1,1,1,1,4,4,2, 1,2,2,3,2,4,2,1,4,4, 3,3,4,0,2,2,4,1,2,0, 1,1,1,2,1,0,0,6,2,4,*)	594
1.5	5	9 (2,2,1,0,4)	13 (3,0,5,1,1, 0,1,1,1)	27 (2,1,0,4,2, 1,3,2,2,4, 2,2,2)	30 (1,0,1,0,0, 1,2,1,0,0, 1,0,1,2,0, 1,3,1,0,3, 5,2,2,2,1, 0,0)	39 (1,2,2,2,0,5,2,2,0,1, 1,1,1,0,5,1,3,3,3,1, 0,1,0,0,0,0,0,3,1,0)	61 (1,2,2,0,0,2,1,2,1,1, 1,1,0,2,1,0,2,4,0,3, 1,0,4,1,2,1,2,1,2,1, 3,2,1,1,3,2,4,2,2)	184
1.1	5	7 (1,1,0,3,2)	4 (0,1,0,0,2, 1,0)	8 (1,4,2,1)	6 (2,1,0,1,0, 2,0,0)	5 (0,3,1,0,0,1)	7 (3,2,1,1,0)	42
0.9	5	6 (4,1,1,0,0)	5 (1,1,0,1,0,2)	6 (0,2,1,2,1)	4 (2,2,0,0,0,0)	2 (0,0,1,1)	1 (0,1)	29

R_0 :the basic reproductive number; X_i : the total number of the *i*-th generation offspring.
*(...,2,2,3,1,3,1,3,0,4,2,5,4,2,1,2,4,1,2,1,2,4,3,1,1,3,1,2,5,5,1,1,3,3,4,2,0,2,2,2,2,2,1,4,1,1,4,2,3,2,2,2,1,3, 1,0,2,3,1,5,3,0,2,0,0,2,1,2,2)

Table 5.4Estimation for simulated data with Binomial distribution (n=1000)

R_0	X_0	Method / Estimate	Non-parametric method	Nbin	Poisson	Bin (n=1000)	Bin (n=100)	Geo
2	1	$\hat{R}$	2.0222 (1.7163~2.3281)	2.0222 (1.7211~2.3233)	2.0222 (1.7284~2.3160)	2.0222 (1.9342~2.1103)	2.0222 (1.7463~2.2981)	2.0222 (1.5115~2.5330)
		$\hat{q}$	0.1869	0.2075	0.1972	0.1968	0.1929	0.4945
1.5	1	$\hat{R}$	1.4894 (1.2235~1.7552)	-*	1.4894 (1.1610~1.8817)	1.4894 (1.4138~1.5649)	1.4894 (1.2520~1.7268)	1.4894 (0.9389~2.0399)
		$\hat{q}$	0.3622		0.4241	0.4237	0.4199	0.6714
1.1	1	$\hat{R}$	1.2000 (0.5836~1.8164)	-*	1.2000 (0.6201~2.0962)	1.2000 (1.1321~1.2679)	1.2000 (0.9866~1.4134)	1.2000 (0.1929~2.2071)
		$\hat{q}$	0.5616		0.6863	0.6860	0.6835	0.8332
0.9	1	$\hat{R}$	0.8750 (0.2843~1.4657)	-*	0.8750 (0.3518~1.8028)	0.8750 (0.8170~0.9330)	0.8750 (0.6925~1.0575)	0.8750 (-0.0126~1.7626)
		$\hat{q}$	1		1	1	1	1
2	5	$\hat{R}$	2.0331 (1.9493~2.1169)	-*	2.0331 (1.8636~2.2025)	2.0331 (1.9448~2.1214)	2.0331 (1.7565~2.3097)	2.0331 (1.7380~2.3282)
		$\hat{q}$	0.1557		0.1944	0.1940	0.1901	0.4919
1.5	5	$\hat{R}$	1.4553 (1.2271~1.6835)	1.4553 (1.2330~1.6775)	1.4553 (1.2421~1.6685)	1.4553 (1.3806~1.5300)	1.4553 (1.2206~1.6900)	1.4553 (1.1212~1.7893)
		$\hat{q}$	0.4688	0.4716	0.4475	0.4471	0.4434	0.6872
1.1	5	$\hat{R}$	1.0571 (0.6883~1.4260)	1.0571 (0.7017~1.4126)	1.0571 (0.7443~1.4571)	1.1 (0.9934~1.1208)	1.0571 (0.8567~1.2576)	1.0571 (0.5686~1.5457)
		$\hat{q}$	0.9016	0.9020	0.8939	0.8237	0.8929	0.9459
0.9	5	$\hat{R}$	0.8571 (0.6348~1.0795)	0.8571 (0.5049~1.2094)	0.8571 (0.5492~1.2754)	0.8571 (0.7998~0.9145)	0.8571 (0.6765~1.0378)	0.8571 (0.3898~1.3245)
		$\hat{q}$	1	1	1	1	1	1

Nbin: Negative Binomial distribution; Bin: Binomial distribution; Geo: Geometric distribution; $\hat{R}$: estimated reproductive number; $\hat{q}$: estimated extinction probability

*: No results because it cannot be estimated

Table 5.5Data generated with Binomial distribution (n=100)

R_0	X_0	X_1	X_2	X_3	X_4	X_5	X_6	Total
2	1	2	1	2	5	10	22	43
		(2)	(1,0)	(2)	(4,1)	(1,2,0,3,4)	(2,4,5,0,1,2,3,1,2,2)	
1.5	1	1	3	8	10	21	24	68
		(1)	(3)	(3,3,2)	(1,1,0,2,1, 1,2,2)	(3,2,0,3,3,3,2,3,1,1)	(1,1,1,1,0,2,1,0,0,1, 0,0,0,1,3,1,4,2,0,4,1)	
1.1	1	3	4	1	1	3	2	15
		(3)	(1,1,2)	(0,0,0,1)	(1)	(3)	(1,0,1)	
0.9	1	1	1	2	1	1	0	7
		(1)	(1)	(2)	(1,0)	(1)	(0)	
2	5	15	32	58	110	214	435	869
		(2,1,2,4,6)	(2,4,1,1,2, 1,3,0,4,3, 3,2,0,4,2, 2,1,1,1,1, 2,2,1,3,4)	(4,3,1,3,2, 2,2,1,2,2, 1,1,2,4,0, 4,1,0,2,1, 1,1,2,2,0, 1,1,3,2,1, 0,2)	(*)	(#)	($)	
1.5	5	6	11	11	16	26	43	118
		(3,1,1,1,0)	(1,0,1,4,3, 2)	(1,2,1,0,0, 3,0,0,1,2,1)	(0,2,1,3,0, 5,1,2,0,2, 0)	(2,1,2,2,0,3,2,3,1,0, 1,1,1,0,3,4)	(4,2,2,1,3,0,0,0,4,3, 1,2,0,3,3,3,3,1,1,1, 2,1,1,0,1,1)	
1.1	5	6	5	6	7	8	9	46
		(1,0,2,0,3)	(1,2,1,0,0,1)	(1,3,0,0,2)	(0,1,1,2,1, 2)	(0,2,2,2,0,1,1)	(1,1,2,2,2,0,1,0)	
0.9	5	7	8	7	2	1	2	32
		(1,2,2,0,2)	(1,1,3,1,0, 1,1,0)	(2,1,1,0,2, 1,0)	(1,0,0,1,0, 0,0)	(0,1)	(2)	

R_0 :the basic reproductive number; X_i : the total number of the *i*-th generation offspring.

*(3,3,2,2,1,1,2,2,1,1,3,3,0,2,3,3,3,3,5,0,3,2,2,0,1,2,1,2,1,2,2,2,2,2,2,0,1,3,1,3,2,2,3,4,5,3,0,0,2,2,2,3,3,2,1,3,0,1,4)

#(2,0,1,2,2,1,3,1,3,0,1,1,2,0,1,1,2,3,4,1,1,3,2,2,3,2,0,0,4,5,1,2,1,1,3,1,2,1,2,0,1,4,2,3,1,0,3,0,1,2,2,0,4,0,1,1,4,3,4,1, 2,1,1,4,3,5,1,1,4,5,2,2,1,3,1,3,3,3,5,0,0,2,4,4,3,1,1,1,2,3,1,3,0,1,3,0,2,4,3,1,0,2,3,1,2,4,5,4,2,3)

$(3,1,1,3,2,1,2,1,1,3,1,2,2,5,2,4,1,1,3,2,4,1,2,3,2,2,0,2,2,1,1,2,3,0,2,3,2,0,2,1,3,4,3,4,2,2,2,5,2,2,1,6,2,2,4,3,0,2,0,1,1,0, 3,1,2,0,1,1,1,1,1,3,1,0,0,0,2,1,1,3,5,2,2,3,3,4,2,6,1,2,4,1,2,1,0,5,3,1,2,2,0,2,2,2,5,0,0,2,0,1,3,3,2,4,0,2,1,2,2,2,2,1,2,3,2, 2,0,0,0,1,1,2,5,1,5,2,4,3,3,2,0,0,3,3,2,2,1,4,3,2,0,2,2,0,2,2,1,2,3,1,5,2,3,3,0,0,0,1,4,1,3,2,3,2,0,4,1,0,4,4,1,1,2,3,0,2,0,2, 2,1,0,4,2,6,3,1,2,4,6,5,0,2,4,3)

Table 5.6 Estimation for simulated data with Binomial distribution (n=100)

R_0	X_0	Method / Estimate	Non-parametric method	NBin	Poisson	Bin (n=1000)	Bin (n=100)	Geo
2	1	$\hat{R}$	2.0000 (1.5942~2.4058)	-*	2.0000 (1.4414~2.7034)	2.0 (1.9124~2.0876)	2.0000 (1.7256~2.2744)	2.0000 (0.9523~3.0477)
		$\hat{q}$	0.2079		0.2032	0.2028	0.1988	0.4999
1.5	1	$\hat{R}$	1.5227 (1.0867~1.9588)	-*	1.5227 (1.1801~1.9338)	1.5227 (1.4463~1.5992)	1.5227 (1.2827~1.7627)	1.5227 (0.9436~2.1019)
		$\hat{q}$	0.3829		0.4028	0.4023	0.3984	0.6567
1.1	1	$\hat{R}$	1.0769 (0.4839~1.6700)	-*	1.0769 (0.5888~1.8069)	1.0769 (1.0126~1.1412)	1.0769 (0.8746~1.2792)	1.0769 (0. 2639~1.8899)
		$\hat{q}$	0.8508		0.8607	0.8605	0.8593	0.9286
0.9	1	$\hat{R}$	0.8571 (0.3948~1.3195)	-*	0.8571 (0.3146~1.8656)	0.8571 (0.7998~0.9145)	0.8571 (0.6765~1.0378)	0.8571 (-0.0775~1.7918)
		$\hat{q}$	1		1	1	1	1
2	5	$\hat{R}$	1.9908 (1.9308~2.0508)	-*	1.9908 (1.8580~2.1235)	1.9908 (1.9034~2.0781)	1.9908 (1.7170~2.2646)	1.9908 (1.7612~2.2204)
		$\hat{q}$	0.2024		0.2057	0.2053	0.2013	0.5024
1.5	5	$\hat{R}$	1.5067 (1.3155~1.6978)	1.5067 (1.2235~1.7898)	1.5067 (1.2289~1.7845)	1.5067 (1.4306~1.5827)	1.5067 (1.2679~1.7454)	1.5067 (1.0668~1.9465)
		$\hat{q}$	0.4394	0.4234	0.4129	0.4124	0.4086	0.6636
1.1	5	$\hat{R}$	1.1081 (1.0127~1.2036)	-*	1.1081 (0.7952~1.5033)	1.1081 (1.0429~1.1733)	1.1081 (0.9029~1.3133)	1.1081 (0.6156~1.6006)
		$\hat{q}$	0.7578		0.8115	0.8113	0.8097	0.9026
0.9	5	$\hat{R}$	0.9000 (0.5870~1.2130)	-*	0.9000 (0.5931~1.3095)	0.9000 (0.8412~0.9588)	0.9000 (0.7149~1.0851)	0.9000 (0.4321~1.3679)
		$\hat{q}$	1		1	1	1	1

Nbin: Negative Binomial distribution; Bin: Binomial distribution; Geo: Geometric distribution; $\hat{R}$: estimated reproductive number; $\hat{q}$: estimated extinction probability

*: No results because it cannot be estimated

Table 5.7 Data generated with Geometric distribution

R_0	X_0	X_1	X_2	X_3	X_4	X_5	X_6	Total
2	1	4 (4)	12 (6,2,2,2)	31 (0,4,0,3,0, 8,3,0,1,0, 3,9)	66 (0,4,2,1,2,0,7,1,2,0, 0,1,2,3,4,5,5,3,3,1, 2,0,2,1,2,0,1,5,2,2, 3)	117 (4,1,2,2,2,2,1,3,3,5, 1,2,0,0,2,3,1,0,0,1, 2,1,3,2,1,3,6,0,2,1, 8,0,3,0,1,0,0,1,3,1, 0,2,0,2,0,0,0,1,0,0, 1,1,0,4,0,6,11,4,2,0, 0,1, 1,1,7,0)	241 (1,0,1,2,1,2,3,4,5,2, 7,2,0,4,5,3,1,1,9,7, 2,4,1,0,3,3,0,4,3,1, 0,0,0,1,0,0,2,8,1,0, 1,1,0,5,0,3,0,0,5,0, 1,1,0,0,6,1,0,8,3,1, 4,5,1,1,3,1,0,1,0,4, 3,2,5,1,4,0,0,0,5,3, 1,5,1,11,1,1,0,1,1,0, 2,3,4,0,0,4,0,0,0,3, 1,2,0,1,1,0,0,7,1,1, 1,11,0,0,0,4,1)	472
1.5	1	3 (3)	4 (2,0,2)	4 (1,1,2,0)	9 (5,1,3,0)	11 (0,1,7,0,0,1,0,1,1)	15 (0,0,2,0,3,4,0,0,5,1,0)	47
1.1	1	3 (3)	2 (1,0,1)	4 (4,0)	4 (1,0,2,1)	3 (0,0,0,3)	3 (3,0,0)	20
0.9	1	4 (4)	4 (3,0,1,0)	3 (0,0,0,3)	3 (1,1,1)	1 (0,0,1)	0 (0)	16
2	5	9 (1,5,1,2,0)	22 (3,4,4,1,0, 8,1,1,0)	58 (2,6,1,3,0, 1,14,3,1,2, 4,0,4,0,0, 4,1,6,2,3,1,0)	85 (3,1,1,0,3,3,1,4,1,0, 4,0,0,0,0,1,1,2,4,1, 2,3,0,0,2,1,2,0,0,0, 1,0,0,4,0,0,0,1,3,0, 0,1,4,6,1,1,0,6,1,0, 7,3,0,3,0,2,1,0)	170 (2,4,0,4,3,0,4,0,2,0, 4,0,6,1,4,0,2,2,3,0, 1,0,0,3,3,9,1,1,6,3, 0,0,0,0,0,5,8,1,4,3, 0,1,0,0,6,0,5,1,0,2, 0,2,3,0,0,1,3,1,2,2, 2,5,3,0,2,4,3,0,5,3, 7,4,0,1,2,5,1,1,2,0, 0,0,0,2,0,)	354 (2,1,0,4,0,12,2,0,0,0, 5,4,0,3,0,1,0,0,0,0, 1,3,1,4,3,0,0,0,2,2, 1,2,0,3,4,0,0,4,0,3, 1,0,4,1,12,2,9,6,1,0, 1,0,2,1,11,1,0,5,0,1, 0,0,6,5,2,0,5,6,0,1, 0,2,1,0,3,2,1,0,2,3, 1,1,1,0,0,7,1,1,0,1,*)	703
1.5	5	8 (1,3,1,1,2)	15 (2,0,1,1,1, 1,0,9)	34 (1,4,0,4,3, 7,1,1,0,4, 0,1,1,1,6,)	55 (2,1,1,0,0,3,0,1,0,2, 3,2,4,3,0,0,3,4,1,1, 0,5,0,3,0,5,0,2,0,1, 1,3,2,2)	84 (3,0,1,0,3,1,1,1,1,0, 0,4,0,1,1,4,1,0,2,0, 2,0,2,0,6,0,3,1,4,0, 1,10,0,2,0,0,0,0,1,2, 3,4,0,2,0,1,1,0,5,0, 2,6,0,2,0)	100 (0,0,2,2,0,0,0,0,0,4, 1,0,1,0,0,0,2,1,0,1, 0,2,3,0,0,1,1,0,0,1, 0,4,1,0,3,6,0,0,1,0, 2,0,3,9,0,4,0,0,6,3, 0,2,0,1,1,1,6,2,0,0, 3,0,0,0,0,0,7,0,0,0, 2,0,3,0,0,0,1,0,0,0, 4,0,0,3)	301
1.1	5	6 (2,2,1,1,0)	13 (3,1,3,1,2,3)	12 (1,0,0,3,0, 0,1,0,1,4,1,0,1)	21 (1,3,0,2,3,0,0,4,1,1, 2,4)	23 (0,1,1,0,0,1,2,1,0,0, 0,3,0,0,0,6,2,3,2,0,1)	14 (4,1,0,0,0,4,0,0,0,0, 1,0,1,1,1,0,0,0,0,0,0,1,0)	94
0.9	5	5 (0,5,0,0,0)	6 (2,0,3,0,1)	12 (3,1,1,0,0,7)	9 (1,4,1,2,1,0,0,0,0,0,0,0)	7 (2,0,4,0,0,0,1,0,0)	1 (0,0,0,0,1,0,0)	45

R_0 :the basic reproductive number; X_i : the total number of the *i*-th generation offspring.

*(1,2,0,2,2,0,0,0,0,0,4,5,0,0,0,4,0,0,0,12,1,1,0,2,0,1,0,0,7,3,2,6,3,1,0,4,0,3,2,0,1,2,1,0,2,0,2,2,5,0,4,0,10,2,5,5,6,2,4,1,4,6,2,1,2,2,0,1,4,0,0,0,0,3,3,2,1,0,3,21)

Table 5.8 Estimation for simulated data with Geometric distribution

R_0	X_0	Method / Estimate	Non-parametric method	NBin	Poisson	Bin (n=1000)	Bin (n=100)	Geo
2	1	$\widehat{R}$	2.0390 (1.8279~2.2500)	2.0390 (1.7412~2.3367)	2.0390 (1.8548~2.2231)	2.0390 (1.9505~2.1274)	2.0390 (1.7620~2.3160	2.0390 (1.7180~2.3600)
		q	0.4477	0.4466	0.1929	0.1924	0.1886	0.4905
1.5	1	$\widehat{R}$	1.4375 (1.0809~1.7942)	1.4375 (0.8005~2.0745)	1.4375 (1.0524~1.9174)	1.4375 (1.3632~1.5118)	1.4375 (1.2042~1.6708)	1.4375 (0.7889~2.0861)
		q	0.6844	0.6875	0.4604	0.4600	0.4563	0.6958
1.1	1	$\widehat{R}$	1.1176 (0.6287~1.6066)	1.1176 (0.4312~1.8040)	1.1176 (0.6729~1.7453)	1.1176 (1.0522~1.1831)	1.1176 (0.9116~1.3237(	1.1176 (0.3863~1,8490)
		q	0.8675	0.8820	0.7972	0.7970	0.7953	0.8947
0.9	1	$\widehat{R}$	0.9375 (0.2585~1.6165)	0.9375 (0.2992~1.5758)	0.9375 (0.5247~1.5463)	0.9375 (0.8775~0.9975)	0.9375 (0.7486~1.1264)	0.9375 (0.2771~1.5979)
		q	1	1	1	1	1	1
2	5	$\widehat{R}$	2.0000 (1.7709~2.2291)	2.0000 (1.7359~2.2641)	2.0000 (1.8516~2.1484)	2.0000 (1.9124~2.0876)	2.0000 (1.7256~2.2744)	2.0000 (1.7430~2.2570)
		q	0.5125	0.5159	0.2032	0.2028	0.1988	0.4999
1.5	5	$\widehat{R}$	1.4726 (1.2305~1.7147)	1.4726 (1.2001~1.7452)	1.4726 (1.3049~1.6404)	1.4726 (1.3975~1.5478)	1.4726 (1.2365~1.7087)	1.4726 (1.2088~1.7364)
		q	0.6904	0.6941	0.4355	0.4351	0.4312	0.6790
1.1	5	$\widehat{R}$	1.1125 (0.7341~1.4909)	1.1125 (0.8090~1.4160)	1.1125 (0.8934~1.3690)	1.1125 (1.0472~1.1778)	1.1125 (0.9069~1.3181)	1.1125 (0.7766~1.4484)
		q	0.8737	0.8785	0.8048	0.8046	0.8030	0.8990
0.9	5	$\widehat{R}$	0.9091 (0.4882~1.3300)	0.9091 (0.4158~1.4024)	0.9091 (0.6495~1.2379)	0.9091 (0.8500~0.9682)	0.9091 (0.7231~1.0951)	0.9091 (0.5198~1.2984)
		q	1	1	1	1	1	1

Nbin: Negative Binomial distribution; Bin: Binomial distribution; Geo: Geometric distribution; R : estimated reproductive number; q: estimated extinction probability
*: No results because it cannot be estimated

Table 5.9 Data generated with Negative Binomial distribution (r=10)

R_0	X_0	X_1	X_2	X_3	X_4	X_5	X_6	Total
2	1	1 (1)	1 (1)	2 (2)	3 (2,1)	9 (3,4,2)	20 (4,8,0,1,1,2,0,3,1)	37
1.5	1	2 (2)	3 (1,2)	3 (0,2,1)	5 (1,0,4)	6 (2,1,1,1,1)	13 (4,3,1,0,2,3)	33
1.1	1	2 (2)	8 (6,2)	6 (1,1,0,1,1, 0,2,0)	7 (2,2,1,1,1,0)	4 (0,1,0,1,1,0,1)	4 (0,2,1,1)	32
0.9	1	1 (1)	3 (3)	2 (0,1,1)	2 (0,2)	2 (2,0)	1 (0,1)	12
2	5	12 (1,1,1,6,3)	29 (2,3,2,4,0, 1,1,5,4,3, 4,0)	55 (1,1,4,1,1, 4,5,0,4,1, 2,0,1,1,2, 0,2,0,3,4, 0,0,2,3,2, 3,4,3,1)	108 (2,2,0,0,2,1,1,1,0,2, 0,3,4,3,2,1,1,4,0,5, 0,2,1,2,3,0,1,1,1,1, 5,0,2,2,0,4,4,2,5,1, 3,1,2,1,6,2,1,2,3,3, 3,3,1,4,2)	221 (*)	443 (#)	873
1.5	5	8 (1,3,1,0,3)	15 (2,4,0,0,2, 3,3,1)	31 (3,1,2,3,0, 2,3,1,4,3, 3,2,1,2,1)	39 (0,0,0,2,2,5,1,0,1,1, 1,0,1,1,1,0,0,0,0,2, 1,0,4,4,2,3,1,0,0,2,4)	54 (0,2,1,1,1,0,0,5,0,1, 1,4,0,1,0,3,0,1,1,1, 1,0,1,4,0,2,3,1,2,1, 0,3,1,1,3,2,3,2,1)	83 (1,0,0,0,1,2,0,0,1,1, 2,1,1,2,3,0,1,1,1,1, 4,1,2,1,4,1,1,1,2,0, 2,3,5,1,0,1,1,1,5,1, 3,2,3,2,3,2,1,0,0,0, 1,2,4,4)	235
1.1	5	8 (1,2,3,1,1)	5 (0,0,2,0,0, 1,1,1)	8 (1,2,2,1,2)	7 (2,2,0,2,0, 0,0,1)	6 (0,1,1,1,1, 1,1)	10 (1,0,4,2,1,2)	49
0.9	5	6 (0,2,0,0,4)	7 (0,1,2,1,3,0)	4 (1,1,1,0,1,0,0)	3 (1,1,1,0)	3 (1,0,2)	3 (2,0,1)	31

R_0 :the basic reproductive number; X_i : the total number of the *i*-th generation offspring.

*(0,1,2,1,0,4,1,0,0,6,1,0,2,2,0,7,3,2,1,2,6,2,3,2,2,2,2,1,4,4,3,1,1,3,1,1,1,2,0,2,2,3,1,2,2,1,2,2,3,3,4,3,2,0,2,1,2,1,0,4,4,2,4,1,1,4,3,1,3,3,2,4,2,1,2,1,2,3,6,0,0,2,4,0,0,1,0,1,3,2,1,1,3,2,4,2,2,3,3,4,0,1,2,4,2,2,4,1)

#(0,4,2,2,1,2,4,3,5,0,2,3,1,6,1,5,1,2,4,2,1,1,3,0,2,1,1,1,0,1,1,0,3,4,2,0,2,1,1,2,2,1,2,3,5,4,3,1,2,2,4,1,2,5,0,3,2,1,1,3,1,1,2,4,1,1,1,1,3,3,0,3,2,2,1,1,3,2,3,3,2,3,3,0,1,1,2,3,0,3,2,2,3,1,1,1,0,1,3,3,1,6,5,2,0,2,4,2,3,1,1,5,0,2,3,4,2,2,1,2,4,1,1,1,3,1,7,0,0,1,1,3,3,2,4,1,1,2,3,5,1,3,1,4,0,2,4,4,2,2,4,2,2,2,2,4,0,1,1,2,3,0,1,3,0,3,1,3,0,2,2,0,3,1,2,1,1,5,3,0,2,3,2,1,0,4,2,0,1,5,1,1,0,1,0,2,3,3,0,0,1,1,2,3,4,3,3,1,2,0,1,4,5,2,0,4,1,2,2,1,4)

Table 5.10 Estimation for simulated data with Negative Binomial distribution (r=10)

R_0	X_0	Method / Estimate	Non-parametric method	Nbin	Poisson	Bin (n=1000)	Bin (n=100)	Geo
2	1	$\hat{R}$	2.1176 (1.6539~2.5814)	2.1176 (1.2709~2.9644)	2.1176 (1.4832~2.9317)	2.1176 (2.0275~2.2077)	2.1176 (1.8355~2.3998)	2.1176 (0.8962~3.3391)
		q	0.1977	0.2658	0.1739	0.1735	0.1697	0.4723
1.5	1	$\hat{R}$	1.6000 (1.2422~1.9578)	-*	1.6000 (1.0944~2.2587)	1.6000 (1.5217~1.6783)	1.6000 (1.3541~1.8459)	1.6000 (0.7061~2.4939)
		q	0.2904		0.3580	0.3576	0.3536	0.6250
1.1	1	$\hat{R}$	1.1071 (0.4203~1.7940)	1.1071 (0.6911~1.5232)	1.1071 (0.7522~1.5715)	1.1071 (1.0420~1.1723)	1.1071 (0.9021~1.3122)	1.1071 (0.5414~1.6729)
		q	0.8339	0.8321	0.8130	0.8128	0.8112	0.9033
0.9	1	$\hat{R}$	1.0000 (0.4696~1.5304)	-*	1.0000 (0.4992~1.7893)	1.0000 (0.9381~1.0619)	1.0000 (0.8050~1.1950)	1.0000 (0.1643~1.8357)
		q	1		1	1	1	1
2	5	$\hat{R}$	2.0186 (1.9479~2.0893)	2.0186 (1.8797~2.1575)	2.0186 (1.8843~2.1529)	2.0186 (1.9306~2.1066)	2.0186 (1.7430~2.2943)	2.0186 (1.7853~2.2519)
		q	0.2177	0.2133	0.1982	0.1978	0.1938	0.4954
1.5	5	$\hat{R}$	1.5132 (1.3254~1.7010)	1.5132 (1.3014~1.7249)	1.5132 (1.3176~1.7087)	1.5132 (1.4370~1.5893)	1.5132 (1.2739~1.7524)	1.5132 (1.2031~1.8232)
		q	0.4590	0.4538	0.4087	0.4083	0.4044	0.6609
1.1	5	$\hat{R}$	1.1282 (0.7888~1.4677)	-*	1.1282 (0.8198~1.5146)	1.1282 (1.0624~1.1940)	1.1282 (0.9212~1.3352)	1.1282 (0.6419~1.6145)
		q	0.7170		0.7817	0.7815	0.7797	0.8864
0.9	5	$\hat{R}$	0.9286 (0.7280~1.1292)	0.9286 (0.5599~1.2973)	0.9286 (0.6066~1.3606)	0.9286 (0.8689~0.9883)	0.9286 (0.7406~1.1166)	0.9286 (0.4329~1.4243)
		q	1	1	1	1	1	1

Nbin: Negative Binomial distribution; Bin: Binomial distribution; Geo: Geometric distribution; $\hat{R}$: estimated reproductive number; $\hat{q}$: estimated extinction probability
*: No results because it cannot be estimated

Table 5.11 Data generated with Negative Binomial distribution (r=5)

R_0	X_0	X_1	X_2	X_3	X_4	X_5	X_6	Total
2	1	3 (3)	8 (2,2,4)	19 (4,2,4,4,1, 1,1,2)	40 (2,1,1,2,3, 1,1,4,1,5, 1,1,2,3,1, 0,5,0,6)	77 (2,1,3,6,4,2,0,2,2,2, 2,1,5,0,0,3,2,0,2,1, 6,2,0,0,3,1,3,1,1,2, 0,0,1,1,1,1,6,3,1,4)	152 (2,3,2,5,3,10,1,0,4,3, 1,0,2,1,1,2,1,0,3,1, 3,0,0,3,0,3,5,2,2,0, 0,1,1,0,0,2,0,2,4,0, 2,3,3,1,1,3,3,4,3,4, 0,2,1,1,3,2,5,0,3,3, 5,1,3,2,0,3,3,2,1,2, 1, 0,3,1,3,1,1)	300
1.5	1	2 (2)	9 (3,6)	14 (0,0,4,1,4, 1,3,1,0)	23 (2,4,1,0,1, 3,0,1,1,1, 2,3,2,2)	31 (1,0,4,0,2,2,1,4,0,2, 0,1,8,1,1,1,0,0,2,0, 1,0,0)	42 (1,1,1,0,3,1,1,4,1,1, 2,1,2,0,1,2,0,1,2,3, 0,0,3,1,4,1,1,2,2,0,1)	122
1.1	1	4 (4)	11 (4,5,1,1)	18 (2,1,1,4,1, 3,1,2,0,3,0)	13 (1,0,1,2,0,0,0,0,2,2, 1,1,1,0,1,0,1,0)	10 (1,2,0,2,0,0,1,0,1,1, 1,0,1)	6 (1,0,0,1,2,0,0,1,0,1)	63
0.9	1	1 (1)	2 (2)	2 (1,1)	1 (1,0)	1 (1)	0 (0)	8
2	5	6 (1,2,2,0,1)	15 (3,0,3,2,4,3)	23 (1,0,2,1,0, 3,2,1,0,1, 1,5,1,2,3)	37 (4,0,0,3,1,0,3,2,1,1, 1,0,2,2,2,2,3,1,3,2, 3,0,1)	78 (3,0,2,4,2,0,2,3,2,0, 2,6,3,1,1,6,2,1,0,3, 1,5,2,3,0,2,2,3,2,2, 2,2,5,0,0,1,3)	175 (*)	339
1.5	5	9 (3,0,4,0,2)	14 (2,0,0,1,3, 4,0,3,1)	22 (1,1,3,0,2, 2,5,1,2,1, 0,0,2,2)	27 (1,1,1,4,0,0,0,4,1,1, 1,2,2,0,1,2,0,1,2,0, 2,1)	41 (1,0,0,4,0,3,1,4,0,0, 0,1,1,4,6,1,1,1,1,2, 2,1,1,4,1,0,1)	61 (1,0,0,1,3,2,2,3,1,1, 1,1,3,0,0,1,1,0,4,3, 3,1,2,0,2,1,1,1,4,1, 1,2,1,4,0,2,0,0,2,4,1)	179
1.1	5	4 (1,0,2,0,1)	5 (0,1,3,1)	6 (1,0,0,3,2)	13 (3,1,2,2,3,2)	8 (2,0,0,1,0,1,0,1,0,1,0,2,0)	9 (0,2,2,0,3,0,1,1)	50
0.9	5	7 (1,0,1,5,0)	5 (0,0,0,1,2, 0,2)	5 (1,3,0,0,1)	5 (3,1,0,0,1)	1 (0,0,0,0,1)	2 (2)	30

R_0 :the basic reproductive number; X_i : the total number of the *i*-th generation offspring.

