國立臺灣大學公共衛生學院流行病學與預防醫學研究所

博士論文

Institute of Epidemiology and Preventive Medicine

College of Public Health

National Taiwan University

Doctoral Dissertation

以寇斯為導向之整合隨機模型評估族群篩檢效益

Consolidated Coxian Phase-type-based Stochastic Model for Evaluation of Effectiveness of Population-based Screening

任小萱

Hsiao-Hsuan Jen

指導教授：陳秀熙　博士

Advisor：Hsiu-Hsi Chen, Ph.D.

中華民國 108 年 7 月

July, 2019

doi:10.6342/NTU201901194

Abstract

Background Effectiveness of population-based colorectal cancer (CRC) screening with fecal immunological test (FIT) in reducing advanced CRC and deaths from CRC is determined by structure, process, multistate disease natural history, and screening policy. Systematic evaluation of how these factors are interplayed with each other to affect the benefit of screening is a thorny issue and cannot rely on traditional experimental and quasi-experimental design and statistical analysis. Therefore, a mathematical modelling is considered an alternative approach.

Aims The consolidated Coxian phase-type Markov (CPHM) process, consisting of three structure-process-based, disease-natural-history-based, and prognostics-based CPHM models, is therefore proposed for evaluating the effectiveness of screening from the components of structure and process to screening policy such as inter-screening interval.

Data Source This thesis is motivated by population-based screening program for CRC with biennial FIT in Taiwan. A total of 5,417,699 eligible population (aged 50-69 years in 2004-2009) composed our study cohort. In 2004-2014, the program covered 56.7% (n=3,074,538) eligible population attending at least one FIT. The positive rate was 6.9% (n=211,888) in the first round. Among them, 67% (n=142,800) underwent confirmatory examination. Data on all screening history together with screening findings were prospectively collected in the central screening monitor system. A total of 120,528 subjects were detected as adenoma. There were

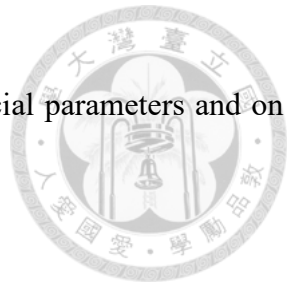
doi:10.6342/NTU201901194

71,543 CRC diagnosed upon screening, including 9,168 and 4,653 screen-detected cases in the first and subsequent rounds of screen, respectively, 11,904 interval cancers, and 45,818 refusers among those never attended. Data on the distribution of AJCC staging were available.

Methods The first CPHM model together with the likelihood functions pertaining to structure and process of population-based service screening was developed to evaluate arrival rate, positive rate, compliance with and waiting time for confirmatory diagnosis. The second series of CPHM models associated with multi-state disease natural history of CRC were developed to model a non-Markovian process with hidden transient states in PCDP and absorbing states reaching to CP classified by tumour staging in order to evaluate the potential of upstaging and to project the effect of inter-screening interval on effectiveness of screening. The prognosis-based CPHM model was further used to assess the benefit of screening based on empirical data.

Results The estimated parameters from structure-process-based CPHM models were used to elucidate the relationships among arrival rate, positive rate, and waiting time. Those from the second series of natural history CPHM models were used to evaluate the force of upstaging (early stage to late stage) within PCDP by age, gender, and location of CRC with the theoretically proven indicators and also assess the effect of inter-screening interval on the effectiveness of screening. Survival of early and late stage in PCDP and CP, respectively, was assessed through prognostics-based CPHM model. Statistical simulation studies on each

doi:10.6342/NTU201901194



CPHM model were conducted to elucidate statistical inference on crucial parameters and on perturbation analyses with respect to influential parameters.

Conclusion The consolidated CPHM model composed of a cascade of CPHM models was proposed to evaluate multi-throng effectiveness of population-based cancer screening from structure, process, and outcome.

doi:10.6342/NTU201901194

中文摘要

研究背景 利用免疫化學法糞便潛血檢查(Fecal immunological test, FIT)進行組織性大規模篩檢計畫在降低晚期大腸直腸癌發生或大腸直腸癌死亡的效益取決於篩檢計畫的結構—結果因素、疾病的多階段進展過程及篩檢策略，這些因素的改變如何彼此影響並決定最終篩檢效益的系統性評估並無法仰賴傳統實驗型或類實驗型研究加以驗證，建構數學模型進行參數改變的驗證則是另一可行之道。

研究目的 本論文旨在建立以寇斯為導向之整合隨機模型，考量結構—結果因素、疾病的多階段進展過程及疾病預後三大面向，用以結構—結果因素或篩檢策略改變時對族群篩檢效益之影響。

資料來源 本論文統計模型發展動機源自於臺灣大腸直腸癌大規模篩檢計畫之評估。臺灣自 2004 年開始由政府提供兩年一次的免疫化學糞便潛血檢驗做為大腸直腸癌篩檢之工具，本論文以 2004 年至 2009 年間篩檢的目標族群(50-60 歲民眾共 5,417,699 人)做為研究世代。在 2004-2014 年間，共計 3,074,538 人參加過至少一次篩檢，篩檢涵蓋率為 56.7%；參加第一次篩檢的個案中有 211,888 人檢查結果為陽性，陽性率為 6.7%；陽性個案中完成轉介者計 142,800 人，轉介完成率為 67%；參與篩檢者在 2004-2014 年間共偵測出 120,528 例大腸腺腫及 71,543 例大腸直腸癌個案(包括 9,168 例第一次篩檢發現個案及 4,653 例後續篩檢個案)，另有 11,904 人為間隔癌個案，而在未參加篩檢的族群中則有 45,818 人被診斷為大腸直腸癌(拒絕個案)。

doi:10.6342/NTU201901194

研究方法 首先建構以寇斯為導向之整合隨機模型，結合不同篩檢模式所因應的概似函數，估計族群的篩檢參與，在參與個案抵達篩檢計畫後，應用結合寇斯多相模式及閾值模式描述陽性個案接受轉介之順從度及等候時間的多相樣態，並且以另一寇斯多相非馬可夫模式探討症狀前期時隱藏狀態的可能性，及抵達症狀期(吸收狀態)依 AJCC 癌症分期分類定義多相態，並據以評估不同篩檢參數對癌症分期分佈之效益，最後，結合疾病預後的寇斯多相模式，評估篩檢組織面、過程面及結果面三類要素在不同情境之下對組織性大規模篩檢於降低晚期大腸直腸癌發生或大腸直腸癌死亡的效益之影響。