*(2,1,2,3,3,2,6,1,0,1,4,5,2,2,1,0,0,0,2,1,2,3,2,1,6,6,4,2,3,5,3,0,4,5,3,1,2,1,0,3,3,2,1,1,1,0,1,1,0,3,3,2,0,0,3,1,1,3,1,5,

4,5,1,0,2,2,2,3,4,3,6,4,0,1,5,3,2,2)

Table 5.12 1Estimation for simulated data with Negative Binomial distribution (r=5)

R_0	X_0	Method / Estimate	Non-parametric method	NBin	Poisson	Bin (n=1000)	Bin (n=100)	Geo
2	1	$\hat{R}$	2.0203 (1.8906~2.1499)	2.0203 (1.7546~2.2860)	2.0203 (1.7913~2.2493)	2.0203 (1.9323~2.1083)	2.0203 (1.7445~2.2960)	2.0203 (1.6223~2.4182)
		$\hat{q}$	0.2667	0.2673	0.1977	0.1973	0.1934	0.4950
1.5	1	$\hat{R}$	1.5125 (1.1151~1.9099)	1.5125 (1.1904~1.8346)	1.5125 (1.2430~1.7820)	1.5125 (1.4363~1.5887)	1.5125 (1.2733~1.7517)	1.5125 (1.0853~1.9397)
		$\hat{q}$	0.5029	0.5107	0.4092	0.4087	0.4048	0.6611
1.1	1	$\hat{R}$	1.0877 (0.5175~1.6579)	1.0877 (0.8339~1.3944)	1.0877 (0.5175~1.6579)	1.0877 (1.0231~1.1523)	1.0877 (0.8844~1.2910)	1.0877 (0.6965~1.4789)
		$\hat{q}$	0.8720	0.8699	0.8432	0.8430	0.8417	0.9194
0.9	1	$\hat{R}$	0.8750 (0.4390~1.3110)	-*	0.8750 (0.3518~1.8028)	0.8750 (0.8170~0.9330)	0.8750 (0.6925~1.0575)	0.8750 (-0.0126~1.7626)
		$\hat{q}$	1		1	1	1	1
2	5	$\hat{R}$	2.0366 (1.7841~2.2891)	2.0366 (1.7963~2.2769)	2.0366 (1.8182~2.2550)	2.0366 (1.9482~2.1249)	2.0366 (1.7597~2.3134)	2.0366 (1.6560~2.4172)
		$\hat{q}$	0.2542	0.2673	0.1935	0.1931	0.1892	0.4910
1.5	5	$\hat{R}$	1.4746 (1.3672~1.5819)	1.4746 (1.2308~1.7183)	1.4746 (1.2555~1.6937)	1.4746 (1.3994~1.5498)	1.4746 (1.2383~1.7108)	1.4746 (1.1299~1.8192)
		$\hat{q}$	0.4950	0.5107	0.4341	0.4377	0.4299	0.6781
1.1	5	$\hat{R}$	1.0976 (0.6931~1.5020)	-*	1.0976 (0.8006~1.4686)	1.0976 (1.0327~1.1625)	1.0976 (0.8934~1.3018)	1.0976 (0.6331~1.5620)
		$\hat{q}$	0.8271		0.8276	0.8274	0.8260	0.9109
0.9	5	$\hat{R}$	0.8929 (0.5464~1.2393)	0.8929 (0.4434~1.3423)	0.8929 (0.5778~1.3180)	0.8929 (0.8343~0.9514)	0.8929 (0.7085~1.0772)	0.8929 (0.4113~1.3744)
		$\hat{q}$	1	1	1	1	1	1

Nbin: Negative Binomial distribution; Bin: Binomial distribution; Geo: Geometric distribution; $\hat{R}$: estimated reproductive number; $\hat{q}$: estimated extinction probability

*: No results because it cannot be estimated

Table 5.13 Data generated with Negative Binomial distribution (r=2)

R_0	X_0	X_1	X_2	X_3	X_4	X_5	X_6	Total
2	1	2 (2)	3 (0,3)	9 (4,4,1)	22 (0,0,6,0,4, 1,0,2,9)	37 (0,1,5,1,0,1,0,1,1,5, 1,5,1,0,1,2,1,0,6,0, 2,3)	78 (0,4,3,1,1,2,3,1,0,3, 2,3,0,7,0,4,0,5,1,2, 4,0,4,5,2,3,0,0,6,1, 0,4,2,1,0,0,4)	152
1.5	1	2 (2)	3 (1,2)	4 (2,2,0)	10 (1,2,1,6)	15 (0,2,3,1,2,0,0,6,1,0)	19 (0,1,1,1,0,0,2,5,0,1, 1,1,1,3,2)	54
1.1	1	2 (2)	2 (1,1)	2 (2,0)	4 (2,2)	2 (0,1,0,1)	3 (3,0)	16
0.9	1	3 (3)	2 (2,0,0)	1 (0,1)	2 (2)	1 (1,0)	0 (0)	10
2	5	5 (0,1,1,1,2)	12 (1,2,1,4,4)	22 (3,1,0,0,1, 2,2,2,3,4, 3,1)	53 (1,2,0,6,3, 2,1,7,2,2, 0,1,0,2,2, 0,1,4,4,0, 8,5)	96 (1,2,0,0,2,1,0,1,3,2, 0,1,2,4,0,0,2,2,1,2, 1,2,0,4,3,2,0,1,2,7, 0,1,2,0,2,1,3,2,0,2, 5,1,4,2,1,0,0,5,3,3, 7,1,3)	193 (*)	386
1.5	5	7 (3,1,0,1,2)	8 (0,0,2,3,1, 0,2)	11 (1,3,1,1,1, 1,1,2)	19 (2,1,2,5,2, 0,0,4,0,3, 0)	36 (3,1,6,0,0,1,1,8,1,4, 1,0,1,0,1,4,1,1,2)	47 (0,1,1,1,1,6,9,2,3,0, 0,1,1,0,3,1,0,2,1,0, 2,3,3,0,3,0,2,3,0,0, 2,0,0,3,1,1)	133
1.1	5	10 (2,0,4,3,1)	14 (2,2,1,1,2, 0,0,0,3,3)	13 (2,1,0,1,2, 3,0,1,1,1, 0,0,1,0)	17 (0,1,0,5,1, 3,2,0,2,0, 1,0,2)	15 (2,0,3,0,0,0,1,0,0,3, 0,0,1,1,0,3,1)	13 (0,0,0,4,0,1,2,0,0,1, 2,2,0,0,1)	87
0.9	5	7 (1,2,3,0,1)	7 (1,1,0,0,1, 4,0)	7 (1,0,1,1,4, 0,0)	3 (1,1,0,0,1, 0,0)	6 (0,2,4)	2 (1,1,0,0,0,0)	37

R_0 :the basic reproductive number; X_i : the total number of the *i*-th generation offspring.

*(1,0,1,4,1,1,4,1,2,0,0,2,5,5,0,0,4,0,4,2,2,2,3,0,6,0,4,0,4,2,0,0,3,3,3,0,3,2,6,0,3,2,0,2,1,0,7,2,2,5,4,0,1,0,3,8,2,3,1,5,
0,1,7,0,2,1,2,4,1,1,0,1,4,0,1,3,0,4,1,3,0,2,6,1,1,2,0,1,0,2,3,0,0,1,3,4)

Table 5.14 Estimation for simulated data with Negative Binomial distribution (r=2)

R_0	X_0	Method / Estimate	Non-parametric method	Nbin	Poisson	Bin (n=1000)	Bin (n=100)	Geo
2	1	$\hat{R}$	2.0405 (1.7805~2.3006)	2.0405 (1.5367~2.5444)	2.0405 (1.7151~2.3660)	2.0405 (1.9521~2.1290)	2.0405 (1.7634~2.3177)	2.0405 (1.4730~2.6081)
		q	0.4328	0.4205	0.1925	0.1921	0.1882	0.4901
1.5	1	$\hat{R}$	1.5143 (1.2076~1.8210)	1.5143 (1.0200~2.0086)	1.5143 (1.1343~1.9807)	1.5143 (1.4381~1.5905)	1.5143 (1.2749~1.7536)	1.5143 (0.8678~2.1607)
		q	0.5053	0.5179	0.4080	0.4076	0.4037	0.6603
1.1	1	$\hat{R}$	1.1538 (0.6994~1.6082)	-*	1.1538 (0.6458~1.9031)	1.1538 (1.0873~1.2204)	1.1538 (0.9445~1.3632)	1.1538 (0.2969~2.0108)
		q	0.7016		0.7458	0.7456	0.7435	0.8667
0.9	1	$\hat{R}$	0.900 (0.2195~1.5805)	0.900 (0.2193~1.5807)	0.900 (0.4115~1.7085)	0.900 (0.8412~0.9588)	0.900 (0.7149~1.0851)	0.900 (0.0895~1.7105)
		q	1	1	1	1	1	1
2	5	$\hat{R}$	1.9741 (1.7781~2.1701)	1.9741 (1.7082~2.2400)	1.9741 (1.7759~2.1723)	1.9741 (1.8871~2.0611)	1.9741 (1.7014~2.2467)	1.9741 (1.6322~2.3159)
		q	0.3692	0.3610	0.2014	0.2100	0.2060	0.5065
1.5	5	$\hat{R}$	1.4884 (1.2780~1.6988)	1.4884 (1.1658~1.8109)	1.4884 (1.2305~1.7462)	1.4884 (1.4128~1.5639)	1.4884 (1.2510~1.7257)	1.4884 (1.0816~1.8951)
		q	0.5489	0.5518	0.4248	0.4244	0.4205	0.6720
1.1	5	$\hat{R}$	1.1081 (0.8527~1.3635)	1.1081 (0.8223~1.3939)	1.1081 (0.8813~1.3755)	1.1081 (1.0429~1.1733)	1.1081 (0.9029~1.3133)	1.1081 (0.7599~1.4563)
		q	0.8556	0.8603	0.81115	0.8113	0.8097	0.9026
0.9	5	$\hat{R}$	0.9129 (0.5201~1.3085)	0.9129 (0.5253~1.3033)	0.9129 (0.6254~1.2907)	0.9129 (0.8550~0.9735)	0.9129 (0.7277~1.1008)	0.9129 (0.4760~1.3526)
		q	1	1	1	1	1	1

Nbin: Negative Binomial distribution; Bin: Binomial distribution; Geo: Geometric distribution; $\hat{R}$: estimated reproductive number; $\hat{q}$: estimated extinction probability

*: No results because it cannot be estimated

Table 5.15 Data generated with Negative Binomial distribution (r=1)

R_0	X_0	X_1	X_2	X_3	X_4	X_5	X_6	Total
2	1	4 (4)	12 (6,2,2, 2)	31 (0,4,0,3, 0, 8,3,0,1,0 , 3,9)	66 (0,4,2,1, 2, 0,7,1,2,0 , 0,1,2,3,4 , 5,5,3,3,1 , 2,0,2,1,2 , 0,1,5,2,2 , 3)	117 (4,1,2,2,2,2,1,3,3,5, 1,2,0,0,2,3,1,0,0,1, 2,1,3,2,1,3,6,0,2,1, 8,0,3,0,1,0,0,1,3,1, 0,2,0,2,0,0,0,1,0,0, 1,1,0,4,0,6,11,4,2,0, 0,1, 1,1,7,0)	241 (1,0,1,2,1,2,3,4,5,2, 7,2,0,4,5,3,1,1,9,7, 2,4,1,0,3,3,0,4,3,1, 0,0,0,1,0,0,2,8,1,0, 1,1,0,5,0,3,0,0,5,0, 1,1,0,0,6,1,0,8,3,1, 4,5,1,1,3,1,0,1,0,4, 3,2,5,1,4,0,0,0,5,3, 1,5,1,11,1,1,0,1,1,0, 2,3,4,0,0,4,0,0,0,3, 1,2,0,1,1,0,0,7,1,1, 1,11,0,0,0,4,1)	472
1.5	1	1 (1)	1 (1)	2 (2)	2 (0,2)	6 (4,2)	9 (0,1,0,1,4,3)	22
1.1	1	8 (8)	10 (0,1,2, 0,3, 4,0,0)	6 (0,0,0,0, 0, 1,1,2,0,2)	8 (6,1,1,0, 0, 0)	6 (0,0,1,0,0, 1,4,0)	3 (1,1,0,1,0,0)	42
0.9	1	3 (3)	3 (1,2,0)	4 (1,2,1)	1 (0,0,1,0)	1 (1)	0 (0)	13
2	5	9 (1,5,1, 2,0)	22 (3,4,4, 1,0, 8,1,1,0)	58 (2,6,1,3, 0, 1,14,3,1, 2, 4,0,4,0,0 , 4,1,6,2,3 ,1,0)	85 (*)	170 (#)	354 (&)	703[#]
1.5	5	14 (4,3,2, 1,4)	12 (0,0,2, 0,0, 0,0,3,2 ,0, 0,3,0,2)	23 (2,2,0,1, 1, 0,4,3,4,0 , 2,4)	42 (0,0,1,1, 1, 9,1,5,2,2 , 1,0,6,2,0 , 0,4,1,3,0 , 0,0,3)	48 (0,0,3,0,0,1,5,0,0,0,0,0, 0,7,0,1,0,2,1,1,2,0,0,0,1 ,0,0,1,2,3,3,0,0,0,0,1,7, 2,1,2,0,2)	71 (4,7,1,2,1,1,6,0,1,1,1,2,0,0 ,3,0,2,2,2,0,5,1,0,0,0,1,4,0 ,0,6,1, 1,3,2,0,0,1,0,0,1,0,1,0,0,2, 0,4,1,1)	215
1.1	5	6 (2,2,1, 1,0)	13 (3,1,3, 1,2, 3)	12 (1,0,0,3, 0, 0,1,0,1,4 , 1,0,1)	21 (1,3,0,2, 3, 0,0,4,1,1 , 2,4)	23 (0,1,1,0,0, 1,2,1,0,0, 0,3,0,0,0, 6,2,3,2,0,1)	14 (4,1,0,0,0,4,0,0,0,0, 1,0,1,1,1,0,0,0,0,0, 0,1,0)	94
0.9	5	5 (0,5,0, 0,0)	6 (2,0,3, 0,1)	12 (3,1,1,0, 0,7)	9 (1,4,1,2, 1, 0,0,0,0,0 ,0,0)	7 (0,2,0,4,0, 0,0,1,0)	1 (0,0,0,0,1,0,0)	45

R_0 :the basic reproductive number; X_i : the total number of the *i*-th generation offspring.

*(3,1,1,0,3,3,1,4,1,0,4,0,0,0,0,1,1,2,4,1,2,3,0,0,2,1,2,0,0,0,1,0,0,0,4,0,0,0,1,3,0,0,1,4,6,1,1,0,6,1,0,7,3,0,3,0,2,1,0,2,4,0,4,3,0,4,0,2,0);#(4,0,6,1,4,0,2,2,3,0,1,0,0,3,3,9,1,1,6,3,0,0,0,0,0,5,8,1,4,3,0,1,0,0,6,0,5,1,0,2,0,2,3,0,0,1,3,1,2,2,5,3,0,2,4,3,0,5,3,7,4,0,1,2,5,1,1,2,0,0,0,0,2,0,2,1,0,4,0,12,2,0,0,0,5);&(4,0,3,0,1,0,0,0,0,1,3,1,4,3,0,0,0,2,2,1,2,0,3,4,0,0,4,0,3,1,0,4,1,12,2,9,6,1,0,1,0,2,1,11,1,0,5,0,1,0,0,6,5,2,0,5,6,0,1,0,2,1,0,3,2,1,0,2,3,1,1,1,0,0,7,1,1,0,1,1,2,0,2,2,0,0,0,0,0,0,4,5,0,0,0,4,0,0,0,12,1,1,0,2,0,1,0,0,7,3,2,6,3,1,0,4,0,3,2,0,1,2,1,0,2,0,2,2,5,0,4,0,10,2,5,5,6,2,4,1,4,6,2,1,2,2,0,1,4,0,0,0,0,3,3,2,1,0,3,21)

Table 5.16 Estimation for simulated data with Negative Binomial distribution (r=1)

R_0	X_0	Method / Estimate	Non-parametric method	Nbin	Poisson	Bin (n=1000)	Bin (n=100)	Geo
2	1	$\hat{R}$	2.0390 (1.8279~2.2501)	2.0390 (1.7412~2.3367)	2.0390 (1.8548~2.2231)	2.0390 (1.9505~2.1274)	2.0390 (1.7620~2.3160)	2.0390 (1.7180~2.3600)
		q	0.4477	0.4466	0.1929	0.1924	0.1886	0.4905
1.5	1	$\hat{R}$	1.6154 (1.0913~2.1395)	1.6154 (0.8842~2.3466)	1.6154 (0.9999~2.4693)	1.6154 (1.5367~1.6941)	1.6154 (1.3683~1.8625)	1.6154 (0.4980~2.7327)
		q	0.3991	0.3812	0.3498	0.3494	0.3454	0.6192
1.1	1	$\hat{R}$	1.0513 (0.1127~1.9899)	1.0513 (0.5096~1.5929)	1.0513 (0.7544~1.4262)	1.0513 (0.9878~1.1148)	1.0513 (0.8514~1.2512)	1.0513 (0.5904~1.5122)
		q	0.9656	0.9642	0.9040	0.9039	0.9031	0.9512
0.9	1	$\hat{R}$	0.9231 (0.3150~1.5311)	-*	0.9231 (0.4770~1.6124)	0.9231 (0.8636~0.9826)	0.9231 (0.7356~1.1105)	0.9231 (0.1988~1.6477)
		q	1		1	1	1	1
2	5	$\hat{R}$	2.0000 (1.7709~2.2291)	2.0000 (1.7359~2.2641)	2.0000 (1.8516~2.1484)	2.0000 (1.9124~2.0876)	2.0000 (1.7256~2.2744)	2.0000 (1.7430~2.2570)
		q	0.5125	0.4466	0.2032	0.2028	0.1988	0.4999
1.5	5	$\hat{R}$	1.4583 (1.1324~1.7843)	1.4583 (1.1528~1.7639)	1.4583 (1.2611~1.6556)	1.4583 (1.3835~1.5331)	1.4583 (1.2234~1.6933)	1.4583 (1.1491~1.7676)
		q	0.6762	0.3812	0.4454	0.4450	0.4412	0.6858
1.1	5	$\hat{R}$	1.1125 (0.7341~1.4909)	1.1125 (0.8090~1.4160)	1.1125 (0.8934~1.3690)	1.1125 (1.0472~1.1778)	1.1125 (0.9069~1.3181)	1.1125 (0.7766~1.4484)
		q	0.8737	0.8785	0.8048	0.8046	0.8030	0.8990
0.9	5	$\hat{R}$	0.9091 (0.4882~1.3300)	0.9091 (0.4158~1.4024)	0.9091 (0.6495~1.2379)	0.9091 (0.8500~0.9682)	0.9091 (0.7231~1.0951)	0.9091 (0.5198~1.2984)
		q	1	1	1	1	1	1

Nbin: Negative Binomial distribution; Bin: Binomial distribution; Geo: Geometric distribution; $\hat{R}$: estimated reproductive number; $\hat{q}$: estimated extinction probability
*: No results because it cannot be estimated

Table 5.17 Average number of offspring in each generation under 10000 times simulations for a given Poisson distribution

R_0	X_0	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	Total	$\widehat{R}$ (SE)	$\hat{q}$
2	1	1.98	3.95	7.87	15.76	31.59	63.18	126.31	252.79	505.39	1009.81	2.00(0.47)	0.209
1.5	1	1.52	2.28	3.40	5.07	7.61	11.46	17.14	25.61	38.46	113.55	1.50(0.28)	0.4176
1.1	1	1.12	1.24	1.39	1.52	1.66	1.84	2.01	2.23	2.45	16.46	1.10(0.09)	0.7314
0.9	1	0.91	0.82	0.72	0.65	0.58	0.52	0.46	0.42	0.37	6.45	0.90(0.13)	0.9037
2	5	10.04	20.10	40.35	80.82	161.68	323.52	647.04	1293.01	2586.91	5168.47	2.00(0.47)	0.0005
1.5	5	7.49	11.20	16.73	25.14	37.78	56.53	84.87	127.39	190.83	562.95	1.50(0.29)	0.013
1.1	5	5.49	6.06	6.70	7.39	8.13	8.88	9.77	10.76	11.82	80.00	1.10(0.13)	0.2192
0.9	5	4.55	4.09	3.68	3.33	2.98	2.68	2.42	2.19	1.99	32.91	0.90(0.04)	0.5964

$\widehat{R}$: Estimated by parametric method; $\hat{q}$: extinction probability calculated by simulation; SE: standard error.

Table 5.18 Average number of offspring in each generation under 10000 times simulations for a given Binomial distribution (n=1000)

R_0	X_0	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	Total	$\hat{R}$ (SE)	$\hat{q}$
2	1	2.00	4.01	8.05	16.07	32.17	64.33	128.73	257.72	515.50	1029.58	2.00(0.0004)	0.2014
1.5	1	1.50	2.25	3.37	5.07	7.64	11.45	17.19	25.81	38.66	113.94	1.50(0.0009)	0.4135
1.1	1	1.10	1.20	1.32	1.45	1.59	1.75	1.93	2.15	2.34	15.82	1.10(0.0002)	0.7418
0.9	1	0.89	0.81	0.73	0.66	0.59	0.53	0.48	0.43	0.40	6.52	0.90(0.0002)	0.9003
2	5	9.99	19.95	39.93	79.99	159.80	319.41	639.15	1278.26	2556.71	5108.19	2.00(0.0003)	0.0006
1.5	5	7.47	11.24	16.84	25.36	37.88	56.81	85.25	127.83	191.79	565.46	1.50(0.0008)	0.0112
1.1	5	5.53	6.06	6.69	7.36	8.10	8.88	9.76	10.78	11.90	80.07	1.10(0.0011)	0.2191
0.9	5	4.47	4.01	3.60	3.24	2.92	2.65	2.41	2.17	1.97	32.44	0.90(0.0018)	0.5945

$\hat{R}$: Estimated by parametric method; $\hat{q}$: extinction probability calculated by simulation.

Table 5.19 Average number of offspring in each generation under 10000 times simulations for a given Binomial distribution (n=100)

R_0	X_0	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	Total	$\widehat{R}$ (SE)	$\hat{q}$
2	1	2.00	3.97	7.95	15.85	31.71	63.49	126.96	253.74	507.29	1013.96	2.00(0.0004)	0.201
1.5	1	1.50	2.22	3.35	4.99	7.48	11.22	16.79	25.19	37.76	111.49	1.50(0.0009)	0.4124
1.1	1	1.11	1.22	1.34	1.47	1.61	1.80	1.99	2.20	2.41	16.15	1.10(0.0002)	0.7366
0.9	1	0.90	0.79	0.72	0.65	0.58	0.51	0.46	0.41	0.36	6.39	0.90(0.0002)	0.9046
2	5	9.99	19.92	39.84	79.62	159.21	318.56	636.90	1273.92	2547.79	5090.76	2.00(0.0003)	0.0003
1.5	5	7.51	11.25	16.96	25.42	38.18	57.32	85.98	128.90	193.44	569.96	1.50(0.0008)	0.0118
1.1	5	5.46	5.97	6.56	7.19	7.87	8.65	9.52	10.43	11.49	78.16	1.10(0.0011)	0.2297
0.9	5	4.50	4.03	3.58	3.20	2.89	2.59	2.36	2.11	1.89	32.14	0.90(0.0018)	0.5982

$\widehat{R}$: Estimated by parametric method; $\hat{q}$: extinction probability calculated by simulation.

Table 5.20 Average number of offspring in each generation under 10000 times simulations for a given Geometric distribution

R_0	X_0	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	Total	$\widehat{R}$ (SE)	$\hat{q}$
2	1	1.98	3.97	7.83	15.67	31.30	62.67	125.58	250.87	502.14	1003.01	2.00(0.0011)	0.5063
1.5	1	1.54	2.30	3.41	5.06	7.54	11.30	16.76	25.13	37.67	111.71	1.50(0.0029)	0.6553
1.1	1	1.10	1.22	1.34	1.47	1.60	1.77	1.94	2.15	2.37	15.95	1.10(0.0021)	0.8511
0.9	1	0.89	0.82	0.72	0.66	0.59	0.55	0.49	0.44	0.38	6.55	0.90(0.0069)	0.9416
2	5	9.98	19.98	39.88	79.43	159.10	318.13	636.66	1273.71	2548.65	5090.53	2.00(0.0005)	0.0339
1.5	5	7.55	11.26	16.82	25.20	37.78	56.80	85.27	127.80	191.54	565.01	1.50(0.0008)	0.1266
1.1	5	5.52	6.07	6.72	7.39	8.06	8.86	9.71	10.75	11.71	79.80	1.10(0.0021)	0.4393
0.9	5	4.53	4.04	3.61	3.24	2.93	2.61	2.35	2.09	1.87	32.26	0.90(0.0018)	0.7419

$\widehat{R}$: Estimated by parametric method; $\hat{q}$: extinction probability calculated by simulation.

Table 5.21 Average number of offspring in each generation under 10000 times simulations for a given Negative Binomial distribution

R_0	X_0	r	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	Total	$\widehat{R}$ (SE)	$\hat{q}$
2	1	10	2.00	4.00	7.98	15.96	31.92	63.88	127.79	255.62	511.23	1021.38	2.00(0.0003)	0.2426
		5	2.03	4.04	8.10	16.18	32.32	64.55	129.10	257.84	515.16	1030.31	2.00(0.0005)	0.2853
		2	2.01	4.00	8.00	16.04	31.95	63.74	127.62	255.35	510.68	1020.39	2.00(0.0008)	0.3894
		1	1.98	3.97	7.83	15.67	31.30	62.67	125.58	250.87	502.14	1003.00	2.00(0.0011)	0.5063
1.5	1	10	1.50	2.28	3.43	5.13	7.68	11.50	17.21	25.81	38.68	114.21	1.50(0.0010)	0.4498
		5	1.50	2.27	3.41	5.08	7.58	11.39	17.10	25.67	38.48	113.46	1.50(0.0012)	0.4858
		2	1.53	2.34	3.49	5.28	7.92	11.85	17.75	26.67	39.98	117.79	1.50(0.0021)	0.5577
		1	1.54	2.30	3.41	5.06	7.54	11.30	16.76	25.13	37.67	111.71	1.50(0.0029)	0.6553
1.1	1	10	1.11	1.23	1.35	1.49	1.65	1.81	1.99	2.18	2.36	16.18	1.10(0.0029)	0.7507
		5	1.11	1.22	1.34	1.48	1.63	1.77	1.94	2.13	2.33	15.94	1.10(0.0020)	0.7657
		2	1.13	1.24	1.37	1.53	1.67	1.84	2.02	2.23	2.46	16.49	1.10(0.0035)	0.8092
		1	1.10	1.20	1.31	1.44	1.59	1.76	1.92	2.10	2.28	15.70	1.10(0.0022)	0.8542
0.9	1	10	0.90	0.82	0.72	0.64	0.57	0.51	0.47	0.42	0.37	6.42	0.90(0.0043)	0.9072
		5	0.90	0.82	0.73	0.65	0.58	0.52	0.46	0.42	0.38	6.46	0.90(0.0027)	0.9129
		2	0.93	0.86	0.79	0.73	0.65	0.57	0.51	0.47	0.41	6.90	0.90(0.0063)	0.9194
		1	0.90	0.82	0.72	0.66	0.59	0.55	0.48	0.44	0.38	6.55	0.90(0.0068)	0.9418

$\widehat{R}$: Estimated by parametric method; $\hat{q}$: extinction probability calculated by simulation

Table 5.22 Average number of offspring in each generation 10000 times simulations for a given Negative Binomial distribution

R_0	X_0	r	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	Total	$\widehat{R}$ (SE)	$\hat{q}$
2	5	10	10.01	19.97	40.10	80.08	160.05	319.88	639.61	1278.84	2558.92	5112.47	2.00(0.0004)	0.0008
		5	9.99	19.95	39.99	79.87	159.68	318.87	638.32	1275.62	2551.19	5098.46	2.00(0.0004)	0.0027
		2	9.99	19.99	39.88	79.69	159.21	318.59	637.45	1275.44	2551.10	5096.34	2.00(0.0003)	0.0094
		1	9.98	19.98	39.88	79.43	159.10	318.13	636.66	1273.71	2548.65	5090.53	2.00(0.0005)	0.0339
1.5	5	10	7.48	11.29	16.92	25.43	38.03	57.07	85.64	128.53	192.87	568.25	1.50(0.0006)	0.0192
		5	7.49	11.27	16.97	25.47	38.19	57.27	85.91	128.85	193.30	569.73	1.50(0.0004)	0.0295
		2	7.49	11.24	16.93	25.43	38.20	57.26	86.11	129.16	193.86	570.68	1.50(0.0006)	0.0603
		1	7.55	11.26	16.82	25.20	37.78	56.80	85.27	127.80	191.54	565.01	1.50(0.0008)	0.1266
1.1	5	10	5.48	6.02	6.61	7.29	7.99	8.87	9.78	10.73	11.80	79.56	1.10(0.0014)	0.2589
		5	5.49	6.05	6.63	7.26	8.00	8.77	9.65	10.61	11.67	79.13	1.10(0.0008)	0.278
		2	5.57	6.10	6.75	7.46	8.23	9.03	9.92	10.89	11.98	80.93	1.10(0.0016)	0.3428
		1	5.53	6.15	6.79	7.40	8.13	8.91	9.78	10.81	11.83	80.34	1.10(0.0020)	0.4451
0.9	5	10	4.50	4.05	3.64	3.29	2.94	2.64	2.37	2.12	1.93	32.50	0.90(0.0014)	0.6217
		5	4.46	4.03	3.63	3.26	2.92	2.64	2.38	2.11	1.91	32.35	0.90(0.0019)	0.6382
		2	4.46	4.06	3.68	3.28	2.93	2.65	2.37	2.10	1.89	32.43	0.90(0.0026)	0.6922
		1	4.53	4.03	3.61	3.25	2.92	2.61	2.37	2.10	1.88	32.31	0.90(0.0021)	0.741

$\widehat{R}$: Estimated by parametric method; $\hat{q}$: extinction probability calculated by simulation

Table 5.23 Tuberculosis Contact Investigation in Long-term Care Facility

		Resident (N=62)	Staff (N=18)
Age	Mean (SD)	75.9 (11.8)	35.5 (12.2)
Gender			
	Male (%)	21 (33.9)	2 (11.1)
	Female (%)	41 (66.1)	16 (88.9)
[a]Positive TST (1st)		21/57 (36.8%)	14/18 (77.8%)
[b]Positive TST (2nd)		11/41 (26.8%)	0/4 (0.0%)
New TST conversion		9/36 (25.0%)	0/4 (0.0%)
Tuberculosis case		4 (6.5%)	1 (5.6%)

TST: tuberculin skin test;

[a] The first TST was done on July 2, 2012.