結果 以寇斯為導向之整合隨機模型應用於臺灣大腸直腸癌大規模篩檢計畫所估計得到的參數為基礎，評估不同篩檢參與抵達率、FIT 陽性率及大腸鏡陽性轉移等待時間之關係。第二個寇斯多相非馬可夫模式所估計得到的疾病多階段病程轉移參數則用來評估篩檢參數對癌症分期分佈的影響，並進一步討論性別、年齡別及部位別(近端大腸及遠端大腸)在晚期癌症因篩檢計畫所產生的分佈差異，並發展相關統計推論指標評估其影響性。依大腸直腸癌在症狀前期或症狀期被診斷時的早期及晚期狀態所對應的存活率建構預後寇斯多相非馬可夫模式。三個寇斯多相非馬可夫模式進一步以電腦模擬方法得到評估篩檢效益重要指標之統計推論性，並討論篩檢組織面、過程面及結果面三類要素在不同情境之下對組織性大規模篩檢於降低晚期大腸直腸癌發生或大腸直腸癌死亡的效益之影響。

doi:10.6342/NTU201901194

結論 本論文利用寇斯多相非馬可夫模式整合隨機模型廣納可能影響篩檢的結構—結果因素、疾病的多階段進展過程及疾病預後三大面向，並用以估計大規模篩檢計畫之效益。

doi:10.6342/NTU201901194

Contents

doi:10.6342/NTU201901194

doi:10.6342/NTU201901194

doi:10.6342/NTU201901194

List of Figures

doi:10.6342/NTU201901194

doi:10.6342/NTU201901194

List of Tables

doi:10.6342/NTU201901194

doi:10.6342/NTU201901194

doi:10.6342/NTU201901194

Chapter 1. Coxian Phase-type Stochastic Process Applied to Population-based Cancer Screening

Summary

This chapter gives a prelude to the context and organization of my thesis. It begins with the provision of the rationale for the use of mathematical modelling approach in evaluation of population-based service screening program. Then, it accounts for why the consolidated Coxian phase-type Markov process is required to model structure, process, and the outcome pertaining to the disease natural history descended from population-based organized service screening program. As such a screening process and disease natural history may be a non-Markovian process the basic concepts of two main methods, supplementary variable and stage device method, are introduced. Theoretical aspect of connecting non-Markovian process with Coxian phase-type Markov process through the stage device method that is the main approach of my thesis is delineated. The titles of the remaining six chapters are provided for the panorama of my thesis.

doi:10.6342/NTU201901194

1.1 Mathematical modelling for evaluation of population-based organized service screening

Although population-based screening for cancer that can be detected with available screening methods has been well applied evaluation of its efficacy and effectiveness often relies on a randomized controlled trial or quasi-experimental design with the adequate control group. The methodology used for such an evaluation follows intention-to-treat analysis for the former and self-section bias adjustment analysis for the latter. Population-based organized service screening program is a classical quasi-experimental pattern. Although the methodology on selection-bias has been successfully applied to estimating the effectiveness of reducing mortality and advanced cancer in population-based screening (Duffy SW et al. (2002), Tabar L et al. (2003), and Wu JC et al. (2010)) two reasons preclude one from doing evaluation in this way. First of all, the unbiased control group in the absence of screening is hardly available when estimating the effectiveness of organized service screening gets involved with self-selection bias adjustment. It may be possible to use the comparator based on the pre-screening period but it becomes intractable when there is an increasing trend in the underlying incidence of cancer of interest due to biological plausibility. Second, there are occasions that may not be appropriate for the use of experimental or quasi-experimental designs for evaluation of effectiveness of population-based organized service screening. Whether the effectiveness of

doi:10.6342/NTU201901194

the screening method in reducing advanced cancer varies with anatomic site cannot rely on such a kind of design but is amenable to the modelling approach. It is argued that a higher rate of interval cancer seen in proximal CRCs in comparison with distal CRCs is mainly due to high potential of upstaging or poor sensitivity. To throw light on this argument may resort to a mathematical modelling rather than the traditional epidemiological design. Third, evaluation of population-based organized service screening with primary endpoint like cause-specific mortality may require long-term follow-up and get involved with enormous costs and complex logistics. The use of surrogate endpoint such as tumour staging for evaluation of primary endpoint is an alternative. The previous study has already demonstrated the advantage of using surrogate endpoint may increase statistical efficiency by reducing approximately one-third variance when the surrogate endpoint outcome as opposed to primary endpoint (Chiu SYH, et al. (2011)). Finally, doing the appraisal of effectiveness in the way indicated above may not decipher how and why key components play important role in the quantification of the benefit of screening. These include coverage rate, positive rate, compliance rate, and inter-screening interval. Analysis of the impact of these complex factors on the effectiveness of population-based screening requires the development of a comprehensive mathematical model for annexing a series of different types of stochastic models including structure, process, and natural history of disease outcome in relation to population-based screening. These influential

doi:10.6342/NTU201901194

factors are characterized by different aspects including pre-determined designed variables such as coverage rate and inter-screening interval and the underlying factors related to the disease process of CRC like the mean sojourn time, the test sensitivity of the screening method, and different potentials of up-staging progression of CRC by anatomic site.

From statistical viewpoint, the use of modelling approach to evaluation of population-based organized service screening program requires perturbation analysis on the crucial parameters and also simulation studies on asymptotic analysis. They have been barely addressed.

1.2 Consolidated Coxian phase-type stochastic model

To deal with the empirical data on these multi-attributes covering structure, process, and the outcome of disease natural history, it is indispensable to integrate different types of stochastic model into a unifying framework. As far as structure is concerned, it begins with the application of a Queue process to deal with coverage rate at different levels of institution in various areas. The Hurdle Coxian phase-type (CPH) model is used to deal with the compliance with and waiting time (WT) for confirmatory diagnosis. As regards the aspect of disease outcome, another CPH is then annexed with the previous CPH tailored for confirmatory diagnosis is therefore developed to model the disease progression of natural history with multi-

doi:10.6342/NTU201901194

state outcomes.

1.3 Coxian phase-type-based model and non-Markovian process

When the CPH models mentioned above are used to model WT for confirmatory diagnosis and model the disease natural history during the pre-clinical detectable phase (PCDP), they often violate the Markov property and become a non-Markovian process. There are three main methods used for solving such a non-Markovian process. These include the stage device, supplementary variables, and embedded Markov process according to the Cox and Miller method (Cox and Miller (1965)). Among three methods, the stage device method is the main subject of my thesis because the classification of cancer in the PCDP can be divided into k stage with TNM information supplemented with other possible biomarkers if available although the two latter methods are still very useful. The accurate and optimal classification of hidden state of PCDP may also determine whether such a classification of staging is a good surrogate endpoint for primary endpoint using cause-specific mortality. For example, the classification of stage II and III of CRC in terms of AJCC is not distinct. If only stage is used for predicting cause-specific mortality the estimate of predicted effectiveness of reducing cause-specific mortality may be conservative.

The basic concepts of supplementary variable approach on the stage device method are

doi:10.6342/NTU201901194

delineated as follows.

Suppose a subject enters the PCDP at age *a* and stay at PCDP with time *y*. If two random variables are denoted by A and Y and they are independent with probability density functions (p.d.f.) $\lambda_1 \exp(-\lambda_1 \times a)$ and f(*y*).

Suppose we have three transition probabilities, $P_{normal}(t)$, $P_{PCDP}(t)$, and $P_{CP}(t)$, representing the probability of staying in normal, PCDP and CP. We then have

$$\frac{d}{dy}P_{PCDP}(y;t) + \frac{d}{dt}P_{PCDP}(y;t) = -\lambda_2(y)P_{PCDP}(y;t) \quad (1\text{-}1)$$

$$P_{PCDP}(0,t) = \lambda_1 e^{-\lambda_1 t} \quad (1\text{-}2)$$

$$\frac{d}{dt}P_{CP} = \int_0^t P_{PCDP}(y;t)\lambda_2(y)dy \quad (1\text{-}3)$$

$$P_{PCDP}(y;t) = \{e^{-\int_0^y \lambda_2(u)du}\}S_{PCDP}(y;t) \quad (1\text{-}4)$$

$$P_{PCDP}(y;t) = \lambda_1 e^{-\lambda_1(t-y)}\exp\{-\int_0^y \lambda_2(u)du\} = \lambda_1 e^{-\lambda_1(t-y)}S(y) \quad (1\text{-}5)$$

$$P_{CP}(t) = \int_0^t du \int_0^u e^{-\lambda_1(u-y)}S(y)\lambda_2(y)dy = \int_0^t \lambda_1 e^{-\lambda_1(t-u)}\{1 - S(u)\}du \quad (1\text{-}6)$$

The equation (1-5) and (1-6) is simply Prob($A + Y < t$). Here, the transition rate (λ_1) from normal to PCDP state follows exponential distribution but the second transition (λ_2) that is not exponential distribution and is dependent on the duration of y, such a process is a non-Markov process. This is often solved by semi-Markov process by the introduction of duration distribution for y. My thesis used the alternative approach with the device stage method to

doi:10.6342/NTU201901194

derive the distribution of Y that is divided into the sum or *K* independent random variables, and its distribution has Laplace transform

$$\Pi \frac{\lambda_{2i}}{\lambda_{2i}+s} \qquad (1\text{-}7)$$

When the λ_{2i}'*s* are all different, (1-7) can be expressed in partial function form

$$\sum \frac{w_i \lambda_{2i}}{\lambda_{2i} + s}$$

So that the p.d.f is

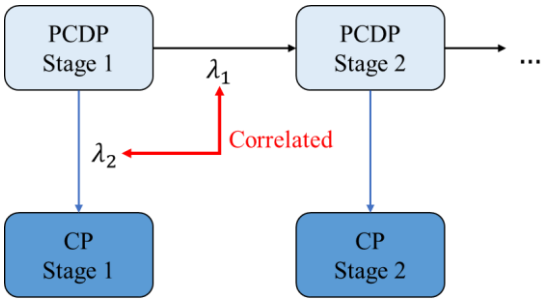
$$\sum w_i \lambda_{2i} e^{-\lambda_{2i} y}$$

The direct application of device stage method to such a non-Markovian process corresponds to the so-called phase-type stochastic process. The transition of λ_2 is the modified parallel form of the stage device method with the p.d.f of each distribution of random variable denoted by x expressed by

$$\sum \pi_i \lambda_2 e^{-\lambda_2 x}.$$

However, when the device stage method is applied to population-based screening data using tumour stage to refine the classification of PCDP there is a complication of direct use of phase-type process as each state of PCDP has a competing pathway leading to clinical phase carrying with each specific tumour stage when staying in the PCDP.

doi:10.6342/NTU201901194



The solution to combine both pathways is the development of Coxian phase-type stochastic process, which becomes the main subject of this thesis. There are several Coxian phase-type-based stochastic process proposed in previous studies (Marshall AH, et al. (2007), Marshall AH, et al. (2009), Titman AC et al. (2010)). However, very few studies have noticed the possibility of having identifiability problem with the estimation of two correlated transition parameters between the transition to adjacent up-staging state and the transition to the absorbing state for such a kind of Coxian phase-type-based model. This is the main subject of my Chapter 6.

1.4 Overall structure of thesis

In the light of the rationales proposed above in Chapter 1, there are the remaining six chapters of this thesis that delineate the development of various kinds of Coxian phase-type stochastic process and their applications to Taiwanese population-based organized service screening program with FIT for CRC. The titles of six chapters are listed as follows.

doi:10.6342/NTU201901194

Chapter 2 Literature review

Chapter 3 Queue, non-compliance, and waiting time of FIT screening service for CRC with hurdle Poisson regression model

Chapter 4 Queue Hurdle Coxian Phase-type Model for Waiting Time in Two-stage Population-based Screening

Chapter 5 Generalized Coxian phase-type Markov process for disease natural history model of CRC

Chapter 6 Statistical indicators for up-staging of CRC assessed by Coxian Phase-type Markov process with device stage method

Chapter 7 Consolidated Coxian phase-type Markov (CPHM) process

doi:10.6342/NTU201901194

Chapter 2. Literature Review

Summary

Statistical methods that are possibly used for modelling structure, and process of organized service screen, and the underlying disease natural history are briefly reviewed. These include hurdle regression model for the first and second part, stochastic models for the disease natural history part, and Coxian phase-type process for three parts. The critiques of these methods in terms of their applications to population-based screening program are several-fold. First, how to connect these models to deal with the underlying screening flowchart from structure, process, until the disease natural history model has not been well studied. Second, statistical properties for modelling the disease natural history haven't been well elucidated. For example, the correlation between upstaging with pre-clinical detectable phase (PCDP) and downstagin from PCDP to clinical phase has not been investigated yet. Third, statistical indicators for evaluating the force of upstaging have not been developed to judge whether the adoption of screening is adequate and how intensive inter-screening interval should be offered. Finally, the consolidation of these statistical models into a unified framework for evaluation of effectiveness of population-based screening program has been barely addressed. .

doi:10.6342/NTU201901194

2.1 Introduction

This chapter is to give a brief literature review of statistical methods with relevance that may be applied to the empirical data on population-based screening program on structure, process, and disease natural history. These include hurdle regression model that is used to model the non-attendance rate and non-compliance rate with confirmatory diagnosis, Markov process for modelling multi-state disease natural history, and Coxian phase-type distribution for modelling relevant transient states and absorbing states.