[b] The second TST was done on September 18, 2012.

Table 5.24 Characteristics of TB cases and TB contacts in Changhua County, Taiwan

Characteristics		TB cases (n=212) N(%)	TB contacts (n=948) N(%)
Location	High incidence area	82(38.7)	311(32.8)
	Medium incidence area	70(33.0)	294(31.0)
	Low incidence area	60(28.3)	343(36.2)
Age	Mean(SD)	133(62.7)	420(44.3)
Age group[@]	<35	12(5.7)	250(26.4)
	35-65	88(41.5)	555(58.5)
	>=65	112(52.8)	143(15.1)
Gender	Male	133(62.7)	420()
	Female	79(37.3)	528()
TST*	Mean(SD)	12.6(7.7)	10.2(6.7)
TST*	<10 mm	65(31.5)	445(46.9)
	10-14 mm	48(23.3)	242(25.5)
	≥15 mm	93(45.1)	260(27.4)
QFT-GIT[#]	Mean(SD)	3.5(3.8)	1.3(2.3)
QFT-GIT	Negative	63(30.1)	541(57.6)
	Positive (≥0.35 IU/ml)	145(69.4)	369(39.3)
	Indeterminate	1(0.5)	38(4.0)
QFT-GIT	<0 IU/ml	10(4.8)	160(17.0)
	0-0.01 IU/ml	8(3.8)	38(4.0)
	0.01-0.35 IU/ml	45(21.5)	325(34.6)
	≥0.35 IU/ml	145(69.4)	387(41.2)
	Indeterminate	1(0.5)	38(4.0)

TST: tuberculin skin test; QFT-GIT: QuantiFERON-TB Gold In-Tube test.

[*] The results of TST were missed in Six TB cases and one TB contact.

[#] The results of QFT-GIT were missed in 3 TB cases and 8 TB contacts.

[@] Five TB cases with age below 30 and 168 TB contacts with age below 30.

Table 5.25 Univariable Analysis for risk factors of TB

Characteristic		Coefficient (95% CI)	OR	(95% CI)
Age Group				
	>=65	2.788(2.140, 3.436)	16.25	8.50, 31.06
	35-65	1.252(0.621, 1.882)	3.50	1.86, 6.57
	<35	*Reference*	-	-
Gender				
	Male	0.816(0.497, 1.135)	2.26	1.64, 3.11
	Female	*Reference*	-	
TST				
	≥10 mm	0.726(0.383, 1.070)	2.07	1.47, 2.92
	< 10 mm	*Reference*	-	
TST				
	≥15 mm	0.850(0.519, 1.180)	2.34	1.68, 3.26
	< 15 mm	*Reference*	-	
QFT-GIT				
	≥0.35 IU/ml	1.265(0.936, 1.594)	3.54	2.55, 4.93
	<0.35 IU/ml	*Reference*	-	
QFT-GIT				
	<0 IU/ml	*Reference*	-	
	0-0.01 IU/ml	1.508(0.490, 2.527)	4.52	1.63, 12.52
	0.01-0.35 IU/ml	0.875(0.150, 1.600)	2.40	1.16, 4.95
	≥0.35 IU/ml	1.958(1.273, 2.644)	7.09	3.57, 14.07

OR: odds ratio; CI: confidence interval; TST: tuberculin skin test; QFT-GIT: QuantiFERON-TB Gold In-Tube test.

Table 5.26 Multivariable Analysis for risk factors of TB (without interaction)

Characteristic		Coefficient (95% CI)	OR	(95% CI)
Age Group				
	>=65	2.652(1.979, 3.326)	14.19	7.23, 27.83
	35-65	1.053(0.408, 1.698)	2.87	1.50, 5.46
	<35	*Reference*	-	-
Gender				
	Male	0.907(0.538, 1.275)	2.48	1.71, 3.58
	Female	*Reference*	-	
TST				
	≥10 mm	0.616(0.220, 1.013)	1.85	1.25, 2.75
	< 10 mm	*Reference*	-	
QFT-GIT				
	≥0.35 IU/ml	0.903(0.542, 1.264)	2.47	1.72, 3.54
	<0.35 IU/ml	*Reference*	-	

OR: odds ratio; CI: confidence interval; TST: tuberculin skin test; QFT-GIT: QuantiFERON-TB Gold In-Tube test.

Table 5.27 Multivariable Analysis for risk factors of TB (with interaction)

Characteristic		Coefficient (95% CI)	OR	(95% CI)
Age Group				
	>=65	2.652(1.977, 3.327)	14.18	7.22, 27.86
	35-65	1.059(0.411, 1.707)	2.88	1.51, 5.51
	<35	*Reference*	-	-
Gender				
	Male	0.930(0.558, 1.303)	2.54	1.75, 3.68
	Female	*Reference*	-	
TST				
	≥10 mm	-0.115 (-0.696, 0.467)		
	< 10 mm	*Reference*	-	
QFT-GIT				
	≥0.35 IU/ml	0.139(-0.414, 0.693)	-	-
	<0.35 IU/ml	*Reference*		
QFT-GIT*TST		1.315(0.566, 2.064)		
	QFT-GIT	Positive TST	4.28	1.16, 15.76
		Negative TST	1.15	0.66, 2.00

OR: odds ratio; CI: confidence interval; TST: tuberculin skin test; QFT-GIT: QuantiFERON-TB Gold In-Tube test.

Table 5.28 Akaike Information Criterion values of models

Models	Number of parameter	AIC
TST(10), age, sex	5	742.064
QFT-GIT age, sex	5	743.594
TST(10), QFT-GIT, age, sex	6	709.523
TST(10), QFT-GIT, age, sex, age*TST(10)	8	710.915
TST(10), QFT-GIT, age, sex, age* QFT-GIT	8	711.976
TST(10), QFT-GIT, age, sex, age*TST(10), age* QFT-GIT	10	713.121
TST(10), QFT-GIT, age, sex, TST(10)*QFT-GIT	7	699.372

Table 5.29 Characteristics of TB cohort study in Changhua County, Taiwan

Data source	Period	Number	Sampling fraction (%)	
TB surveillance system	2009-2011			calculation
Population(a)	(person-year)	2,009,183*		
All TB cases (T)		2,420		
TB contacts database	2005-2011			
All contacts (C)		22,510		
Receiving TST (C_)		6,259		
Negative TST(C-)		2,758		
Positive TST(C+)		3,501		
Prevalent TB case		28		
Incident TB cases		45		
IGRA survey for general population	2011-2013			
Non-contacts (n-c)				
Negative TST		218	0.07152 (f_0)	3*n-c/a
Positive TST		261		
Contacts				
Negative TST(c-)		3	0.21171	7*c-/[C*(C+/C_)]
Positive TST(c+)		10	0.55595	7*c+/(C)(C+/C_)
TB cases (ct)		1	0.12397	3*ct/T
Matched case-control study	2012-2014			
TB cases (t)		207	25.66116	3*t/T
Contacts		780		
Negative TST (Cg-)		346	24.41796	7*Cg-/(C-/C_)
Positive TST (Cg+)		434	24.12821	7*Cg+/C(C+/C_)
Incident TB cases		2		

All persons with age ≥30 were enrolled in this study. Prevalent TB case: diagnosed as TB at survey periods; Incident TB cases: diagnosed as TB cases during follow-up period.

Table 5.30 **Univariate analysis for infection rate (λ) and conversion rate (μ) of three-state Markov model with and without covariates**

		Without Covariate	One covariate affecting λ or μ		One covariate affecting both λ and μ	
Transition /Covariates		*Estimate (95% CI)*	*Estimate (95% CI)*	*RR (95%CI)*	*Estimate (95% CI)*	*RR (95%CI)*
Free of TB infection → LTBI						
Infection rate (λ)		0.0168 (0.0157,0.0180)				
Age	30-44		reference		reference	
	45-64		-0.5316 (-0.6911 , -0.3637)	0.59 (0.50, 0.69)	-0.5348 (-0.7016 , -0.3653)	0.59 (0.50, 0.69)
	≥65		-0.6016 (-0.7975 , -0.4031)	0.55 (0.45, 0.67)	-0.6302 (-0.839 , -0.44)	0.53 (0.43, 0.64)
Sex	Male		0.403 (0.2656 , 0.5465)	1.50 (1.30, 1.73)	0.4007 (0.2631 , 0.5449)	1.49 (1.30, 1.72)
IGRA	Positive		0.4662 (0.3261 , 0.6149)	1.59 (1.39, 1.85)	0.4671 (0.3274 , 0.6107)	1.59 (1.39, 1.84)
LTBI →TB						
Conversion rate* (μ)		0.0113 (0.0098,0.0129)				
Age	30-44		reference		reference	
	45-64		0.7315 (0.2298 , 1.2374)	2.08 (1.26, 3.45)	0.7074 (0.2155 , 1.235)	2.03 (1.24, 3.44)
	≥65		1.8319 (1.3368 , 2.3399)	6.25 (3.81, 10.38)	1.846 (1.3505 , 2.354)	6.33 (3.86, 10.5)
Sex	Male		0.5283 (0.2405 , 0.8081)	1.70 (1.27, 2.24)	0.5277 (0.2427 , 0.8139)	1.70 (1.28, 2.26)
IGRA	Positive		0.7886 (0.4719 , 1.0822)	2.20 (1.60, 2.95)	0.7495 (0.4533 , 1.0474)	2.12 (1.57, 2.85)

CI: credible interval; **TB:** tuberculosis**; LTBI :** Latent TB infection; **IGRA:** Interferon-Gamma Release Assays; * Three-state Markov model without incorporating any covariates.

Table 5.31 **Multivariate analysis for estimating infection rates and conversion rate of three-state Markov model**

		Three covariates affecting λ or μ		Three covariates affecting both λ and μ	
Parameter		*Estimate (95% CI)*	*RR (95%CI)*	*Estimate (95% CI)*	*RR (95% CI)*
Free of TB infection → LTBI					
baseline (λ)				0.01883 (0.0163, 0.0220)	
Age	30-44	reference		reference	
	45-64	-0.5825 (-0.749 , -0.4193)	0.56 (0.47,1.51)	-0.5799 (-0.7515,-0.4194)	0.56 (0.47, 0.66)
	≥65	-0.7585 (-0.9606, -0.5481)	0.47 (0.38, 0.58)	-0.7896 (-0.9993,-0.5869)	0.45 (0.37, 0.56)
Sex	Male	0.3576 (0.2164, 0.495)	1.43 (1.24, 1.64)	0.3583 (0.2124,0.4975)	1.24 (1.24, 1.64)
IGRA	Positive	0.5307 (0.3776, 0.6699)	1.70 (1.46, 1.95)	0.5371 (0.3961,0.6933)	1.71 (1.49, 2.00)
LTBI →TB					
baseline (μ)				0.0032 (0.0020, 0.0051)	
Age	30-44	reference		reference	
	45-64	0.6126 (0.0956,1.1364)	1.85 (1.10, 3.12)	0.4256 (-0.061, 0.9224)	1.53 (0.94, 2.52)
	≥65	1.6308 (1.1248,2.176)	5.11 (3.08, 8.81)	1.4886 (0.9787,1.9605)	4.43 (2.66, 7.10)
Sex	Male	0.4233 (0.1038,0.719)	1.53 (1.11, 2.05)	0.4104 (0.105, 0.7157)	1.51 (1.11, 2.05)
IGRA	Positive	0.4934 (0.1716,0.8105)	1.64 (1.19, 2.25)	0.4578 (0.1428, 0.7743)	1.58 (1.15, 2.17)

CI: credible interval; **TB:** tuberculosis**; LTBI :** Latent TB infection; **IGRA:** Interferon-Gamma Release Assays

Table 5.32 **Model selection for multivariate analysis in three states Markov model.**

Model	$\bar{D}$	**pD**	**DIC**
Incorporating covariates on infection rate: $\lambda_i = \lambda_0 \exp\left(\beta_{11} \times age\ group1_i + \beta_{12} \times age\ group2_i + \beta_{13} \times sex_i + \beta_{14} \times IGRA_i\right)$	4387.77	6.97	4394.73
Incorporating covariates on conversion rate: $\mu_i = \mu_0 \exp\left(\beta_{21} \times age\ group1_i + \beta_{22} \times age\ group2_i + \beta_{23} \times sex_i + \beta_{24} \times IGRA_i\right)$	4415.11	6.93	4422.04
Incorporating covariates on rates of infection and conversion: $\lambda_i = \lambda_0 \exp\left(\beta_{11} \times age\ group1_i + \beta_{12} \times age\ group2_i + \beta_{13} \times sex_i + \beta_{14} \times IGRA_i\right)$ $\mu_i = \mu_0 \exp\left(\beta_{21} \times age\ group1_i + \beta_{22} \times age\ group2_i + \beta_{23} \times sex_i + \beta_{24} \times IGRA_i\right)$	4298.77	10.92	4309.68

DIC: Deviance Information Criterion; $\bar{D}$ **:** Expectation of deviance**; pD** : the effective number of parameters

Table 5.33 Estimated time to the LTB size of 10 according to TB cohort study with simulation of birth-death process

Covariates			Parameters		$\hat{T}$ Mean (SD) days			
Gender	**Age group**	**IGRA status**	λ	μ	N_0=1	N_0=2	N_0=3	N_0=5
Female	<45	Negative	0.0181	0.0034	51.53(33.38)	41.58(17.15)	30.32(12.90)	16.50(7.87)
Female	45-65	Negative	0.0101	0.00524	56.01(63.2)	66.41(42.28)	47.28(28.55)	25.73(17.77)
Female	>=65	Negative	0.00828	0.0151	33.87 (72.54)	62.85(69.76)	92.92(83.10)	66.61(61.37)
Male	<45	Negative	0.0264	0.00505	34.82 (22.33)	28.51(11.75)	20.78(8.84)	11.31(5.39)
Male	45-65	Negative	0.0146	0.00777	38.13(42.10)	44.70(28.34)	32.68(19.49)	18.27(12.21)
Male	>=65	Negative	0.0121	0.0224	19.64(44.34)	38.00(50.41)	58.44(51.37)	51.15(44.95)
Female	<45	Positive	0.032	0.00522	30.31(18.09)	23.93(10.12)	16.58(6.87)	9.45(4.14)
Female	45-65	Positive	0.0177	0.00801	38.56(35.46)	40.61(23.64)	27.60(15.48)	15.95(9.97)
Female	>=65	Positive	0.0146	0.023	16.08(35.17)	37.88(39.51)	28.43(28.80)	20.59(30.23)
Male	<45	Positive	0.0465	0.00892	19.77(12.68)	16.19(6.67)	11.80(5.02)	6.42(3.06)
Male	45-65	Positive	0.0258	0.0119	27.59(24.10)	26.28(15.50)	18.95(10.71)	10.55(6.01)
Male	>=65	Positive	0.0213	0.0341	13.33(27.00)	26.83(27.91)	23.91(20.81)	16.00(19.78)
Overall			0.0168	0.0113	27.86 (33.73)	36.85(28.63)	25.01(16.87)	15.9(15.15)

$\hat{T}$: Time to size of 10 cases; N_0 : initial number of cases; SD: standard deviation.

Table 5.34 **Estimated time to the TB size of 30 according to TB cohort study with simulation of birth-death process**

Covariates			Parameters		$\hat{T}$ Mean (SD) days			
Gender	**Age group**	**IGRA status**	λ	μ	N_0=1	N_0=5	N_0=10	N_0=15
Female	<45	Negative	0.0181	0.0034	72.02(39.50)	42.76(10.06)	25.63(6.19)	15.94(4.22)
Female	45-65	Negative	0.0101	0.00524	95.18(77.70)	68.56(25.95)	41.22(16.03)	25.92(9.29)
Female	>=65	Negative	0.00828	0.0151	31.37(76.42)	72.27(60.40)	48.51(31.44)	34.82(36.44)
Male	<45	Negative	0.0264	0.00505	49.78(27.27)	29.25(6.93)	17.55(4.27)	10.91(2.91)
Male	45-65	Negative	0.0146	0.00777	61.59(56.16)	47.55(17.82)	28.56(10.90)	18.68(5.46)
Male	>=65	Negative	0.0121	0.0224	18.95(48.27)	56.95(55.13)	19.41(24.63)	19.13(10.52)
Female	<45	Positive	0.032	0.00522	42.19(22.26)	23.89(5.51)	14.44(3.62)	9.00(2.37)
Female	45-65	Positive	0.0177	0.00801	51.53(45.68)	41.50(14.49)	25.23(6.52)	15.21(5.54)
Female	>=65	Positive	0.0146	0.023	23.55(49.62)	43.65(34.45)	27.00(24.52)	18.86(24.37)
Male	<45	Positive	0.0465	0.00892	28.26(15.48)	16.61(3.94)	9.97(2.43)	6.19(1.65)
Male	45-65	Positive	0.0258	0.0119	35.40(31.47)	28.29(9.77)	17.38(4.48)	10.32(3.79)
Male	>=65	Positive	0.0213	0.0341	18.00(34.46)	29.40(22.81)	17.90(14.63)	11.06(11.46)
Overall			0.0168	0.0113	48.47(47.57)	43.48(13.32)	25.15(11.22)	15.92(7.57)

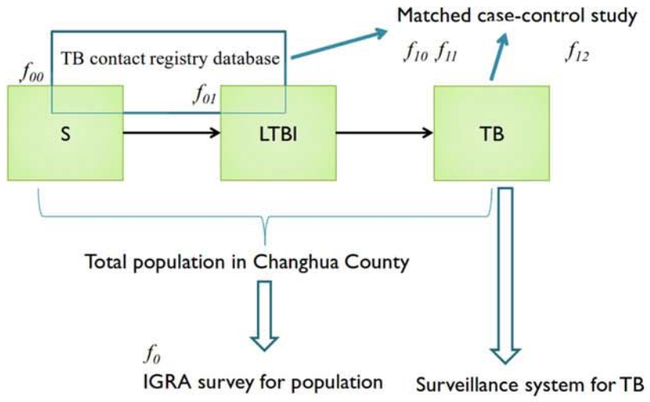
$\hat{T}$: Time to state10; SD: standard deviation.

Table 5.35 **Estimated extinct probability of TB by generating function of birth-death process**

Covariates			$\hat{q}$									
Gender	**Age group**	**IGRA status**	N_0=1	N_0=2	N_0=3	N_0=4	N_0=5	N_0=6	N_0=7	N_0=8	N_0=9	N_0=10
Female	<45	Negative	0.1878	0.0353	0.0066	0.0012	$2.3*10^{-4}$	$4.4*10^{-5}$	$8.3*10^{-6}$	$1.6*10^{-6}$	$1.55*10^{-6}$	$5.47*10^{-8}$
Female	45-65	Negative	0.5188	0.2678	0.1383	0.0714	0.0369	0.0191	0.0099	0.0051	0.0026	0.0014
Female	>=65	Negative	1	1	1	1	1	1	1	1	1	1
Male	<45	Negative	0.1913	0.0366	0.007	0.0013	$2.6*10^{-4}$	$4.9*10^{-5}$	$9.4*10^{-6}$	$1.8*10^{-6}$	$3.4*10^{-7}$	$6.6*10^{-8}$
Male	45-65	Negative	0.5322	0.2830	0.1505	0.0086	0.0426	0.0227	0.0039	0.0018	$7.9*10^{-4}$	$3.6*10^{-4}$
Male	>=65	Negative	1	1	1	1	1	1	1	1	1	1
Female	<45	Positive	0.1631	0.0266	0.0043	$7.2*10^{-4}$	$1.2*10^{-4}$	$1.9*10^{-5}$	$3.1*10^{-6}$	$5.0*10^{-7}$	$8.2*10^{-8}$	$1.3*10^{-8}$
Female	45-65	Positive	0.4525	0.2048	0.0927	0.0419	0.019	0.0086	0.0039	0.0018	$8.0*10^{-4}$	$3.6*10^{-4}$
Female	>=65	Positive	1	1	1	1	1	1	1	1	1	1
Male	<45	Positive	0.1918	0.0368	0.0071	0.0014	$2.6*10^{-4}$	$5.0*10^{-5}$	$9.6*10^{-6}$	$1.8*10^{-6}$	$3.5*10^{-7}$	$6.7*10^{-8}$
Male	45-65	Positive	0.4612	0.2127	0.0981	0.0453	0.0209	0.0096	0.0044	0.0020	$9.4*10^{-4}$	$4.4*10^{-4}$
Male	>=65	Positive	1	1	1	1	1	1	1	1	1	1
		Overall	0.6726	0.4517	0.3033	0.2037	0.1368	0.0919	0.0617	0.0415	0.0278	0.0187

N_0 : initial number of cases: $\hat{q}$: estimated extinction probability by generating function

Figure3.1The framework of our TB case-cohort study design



f_0 : sampling fraction of subjects belong to general cohort without contact history

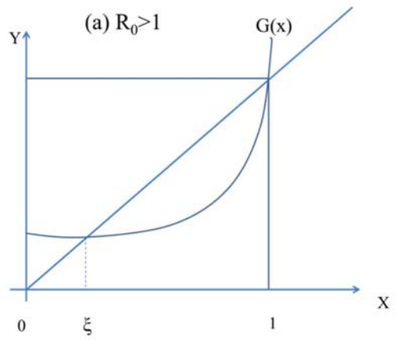
f_{00} : sampling fraction of subjects belong to the contact cohort free of TB infection

f_{01} : sampling fraction of subjects belong to the contact cohort with LTBI

f_{02} : sampling fraction of subjects belong to the cohort of TB cases

Figure3.2 Graph of the root of generating function

(a) $R_0 = \mu > 1$; (b) $R_0 = \mu \leq 1$



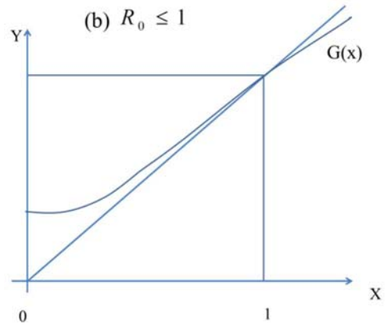
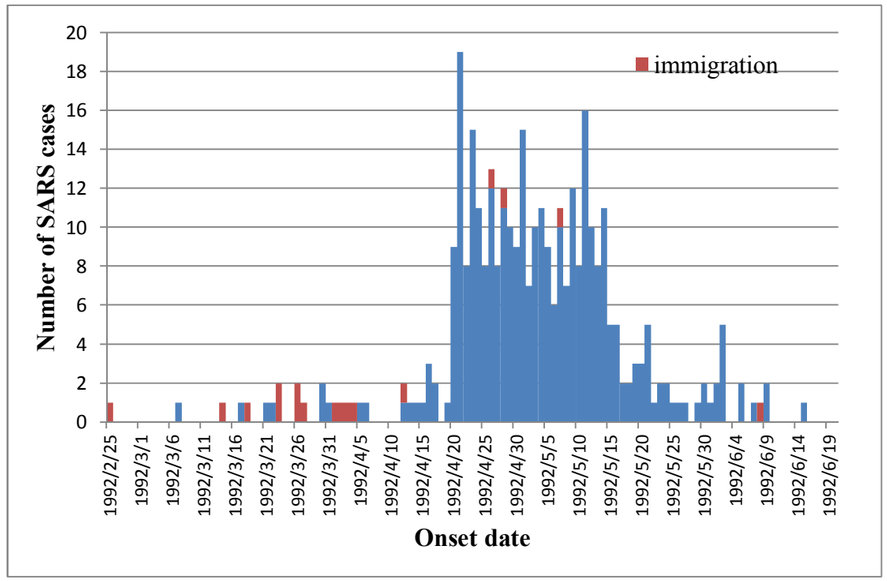
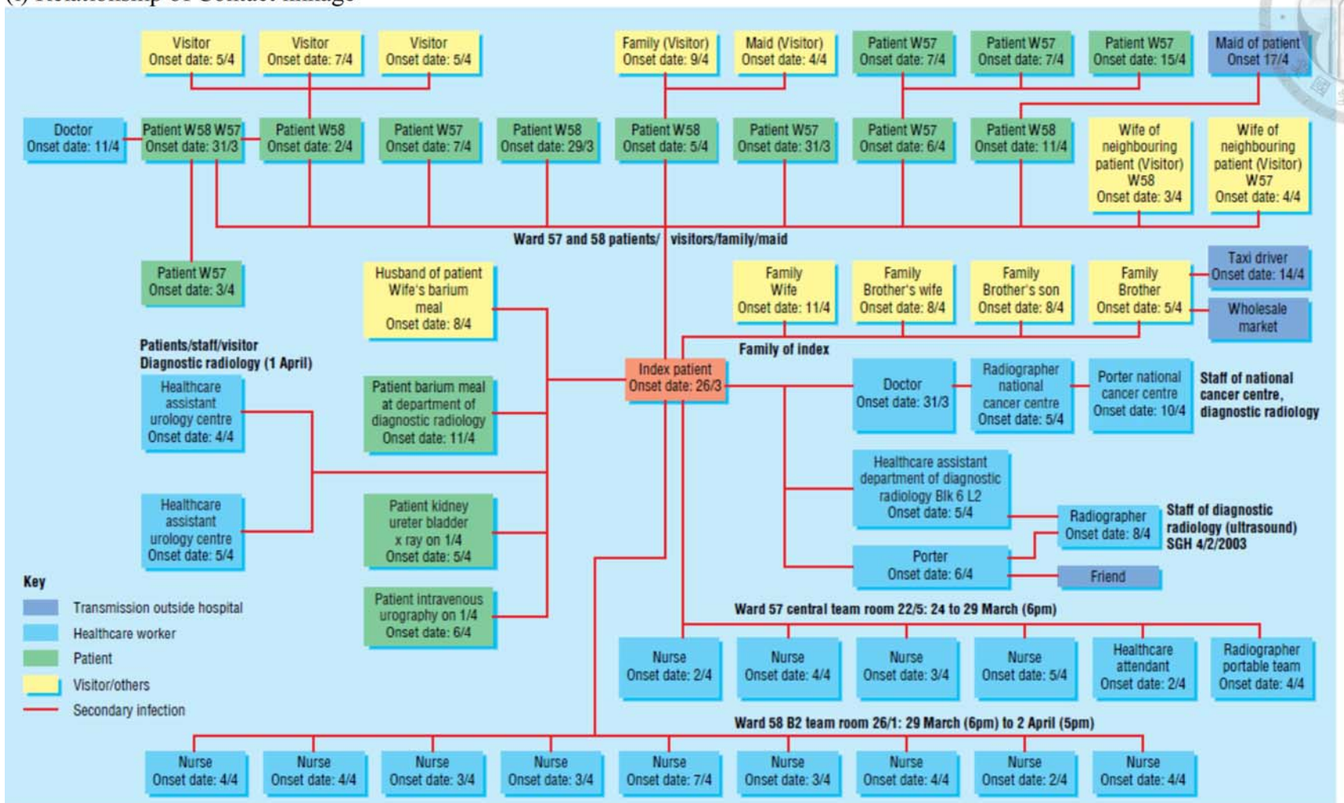


Figure 4. 1**The outbreak of SARS**

(a) The number of patients with SARS of Taiwan in 2003



(b) The ooutbreak of SARS in a hospital of Singapore
(i) Relationship of Contact linkage



(ii) Epidemiologic Curve for SRAS Cases in a hospital of Singapore

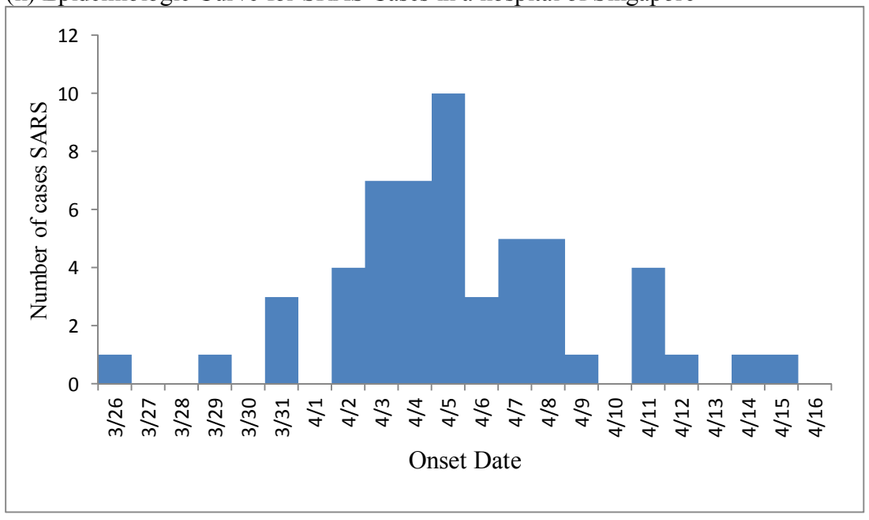
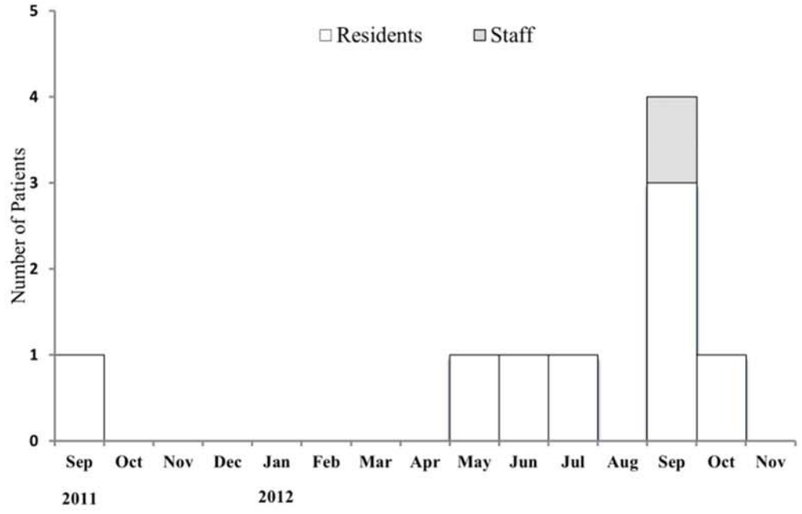


Figure 4. 2**Pulmonary Tuberculosis in Taiwan**

(a) Epidemiologic Curve for Pulmonary Tuberculosis Cases in a Long-term Care Facility