2.2 Hurdle Regression Model

Hurdle regression model is often used to model mixture outcomes. The hurdle model consists of binary part and count part. In the binary part, we firstly determine whether the event occur (count greater than or equal to one) or not (count equal to zero) by using the logistic regression model. Then, we model the number of event given count greater than zero by using the Poisson or negative binomial regression model. In the past two decades, the hurdle model has been widely applied in healthcare science (Winkelmann, R. (2004), Deb P et al. (2018), Langsetmo L et al. (2017), and Neelon B et al. (2013)), and more recently used in other fields such as environmental heath, non-communicable disease, and genetics. In 2018, Zhen et al. used the negative binomial hurdle model with spatial random effects to predict the risk of

doi:10.6342/NTU201901194

children's lead poisoning by socio-economic variables and environmental factors. The outcome is the number of children whose blood lead levels ≧ 10μg/dL in each census block. Because some area did not have any cases, which shows the issue of excessive zeros, so the traditional Poisson model was not suitable in this situation. Therefore, the hurdle model is an appropriate method to overcome this problem and separate zero count and positive counts by two-stage approach. The results of model selection demonstrate that the negative binomial hurdle model was better than the Poisson hurdle model because it has a dispersion parameter that can further account for the overdispersion from the excessive zeros. Lin, et al. (2018) applied hurdle regression model to explore the association between *Aedes aegypti* and characteristics of water holding containers in southern Taiwan. The survey data show that most of containers has zero *Aedes aegypti*. Thus, they firstly use a logistic regression model to model the possibility of the presence of *Aedes aegypti* in the containers. The association between location, seasonal, and functional characteristics and the number of *Aedes aegypti* among positive containers were further modelled by using Poisson or negative binomial distributions. They also compared hurdle model with zero-inflated model which is also commonly used to deal with excess zeros. The main difference is that the binary part in zero-inflated model often include false zeros and true zeros but that in hurdle model only have true zeros. Hurdle model was selected in this study because that all zero containers are all due to the inappropriate environment which means

doi:10.6342/NTU201901194

true zeros.

2.3 Multi-state Model for Natural History of Cancer

Multi-state models are widely used to establish the disease natural history of chronic diseases and cancer. Once the natural history in which the disease progression would not be interrupted by intervention or treatment has been established, a pseudo-control group can be developed to compare with intervention (i.e, aspirin, screening, polypectomy) and then to evaluate the efficacy and effectiveness of the intervention. For example, in population-based mass screening program, the evaluation of screening effectiveness is difficult without a control group. As a result, the three-state model is a typical approach for describing the natural history of cancer by screening. In 2000, Chen et al. proposed a three-state model to estimate the mean sojourn time on the PCDP in breast cancer screening. The state space was defined as X(t)={0: normal, 1: PCDP, 2: clinical phase (CP)}. The definition of PCDP is for those asymptomatic individuals detected by screening modality, and CP is for those symptomatic subjects with clinical appearance (Day, and Walter, (1984)).The three-state model Markov model can be expressed as:

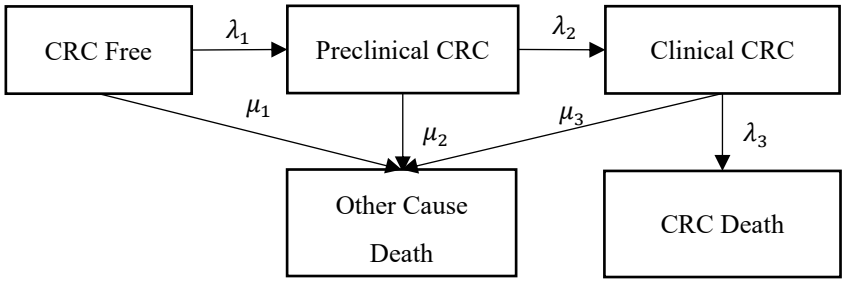
doi:10.6342/NTU201901194

In their study, they tackle data with irregular screening interval. Therefore, they used continuous time three-state Markov model. Such model was governed by the intensity matrix (Q) which can be given as follows,

$$Q = \begin{bmatrix} -\lambda_1 & \lambda_1 & 0 \\ 0 & -\lambda_2 & \lambda_2 \\ 0 & 0 & 0 \end{bmatrix}$$

where λ_1 is the transition rate from no disease to the PCDP and λ_2 is the transition rate from the PCDP to CP.

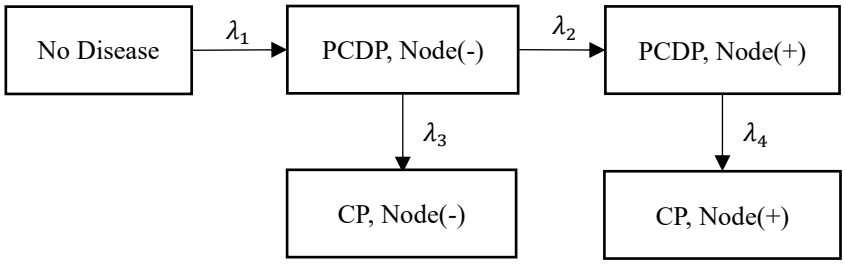
In 1999, Chen et al. extended the three-state model to the five-state model for the prognosis of CRC death. The additional two states are CRC death and other cause of death. The model was depicted below.