(b) The number of TB cases in Changhua County from 1982 to 2012

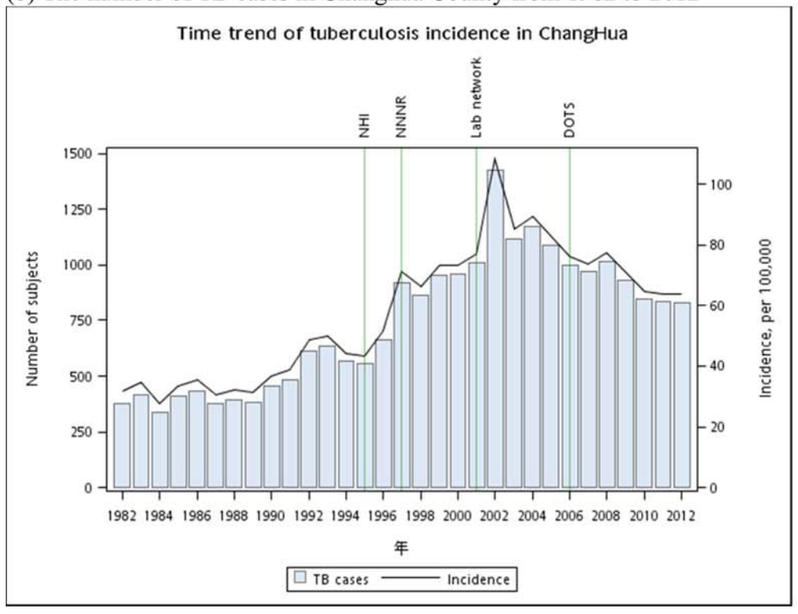


Figure 4. 3 **Flowchart of subjects IGRA survey for general population from 2011 to 2013 in Changhua County.**

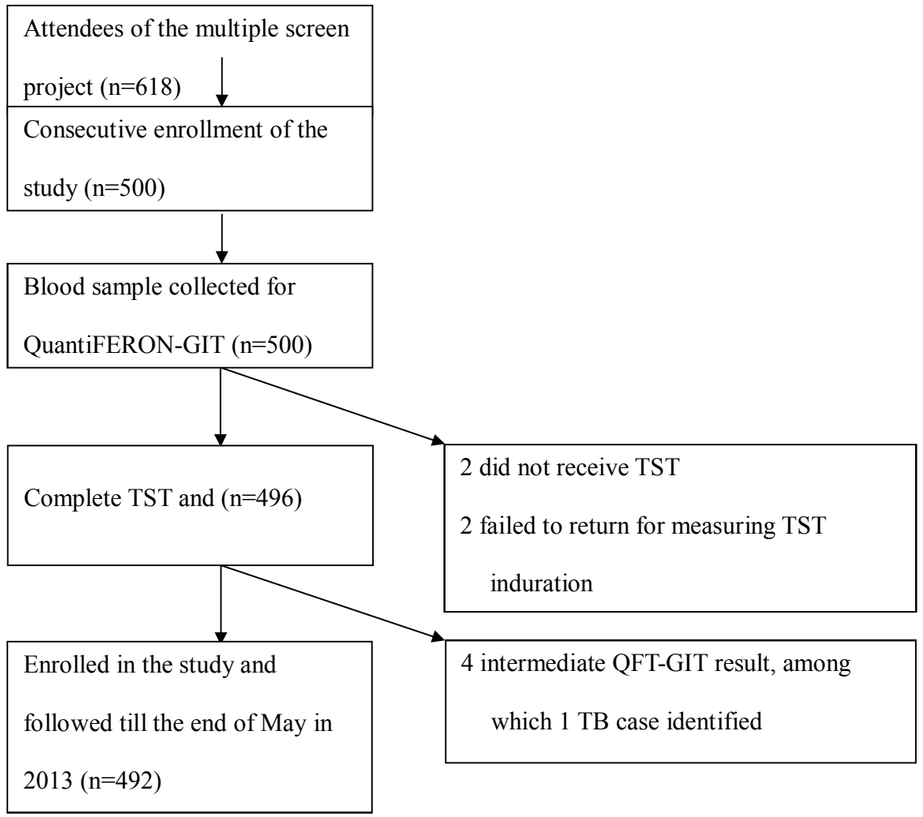
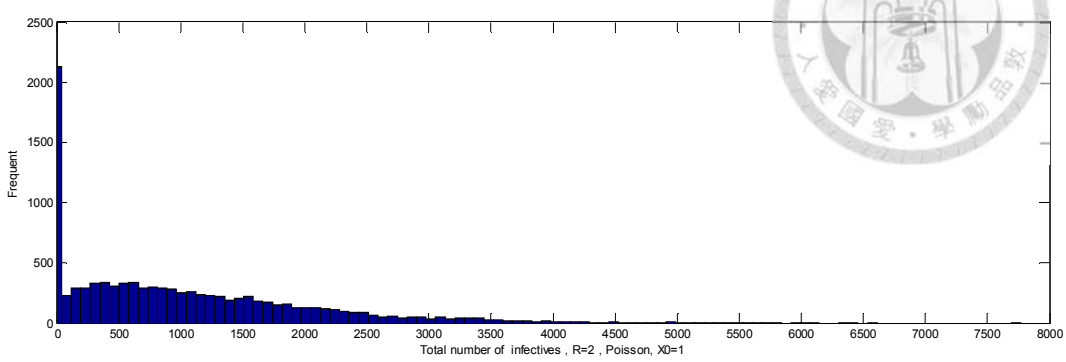


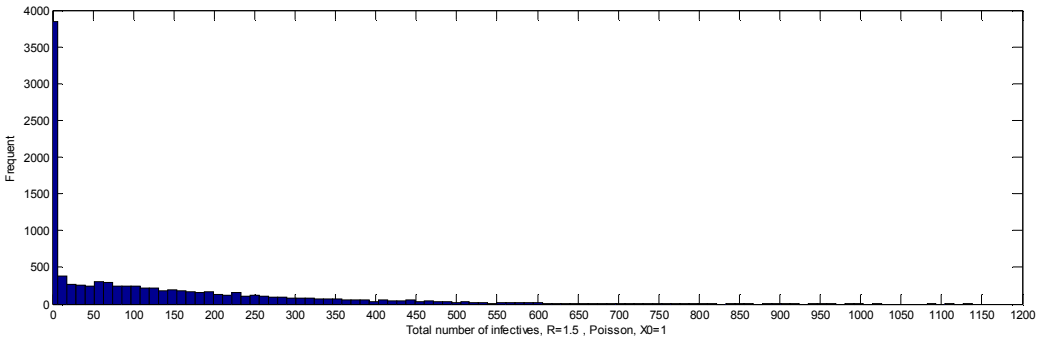
Figure 5. 1

Total number of all infectives after 9 generations, Poisson offspring distribution

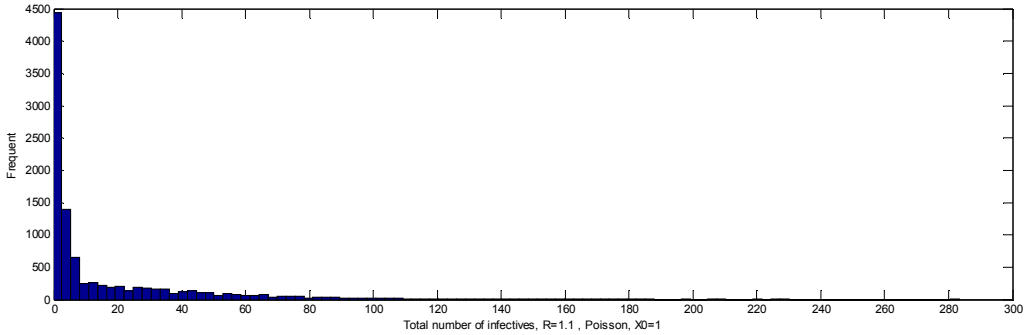
(a) R=2 and X_0=1



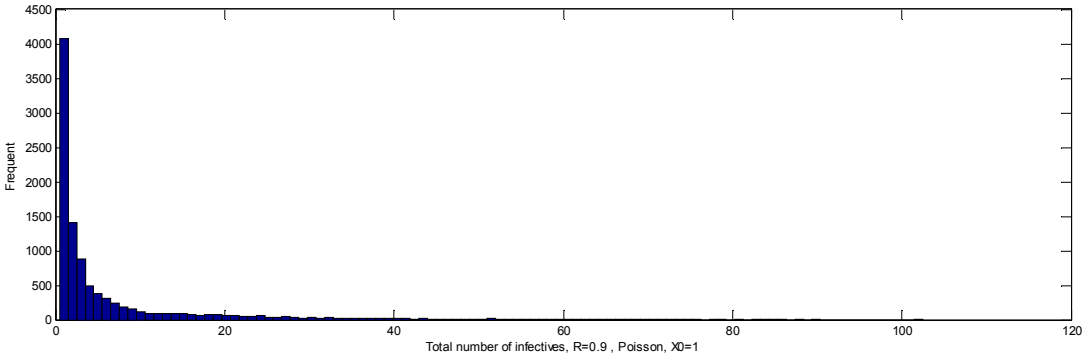
(b) R=1.5 and X_0=1



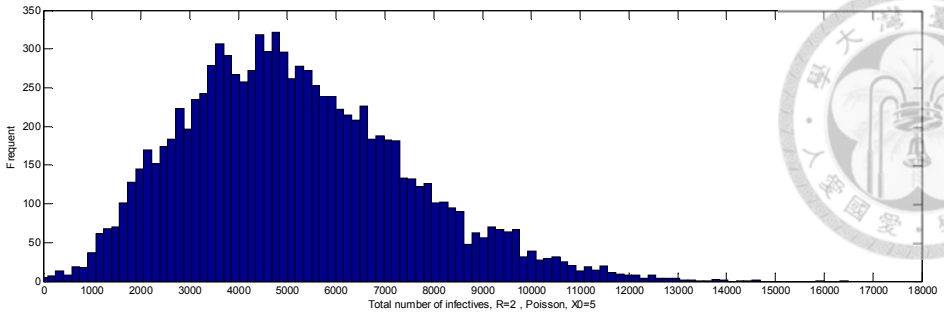
(c) R=1.1 and X_0=1



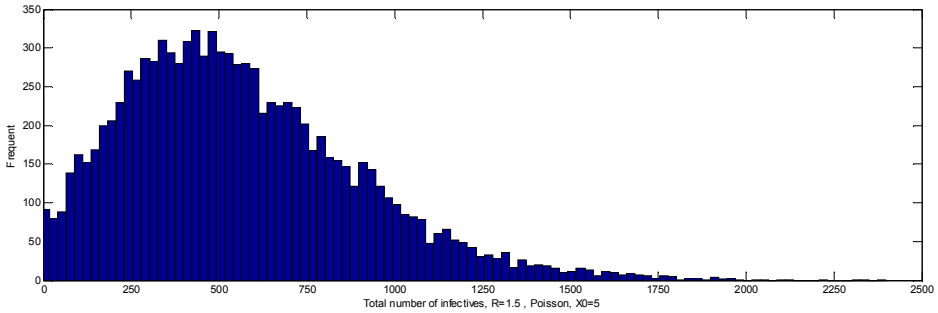
(d) R=0.9 and X_0=1



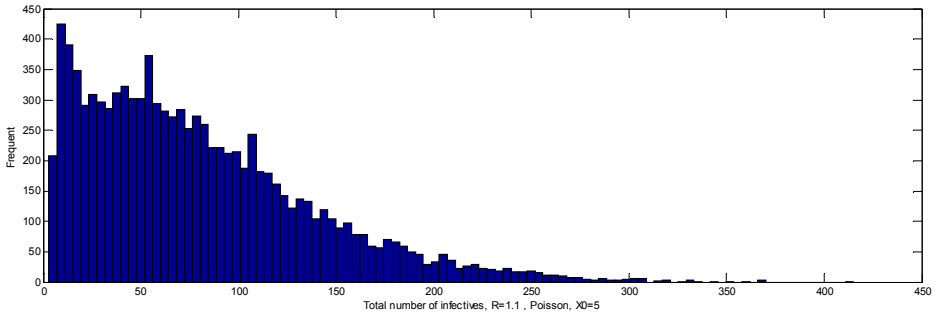
(e) R=2 and X_0=5



(f) R=1.5 and X_0=5



(g) R=1.1 and X_0=5



(h) R=0.9 and X_0=5

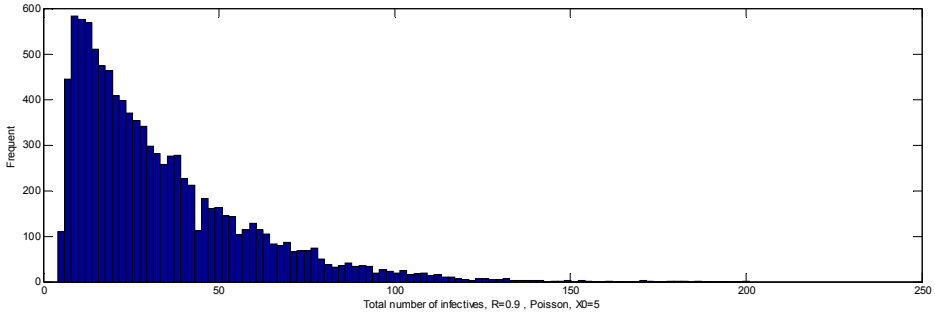
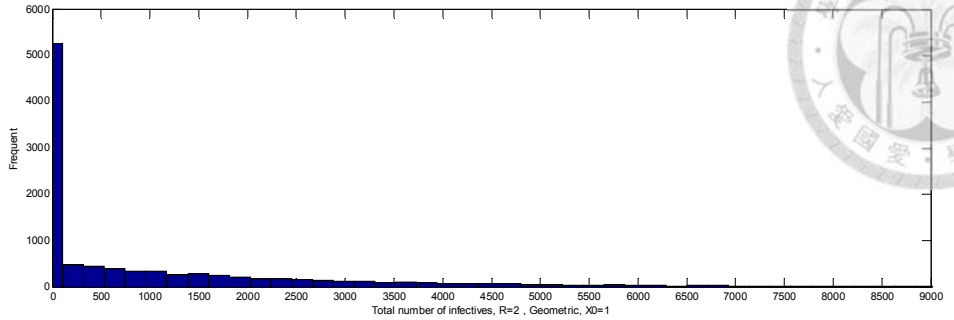
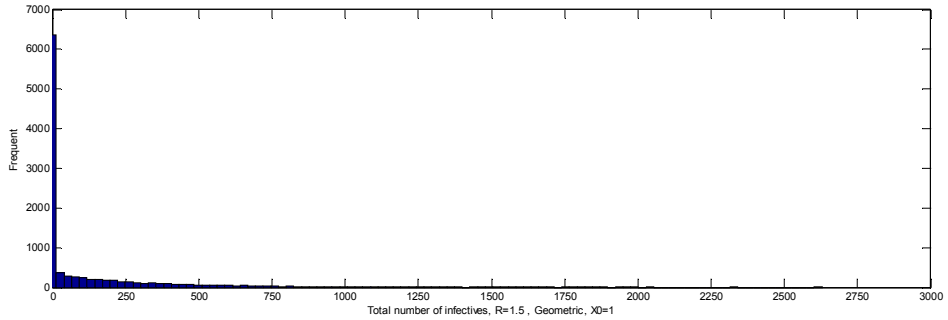


Figure 5. 2 Total number of all infectives after 9 generations, Geometric offspring distribution

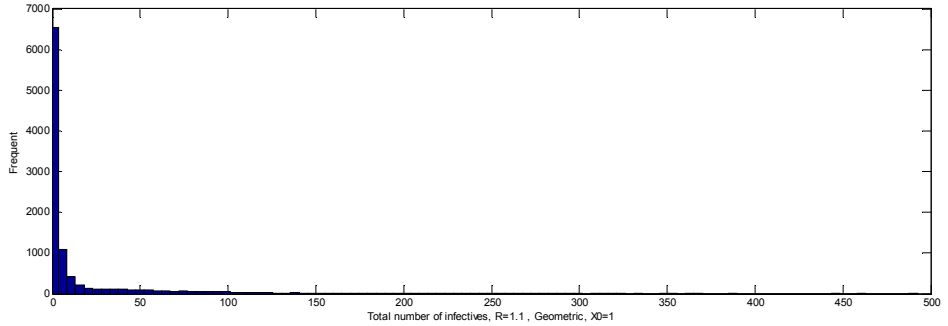
(a) R=2 and X_0=1



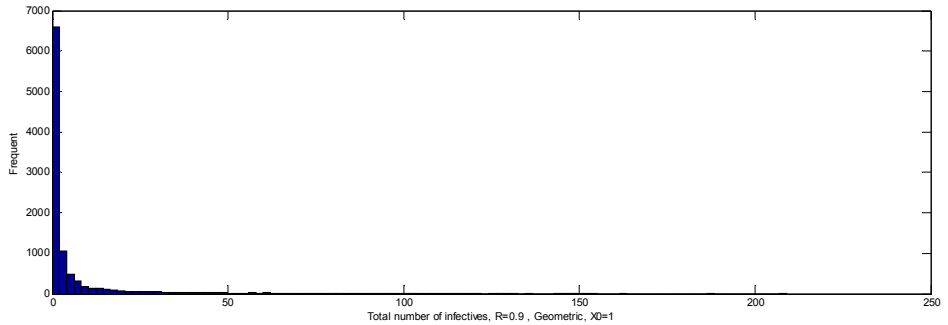
(b) R=1.5 and X_0=1



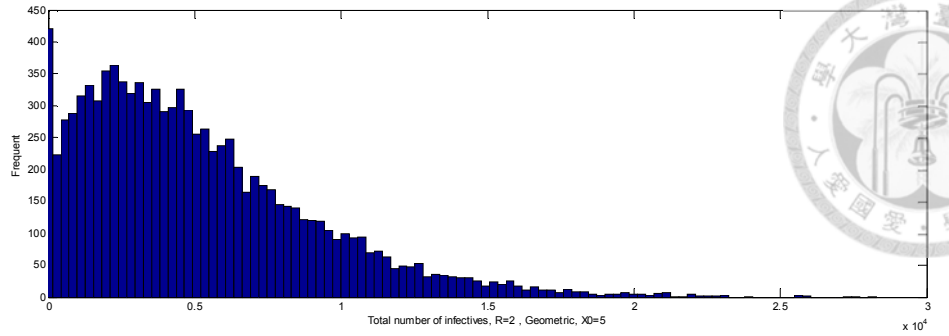
(c) R=1.1 and X_0=1



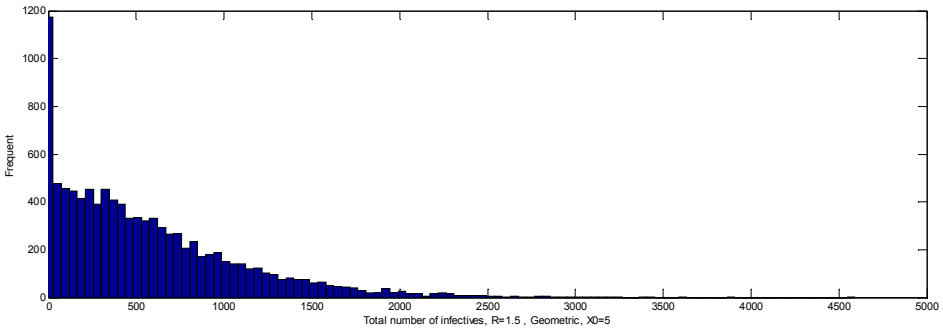
(d) R=0.9 and X_0=1



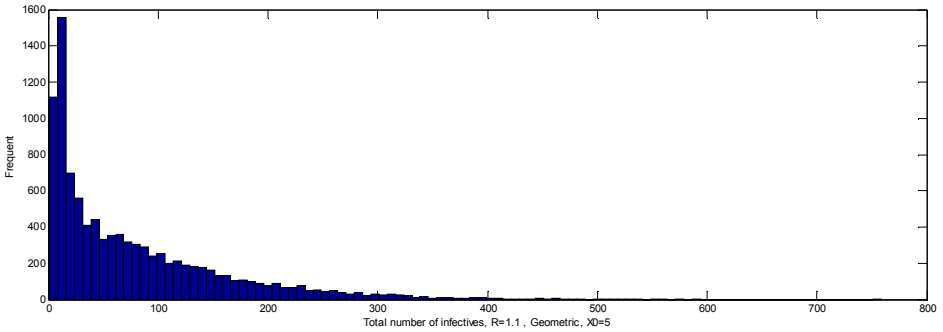
(e) R=2 and X_0=5



(f) R=1.5 and X_0=5



(g) R=1.1 and X_0=5



(h) R=0.9 and X_0=5

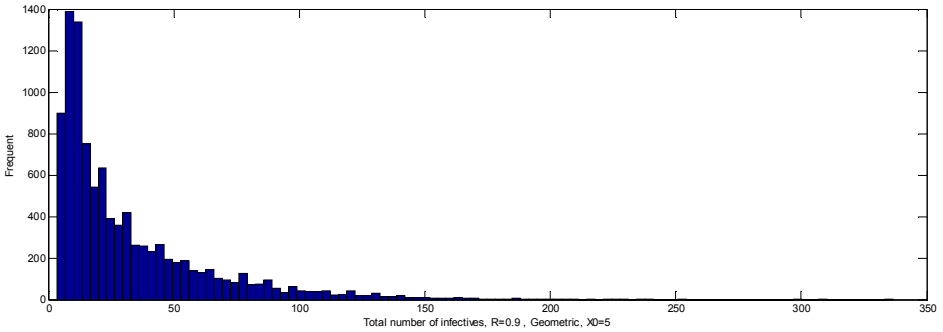
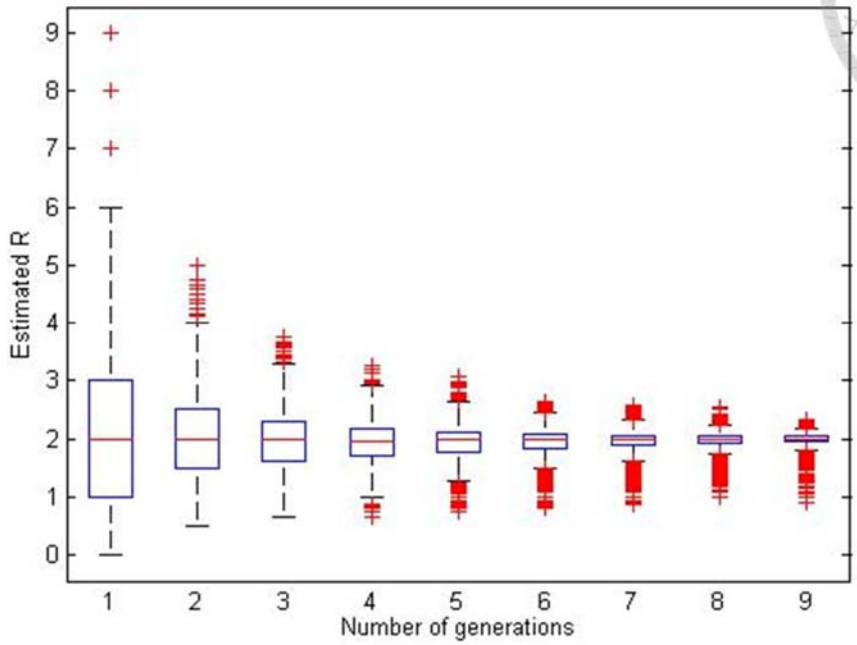
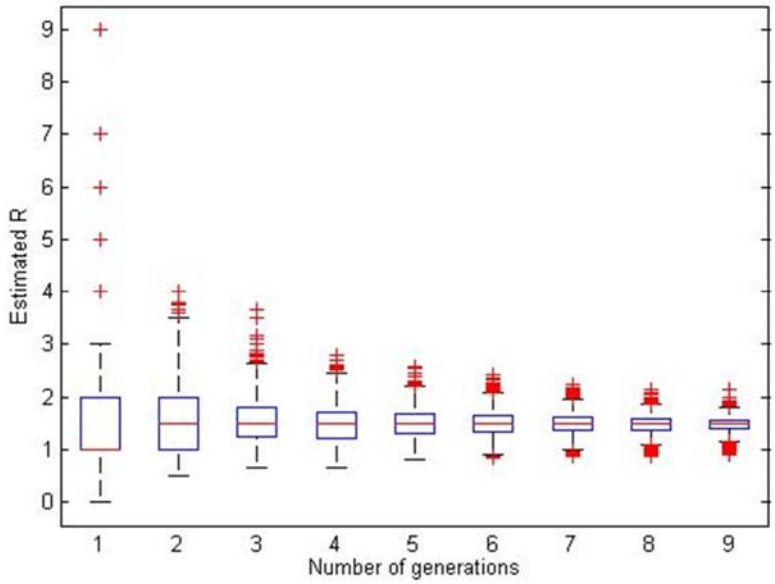


Figure 5. 3 The relations between estimated R_0 and the number of generations which was obtained from the simulated data with differential offspring distributions.

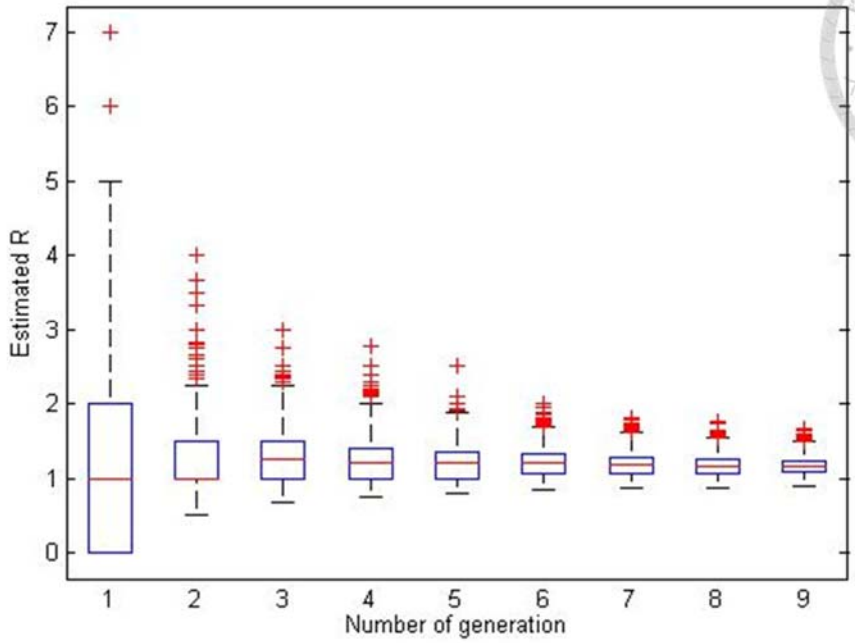
(a) Poisson distribution (λ=2) with initial one infected case.



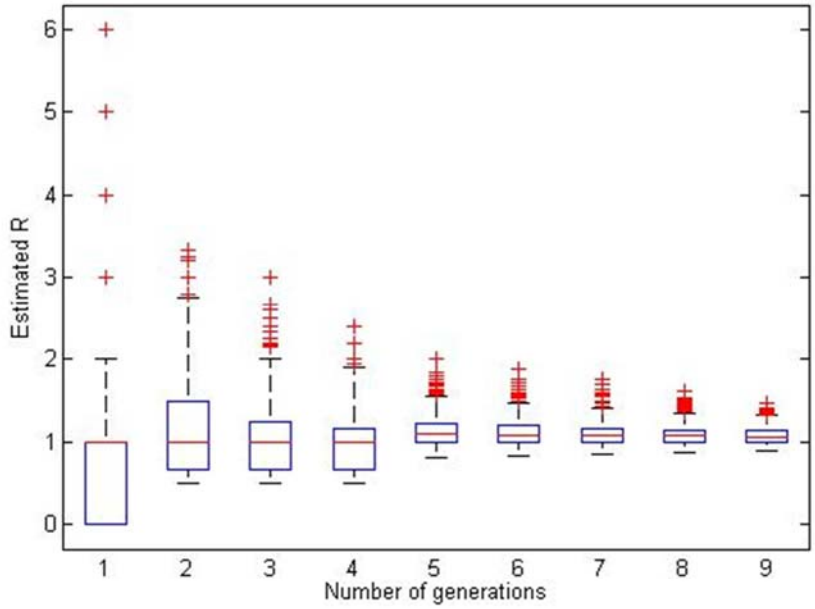
(b) Poisson distribution (λ=1.5) with initial one infected case.



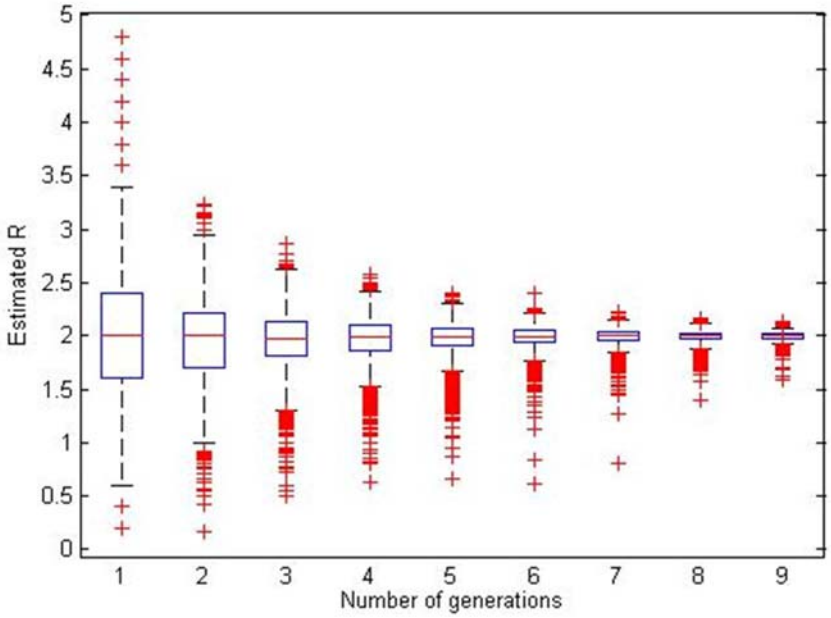
(b)Poisson distribution (λ=1.1) with initial one infected case.



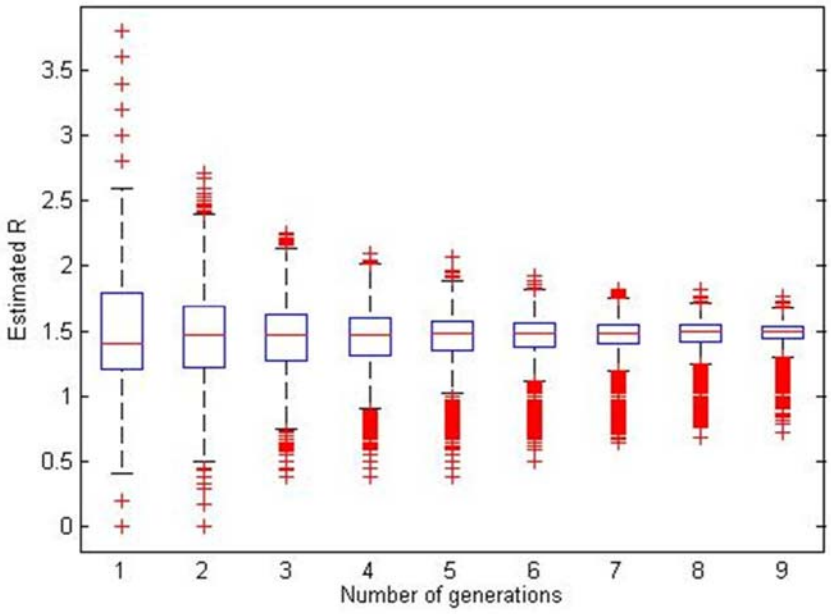
(c) Poisson distribution (λ=0.9) with initial one infected case.



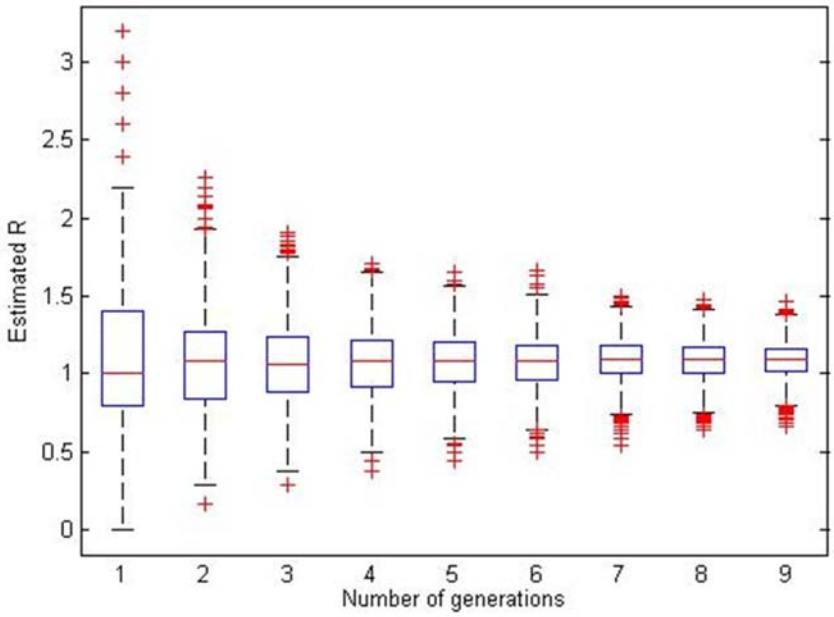
(d) Poisson distribution (λ=2) with initial five infected case.



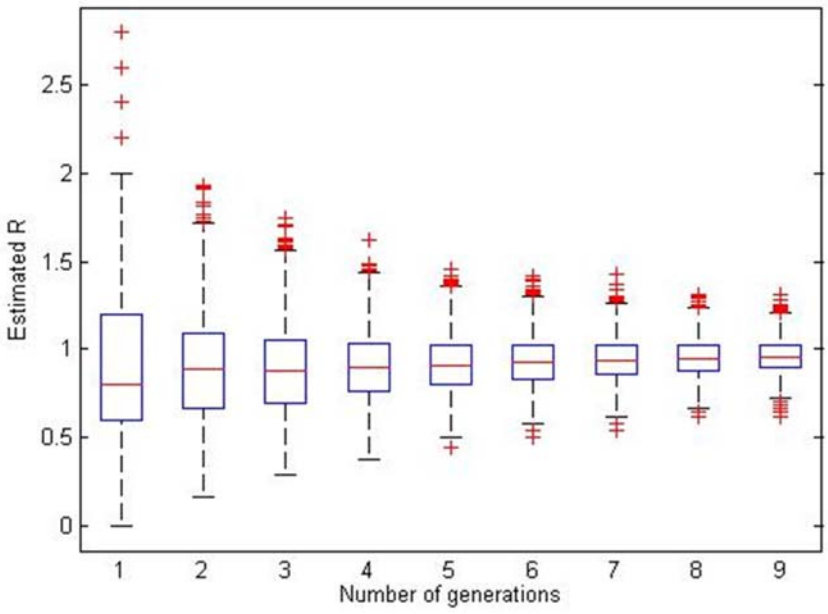
(e) Poisson distribution (λ=1.5) with initial five infected case.



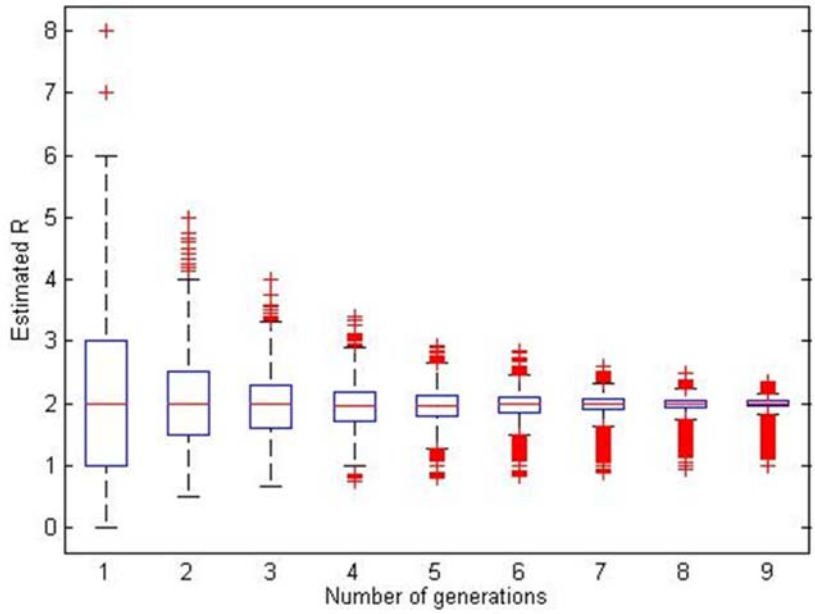
(f)Poisson distribution (λ=1.1) with initial five infected case.



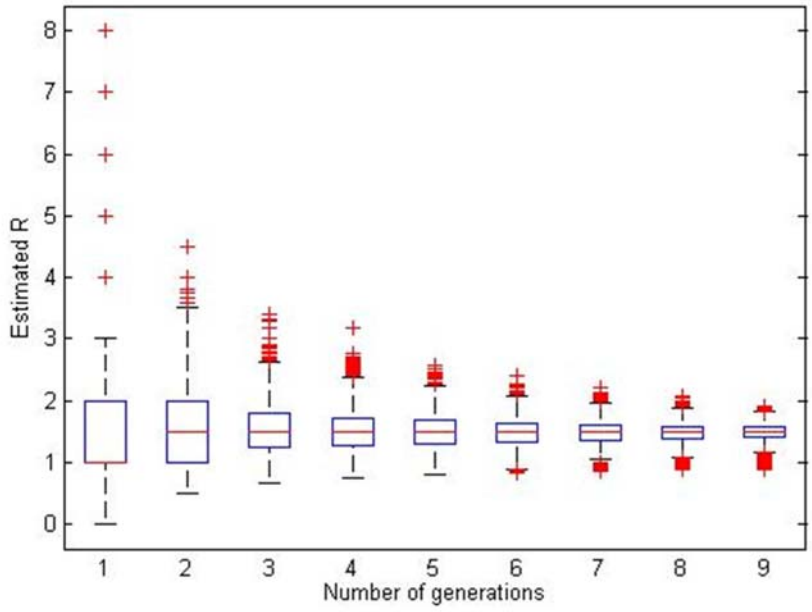
(g) Poisson distribution (λ=0.9) with initial five infected case.



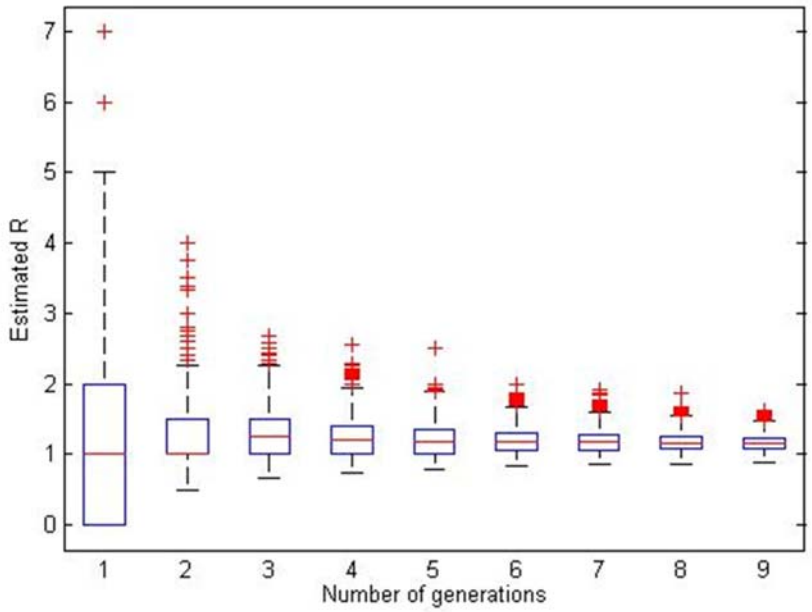
(h) Binomial distribution (n=1000, p=0.002) with initial one infected case.



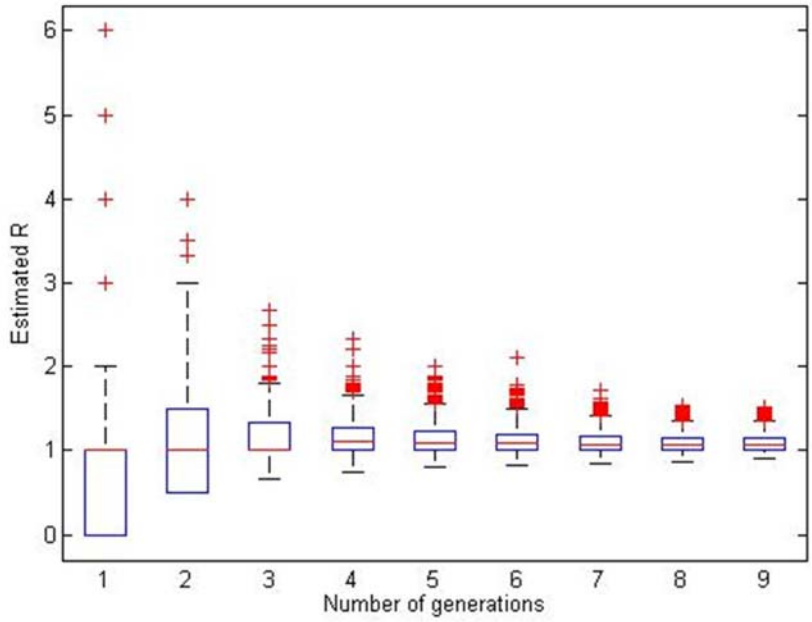
(i) Binomial distribution (n=1000, p=0.0015) with initial one infected case.



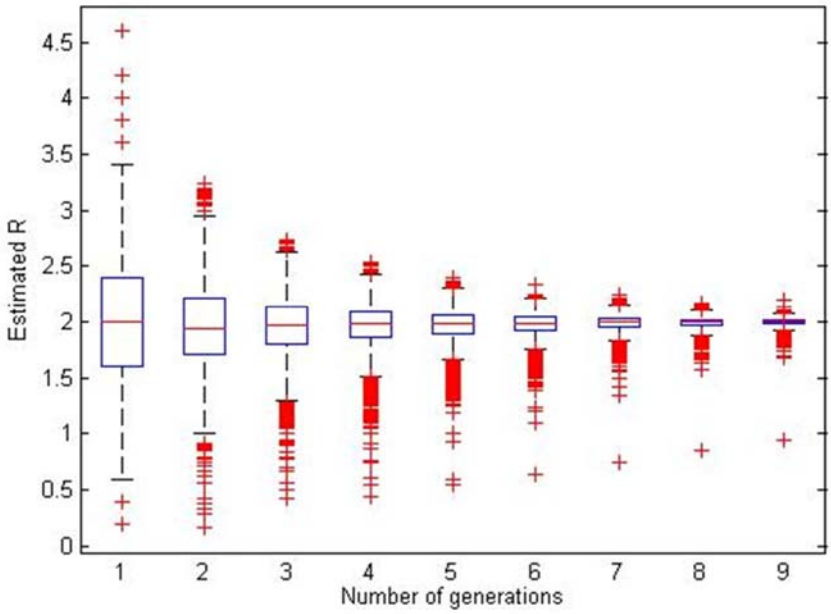
(j) Binomial distribution (n=1000, p=0.0011) with initial one infected case.



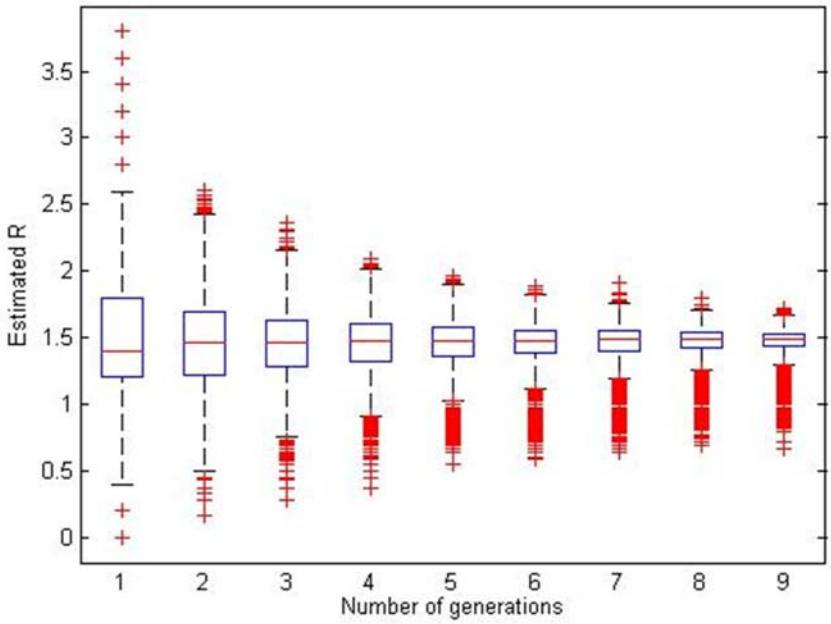
(k) Binomial distribution (n=1000, p=0.0009) with initial one infected case.



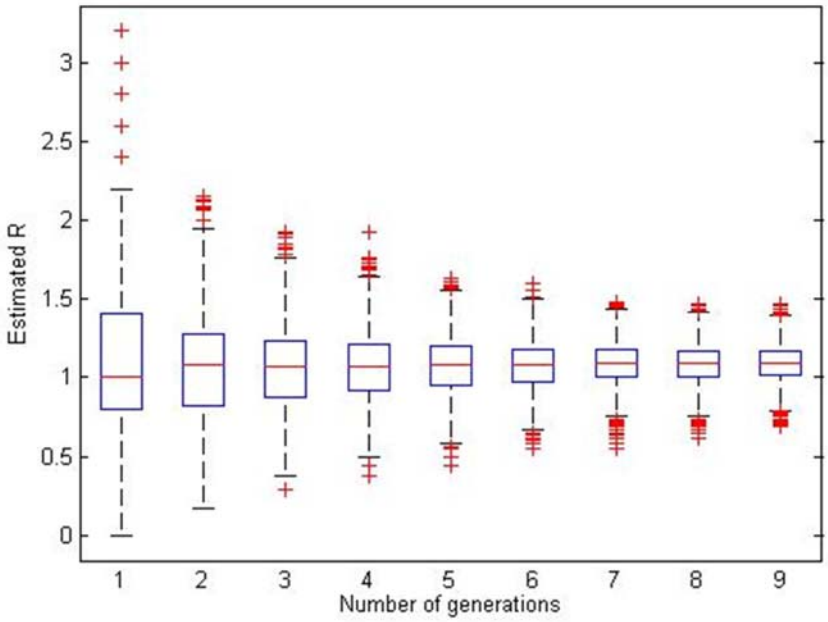
(l) Binomial distribution (n=1000, p=0.002) with initial five infected case.



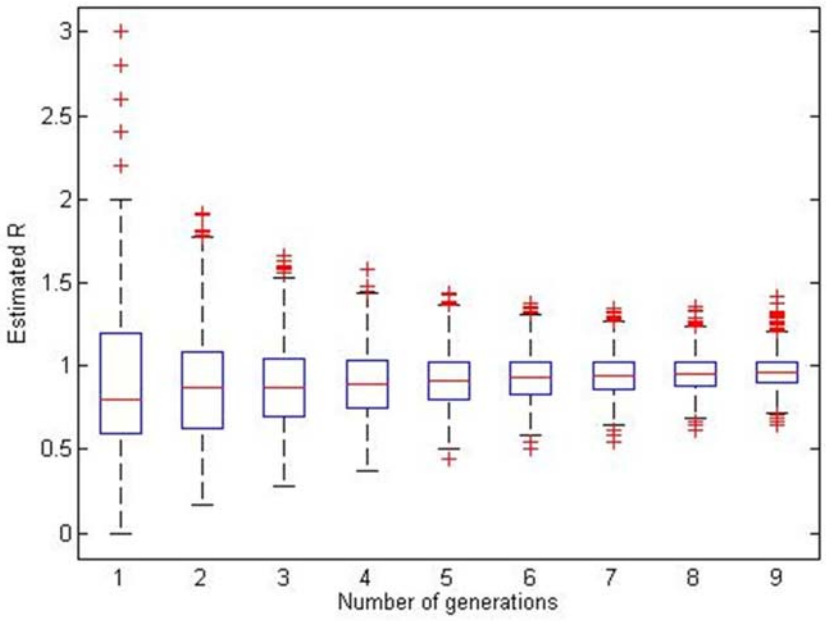
(m) Binomial distribution (n=1000, p=0.0015) with initial five infected case.



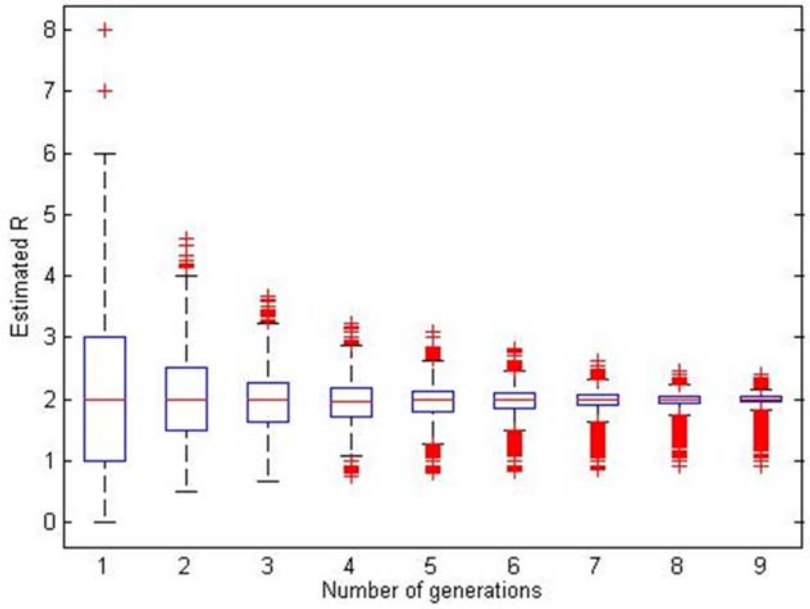
(n) Binomial distribution (n=1000, p=0.0011) with initial five infected case.



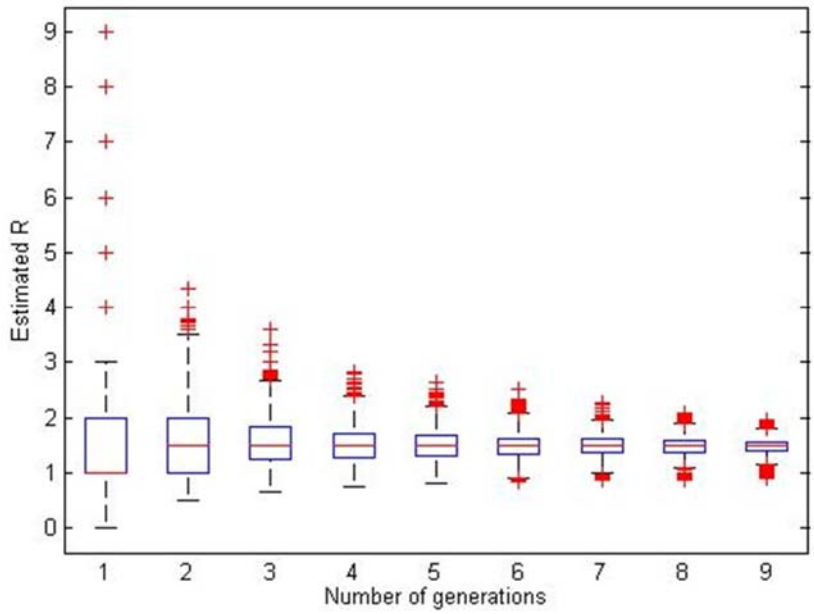
(o) Binomial distribution (n=1000, p=0.0009) with initial five infected case.



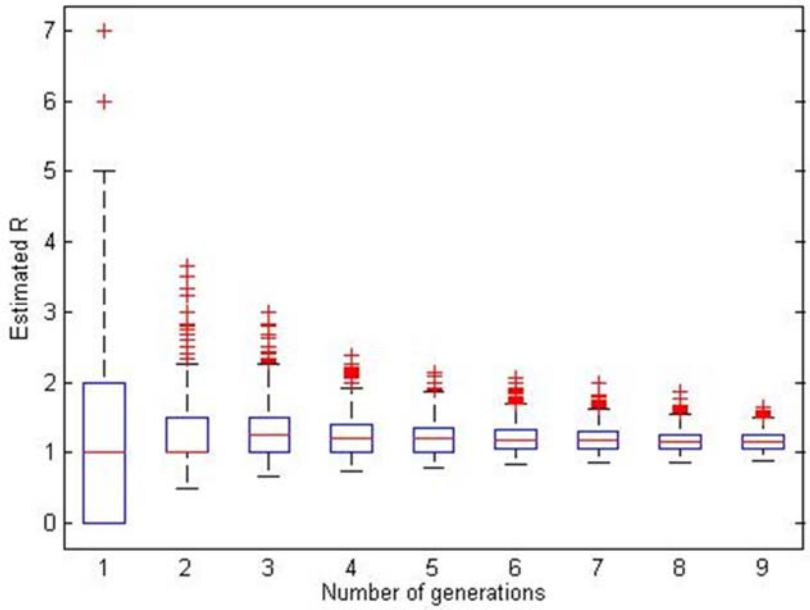
(p) Binomial distribution (n=100, p=0.02) with initial one infected case.



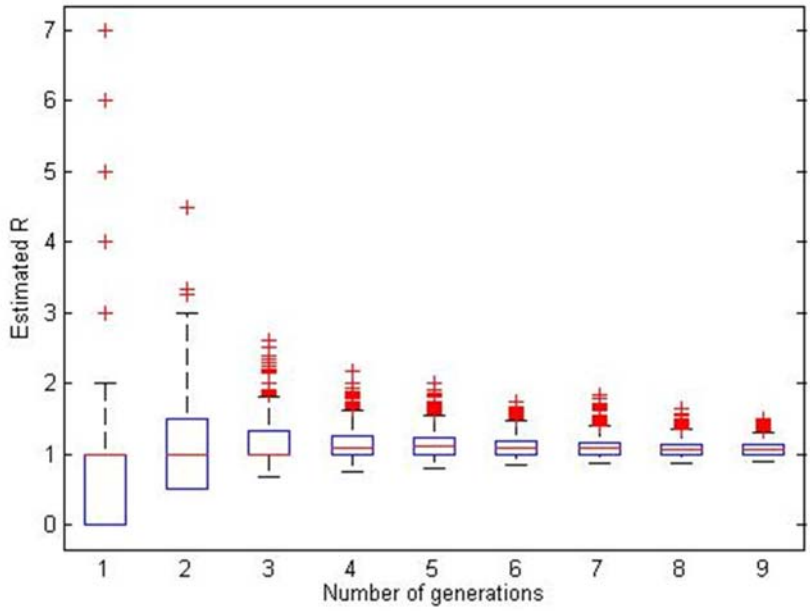
(q) Binomial distribution (n=100, p=0.015) with initial one infected case.



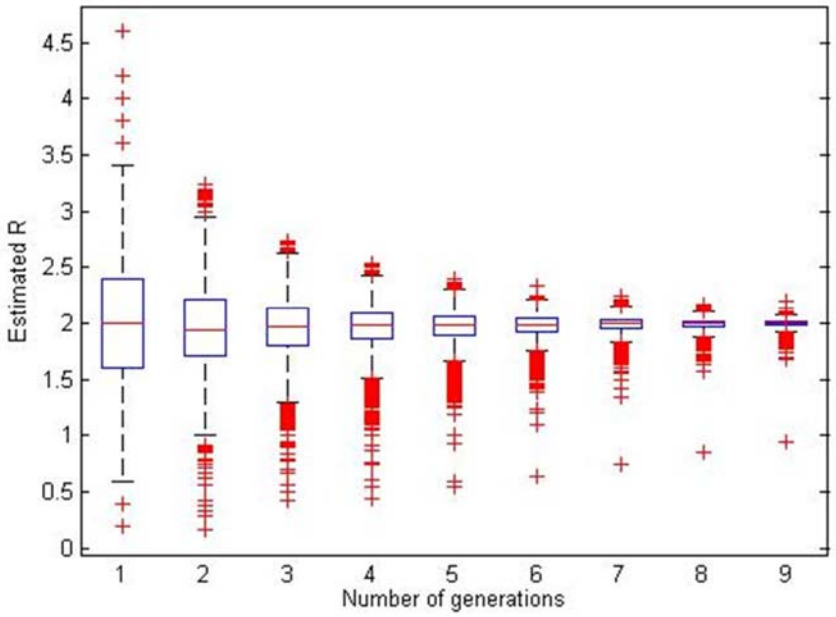
(r) Binomial distribution (n=100, p=0.011) with initial one infected case.



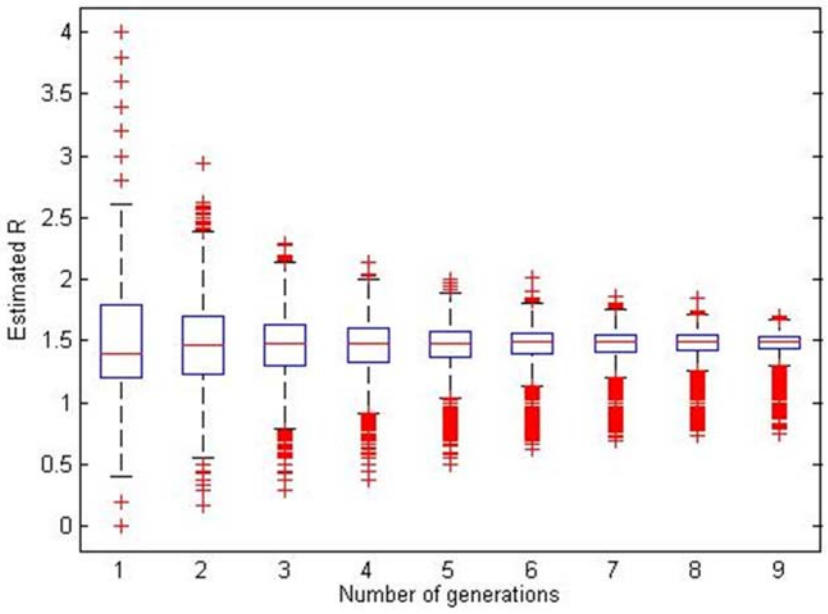
(s) Binomial distribution (n=100, p=0.009) with initial one infected case.



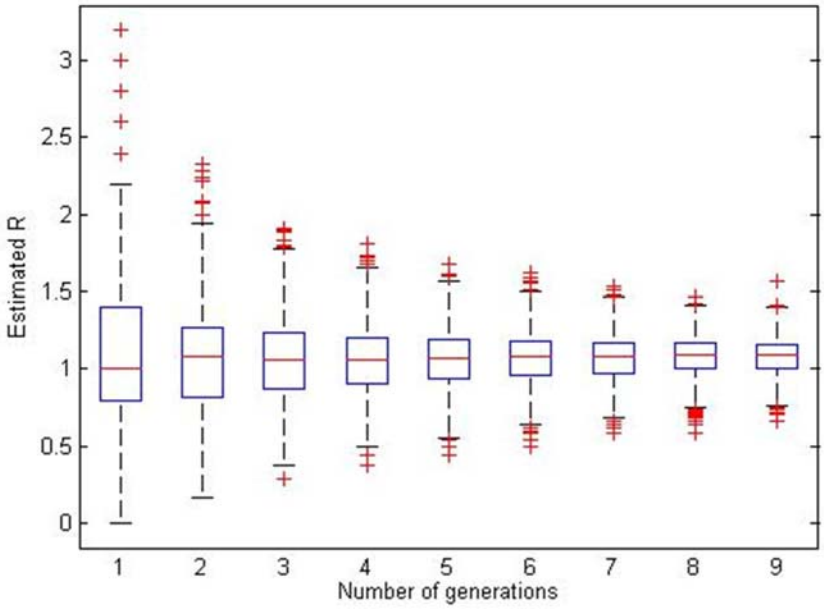
(t) Binomial distribution (n=100, p=0.02) with initial five infected case.



(u) Binomial distribution (n=100, p=0.015) with initial five infected case.



(v) Binomial distribution (n=100, p=0.011) with initial five infected case.



(w) Binomial distribution (n=100, p=0.009) with initial five infected case.