where λ_1 is the transition rate from CRC free to the preclinical CRC, λ_2 is the transition rate from the preclinical CRC to clinical CRC, λ_3 is the transition rate from the clinical CRC to death from CRC; μ_1, μ_2, μ_3 are the absorbing rate of other cause death from CRC free, preclinical CRC, and clinical CRC, respectively. By using this five-state Markov model, they did a simulation to project the efficacy of CRC screening for high-risk group in different

doi:10.6342/NTU201901194

interscreening intervals so that it can provide decision makers a reference of screening policy. However, it is not a good model for prediction of death from other disease and evaluation on the effectiveness of screening program, because it did not contain any information about treatment after diagnosis. In addition, someone might argue that the transition from preclinical CRC to CRC death is possible. This model can only be used to evaluate the efficacy of prognosis, which can be obtained by survival analysis as well. The difference between two approaches is that the multi-state model is a parametric method and the survival analysis is a non-parametric method. Therefore, if the state space of five-state model can be defined by tumour attributes, cancer stage, or clinical measures, which correlate with prognosis of diagnosis outcome, the pathway of disease progression can be predicted accurately. For example, Chen et al. (2000) extended the three-state model to a five-state model relating to lymph node spread of breast cancer as well as tumour size. Taking lymph node for example, the state space of five-state model was no disease (0), PCDP without lymph node involvement (1), PCDP with lymph node involvement (2), CP without lymph node involvement (3), and CP with lymph node involvement (4) as shown below



doi:10.6342/NTU201901194

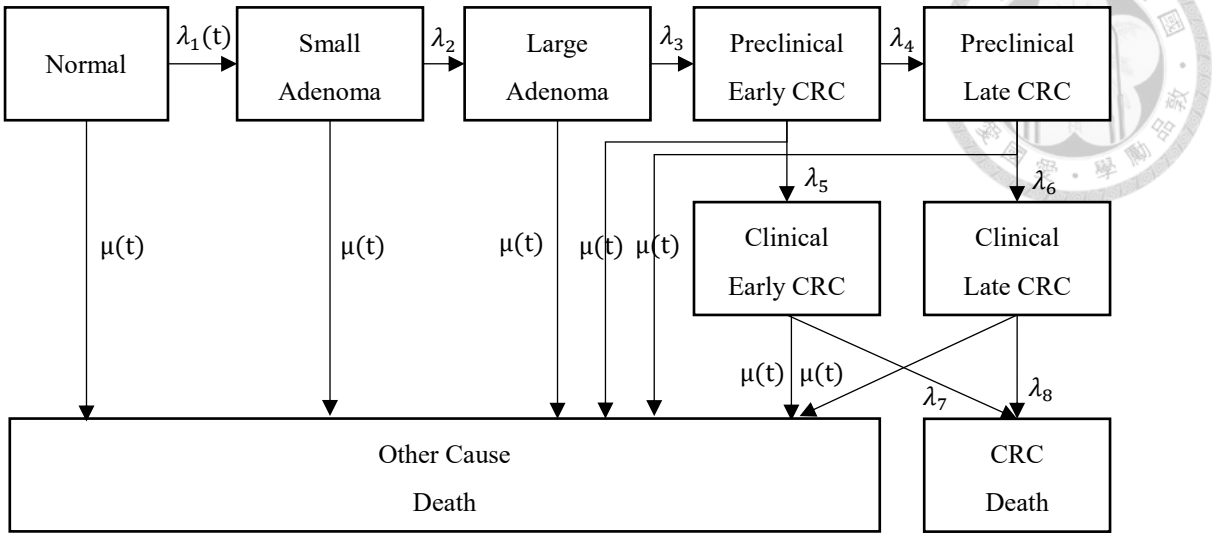
The intensity matrix can be expressed as

$$\begin{bmatrix} -\lambda_1 & \lambda_1 & 0 & 0 & 0 \\ 0 & -(\lambda_2+\lambda_3) & \lambda_2 & \lambda_3 & 0 \\ 0 & 0 & -\lambda_4 & 0 & \lambda_4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

where λ_1 is the transition rate from no disease to the PCDP, λ_2 is the transition rate from the PCDP without lymph node involvement to PCDP with lymph node involvement, λ_3 is the transition rate from the PCDP without lymph node involvement to CP without lymph node involvement, and λ_4 is the transition rate from the PCDP with lymph node involvement to CP with lymph node involvement.

In 2006, Wu et al. aimed to evaluate the cost-effectiveness of CRC screening modalities. They proposed a nine-state Markov model to simulate the natural history of CRC. The state space is X(t)={0:normal, 1:small adenoma (adenoma size<1 cm), 2:large adenoma (adenoma size≧1 cm), 3:pre-clinical early CRC (Duke's stage A and B), 4:pre-clinical late CRC (Duke's stage C and D), 5:clinical early CRC (Duke's stage A and B), 6:clinical late CRC (Duke's stage C and D), 7:CRC death, 8:other cause of death}, and the model can be illustrated by

doi:10.6342/NTU201901194



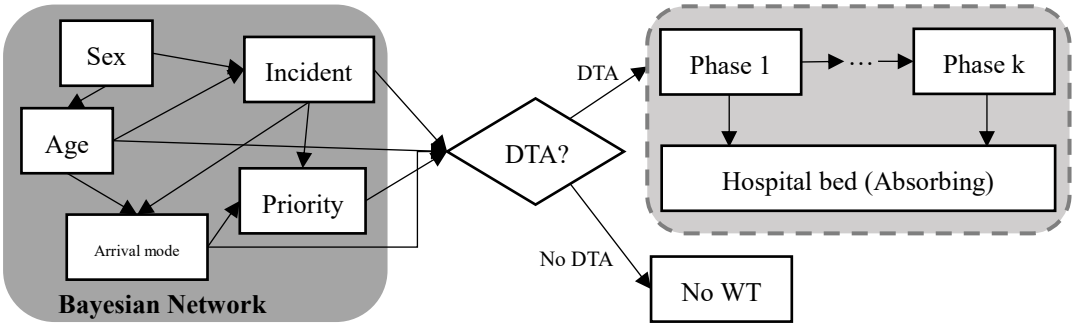
where $\lambda_1(t)$ represents the annual incidence rate allowing to vary with time, following the Weibull distribution; $\lambda_2 - \lambda_8$ are the transition rate with constant over time; $\mu(t)$ is the annual age-specific mortality from other causes.