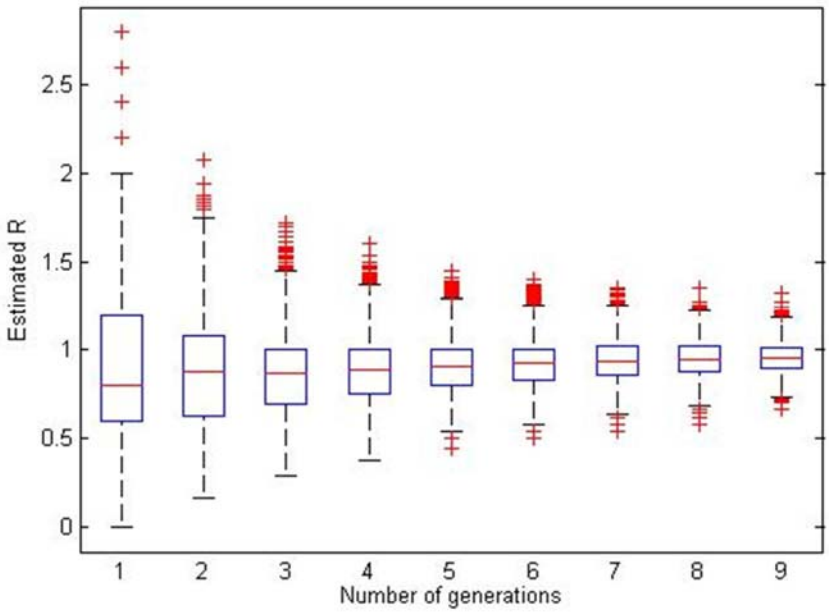
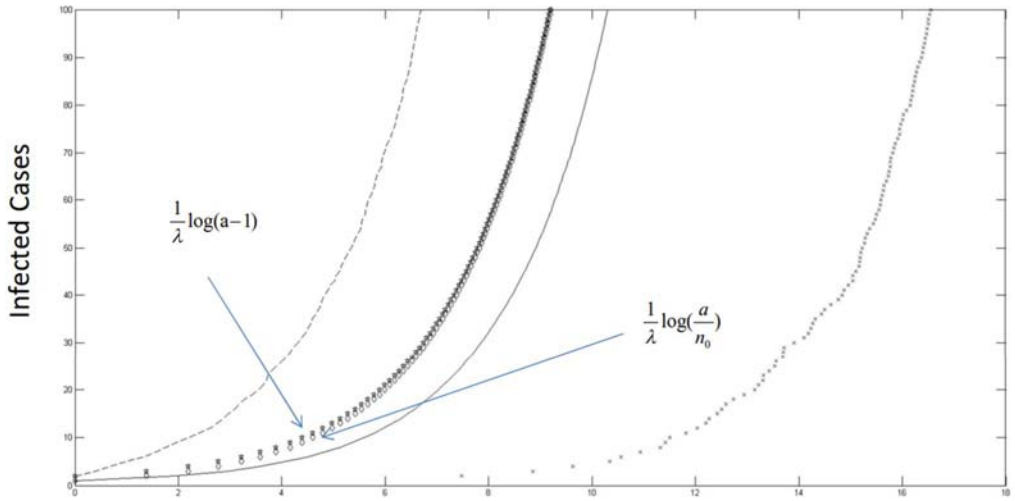


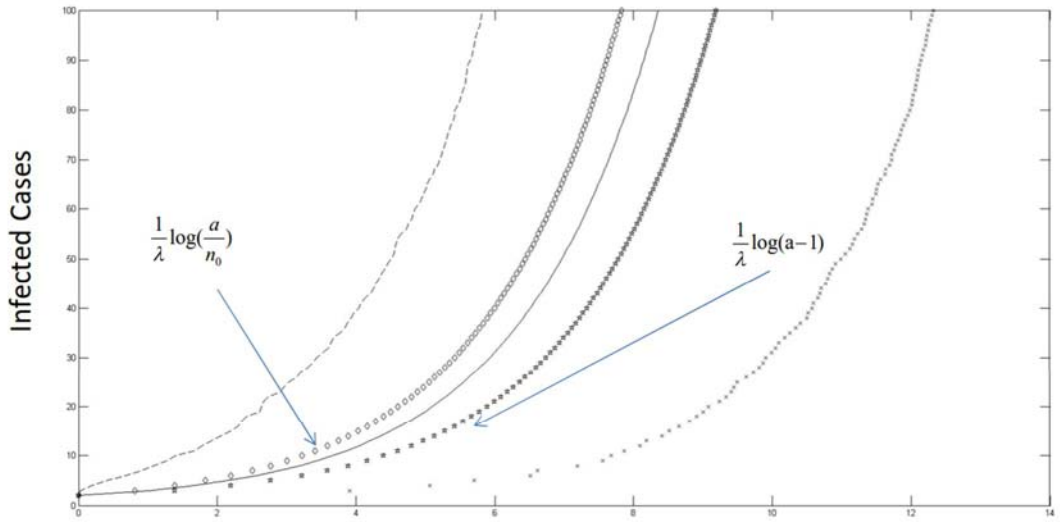
Figure 5. 4 1000 times of Simulation with pure birth process with λ=0.5, estimated by equations: 'solid line' : mean of simulation data, 'dashed line': 0.25% quantile, 'x': 97.5% quantile, 'diamond': approximation equation 1: $\frac{1}{\lambda}\log(\frac{a}{n_0})$, a: final infected case number, n_0: the number of initial case; 'pentagram': approximation equation 2: $\sim\frac{1}{\lambda}\log(a-1)$

(a)The number of initial case is one (N_0=1)



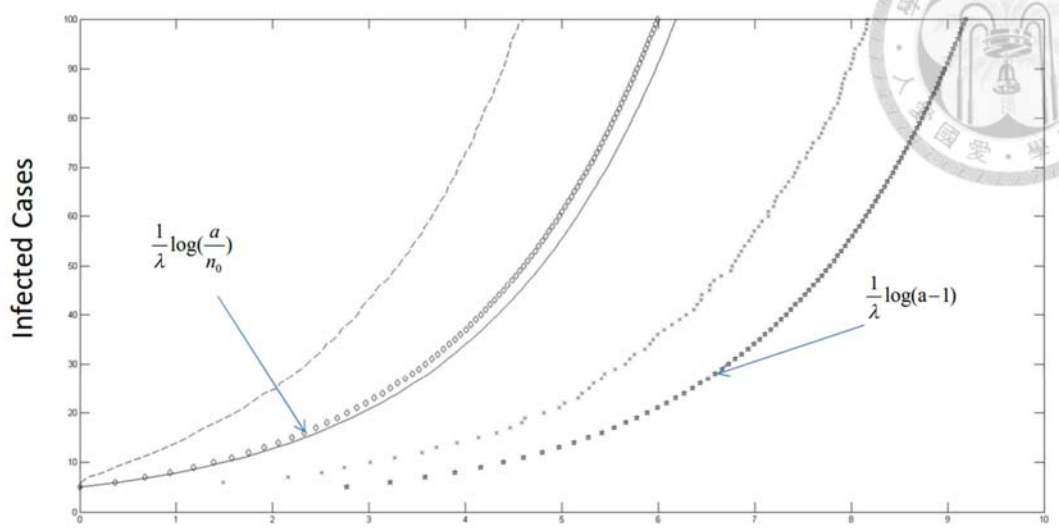
λ=0.5 N_0=1

(b) The number of initial case is 2 (N_0=2)



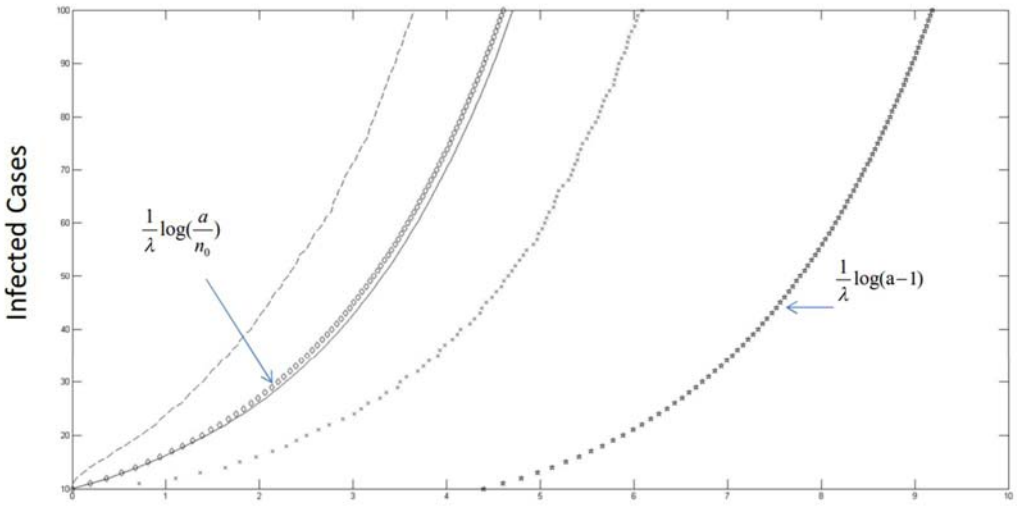
λ=0.5 N_0=2

(c) The number of initial case is 5 (N_0=5)



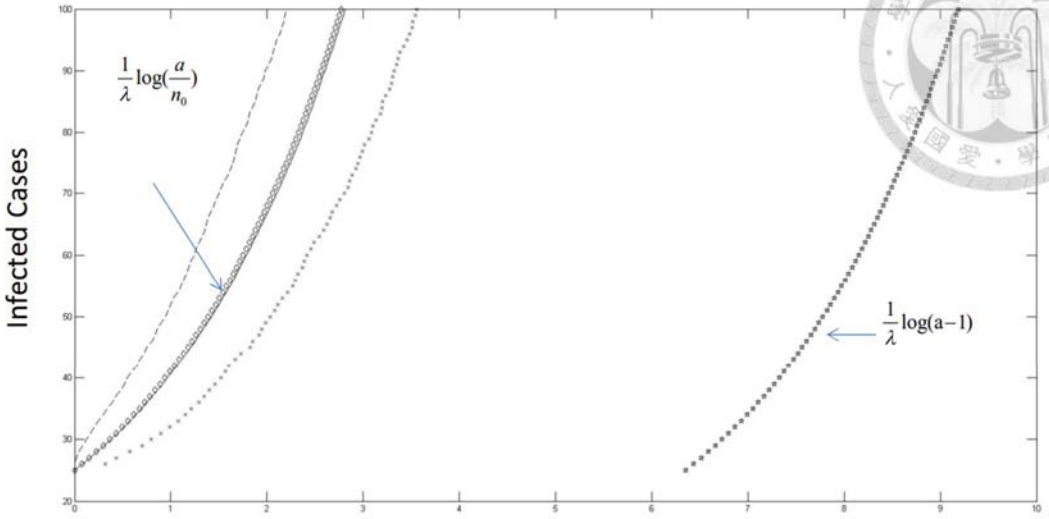
λ=0.5 N_0=5

(d) The number of initial case is 10 (N_0=10)



λ=0.5 N_0=10

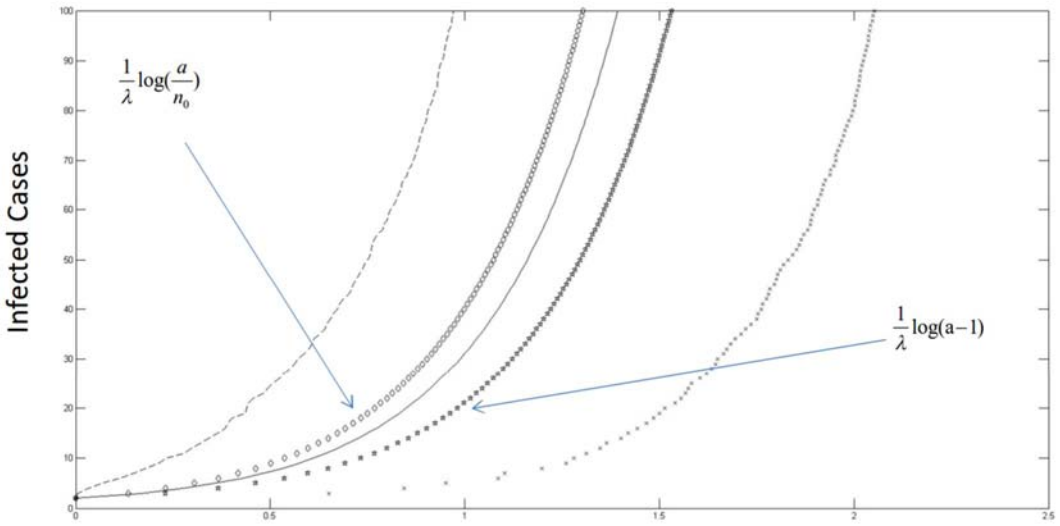
(e) The number of initial case is 10 (N_0=25)



λ=0.5 N_0=25

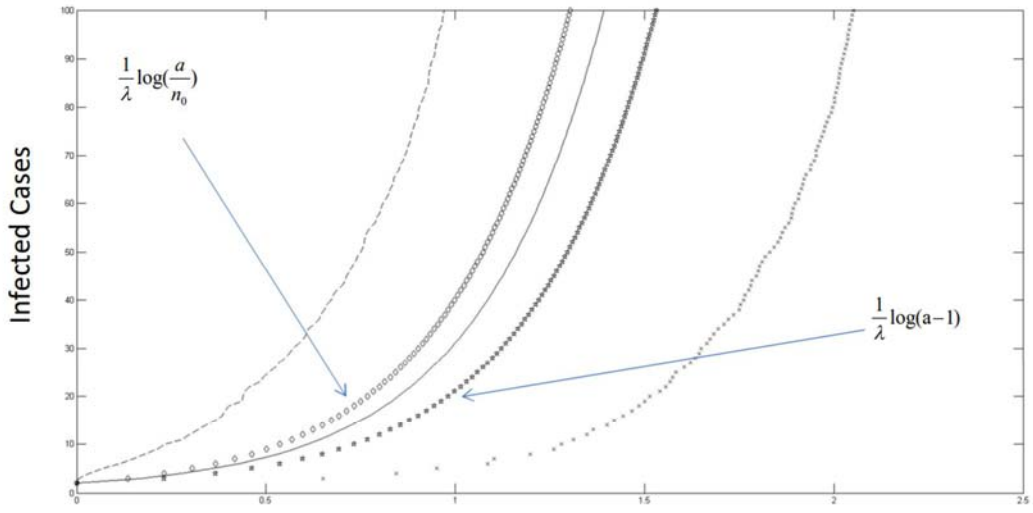
Figure 5. 5 1000 times of Simulation with pure birth process with λ=3, estimated by equations. 'solid line' : mean of simulation, 'dashed line': 0.25% quantile, 'x': 97.5% quantile, 'diamond': approximation equation1: $\frac{1}{\lambda}\log(\frac{a}{n_0})$, a: final infected cases, n_0: the number of initial case; 'pentagram': approximation equation2: $\frac{1}{\lambda}\log(a-1)$.

(a)The number of initial case is one (N_0=1)



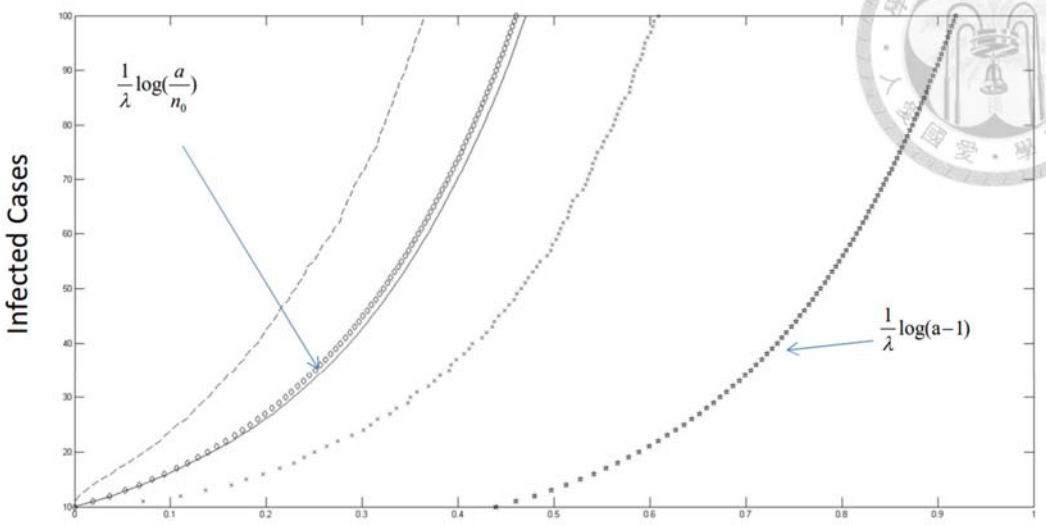
λ=3 N_0=1

(b) The number of initial case is 2 (N_0=2)



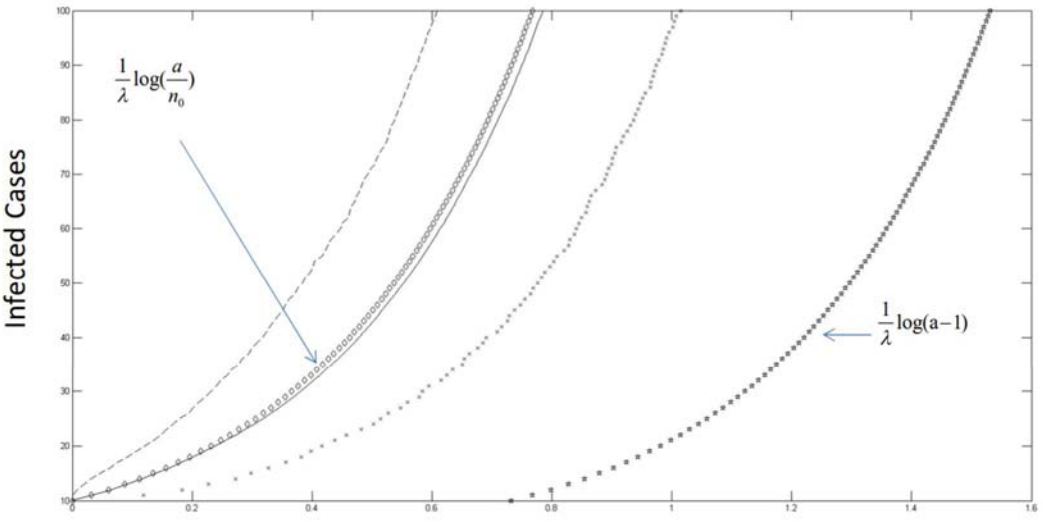
λ=3 N_0=2

(c) The number of initial case is 5 (N_0=5)



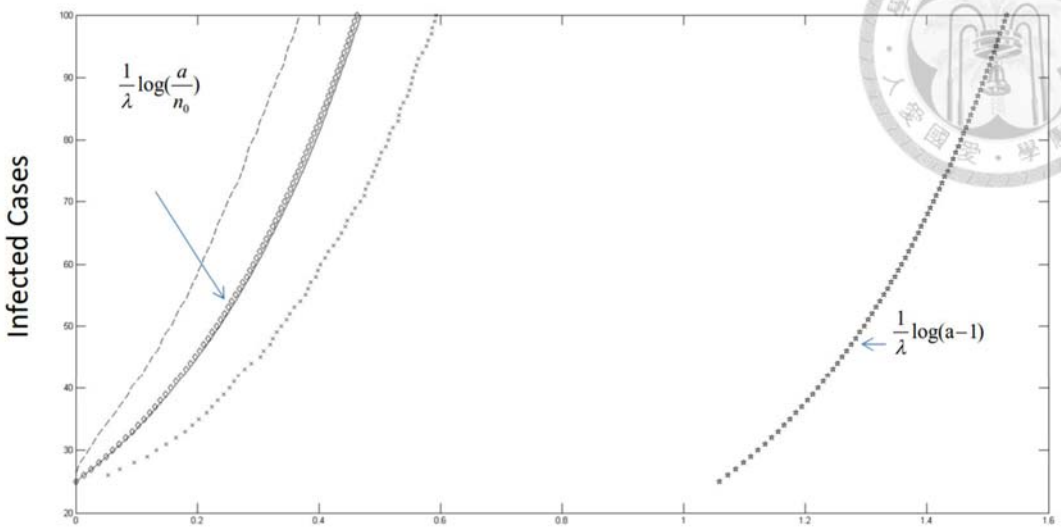
λ=3 N_0=5

(d) The number of initial case is 10 (N_0=10)



λ=3 N_0=10

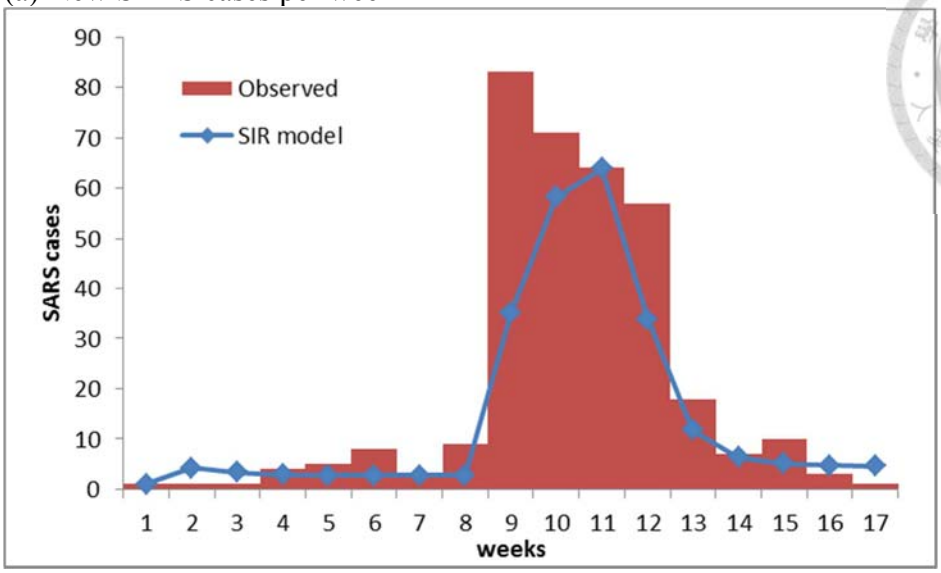
(e) The number of initial case is 10 (N_0=25)



λ=3 N_0=25

Figure 5. 6 The observed SARS cases and simulated cases with deterministic model in Taiwan, 2003.

(a) New SARS cases per week



(b) Cumulated SARS cases per week

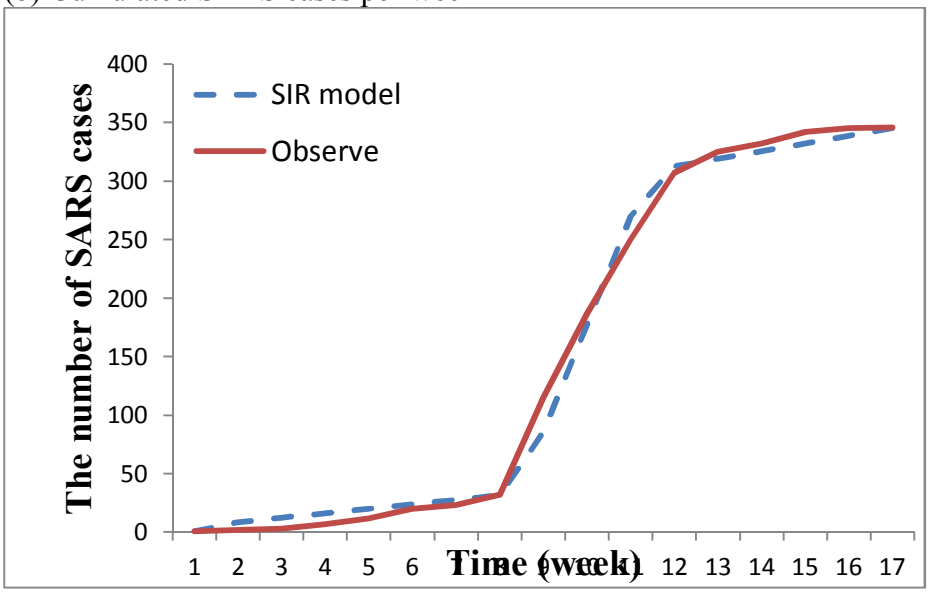
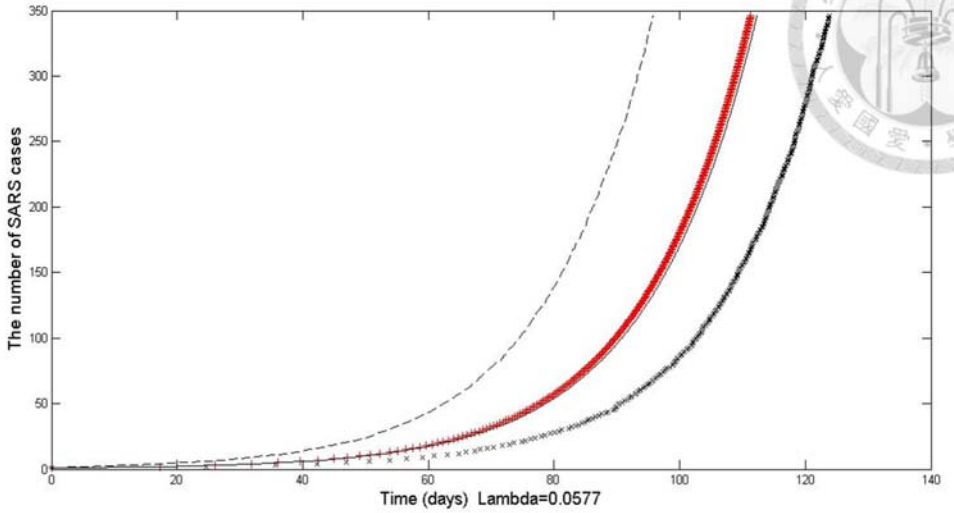
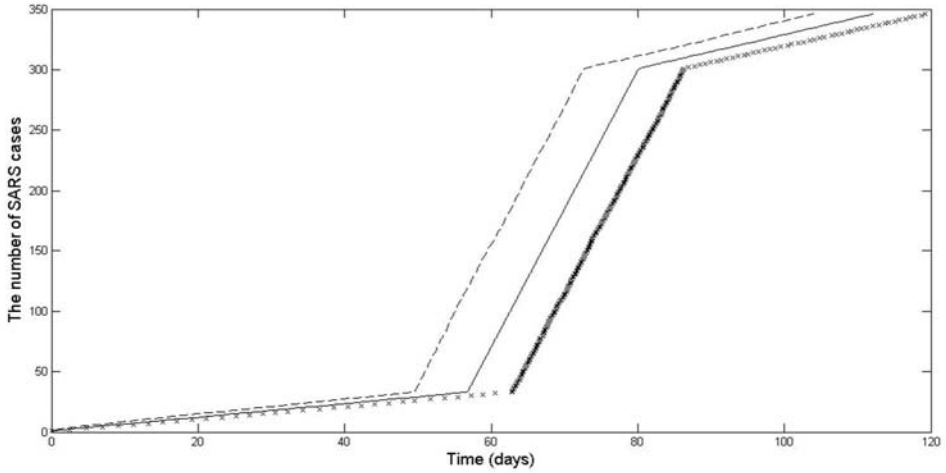


Figure 5.7 Simulated cases with birth death process in Taiwan, 2003

(a) Simulation with pure birth process



(b) Simulation with general birth death process



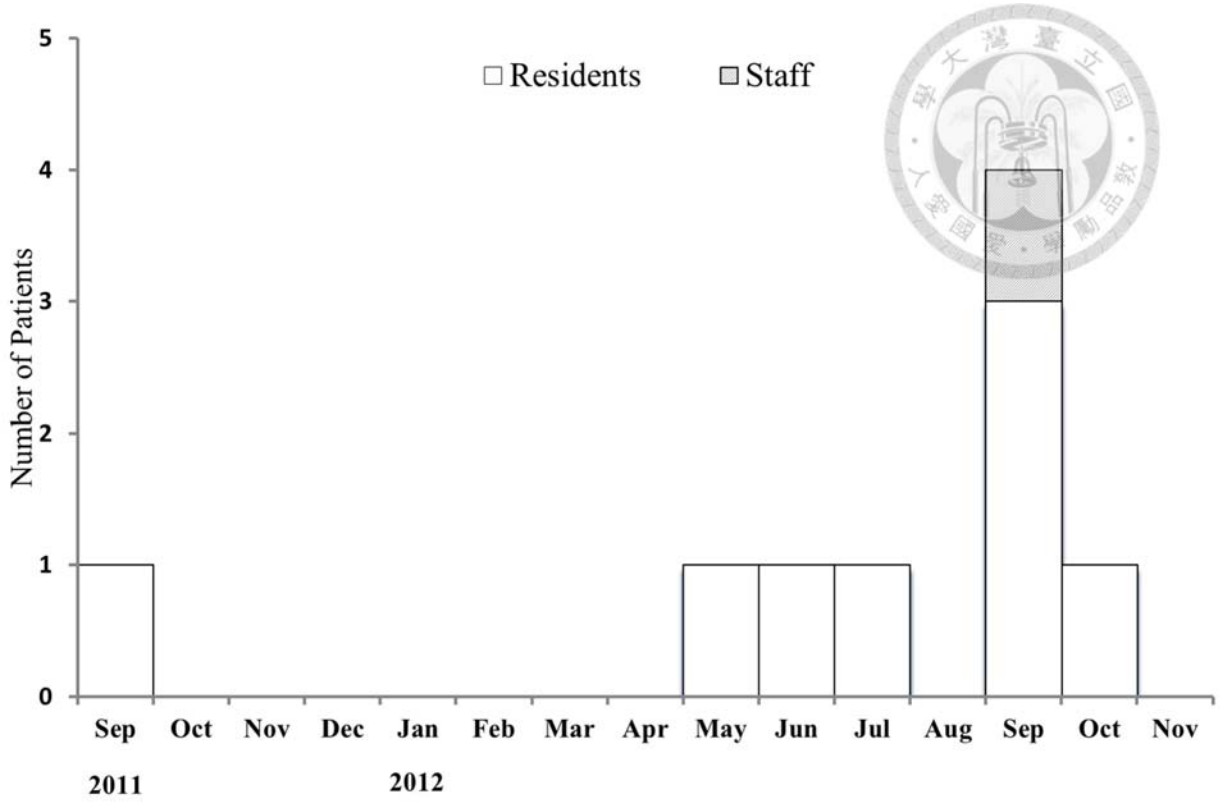


Figure 5 8 Epidemic Data in Chronological Order for Pulmonary Tuberculosis Cases in a Long-term Care Facility

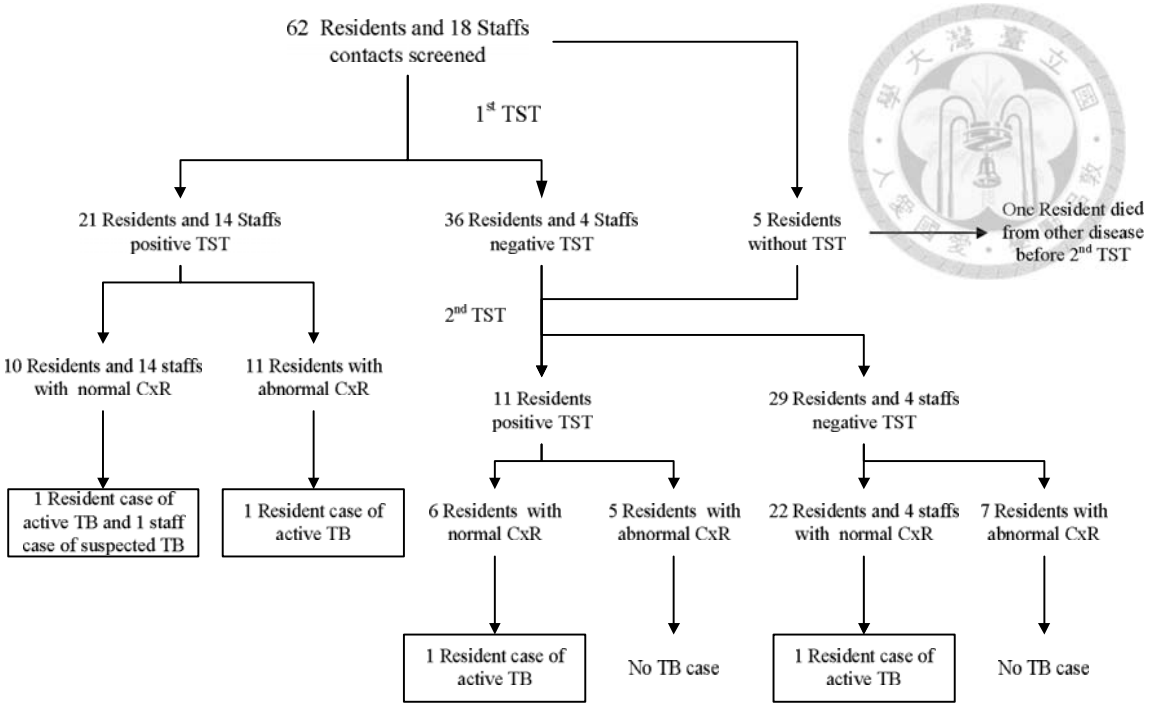


Figure 5.9 Flow-Chart Overview of the Contacts Investigation.

One resident with 1st TST negative died from other illness before 2nd TST. Three confirmed cases and one suspected case were noted before or during the contact investigation period.

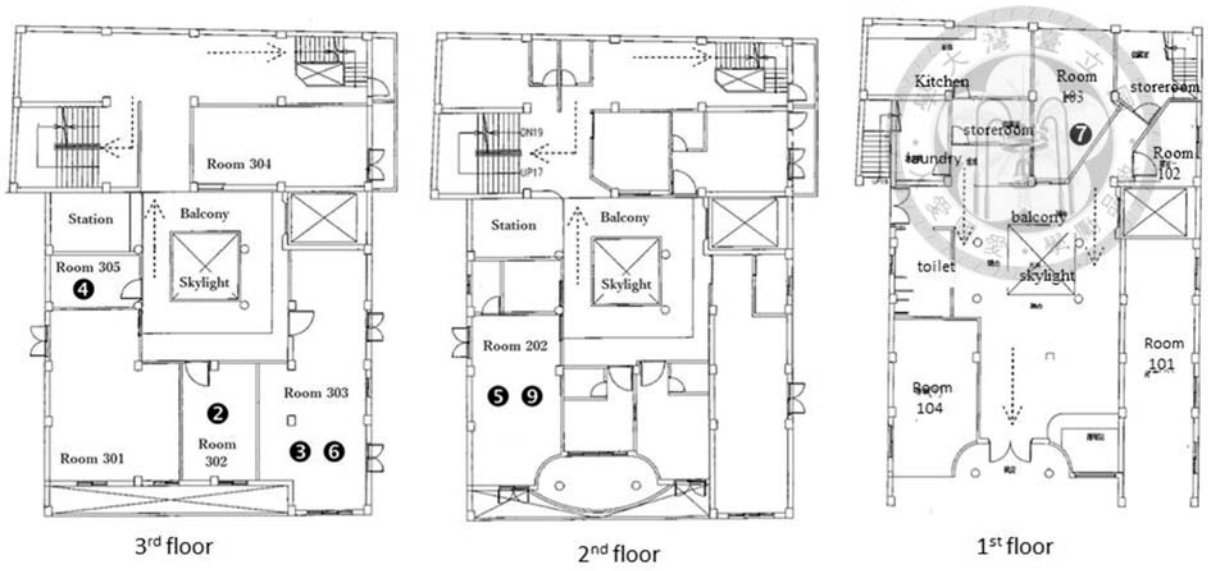


Figure 5.10 Room locations of residents with confirmed active cases during investigation period in the Long-term Care Facility. The index case had been in room 303; the eighth case was a suspected staff-case.

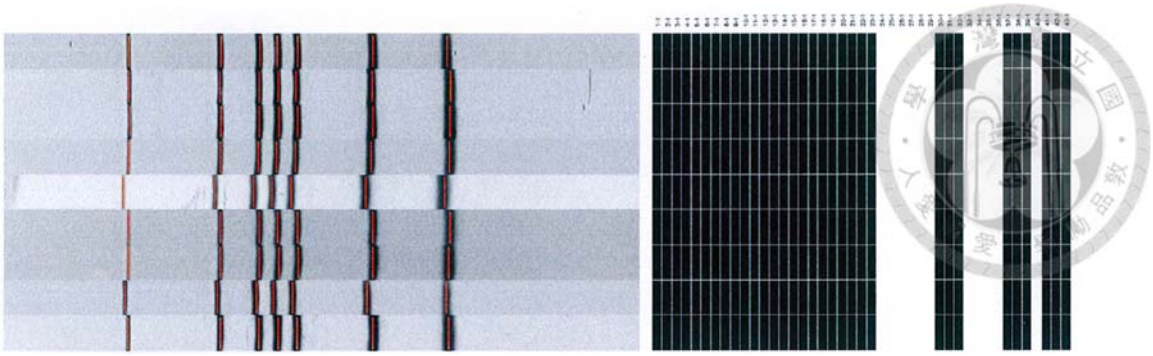
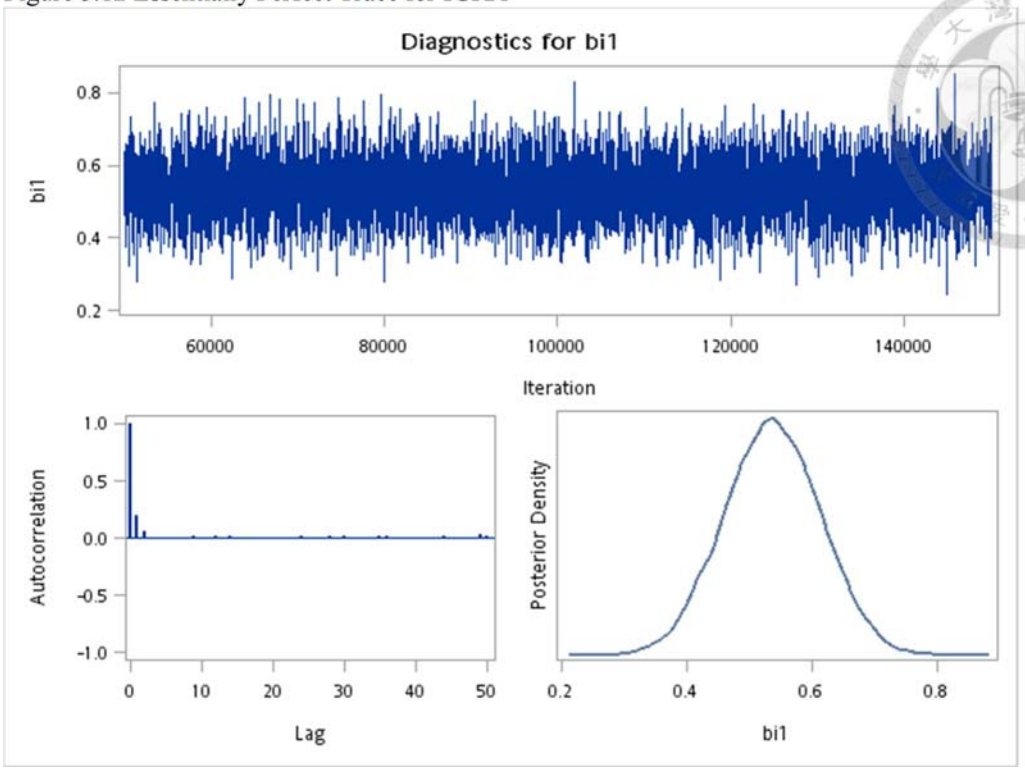


Figure 5. 11 Patients with TB in the cluster had identical RFLP and spoligotyping pattents

Figure 5.12 Essentially Perfect Trace for IGRA



Appendix

Appendix Table

Appendix Table 1 Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=0.5 given different initial infected cases. (Continue)

		Infected size														
Initial cases	Ta	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Mean	**0**	1.9640	2.9356	3.5895	4.0826	4.5069	4.8322	5.1150	5.3567	5.5852	5.7950	5.9716	6.1498	6.3019	6.4329
	SD	-	1.9247	2.1348	2.2221	2.2710	2.2800	2.3013	2.3183	2.3383	2.3476	2.3554	2.3690	2.3804	2.3872	2.3852
	Q0.25	-	0.0570	0.3786	0.6873	0.9971	1.3306	1.5704	1.7432	1.9698	2.2002	2.3985	2.6153	2.7553	2.8700	2.9722
	Q97.5	-	7.4802	8.8542	9.6274	10.3453	10.5578	10.9327	11.3224	11.4233	11.5039	11.8130	12.0349	12.2080	12.2657	12.4157
2	Mean		0	0.9964	1.6520	2.1551	2.5465	2.8684	3.1585	3.4030	3.6309	3.8329	4.0079	4.1704	4.3240	4.4684
	SD		-	1.0097	1.2047	1.3169	1.3893	1.4129	1.4335	1.4518	1.4762	1.4960	1.5087	1.5255	1.5263	1.5356
	Q0.25		-	0.0259	0.2035	0.4004	0.6665	0.8754	1.1001	1.2371	1.3889	1.6130	1.6866	1.8292	1.9387	2.0845
	Q97.5		-	3.7096	4.7580	5.2798	5.7375	6.2250	6.4820	6.8559	7.1080	7.3642	7.4235	7.6819	7.8180	8.0010
5	Mean					0	0.4147	0.7397	1.0337	1.2879	1.5125	1.7086	1.8877	2.0550	2.2084	2.3565
	SD					-	0.4279	0.5285	0.6050	0.6490	0.6836	0.7069	0.7348	0.7576	0.7726	0.7882
	Q0.25					-	0.0067	0.0945	0.1764	0.3359	0.4954	0.6062	0.7026	0.8244	0.9456	1.0116
	Q97.5					-	1.5949	2.0531	2.4528	2.8931	3.1819	3.3801	3.5981	3.8508	4.0386	4.1905

Ta: Expected time to infected size a; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data.

Appendix Table 1. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=0.5 given different initial infected cases. (Continue)

		Infected size														
Initial cases	Ta	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1	Mean	6.5675	6.6967	6.8129	6.9216	7.0264	7.1260	7.2236	7.3144	7.3985	7.4843	7.5646	7.6404	7.7168	7.7850	7.8532
	SD	2.3806	2.3813	2.3812	2.3871	2.3921	2.3943	2.3961	2.3963	2.4049	2.4134	2.4159	2.4211	2.4214	2.4229	2.4236
	Q0.25	3.0930	3.1758	3.2786	3.4304	3.5492	3.6437	3.7035	3.7316	3.8050	3.8952	4.0053	4.0923	4.1595	4.2179	4.3037
	Q97.5	12.5058	12.5920	12.7352	12.9359	13.1465	13.2214	13.3036	13.3119	13.4211	13.5382	13.5558	13.6858	13.7068	13.7168	13.9063
2	Mean	4.5966	4.7244	4.8400	4.9507	5.0527	5.1525	5.2474	5.3379	5.4312	5.5129	5.5943	5.6754	5.7501	5.8242	5.8957
	SD	1.5399	1.5444	1.5459	1.5519	1.5550	1.5571	1.5663	1.5662	1.5665	1.5686	1.5674	1.5725	1.5705	1.5702	1.5699
	Q0.25	2.1517	2.3123	2.4774	2.5678	2.6916	2.7858	2.8295	2.9189	3.0169	3.1305	3.2041	3.2653	3.3011	3.3815	3.4622
	Q97.5	8.2846	8.4080	8.4788	8.5291	8.6955	8.7662	8.8221	8.9581	9.0059	9.0807	9.1333	9.1616	9.2679	9.3380	9.3571
5	Mean	2.4879	2.6189	2.7359	2.8551	2.9694	3.0668	3.1651	3.2565	3.3451	3.4265	3.5044	3.5830	3.6559	3.7226	3.7891
	SD	0.7960	0.8088	0.8189	0.8286	0.8388	0.8428	0.8420	0.8502	0.8539	0.8565	0.8599	0.8643	0.8689	0.8714	0.8721
	Q0.25	1.1509	1.2735	1.3636	1.4592	1.5498	1.6072	1.7330	1.8237	1.9107	2.0394	2.0780	2.1525	2.1998	2.2300	2.3110
	Q97.5	4.2967	4.4496	4.6137	4.7846	4.8603	4.9582	5.0689	5.1947	5.2575	5.3489	5.4334	5.4877	5.5799	5.7056	5.7865

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data.

Appendix Table 1. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=0.5 given different initial infected cases. (Continue)

		Infected size														
Initial cases	Ta	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
1	Mean	7.9213	7.9874	8.0492	8.1116	8.1709	8.2316	8.2877	8.3421	8.3988	8.4530	8.5043	8.5553	8.6042	8.6523	8.6990
	SD	2.4214	2.4194	2.4187	2.4201	2.4221	2.4230	2.4281	2.4284	2.4307	2.4327	2.4300	2.4315	2.4316	2.4314	2.4306
	Q0.25	4.3848	4.4523	4.5067	4.5627	4.5890	4.6558	4.7049	4.7564	4.7818	4.8353	4.8953	4.9374	5.0192	5.0423	5.1139
	Q97.5	14.1047	14.1889	14.2285	14.2699	14.3133	14.4432	14.5058	14.6185	14.7703	14.8336	14.8796	14.9653	15.0405	15.0481	15.1073
2	Mean	5.9575	6.0284	6.0888	6.1541	6.2140	6.2712	6.3278	6.3834	6.4338	6.4865	6.5379	6.5882	6.6344	6.6784	6.7243
	SD	1.5699	1.5727	1.5741	1.5749	1.5740	1.5742	1.5724	1.5746	1.5768	1.5793	1.5797	1.5825	1.5841	1.5856	1.5856
	Q0.25	3.5325	3.6195	3.6574	3.7142	3.7821	3.8071	3.8671	3.9777	4.0372	4.0708	4.1041	4.1197	4.1581	4.1877	4.2136
	Q97.5	9.3571	9.4295	9.4847	9.5816	9.6030	9.6620	9.7344	9.8171	9.8956	9.9500	10.0052	10.0419	10.1116	10.1252	10.1633
5	Mean	3.8578	3.9207	3.9838	4.0426	4.1015	4.1593	4.2164	4.2719	4.3279	4.3817	4.4307	4.4790	4.5255	4.5715	4.6181
	SD	0.8731	0.8761	0.8816	0.8832	0.8863	0.8902	0.8912	0.8945	0.8985	0.9001	0.9016	0.9013	0.9036	0.9048	0.9036
	Q0.25	2.3716	2.4290	2.4733	2.5162	2.5774	2.6022	2.6662	2.7178	2.7754	2.7994	2.8502	2.9114	2.9512	2.9738	3.0482
	Q97.5	5.8585	5.8934	5.9712	6.0294	6.1289	6.2193	6.2905	6.3424	6.3999	6.4275	6.4407	6.5181	6.5421	6.5549	6.5985

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data.

Appendix Table 1. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=0.5 given different initial infected cases. (Continue)

		Infected size														
Initial cases	Ta	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
1	Mean	8.7425	8.7861	8.8300	8.8727	8.9133	8.9541	8.9914	9.0304	9.0683	9.1060	9.1414	9.1766	9.2120	9.2474	9.2810
	SD	2.4291	2.4273	2.4264	2.4260	2.4258	2.4265	2.4269	2.4256	2.4261	2.4244	2.4257	2.4242	2.4232	2.4253	2.4253
	Q0.25	5.1472	5.1960	5.2340	5.2889	5.3350	5.3813	5.4209	5.4748	5.5264	5.5409	5.5576	5.6154	5.6211	5.6535	5.6670
	Q97.5	15.1613	15.1733	15.1869	15.1984	15.2199	15.2691	15.2845	15.3164	15.3767	15.4466	15.4787	15.5026	15.5560	15.5720	15.5913
2	Mean	6.7694	6.8129	6.8562	6.8992	6.9399	6.9777	7.0183	7.0581	7.0958	7.1317	7.1667	7.2004	7.2356	7.2702	7.3034
	SD	1.5862	1.5887	1.5902	1.5919	1.5930	1.5943	1.5926	1.5946	1.5950	1.5955	1.5959	1.5957	1.5969	1.5969	1.5952
	Q0.25	4.2661	4.3238	4.3483	4.3768	4.3931	4.4345	4.4465	4.4868	4.5413	4.5742	4.5887	4.6136	4.6553	4.6962	4.7141
	Q97.5	10.1972	10.2475	10.2944	10.3419	10.4059	10.4796	10.5285	10.5544	10.5730	10.5808	10.6107	10.6573	10.7356	10.7486	10.7835
5	Mean	4.6620	4.7055	4.7482	4.7895	4.8293	4.8668	4.9051	4.9442	4.9821	5.0184	5.0551	5.0899	5.1261	5.1602	5.1939
	SD	0.9031	0.9018	0.9032	0.9040	0.9049	0.9084	0.9097	0.9117	0.9104	0.9096	0.9118	0.9126	0.9114	0.9110	0.9116
	Q0.25	3.0898	3.1283	3.1970	3.2180	3.2414	3.2943	3.3172	3.3558	3.3888	3.4197	3.4645	3.4938	3.5181	3.5681	3.5886
	Q97.5	6.6657	6.7266	6.7785	6.8459	6.8900	6.9063	6.9742	6.9982	7.0234	7.0641	7.0778	7.1224	7.1332	7.1472	7.1670

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data.

Appendix Table 1. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=0.5 given different initial infected cases. (Continue)

		Infected size														
Initial cases	Ta	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
1	Mean	9.3141	9.3457	9.3790	9.4102	9.4408	9.4732	9.5035	9.5357	9.5646	9.5932	9.6220	9.6486	9.6762	9.7029	9.7287
	SD	2.4255	2.4236	2.4251	2.4235	2.4244	2.4239	2.4250	2.4264	2.4263	2.4268	2.4264	2.4269	2.4265	2.4293	2.4296
	Q0.25	5.7279	5.7653	5.7917	5.8059	5.8311	5.8637	5.9300	5.9304	5.9420	5.9769	6.0179	6.0353	6.1040	6.1228	6.1641
	Q97.5	15.6147	15.6427	15.6497	15.6929	15.7289	15.7465	15.7690	15.7717	15.7836	15.8206	15.8417	15.8664	15.9182	15.9454	15.9504
2	Mean	7.3383	7.3721	7.4042	7.4373	7.4665	7.4958	7.5268	7.5555	7.5837	7.6120	7.6392	7.6666	7.6954	7.7231	7.7505
	SD	1.5949	1.5963	1.5977	1.5978	1.5974	1.5985	1.5980	1.5993	1.5999	1.6001	1.6009	1.6015	1.6016	1.6029	1.6032
	Q0.25	4.7668	4.7917	4.8156	4.8622	4.9112	4.9462	4.9923	5.0118	5.0414	5.0996	5.1274	5.1579	5.1838	5.2166	5.2372
	Q97.5	10.8069	10.8456	10.9034	10.9108	10.9493	10.9668	10.9772	11.0116	11.0648	11.0945	11.1071	11.1178	11.1287	11.1507	11.1631
5	Mean	5.2293	5.2620	5.2950	5.3273	5.3582	5.3887	5.4188	5.4494	5.4785	5.5077	5.5365	5.5649	5.5915	5.6198	5.6467
	SD	0.9128	0.9137	0.9142	0.9154	0.9156	0.9152	0.9136	0.9135	0.9146	0.9160	0.9160	0.9171	0.9184	0.9192	0.9198
	Q0.25	3.6291	3.6568	3.6914	3.7229	3.7422	3.7800	3.8146	3.8333	3.8656	3.9003	3.9316	3.9514	3.9818	3.9885	4.0423
	Q97.5	7.2347	7.2471	7.3016	7.3282	7.3783	7.4017	7.4277	7.4608	7.4782	7.5213	7.5271	7.5606	7.5883	7.6352	7.6462

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data.

Appendix Table 1. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=0.5 given different initial infected cases. (Continue)

		Infected size														
Initial cases	Ta	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
1	Mean	9.7560	9.7820	9.8073	9.8329	9.8584	9.8833	9.9075	9.9318	9.9548	9.9785	10.0023	10.0255	10.0483	10.0698	10.0921
	SD	2.4307	2.4320	2.4320	2.4316	2.4322	2.4305	2.4312	2.4314	2.4317	2.4315	2.4316	2.4314	2.4315	2.4329	2.4348
	Q0.25	6.1923	6.2088	6.2266	6.2545	6.2779	6.3186	6.3384	6.3461	6.3686	6.4025	6.4424	6.4499	6.4743	6.4936	6.5129
	Q97.5	15.9731	16.0065	16.0295	16.0849	16.1557	16.1654	16.1916	16.2046	16.2135	16.2409	16.2552	16.2818	16.3061	16.3378	16.3723
2	Mean	7.7788	7.8067	7.8323	7.8574	7.8843	7.9104	7.9351	7.9609	7.9839	8.0077	8.0313	8.0549	8.0782	8.1021	8.1241
	SD	1.6047	1.6043	1.6045	1.6053	1.6058	1.6059	1.6055	1.6041	1.6041	1.6038	1.6043	1.6040	1.6029	1.6037	1.6039
	Q0.25	5.2527	5.2804	5.2922	5.3068	5.3586	5.3613	5.4074	5.4372	5.4431	5.4691	5.5135	5.5176	5.5389	5.5722	5.5918
	Q97.5	11.1841	11.2186	11.2653	11.2666	11.3388	11.3510	11.3983	11.4047	11.4336	11.4681	11.5191	11.5247	11.5309	11.5713	11.6000
5	Mean	5.6749	5.7006	5.7252	5.7511	5.7770	5.8018	5.8267	5.8501	5.8754	5.9005	5.9234	5.9460	5.9692	5.9914	6.0150
	SD	0.9191	0.9209	0.9215	0.9216	0.9209	0.9207	0.9217	0.9216	0.9217	0.9215	0.9213	0.9218	0.9213	0.9210	0.9211
	Q0.25	4.0578	4.0680	4.0816	4.1093	4.1488	4.1716	4.2223	4.2558	4.2632	4.3071	4.3456	4.3575	4.3734	4.3965	4.4248
	Q97.5	7.7019	7.7413	7.7564	7.7861	7.8016	7.8266	7.8727	7.8841	7.9051	7.9228	7.9437	7.9692	7.9951	8.0160	8.0558

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data.

Appendix Table 1. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=0.5 given different initial infected cases. (Continue)

		Infected size									
Initial cases	Ta	91	92	93	94	95	96	97	98	99	100
1	Mean	10.1145	10.1365	10.1595	10.1814	10.2018	10.2212	10.2412	10.2627	10.2827	10.3030
	SD	2.4342	2.4340	2.4342	2.4344	2.4340	2.4341	2.4345	2.4342	2.4344	2.4333
	Q0.25	6.5200	6.5487	6.5638	6.5813	6.5977	6.6119	6.6342	6.6565	6.6682	6.6896
	Q97.5	16.3956	16.4066	16.4290	16.4454	16.4681	16.4769	16.5047	16.5095	16.5267	16.5557
2	Mean	8.1457	8.1676	8.1889	8.2113	8.2324	8.2537	8.2746	8.2961	8.3169	8.3376
	SD	1.6038	1.6046	1.6054	1.6057	1.6054	1.6051	1.6061	1.6062	1.6059	1.6053
	Q0.25	5.6056	5.6275	5.6524	5.6696	5.7035	5.7252	5.7486	5.7692	5.7763	5.8111
	Q97.5	11.6136	11.6234	11.6528	11.6704	11.6981	11.7026	11.7222	11.7319	11.7514	11.8017
5	Mean	6.0370	6.0595	6.0808	6.1027	6.1235	6.1442	6.1637	6.1844	6.2031	6.2236
	SD	0.9216	0.9221	0.9225	0.9222	0.9211	0.9217	0.9227	0.9236	0.9237	0.9263
	Q0.25	4.4475	4.4619	4.5079	4.5249	4.5497	4.5752	4.5827	4.6021	4.6247	4.6373
	Q97.5	8.0736	8.1034	8.1274	8.1559	8.1647	8.1997	8.2070	8.2135	8.2222	8.3246

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data.

Appendix Table 1. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=0.5 given different initial infected cases. (Continue)

		Infected size													
Initial cases	Ta	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
10	Mean	**0**	0.2021	0.3846	0.5492	0.7132	0.8522	0.9893	1.1164	1.2361	1.3469	1.4560	1.5488	1.6453	1.7327	1.8208
	SD	-	0.1955	0.2695	0.3126	0.3516	0.3832	0.4097	0.4222	0.4338	0.4495	0.4552	0.4681	0.4772	0.4833	0.4929
	Q0.25	-	0.0056	0.0469	0.1208	0.2129	0.2949	0.3670	0.4623	0.5404	0.6047	0.7110	0.7538	0.8308	0.9039	0.9718
	Q97.5	-	0.7495	1.0999	1.3202	1.5640	1.7546	1.9183	2.0636	2.2286	2.3674	2.4463	2.5317	2.6409	2.7571	2.8416
Initial cases	Ta	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
25	Mean	0	0.0783	0.1560	0.2291	0.3021	0.3727	0.4383	0.5053	0.5689	0.6310	0.6911	0.7504	0.8097	0.8631	0.9124
	SD	-	0.0789	0.1111	0.1343	0.1534	0.1665	0.1767	0.1858	0.1953	0.2024	0.2113	0.2171	0.2284	0.2335	0.2376
	Q0.25	-	0.0015	0.0206	0.0425	0.0794	0.1218	0.1600	0.2055	0.2503	0.2960	0.3387	0.3782	0.4363	0.4615	0.4948
	Q97.5	-	0.2919	0.4483	0.5519	0.6710	0.7709	0.8692	0.9349	1.0108	1.0837	1.1492	1.2333	1.3144	1.3893	1.4521

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data.

Appendix Table 1. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=0.5 given different initial infected cases. (Continue)

		Infected size															
Initial cases	Ta	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
10	Mean	1.9040	1.9827	2.0575	2.1304	2.2012	2.2707	2.3358	2.3991	2.4598	2.5213	2.5832	2.6421	2.6990	2.7509	2.8034	2.8538
	SD	0.5020	0.5072	0.5160	0.5247	0.5313	0.5385	0.5410	0.5439	0.5486	0.5528	0.5549	0.5572	0.5597	0.5619	0.5648	0.5664
	Q0.25	1.0126	1.1036	1.1699	1.2379	1.2850	1.3662	1.4068	1.4622	1.5152	1.5612	1.6069	1.6467	1.7068	1.7556	1.7906	1.8361
	Q97.5	2.9363	3.0238	3.1243	3.2459	3.3226	3.4149	3.4787	3.5510	3.6101	3.6506	3.6969	3.7765	3.8246	3.8975	3.9788	4.0229
Initial cases	Ta	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55
25	Mean	0.9663	1.0166	1.0648	1.1104	1.1566	1.2035	1.2484	1.2927	1.3354	1.3754	1.4165	1.4560	1.4949	1.5313	1.5696	1.6058
	SD	0.2435	0.2473	0.2503	0.2563	0.2609	0.2662	0.2728	0.2777	0.2821	0.2859	0.2886	0.2907	0.2925	0.2960	0.2994	0.3007
	Q0.25	0.5434	0.5891	0.6339	0.6685	0.7122	0.7448	0.7810	0.8064	0.8388	0.8573	0.9153	0.9509	0.9873	1.0153	1.0345	1.0712
	Q97.5	1.4947	1.5497	1.6018	1.6615	1.7123	1.7654	1.8430	1.8777	1.9342	1.9834	2.0313	2.0871	2.1194	2.1649	2.2133	2.2418

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data

Appendix Table 1. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=0.5 given different initial infected cases. (Continue)

		Infected size															
Initial cases	Ta	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56
10	Mean	2.9048	2.9537	3.0005	3.0482	3.0939	3.1395	3.1824	3.2268	3.2700	3.3120	3.3517	3.3936	3.4327	3.4699	3.5065	3.5420
	SD	0.5680	0.5724	0.5738	0.5781	0.5832	0.5852	0.5849	0.5844	0.5865	0.5860	0.5865	0.5896	0.5925	0.5939	0.5930	0.5941
	Q0.25	1.8775	1.9378	1.9878	2.0092	2.0428	2.0857	2.1295	2.1736	2.2131	2.2345	2.2804	2.3248	2.3475	2.3812	2.4183	2.4625
	Q97.5	4.0746	4.1293	4.2058	4.2510	4.3365	4.3688	4.4147	4.4408	4.4712	4.5246	4.5799	4.6217	4.7020	4.7312	4.7593	4.7986
Initial cases	Ta	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71
25	Mean	1.6402	1.6738	1.7073	1.7415	1.7757	1.8087	1.8403	1.8711	1.9026	1.9341	1.9644	1.9960	2.0263	2.0543	2.0823	2.1125
	SD	0.3036	0.3051	0.3073	0.3103	0.3126	0.3144	0.3165	0.3192	0.3213	0.3233	0.3234	0.3250	0.3271	0.3291	0.3300	0.3325
	Q0.25	1.1110	1.1267	1.1509	1.1841	1.2152	1.2371	1.2602	1.2806	1.3020	1.3336	1.3676	1.3930	1.4124	1.4335	1.4628	1.5153
	Q97.5	2.3158	2.3357	2.3701	2.4145	2.4579	2.5005	2.5164	2.5426	2.5726	2.6111	2.6482	2.6943	2.7147	2.7438	2.7871	2.8181

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data

Appendix Table 1. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=0.5 given different initial infected cases. (Continue)

		Infected size															
Initial cases	Ta	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72
10	Mean	3.5787	3.6137	3.6482	3.6824	3.7150	3.7461	3.7800	3.8121	3.8429	3.8738	3.9062	3.9355	3.9637	3.9936	4.0226	4.0523
	SD	0.5970	0.5972	0.5983	0.5995	0.5998	0.6008	0.6032	0.6055	0.6052	0.6049	0.6045	0.6063	0.6064	0.6065	0.6089	0.6097
	Q0.25	2.4802	2.5237	2.5469	2.5871	2.6227	2.6590	2.7018	2.7173	2.7594	2.7875	2.8125	2.8613	2.8985	2.9280	2.9480	2.9581
	Q97.5	4.8315	4.8729	4.8837	4.9067	4.9419	4.9815	5.0041	5.0366	5.0544	5.0885	5.1181	5.1782	5.1852	5.2096	5.2346	5.2894
Initial cases	Ta	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87
25	Mean	2.1418	2.1707	2.1977	2.2247	2.2507	2.2768	2.3016	2.3270	2.3519	2.3772	2.4020	2.4265	2.4508	2.4740	2.4969	2.5202
	SD	0.3334	0.3349	0.3362	0.3361	0.3368	0.3390	0.3404	0.3404	0.3417	0.3428	0.3443	0.3455	0.3452	0.3459	0.3478	0.3476
	Q0.25	1.5384	1.5550	1.5959	1.6250	1.6429	1.6631	1.6772	1.7085	1.7254	1.7453	1.7795	1.7959	1.8362	1.8493	1.8633	1.8848
	Q97.5	2.8494	2.9011	2.9234	2.9394	2.9799	3.0080	3.0375	3.0693	3.0844	3.1146	3.1404	3.1625	3.1863	3.1968	3.2424	3.2635

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data

Appendix Table 1. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=0.5 given different initial infected cases. (Continue)

		Infected size															
Initial cases	Ta	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88
10	Mean	4.0812	4.1085	4.1372	4.1631	4.1893	4.2164	4.2429	4.2679	4.2930	4.3173	4.3420	4.3660	4.3898	4.4139	4.4376	4.4615
	SD	0.6106	0.6113	0.6138	0.6128	0.6135	0.6157	0.6162	0.6163	0.6158	0.6157	0.6159	0.6155	0.6153	0.6147	0.6140	0.6155
	Q0.25	2.9764	3.0170	3.0256	3.0561	3.0821	3.0953	3.1192	3.1558	3.1705	3.1956	3.2150	3.2411	3.2567	3.2930	3.3142	3.3256
	Q97.5	5.3156	5.3592	5.3883	5.4150	5.4316	5.4585	5.4880	5.5440	5.5648	5.5684	5.5755	5.5829	5.6028	5.6401	5.6837	5.7050
	Ta	89	90	91	92	93	94	95	96	97	98	99	100				
	Mean	4.4848	4.5069	4.5289	4.5519	4.5733	4.5945	4.6164	4.6382	4.6596	4.6805	4.6998	4.7198				
	SD	0.6150	0.6158	0.6144	0.6151	0.6160	0.6164	0.6173	0.6170	0.6172	0.6166	0.6167	0.6165				
	Q0.25	3.3511	3.3672	3.4112	3.4242	3.4320	3.4637	3.4826	3.4974	3.5362	3.5553	3.5813	3.5974				
	Q97.5	5.7421	5.7576	5.7854	5.7939	5.8053	5.8303	5.8357	5.8525	5.8618	5.8793	5.8911	5.9178				
Initial cases	Ta	88	89	90	91	92	93	94	95	96	97	98	99	100			
25	Mean	2.5437	2.5654	2.5863	2.6097	2.6322	2.6546	2.6754	2.6967	2.7184	2.7393	2.7602	2.7803	2.7994			
	SD	0.3478	0.3495	0.3499	0.3512	0.3522	0.3520	0.3521	0.3527	0.3537	0.3536	0.3555	0.3565	0.3568			
	Q0.25	1.9176	1.9289	1.9516	1.9865	2.0022	2.0219	2.0452	2.0662	2.0877	2.1010	2.1197	2.1371	2.1572			
	Q97.5	3.2916	3.3270	3.3507	3.3722	3.3932	3.4078	3.4443	3.4615	3.4895	3.5089	3.5168	3.5395	3.5510			

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data

Appendix Table 2 Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=3 given different initial infected cases. (Continue)

		Infected size														
Initial cases	Ta	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Mean	0	0.3187	0.4780	0.5848	0.6715	0.7352	0.7891	0.8353	0.8772	0.9150	0.9470	0.9772	1.0056	1.0316	1.0554
	SD	-	0.3208	0.3558	0.3704	0.3785	0.3800	0.3836	0.3864	0.3897	0.3913	0.3926	0.3948	0.3967	0.3979	0.3975
	Q0.25	-	0.0081	0.0490	0.1130	0.1710	0.2227	0.2663	0.2958	0.3296	0.3729	0.4163	0.4339	0.4542	0.4818	0.5034
	Q97.5	-	1.1056	1.3383	1.4488	1.5725	1.6214	1.6715	1.7304	1.7741	1.7961	1.8183	1.9166	1.9475	1.9584	1.9736
2	Mean		0	0.1661	0.2753	0.3592	0.4244	0.4781	0.5264	0.5672	0.6051	0.6388	0.6680	0.6951	0.7207	0.7447
	SD		-	0.1683	0.2008	0.2195	0.2316	0.2355	0.2389	0.2420	0.2460	0.2493	0.2515	0.2543	0.2544	0.2559
	Q0.25		-	0.0043	0.0339	0.0667	0.1111	0.1459	0.1834	0.2062	0.2315	0.2688	0.2811	0.3049	0.3231	0.3474
	Q97.5		-	0.6183	0.7930	0.8800	0.9563	1.0375	1.0803	1.1427	1.1847	1.2274	1.2372	1.2803	1.3030	1.3335
5	Mean					0	0.0691	0.1233	0.1723	0.2147	0.2521	0.2848	0.3146	0.3425	0.3681	0.3927
	SD					-	0.0713	0.0881	0.1008	0.1082	0.1139	0.1178	0.1225	0.1263	0.1288	0.1314
	Q0.25					-	0.0011	0.0158	0.0294	0.0560	0.0826	0.1010	0.1171	0.1374	0.1576	0.1686
	Q97.5					-	0.2658	0.3422	0.4088	0.4822	0.5303	0.5633	0.5997	0.6418	0.6731	0.6984

Ta: Expected time to infected size a; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data.