2.4 Coxian Phase-type Process

Coxian Phase-type (CPH) model is a subclass of Phase-type distribution, which is a continuous time Markov process and has been widely used in the healthcare industry to model the positive-skewed time-to-event data. Using the Phase-type distribution to estimate parameters might encounter the identifiability problem staying in a transient phase can transit to any other transient phase at the next step. However, the CPH has the ordered phases and can only start in the first phase and transit to the sequential transient phase or to the absorbing state. Therefore, the parameters need to be estimated reduce to 2k-1 (if the number of phases is k).

doi:10.6342/NTU201901194

In 1999, Faddy and McClean used it to estimate the length of stay (LOS) of hospital and also incorporated covariates (age of patient at admission and year of admission) into the exponential regression. The phases can be explained by the severity of disease and identify the patient's characteristics in the variation of duration time. Afterwards, the extension of CPH was increased to deal with the complicated information in the era of big data. Marshall et al. (2007) proposed the Discrete Conditional Coxian Phase-type (DC-Ph) model applying to the trolley waiting time (WT) at Accident and Emergency (A&E) departments. The first part of this model was to identify what kind of A&E patients have an emergency admission requiring further therapy. If he/she receives an emergency admission, then the second part of this model will estimate the WT from the clinician's decision to admit (DTA) until hospital bed allocation with CPH distribution. The illustration of the DC-Ph model can be delineated as



Those who had non-trauma case (incident type), severe level assigned by hospital staff (priority level), or arrived at hospital via ambulance or private transport (arrival mode) would have

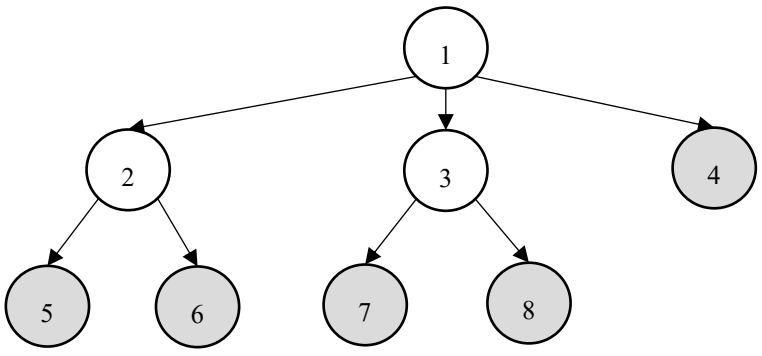
doi:10.6342/NTU201901194

higher likelihood of hospital admission. The DC-Ph model based on patient's information can provide assessment model to predict patient's risk and improve the efficiency of healthcare management.

As there are a great heterogeneity of patient pathways, McClean et al. (2010) proposed the mixed Coxian phase-type model (MC-PH) to deal with this problem and used the Coxian phase-type survival tree to estimate parameters. They used survival tree to cluster patients for homogeneity by covariates and define more specific pathway for each patient. Then, applying the MC-PH distribution, which combines two or more CPH distribution, model the skewed time data. The p.d.f is derived by

$$\mathrm{f(t)} = \sum_{i=1}^{c} \boldsymbol{\pi_c} \exp(\boldsymbol{Q_c} t) \boldsymbol{q_c}. \qquad (2.1)$$

In these c mixture components, it assumes that the transition probability of moving to another components is zero. The construction of CPH survival tree can be depicted as



In this example, patients can be clustered into five groups by covariates with survival tree. Each of terminal nodes (4-8) have different CPH distribution. The MC-PH can provide more precise

doi:10.6342/NTU201901194

information in comparison with C-Ph model.

The other method to deal with heterogeneity was using the Coxian phase-type regression (Tang et al., (2012)). They incorporated covariates directly into the intensity matrix of the CPH distribution. The transition rate and absorbing rate are replaced by

$$\lambda_k = \lambda_{0k}\exp(\boldsymbol{X\beta}) \tag{2.2}$$

and

$$\mu_k = \mu_{0k}\exp(\boldsymbol{X\beta}), \tag{2.3}$$

where λ_{0k} and μ_{0k} are baseline rate, separately. Then, the p.d.f can be derived as

$$\mathrm{f(t)} = \boldsymbol{\pi}\exp(\exp(\boldsymbol{X\beta})\,\mathbf{Q}\mathrm{t})\,(\exp(\boldsymbol{X\beta})\,\mathbf{q}) = \boldsymbol{\pi}\exp\left(\widetilde{\boldsymbol{Q}}\mathrm{t}\right)\widetilde{\boldsymbol{q}} \tag{2.4}$$

where

$$\widetilde{\boldsymbol{Q}} = \exp(\boldsymbol{X\beta})\,\mathbf{Q} \tag{2.5}$$

$$\widetilde{\boldsymbol{q}} = \exp(\boldsymbol{X\beta})\,\mathbf{q}. \tag{2.6}$$

In 2017, Donnelly et al. applied the CPH regression model within a two-stage approach that using the linear mixed-effect model described how responses changed with time in the first stage (Stage I), and using the CPH distribution model the survival process in the second stage (Stage II). The random effects considering in Stage I would be counted as covariates within the CPH distribution. In Stage I, the response of interest would be modelled by the generalized

doi:10.6342/NTU201901194

linear mixed model as below

$$\boldsymbol{Y_i} = \boldsymbol{X_i\beta} + \boldsymbol{Z_i b_i} + \boldsymbol{\epsilon_i} \tag{2.7}$$

where $\boldsymbol{\beta}$ is a vector of fixed effects, $\boldsymbol{b_i} = (b_{i0} \;\; b_{i1})^{\mathrm{T}}$ is a vector of individual-specific random effects following $\mathrm{BN}(\mathbf{0}, \boldsymbol{\Sigma})$ that b_{i0} and b_{i1} are random intercept and random slope (rate change with time), and $\boldsymbol{\epsilon_i}$ is a vector of residual errors following $\mathrm{N}(\mathbf{0}, \boldsymbol{R_i})$ with $\boldsymbol{R_i} = \sigma^2 \boldsymbol{I}$. After estimating the random effects in Stage I, the predicted value would be used within the CPH regression. Then, the p.d.f is given by

$$\mathrm{f(t)} = \boldsymbol{\pi} \exp(\exp(\boldsymbol{b_i\beta})\, \mathbf{Q}\mathrm{t})\, (\exp(\boldsymbol{b_i\beta})\, \mathbf{q}). \tag{2.8}$$

The proposed model can be used to derive individual transition rates for developing a technique of personalized medicine.

2.5 Statistical applications to systematic evaluation of population-based screening data

While these statistical methods are useful on certain occasion their applications to population-based screening data are still very limited. First, there is lacking of annexation of these proposed models when applied to population-based screening. How to connect these models to deal with the underlying screening flowchart from structure, process, until the disease natural history model has not been well studied. Second, statistical properties for modelling the disease natural history haven't been well elucidated. For example, the correlation

doi:10.6342/NTU201901194

between upstaging with pre-clnical detectable phase (PCDP) and downstagin from PCDP to clinical phase has not been investigated yet. Third, statistical indicators for evaluating force of upstaging have not been developed to judge whether the adoption of screening is adequate and how intensive inter-screening interval should be offered. Finally, the consolidation of these statistical models into an unified framework is fundamental to evaluation of effectiveness of population-based screening program.

doi:10.6342/NTU201901194

Chapter 3. Queue, Non-compliance, and Waiting Time of FIT Screening Service for CRC with Hurdle Poisson Regression Model

Summary

Introduction Population-based colorectal cancer (CRC) screening program with Fecal Immunochemical Test (FIT) is often faced with a non-compliance issue and its subsequent waiting time (WT) for those FIT positives complying with confirmatory diagnosis.

Aims To identify factors associated with both of two correlated problems in the same model.

Methods A total of 294,469 subjects either with positive FIT tests or having family history collected from 2004 to 2013 were enrolled for analysis. We applied a hurdle Poisson regression model to accommodate the hurdle of compliance and also its related WT for undergoing colonoscopy while assessing factors responsible for the mixture of two outcomes.

Results The effect on compliance and WT varied with contextual factors such as geographic areas, type of screening units, and level of urbanization. The hurdle score, representing the risk score in association with non-compliance, and the WT score, reflecting the rate of taking colonoscopy, were used to classify subjects into each of three groups representing the degree of compliance and the level of heath awareness.

Conclusion Our model was not only successfully applied to evaluating factors associated with the compliance and the WT distribution, but also developed into a useful assessment model for

doi:10.6342/NTU201901194

stratifying the risk and predicting whether and when screenees comply with the procedure of receiving confirmatory diagnosis given contextual factors and individual characteristics.

doi:10.6342/NTU201901194

3.1 Introduction

Hurdle Poisson regression model dealing with non-compliance and waiting time in population-based cancer screening

Population-based screening program is often faced with a non-compliance issue among screen-positive waiting for confirmatory examination similar to the scenario seen in most of phase III clinical trials (Cuzick J et al. (1997)). It is therefore of great interest to identify factors associated with the causes of non-compliance (Duffy SW et al (2002), Giorgi RP et al. (2005), and Siddiqui AA et al. (2006)). Taking colorectal cancer (CRC) screening with fecal immunological test (FIT) for example, whether to comply with colonoscopy examination among those FIT positives is a thorny issue and identifying factors associated with non-compliance is worthy of being investigated.

However, in addition to the compliance issue, the subsequent waiting time (WT) for those complying with colonoscopy is also a great concern as it is an indicator for the quality assurance of organized service screening. The longer the WT, the more likely the quality of screening program would be compromised due to the occurrence of interval cancers and also the influence of increasing patient's anxiety (Yu D et al. (2008), and Denters MJ et al. (2013)).

It is customary to deal with non-compliance and WT in two separate processes. The conventional binary model such as logistic regression model is often used for non-compliance

doi:10.6342/NTU201901194

whereas the simple Poisson regression model or Cox proportional hazards regression model is employed to model the WT process. It should be noted that both non-compliance and its related WT issues are correlated, the higher compliance rate, the greater number of positive cases to be referred to undergo confirmatory examination, and the longer WT for confirmatory diagnosis would be expected. The application of methodology to considering non-compliance and WT simultaneously, which has been barely addressed, is therefore required. One of candidate models is the use of a hurdle Poisson regression model to elucidate the factors responsible for non-compliance and WT for undergoing confirmatory diagnosis i.e. colonoscopy. The hurdle regression model proposed by Mullahy (2007) has been widely used in health care to solve the problem of heterogeneity in evaluation of hospital visits and health care expenditures in relation to zero cost resulting from censoring or no requirement for visit (Winkelmann R et al. (2004), and Deb P et al. (2018)) and also in the assessment of health care utilization with respect to whether to be hospitalized and the subsequent length of stay for hospitalization (Langsetmo L et al. (2017)).

It is therefore interesting to make better use of hurdle Poisson regression model, taking two issues into account, to ascertain contextual factors in association with the infrastructure of health care, geographic area and individual characteristics in order to predict whether and when is the probability of complying with the procedure of undergoing confirmatory diagnosis. The

doi:10.6342/NTU201901194

proposed model was applied to the empirical data from Taiwanese nationwide population-based CRC screening program.

3.2 Methods

3.2.1 Study framework and subjects

The application of both the hurdle model and the CPH model by using Taiwan nationwide CRC screening program with FIT. In brief, the program was offered for subjects aged 50 to 69 years biennially. After being detected as positive cases, screenees would be referred for confirmatory diagnosis. Details about the program have been described previously (Chiu HM et al. (2015)).

In this population-based cancer screening program, the eligible subjects pursuant to inclusion and exclusion criteria such as age are invited to have the uptake of screen with an appropriate screening modality. Some would have the uptake of screen whereas others would refuse to take it. This is first level of compliance issue in population-based screening but we are not tempted to deal with this issue in the current study. After the uptake of screening, attendees are classified into positive and negative results. Both positive and those who have family history of CRC among negative subject that are defined as high-risk subjects were referred to undergo confirmatory diagnosis such as colonoscopy, sigmoidoscopy or barium

doi:10.6342/NTU201901194

enema, but colonoscopy was the primary confirmation tool for CRC in our program. We therefore only focused on this instrument. Confirmation with colonoscopy defined by a binary outcome is the second level of compliance with the referral of undergoing confirmatory diagnosis, yielding non-complier and complier corresponding to the hurdle part and non-hurdle part. The overall study framework is shown in Figure 3.1

All the subjects who had ever attended this nationwide screening program between 2004 and 2013 with positive FIT and high risk group as defined above were enrolled in this study. Those who undertook at unspecified screening unit or non-qualified laboratory, received unknown kit to evaluate test characteristics, had missing fecal hemoglobin concentration or had been referred for confirmation diagnosis with non-colonoscopy modality were excluded from the study. As a result, there were 294,469 referrals either due to FIT positives or high-risk subjects used in the following analysis.

From 2004 to 2013, a total of 3,030,454 subjects participated in at least one time of FIT CRC screening. During the inaugural period, there were 1,258,560 screening tests and 46,235 referrals of whom 41,591 underwent colonoscopy. In the rolling-out period, there were 3,723,789 screening tests and 270,700 referrals of whom 252,878 underwent colonoscopy.

doi:10.6342/NTU201901194

3.2.2 Study variables and definition

Individual characteristics include gender and age at entry of screen. Contextual factors include calendar periods, round of screening, geographic areas (northern, central, southern, eastern and offshore islands of Taiwan), type of screening units (hospital, public health center of health bureau in individual municipality, or local clinic), and level of urbanization (metropolitan, sub-metropolitan and non-metropolitan). In the inaugural 5 years (2004-2009) of the screening program, we offered screening service mainly at the public health center of the health bureau in the individual municipality. Public health nursing staffs helped the referral of subjects to go for confirmatory colonoscopy at the hospitals. From 2010 onwards, for the purpose of increasing coverage rate, hospitals and local clinics joined to actively invite people for FIT and underwent colonoscopy at hospitals among referrals. Due to the change of screening policy, calendar periods were divided into inaugural period (2004-2009) and rolling-out period (2010-2013). Besides, screenees with positive FIT for first-time screening were defined as 'prevalent screen' and those for later screening rounds were defined as 'subsequent screen'.