Appendix Table 2. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=3 given different initial infected cases. (Continue)

		Infected size														
Initial cases	Ta	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1	Mean	1.0767	1.0980	1.1183	1.1372	1.1539	1.1703	1.1859	1.2011	1.2167	1.2308	1.2438	1.2569	1.2690	1.2810	1.2924
	SD	0.3968	0.3969	0.3969	0.3978	0.3987	0.3991	0.3993	0.3994	0.4008	0.4022	0.4026	0.4035	0.4036	0.4038	0.4039
	Q0.25	0.5193	0.5317	0.5479	0.5756	0.5908	0.6073	0.6324	0.6433	0.6503	0.6599	0.6745	0.6890	0.7017	0.7160	0.7230
	Q97.5	2.0004	2.0346	2.0403	2.0501	2.0868	2.1010	2.1136	2.1195	2.1371	2.1596	2.1664	2.1759	2.1841	2.1895	2.2147
2	Mean	0.7661	0.7874	0.8067	0.8251	0.8421	0.8588	0.8746	0.8897	0.9052	0.9188	0.9324	0.9459	0.9584	0.9707	0.9826
	SD	0.2572	0.2566	0.2574	0.2576	0.2586	0.2592	0.2595	0.2610	0.2610	0.2611	0.2614	0.2612	0.2621	0.2617	0.2617
	Q0.25	0.3586	0.3854	0.4129	0.4280	0.4486	0.4643	0.4716	0.4865	0.5028	0.5218	0.5340	0.5442	0.5502	0.5636	0.5770
	Q97.5	1.3525	1.3808	1.4013	1.4131	1.4215	1.4492	1.4610	1.4703	1.4930	1.5010	1.5135	1.5222	1.5269	1.5446	1.5563
5	Mean	0.4146	0.4365	0.4560	0.4759	0.4949	0.5111	0.5275	0.5427	0.5575	0.5711	0.5841	0.5972	0.6093	0.6204	0.6315
	SD	0.1327	0.1348	0.1365	0.1381	0.1398	0.1405	0.1403	0.1417	0.1423	0.1428	0.1433	0.1440	0.1448	0.1452	0.1453
	Q0.25	0.1918	0.2122	0.2273	0.2432	0.2583	0.2679	0.2888	0.3040	0.3185	0.3399	0.3463	0.3587	0.3666	0.3717	0.3852
	Q97.5	0.7161	0.7416	0.7689	0.7974	0.8100	0.8264	0.8448	0.8658	0.8762	0.8915	0.9056	0.9146	0.9300	0.9509	0.9644

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data.

Appendix Table 2. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=3 given different initial infected cases. (Continue)

		Infected size														
Initial cases	Ta	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
1	Mean	1.3037	1.3147	1.3250	1.3348	1.3449	1.3542	1.3635	1.3724	1.3814	1.3904	1.3987	1.4068	1.4144	1.4226	1.4303
	SD	0.4036	0.4032	0.4031	0.4033	0.4037	0.4038	0.4047	0.4047	0.4051	0.4054	0.4050	0.4053	0.4053	0.4052	0.4051
	Q0.25	0.7317	0.7436	0.7528	0.7684	0.7772	0.7822	0.7907	0.8031	0.8079	0.8143	0.8230	0.8370	0.8449	0.8561	0.8632
	Q97.5	2.2209	2.2357	2.2462	2.2471	2.2657	2.2691	2.2836	2.3004	2.3124	2.3159	2.3244	2.3281	2.3359	2.3396	2.3448
2	Mean	0.9929	1.0047	1.0148	1.0257	1.0357	1.0452	1.0546	1.0639	1.0723	1.0811	1.0896	1.0980	1.1057	1.1131	1.1207
	SD	0.2616	0.2621	0.2624	0.2625	0.2623	0.2624	0.2621	0.2624	0.2628	0.2632	0.2633	0.2637	0.2640	0.2643	0.2643
	Q0.25	0.5887	0.6032	0.6096	0.6190	0.6303	0.6345	0.6445	0.6629	0.6729	0.6785	0.6840	0.6866	0.6930	0.6979	0.7023
	Q97.5	1.5595	1.5716	1.5808	1.5969	1.6005	1.6103	1.6224	1.6362	1.6493	1.6583	1.6675	1.6736	1.6853	1.6875	1.6939
5	Mean	0.6430	0.6535	0.6640	0.6738	0.6836	0.6932	0.7027	0.7120	0.7213	0.7303	0.7385	0.7465	0.7543	0.7619	0.7697
	SD	0.1455	0.1460	0.1469	0.1472	0.1477	0.1484	0.1485	0.1491	0.1497	0.1500	0.1503	0.1502	0.1506	0.1508	0.1506
	Q0.25	0.3953	0.4048	0.4122	0.4194	0.4296	0.4337	0.4444	0.4530	0.4626	0.4666	0.4750	0.4852	0.4919	0.4956	0.5080
	Q97.5	0.9764	0.9822	0.9952	1.0049	1.0215	1.0366	1.0484	1.0571	1.0666	1.0713	1.0735	1.0864	1.0903	1.0925	1.0998

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data.

Appendix Table 2. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=3 given different initial infected cases. (Continue)

		Infected size														
Initial cases	Ta	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
1	Mean	1.4376	1.4448	1.4519	1.4587	1.4655	1.4722	1.4790	1.4853	1.4918	1.4978	1.5039	1.5100	1.5160	1.5219	1.5275
	SD	0.4049	0.4046	0.4044	0.4043	0.4043	0.4044	0.4045	0.4043	0.4043	0.4041	0.4043	0.4040	0.4039	0.4042	0.4042
	Q0.25	0.8726	0.8779	0.8806	0.8867	0.8927	0.8991	0.9051	0.9180	0.9218	0.9333	0.9383	0.9482	0.9517	0.9534	0.9548
	Q97.5	2.3569	2.3638	2.3765	2.3832	2.3869	2.3967	2.3986	2.4040	2.4077	2.4130	2.4244	2.4278	2.4327	2.4388	2.4423
2	Mean	1.1282	1.1355	1.1427	1.1499	1.1567	1.1629	1.1697	1.1763	1.1826	1.1886	1.1944	1.2001	1.2059	1.2117	1.2172
	SD	0.2644	0.2648	0.2650	0.2653	0.2655	0.2657	0.2654	0.2658	0.2658	0.2659	0.2660	0.2660	0.2662	0.2661	0.2659
	Q0.25	0.7110	0.7206	0.7247	0.7295	0.7322	0.7391	0.7411	0.7478	0.7569	0.7624	0.7648	0.7689	0.7759	0.7827	0.7857
	Q97.5	1.6995	1.7079	1.7157	1.7236	1.7343	1.7466	1.7548	1.7591	1.7622	1.7635	1.7685	1.7762	1.7893	1.7914	1.7972
5	Mean	0.7770	0.7843	0.7914	0.7982	0.8049	0.8111	0.8175	0.8240	0.8303	0.8364	0.8425	0.8483	0.8543	0.8600	0.8657
	SD	0.1505	0.1503	0.1505	0.1507	0.1508	0.1514	0.1516	0.1519	0.1517	0.1516	0.1520	0.1521	0.1519	0.1518	0.1519
	Q0.25	0.5150	0.5214	0.5328	0.5363	0.5402	0.5490	0.5529	0.5593	0.5648	0.5699	0.5774	0.5823	0.5863	0.5947	0.5981
	Q97.5	1.1109	1.1211	1.1298	1.1410	1.1483	1.1511	1.1624	1.1664	1.1706	1.1773	1.1796	1.1871	1.1889	1.1912	1.1945

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data.

Appendix Table 2. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=3 given different initial infected cases. (Continue)

		Infected size														
Initial cases	Ta	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
1	Mean	1.5331	1.5386	1.5436	1.5488	1.5539	1.5591	1.5643	1.5692	1.5740	1.5788	1.5836	1.5882	1.5928	1.5974	1.6023
	SD	0.4043	0.4039	0.4042	0.4039	0.4041	0.4040	0.4042	0.4044	0.4044	0.4045	0.4044	0.4045	0.4044	0.4049	0.4049
	Q0.25	0.9604	0.9632	0.9717	0.9771	0.9817	0.9865	0.9933	0.9944	0.9977	1.0028	1.0074	1.0150	1.0218	1.0277	1.0295
	Q97.5	2.4457	2.4546	2.4586	2.4623	2.4654	2.4684	2.4719	2.4836	2.4944	2.5012	2.5031	2.5083	2.5117	2.5139	2.5171
2	Mean	1.2230	1.2287	1.2340	1.2395	1.2444	1.2493	1.2545	1.2593	1.2640	1.2687	1.2732	1.2778	1.2826	1.2872	1.2918
	SD	0.2658	0.2661	0.2663	0.2663	0.2662	0.2664	0.2663	0.2665	0.2667	0.2667	0.2668	0.2669	0.2669	0.2671	0.2672
	Q0.25	0.7945	0.7986	0.8026	0.8104	0.8185	0.8244	0.8321	0.8353	0.8402	0.8499	0.8546	0.8597	0.8640	0.8694	0.8729
	Q97.5	1.8012	1.8076	1.8172	1.8185	1.8249	1.8278	1.8295	1.8353	1.8441	1.8491	1.8512	1.8530	1.8548	1.8585	1.8605
5	Mean	0.8716	0.8770	0.8825	0.8879	0.8930	0.8981	0.9031	0.9082	0.9131	0.9179	0.9227	0.9275	0.9319	0.9366	0.9411
	SD	0.1521	0.1523	0.1524	0.1526	0.1526	0.1525	0.1523	0.1522	0.1524	0.1527	0.1527	0.1528	0.1531	0.1532	0.1533
	Q0.25	0.6048	0.6095	0.6152	0.6205	0.6237	0.6300	0.6358	0.6389	0.6443	0.6500	0.6553	0.6586	0.6636	0.6648	0.6737
	Q97.5	1.2058	1.2078	1.2169	1.2214	1.2297	1.2336	1.2379	1.2435	1.2464	1.2536	1.2545	1.2601	1.2647	1.2725	1.2744

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data.

Appendix Table 2. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=3 given different initial infected cases. (Continue)

		Infected size														
Initial cases	Ta	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
1	Mean	1.6069	1.6114	1.6158	1.6199	1.6240	1.6283	1.6323	1.6364	1.6405	1.6443	1.6482	1.6521	1.6561	1.6599	1.6636
	SD	0.4051	0.4053	0.4053	0.4053	0.4054	0.4051	0.4052	0.4052	0.4053	0.4053	0.4053	0.4052	0.4052	0.4055	0.4058
	Q0.25	1.0347	1.0372	1.0424	1.0474	1.0498	1.0539	1.0583	1.0637	1.0681	1.0718	1.0738	1.0766	1.0793	1.0851	1.0897
	Q97.5	2.5220	2.5290	2.5303	2.5394	2.5414	2.5458	2.5474	2.5522	2.5569	2.5662	2.5689	2.5723	2.5742	2.5900	2.5934
2	Mean	1.2965	1.3011	1.3054	1.3096	1.3141	1.3184	1.3225	1.3268	1.3306	1.3346	1.3385	1.3425	1.3464	1.3503	1.3540
	SD	0.2674	0.2674	0.2674	0.2675	0.2676	0.2677	0.2676	0.2673	0.2674	0.2673	0.2674	0.2673	0.2672	0.2673	0.2673
	Q0.25	0.8755	0.8801	0.8820	0.8845	0.8931	0.8936	0.9012	0.9062	0.9072	0.9115	0.9189	0.9196	0.9231	0.9287	0.9320
	Q97.5	1.8640	1.8698	1.8775	1.8778	1.8898	1.8918	1.8997	1.9008	1.9056	1.9113	1.9199	1.9208	1.9218	1.9285	1.9333
5	Mean	0.9458	0.9501	0.9542	0.9585	0.9628	0.9670	0.9711	0.9750	0.9792	0.9834	0.9872	0.9910	0.9949	0.9986	1.0025
	SD	0.1532	0.1535	0.1536	0.1536	0.1535	0.1534	0.1536	0.1536	0.1536	0.1536	0.1536	0.1536	0.1536	0.1535	0.1535
	Q0.25	0.6763	0.6780	0.6803	0.6849	0.6915	0.6953	0.7037	0.7093	0.7105	0.7179	0.7243	0.7262	0.7289	0.7327	0.7375
	Q97.5	1.2837	1.2902	1.2927	1.2977	1.3003	1.3044	1.3121	1.3140	1.3175	1.3205	1.3239	1.3282	1.3325	1.3360	1.3426

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data.

Appendix Table 2. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=3 given different initial infected cases. (Continue)

		Infected size									
Initial cases	Ta	91	92	93	94	95	96	97	98	99	100
1	Mean	1.6673	1.6710	1.6747	1.6782	1.6818	1.6852	1.6884	1.6920	1.6952	1.6983
	SD	0.4057	0.4057	0.4057	0.4057	0.4057	0.4057	0.4058	0.4057	0.4057	0.4056
	Q0.25	1.0914	1.0955	1.0984	1.0995	1.1072	1.1078	1.1088	1.1146	1.1185	1.1220
	Q97.5	2.5958	2.5989	2.6007	2.6017	2.6040	2.6119	2.6194	2.6228	2.6256	2.6282
2	Mean	1.3576	1.3613	1.3648	1.3685	1.3721	1.3756	1.3791	1.3827	1.3862	1.3896
	SD	0.2673	0.2674	0.2676	0.2676	0.2676	0.2675	0.2677	0.2677	0.2676	0.2676
	Q0.25	0.9343	0.9379	0.9421	0.9449	0.9506	0.9542	0.9581	0.9615	0.9627	0.9685
	Q97.5	1.9356	1.9372	1.9421	1.9451	1.9497	1.9504	1.9537	1.9553	1.9586	1.9669
5	Mean	1.0062	1.0099	1.0135	1.0171	1.0206	1.0240	1.0273	1.0307	1.0338	1.0373
	SD	0.1536	0.1537	0.1538	0.1537	0.1535	0.1536	0.1538	0.1539	0.1540	0.1544
	Q0.25	0.7412	0.7436	0.7513	0.7541	0.7583	0.7625	0.7638	0.7670	0.7708	0.7729
	Q97.5	1.3456	1.3506	1.3546	1.3593	1.3608	1.3666	1.3678	1.3689	1.3704	1.3874

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data.

Appendix Table 2. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=3 given different initial infected cases. (Continue)

		Infected size													
Initial cases	Ta	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
10	Mean	0	0.0337	0.0641	0.0915	0.1189	0.1420	0.1649	0.1861	0.2060	0.2245	0.2427	0.2581	0.2742	0.2888	0.3035
	SD	-	0.0326	0.0449	0.0521	0.0586	0.0639	0.0683	0.0704	0.0723	0.0749	0.0759	0.0780	0.0795	0.0806	0.0822
	Q0.25	-	0.0009	0.0078	0.0201	0.0355	0.0492	0.0612	0.0770	0.0901	0.1008	0.1185	0.1256	0.1385	0.1507	0.1620
	Q97.5	-	0.1249	0.1833	0.2200	0.2607	0.2924	0.3197	0.3439	0.3714	0.3946	0.4077	0.4219	0.4402	0.4595	0.4736
		Infected size														
Initial cases	Ta	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39
25	Mean	0	0.0131	0.0260	0.0382	0.0503	0.0621	0.0730	0.0842	0.0948	0.1052	0.1152	0.1251	0.1350	0.1438	0.1521
	SD	-	0.0132	0.0185	0.0224	0.0256	0.0277	0.0295	0.0310	0.0326	0.0337	0.0352	0.0362	0.0381	0.0389	0.0396
	Q0.25	-	0.0003	0.0034	0.0071	0.0132	0.0203	0.0267	0.0343	0.0417	0.0493	0.0565	0.0630	0.0727	0.0769	0.0825
	Q97.5	-	0.0486	0.0747	0.0920	0.1118	0.1285	0.1449	0.1558	0.1685	0.1806	0.1915	0.2056	0.2191	0.2315	0.2420

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data.

Appendix Table 2. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=3 given different initial infected cases. (Continue)

		Infected size															
Initial cases	Ta	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
10	Mean	0.3173	0.3304	0.3429	0.3551	0.3669	0.3784	0.3893	0.3999	0.4100	0.4202	0.4305	0.4403	0.4498	0.4585	0.4672	0.4756
	SD	0.0837	0.0845	0.0860	0.0874	0.0886	0.0898	0.0902	0.0906	0.0914	0.0921	0.0925	0.0929	0.0933	0.0936	0.0941	0.0944
	Q0.25	0.0837	0.0845	0.0860	0.0874	0.0886	0.0898	0.0902	0.0906	0.0914	0.0921	0.0925	0.0929	0.0933	0.0936	0.0941	0.0944
	Q97.5	0.4894	0.5040	0.5207	0.5410	0.5538	0.5692	0.5798	0.5918	0.6017	0.6084	0.6162	0.6294	0.6374	0.6496	0.6631	0.6705
		Infected size															
Initial cases	Ta	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55
25	Mean	0.1610	0.1694	0.1775	0.1851	0.1928	0.2006	0.2081	0.2155	0.2226	0.2292	0.2361	0.2427	0.2492	0.2552	0.2616	0.2676
	SD	0.0406	0.0412	0.0417	0.0427	0.0435	0.0444	0.0455	0.0463	0.0470	0.0477	0.0481	0.0484	0.0488	0.0493	0.0499	0.0501
	Q0.25	0.0906	0.0982	0.1057	0.1114	0.1187	0.1241	0.1302	0.1344	0.1398	0.1429	0.1525	0.1585	0.1646	0.1692	0.1724	0.1785
	Q97.5	0.2491	0.2583	0.2670	0.2769	0.2854	0.2942	0.3072	0.3130	0.3224	0.3306	0.3385	0.3478	0.3532	0.3608	0.3689	0.3736

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data

Appendix Table 2. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=3 given different initial infected cases. (Continue)

		Infected size															
Initial cases	Ta	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56
10	Mean	0.4841	0.4923	0.5001	0.5080	0.5157	0.5232	0.5304	0.5378	0.5450	0.5520	0.5586	0.5656	0.5721	0.5783	0.5844	0.5903
	SD	0.0947	0.0954	0.0956	0.0963	0.0972	0.0975	0.0975	0.0974	0.0978	0.0977	0.0977	0.0983	0.0988	0.0990	0.0988	0.0990
	Q0.25	0.3129	0.3230	0.3313	0.3349	0.3405	0.3476	0.3549	0.3623	0.3688	0.3724	0.3801	0.3875	0.3913	0.3969	0.4030	0.4104
	Q97.5	0.6791	0.6882	0.7010	0.7085	0.7228	0.7281	0.7358	0.7401	0.7452	0.7541	0.7633	0.7703	0.7837	0.7885	0.7932	0.7998
Initial cases	Ta	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71
25	Mean	0.2734	0.2790	0.2845	0.2902	0.2959	0.3015	0.3067	0.3118	0.3171	0.3224	0.3274	0.3327	0.3377	0.3424	0.3471	0.3521
	SD	0.0506	0.0508	0.0512	0.0517	0.0521	0.0524	0.0527	0.0532	0.0535	0.0539	0.0539	0.0542	0.0545	0.0549	0.0550	0.0554
	Q0.25	0.1852	0.1878	0.1918	0.1974	0.2025	0.2062	0.2100	0.2134	0.2170	0.2223	0.2279	0.2322	0.2354	0.2389	0.2438	0.2525
	Q97.5	0.3860	0.3893	0.3950	0.4024	0.4096	0.4167	0.4194	0.4238	0.4288	0.4352	0.4414	0.4491	0.4524	0.4573	0.4645	0.4697

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data

Appendix Table 2. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=3 given different initial infected cases. (Continue)

		Infected size															
Initial cases	Ta	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72
10	Mean	0.5964	0.6023	0.6080	0.6137	0.6192	0.6243	0.6300	0.6353	0.6405	0.6456	0.6510	0.6559	0.6606	0.6656	0.6704	0.6754
	SD	0.0995	0.0995	0.0997	0.0999	0.1000	0.1001	0.1005	0.1009	0.1009	0.1008	0.1008	0.1010	0.1011	0.1011	0.1015	0.1016
	Q0.25	0.4134	0.4206	0.4245	0.4312	0.4371	0.4432	0.4503	0.4529	0.4599	0.4646	0.4687	0.4769	0.4831	0.4880	0.4913	0.4930
	Q97.5	0.8052	0.8121	0.8139	0.8178	0.8236	0.8303	0.8340	0.8394	0.8424	0.8481	0.8530	0.8630	0.8642	0.8683	0.8724	0.8816
Initial cases	Ta	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87
25	Mean	0.3570	0.3618	0.3663	0.3708	0.3751	0.3795	0.3836	0.3878	0.3920	0.3962	0.4003	0.4044	0.4085	0.4123	0.4161	0.4200
	SD	0.0556	0.0558	0.0560	0.0560	0.0561	0.0565	0.0567	0.0567	0.0570	0.0571	0.0574	0.0576	0.0575	0.0576	0.0580	0.0579
	Q0.25	0.2564	0.2592	0.2660	0.2708	0.2738	0.2772	0.2795	0.2847	0.2876	0.2909	0.2966	0.2993	0.3060	0.3082	0.3105	0.3141
	Q97.5	0.4749	0.4835	0.4872	0.4899	0.4967	0.5013	0.5062	0.5116	0.5141	0.5191	0.5234	0.5271	0.5311	0.5328	0.5404	0.5439

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data

Appendix Table 2. Expected time to infected size ‘a’ after 1000 times of simulation with simulation with pure birth process with λ=3 given different initial infected cases.

		Infected size															
Initial cases	Ta	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88
10	Mean	0.6802	0.6848	0.6895	0.6938	0.6982	0.7027	0.7071	0.7113	0.7155	0.7196	0.7237	0.7277	0.7316	0.7356	0.7396	0.7436
	SD	0.1018	0.1019	0.1023	0.1021	0.1022	0.1026	0.1027	0.1027	0.1026	0.1026	0.1026	0.1026	0.1025	0.1025	0.1023	0.1026
	Q0.25	0.4961	0.5028	0.5043	0.5094	0.5137	0.5159	0.5199	0.5260	0.5284	0.5326	0.5358	0.5402	0.5428	0.5488	0.5524	0.5543
	Q97.5	0.8859	0.8932	0.8980	0.9025	0.9053	0.9097	0.9147	0.9240	0.9275	0.9281	0.9293	0.9305	0.9338	0.9400	0.9473	0.9508
	Ta	89	90	91	92	93	94	95	96	97	98	99	100				
	Mean	0.7475	0.7512	0.7548	0.7587	0.7622	0.7657	0.7694	0.7730	0.7766	0.7801	0.7833	0.7866				
	SD	0.1025	0.1026	0.1024	0.1025	0.1027	0.1027	0.1029	0.1028	0.1029	0.1028	0.1028	0.1028				
	Q0.25	0.5585	0.5612	0.5685	0.5707	0.5720	0.5773	0.5804	0.5829	0.5894	0.5926	0.5969	0.5996				
	Q97.5	0.9570	0.9596	0.9642	0.9657	0.9676	0.9717	0.9726	0.9754	0.9770	0.9799	0.9819	0.9863				
Initial cases	Ta	88	89	90	91	92	93	94	95	96	97	98	99	100			
25	Mean	0.4240	0.4276	0.4311	0.4350	0.4387	0.4424	0.4459	0.4494	0.4531	0.4566	0.4600	0.4634	0.4666			
	SD	0.0580	0.0583	0.0583	0.0585	0.0587	0.0587	0.0587	0.0588	0.0589	0.0589	0.0593	0.0594	0.0595			
	Q0.25	0.3196	0.3215	0.3253	0.3311	0.3337	0.3370	0.3409	0.3444	0.3480	0.3502	0.3533	0.3562	0.3595			
	Q97.5	0.5486	0.5545	0.5585	0.5620	0.5655	0.5680	0.5740	0.5769	0.5816	0.5848	0.5861	0.5899	0.5918			

Ta: Expected time to infected size ‘a’; SD: standard deviation; Q2.5: 2.5% quantile of simulation data; Q97.5: 97.5% quantile of simulation data

IRB

臺北市立聯合醫院人體試驗委員會
Taipei City Hospital Institutional Review Board
計畫執行許可書

民國 103 年 12 月 31 日
聯絡人：李玉菁
電話:(02)27093600 分機 3846
傳真:(02)27079021

案件編號：TCHIRB-1031201-E
計畫名稱：彰化地區肺結核群聚事件之分析
計畫主持人：葉彥伯(彰化縣縣政府)/賴昭智(仁愛院區)
計畫書版本/日期: Ver2.0_1031212
同意書版本/日期: 免受試者同意書
計畫期間為民國 103 年 12 月 01 日起至民國 104 年 11 月 30 日
上述計畫業經本院 104 年第 2 次人體試驗委員會會議審查(追認核備)，民國 103 年 12 月 25 日審查通過，特此證明。

本許可書自民國 103 年 12 月 25 日起至民國 104 年 11 月 30 日止。

※追認核備之案件，若會議有疑義時，本會有權撤銷此計畫執行許可書。
※未完成期中報告或結案報告者不得申請新案。
※請於計畫執行許可書有效期限到期前一個月內繳交期中報告。
※請於計畫執行許可書到期前繳交結案報告。
※計畫內容若有任何修改或增減，計畫主持人或計畫委託單位需於許可書有效期限到期日 6 週前檢送修正案至本會，經審查通過後方可實施。

田麗珠 田麗珠

臺北市立聯合醫院人體試驗委會主任委員

CERTIFICATE

Date：December 31, 2014

The project entitled "Outbreak of Pulmonary Tuberculosis in Changhua County"submitted by investigator Yen-Po Yeh/Chao-Chih Lai has been approved by Institutional Review Board of the Taipei City Hospital.
Protocol Version Date: Ver2.0_1031212
ICF Version Date: Waiver

Study period: since 12/01/2014 to 11/30/2015
Above project is approved by the TCHIRB on 12/25/2014 and valid till 11/30/2015.

※ All protocols should been subject to final endorsement by the board. TCHIRB has the right to revoke the approval.
※ No new protocol can be applied if the midterm report or the final report is not handed in.
※ The midterm report should be handed in one month before the expiry date of this certificate.
※ The final report should be handed in before the expiry date of this certificate or project end.
※ The PI should send the amendments to TCHIRB at least six weeks before the expiry date of this certificate if there are changes or modification related the protocol. Any change should not be executed until being approved by TCHIRB.

Lih-Chu Tien
Chairman
Taipei City Hospital Institutional Review Board

The committee is organized and operates in accordance with ICH-GCP regulations and guideline.
本委員會組織與運作皆遵守 ICH-GCP 規定

臺北市立聯合醫院人體試驗委員會
Taipei City Hospital Institutional Review Board
計畫執行許可書

民國 103 年 07 月 28 日
聯絡人：高凡
電話:(02)27093600 分機 3828
傳真:(02)27079021

案件編號：TCHIRB-1030705-E
計畫名稱：使用隨機模型建構傳染病模型
計畫主持人：陳立昇(臺北醫學大學)/賴昭智(仁愛院區)
計畫書版本/日期: Ver1.0_1030627
同意書版本/日期: 免受試者同意書
計畫期間為民國 103 年 07 月 01 日起至民國 104 年 06 月 30 日
上述計畫業經本院 103 年第 9 次人體試驗委員會會議審查(追認核備)，民國 103 年 07 月 25 日審查通過，特此證明。

本許可書自民國 103 年 07 月 25 日起至民國 104 年 06 月 30 日止。

※追認核備之案件，若會議有疑義時，本會有權撤銷此計畫執行許可書。
※未完成期中報告或結案報告者不得申請新案。
※請於計畫執行許可書有效期限到期前一個月內繳交期中報告。
※請於計畫執行許可書到期前繳交結案報告。
※計畫內容若有任何修改或增減，計畫主持人或計畫委託單位需於許可書有效期限到期日 6 週前檢送修正案至本會，經審查通過後方可實施。

田麗珠 田麗珠

臺北市立聯合醫院人體試驗委會主任委員

CERTIFICATE

Date：July 28, 2014

The project entitled "Stochastic Model in Epidemic Outbreak "submitted by investigatorLi -Sheng Chen/Chaochih Lai has been approved by Institutional Review Board of the Taipei City Hospital.
Protocol Version Date: Ver1.0_1030627
ICF Version Date:Waiver

Study period: since 07/01/2014 to 06/30/2015
Above project is approved by the TCHIRB on 07/25/2014 and valid till 6/30/2015.

※ All protocols should been subject to final endorsement by the board. TCHIRB has the right to revoke the approval.
※ No new protocol can be applied if the midterm report or the final report is not handed in.
※ The midterm report should be handed in one month before the expiry date of this certificate.
※ The final report should be handed in before the expiry date of this certificate or project end.
※ The PI should send the amendments to TCHIRB at least six weeks before the expiry date of this certificate if there are changes or modification related the protocol. Any change should not be executed until being approved by TCHIRB.

Lih-Chu Tien
Lih-Chu Tien
Chairman
Taipei City Hospital Institutional Review Board

The committee is organized and operates in accordance with ICH-GCP regulations and guideline.
本委員會組織與運作皆遵守 ICH-GCP 規定