Table 3.1 indicates most of attendees would undergo confirmatory colonoscopy after being detected as positive FIT. The compliance rate with colonoscopy was around 68% and the median WT for colonoscopy was 46 days. Individual characteristics had small influence on the

doi:10.6342/NTU201901194

compliance rate or the median WT, although male and the elderly had slightly lower compliance rate and longer median WT. Significant contextual factors included calendar periods, round of screening, and type of screening units. Screenees attending screening at the inaugural period had better compliance rate (82%) and shorter median WT (29 days). Subsequent cases had higher compliance rate and shorter median WT than prevalent screen cases. Among type of screening units, attendees undergoing screening at public health centers had the best compliance rate and only needed to wait 35 days for colonoscopy. The differences across geographic areas and level of urbanization were not substantial.

3.2.3 Hurdle Poisson regression model analysis

To accommodate the non-compliance of undergoing colonoscopy (non-complier) and also WT for undergoing colonoscopy among the compliers, the hurdle Poison model is proposed to consider both the non-complier and the WT distribution for the complier. Note that the time waiting for the appointment with colonoscopy administered by gastroenterologist was also included in the calculation of WT until the date of confirmatory diagnosis. The WT for colonoscopy here was defined from the date of reporting FIT result to the date of confirmatory diagnosis with colonoscopy. Because the change in the structure of screening program might result in heterogeneity between the inaugural period and rolling out period, we also evaluated

doi:10.6342/NTU201901194

the interaction between factors and periods of screening program.

The hurdle model, proposed by Mullahy (1986), consists of two components: the binary part is to measure whether the response crosses over the hurdle or not and the truncated Poisson part is to elucidate the response above the hurdle. These two components tackle both zeros and the positively skewed non-zero counts (Neelon B et al. (2013)). The hurdle part is the application of logistic regression model to identify factors affecting non-compliance with colonoscopy and non-hurdle (the progressive) part is modelled by the truncated Poisson regression model given the count greater than one to identify factors affecting WT for undergoing colonoscopy. In the zero part of truncated hurdle Poisson model, it consists of two types of subjects, illustrated by Figure 3.2. One is those who are not willing to undertake the confirmatory examination like Subject B and the other is those who are still waiting for confirmatory examination in a queue at 30th day (censoring) while the final WT is 37 days for Subject A.

In addition, Figure 3.2 accounts for why we need to use hurdle Poisson regression model instead of simple Poisson regression model and logistic regression model. Let the WT start from the date of positive FIT until the referral of colonoscopy examination. Suppose the WT for colonoscopy for the subject A is 37 days and the subject B is a non-complier, namely the infinity of WT, which cannot be captured by simple Poisson model or even negative binominal

doi:10.6342/NTU201901194

model. Thus, regarding the WT at 30th day, the waiting status for the subject A cannot be distinguished from the subject B although both are distinct. The zero of the subject A is captured by the truncated Poisson regression and the zero of the subject B is captured by the logistic regression, both of which form hurdle Poisson regression model. Therefore, if we use the traditional Poisson regression model, we cannot identify what sort of zero is derived from given time t and which factors will affect their health behavior of non-compliance.

Therefore, in our screening scenario, the hurdle was for those with FIT positive test or family history to determine whether he/she received the follow-up confirmation diagnosis. After passing through the hurdle, the subsequent truncated model dealt with how many subjects undertook the confirmatory examination given a specific period of time so that it can identify which factors are responsible for the WT for confirmatory examination.

3.2.4 Likelihood Function for Estimation of Parameters

After aggregating data by relevant covariates, we assume there are G subsets and define confirmation with colonoscopy as a binary outcome, y_{ij}, which represents whether the j-th screenee of subset i had underwent colonoscopy or not so that $y_i = \sum_j y_{ij}$ is the number of screenees had underwent colonoscopy in subset i and N_i is the total number of screenee in subset $i \in \{1,\ldots,G\}$. t_{ij} is time to undergo colonoscopy of the j-th screenee in subset i so that

doi:10.6342/NTU201901194

$t_i = \sum_j t_{ij}$ represents total waiting time in subset *i*. p_i is the probability of refusing to undergo colonoscopy (non-complier), and λ_i is the mean arrival rate of receiving colonoscopy. To easily express the likelihood function for the hurdle model, we divided each subset *i* into two groups by non-compliance or compliance, symbolizing i_1 and i_2 so that $i^* = \{I_1, I_2\} \cup ... \cup \{G_1, G_2\}$.

The hurdle is crossed if a count is greater than zero. Therefore, the process generates a binary response (whether the number of screenees undergoing colonoscopy takes on the value zero or a positive value), and the probability mass function (p.m.f) is

$$\mathrm{P}(Y_{i^*} = y_{i^*}) = \begin{cases} {p_{i^*}}^{N_{i^*} - y_{i^*}} & , y_{i^*} = 0 \ (i^* \epsilon i_1) \\ (1 - p_{i^*})^{y_{i^*}}, & y_{i^*} > 0 \ (i^* \epsilon i_2) \end{cases} \tag{3-1}$$

Because Y_{i^*} is a count data, which follows Poisson distribution, the probability of zero counts is

$$\mathrm{P}(Y_{i^*} = 0) = \frac{e^{-\lambda_{i^*} t_{i^*}} (\lambda_{i^*} t_{i^*})^{y_{i^*}}}{y_{i^*}!} = e^{-\lambda_{i^*} t_{i^*}} \tag{3-2}$$

and the probability of no zero counts equals $1 - e^{-\lambda_{i^*} t_{i^*}}$. As result, we can obtain the p.m.f of truncated Poisson process (given the count greater than one)

$$\mathrm{P}(Y_{i^*} = y_{i^*} | Y_{i^*} > 0) = \frac{e^{-\lambda_{i^*} t_{i^*}} (\lambda_{i^*} t_{i^*})^{y_{i^*}}}{y_{i^*}! \left(1 - e^{-\lambda_{i^*} t_{i^*}}\right)} \tag{3-3}$$

Thus, the p.m.f of hurdle model can be expressed as

$$\mathrm{P}(Y_{i^*} = y_{i^*}) = \begin{cases} {p_{i^*}}^{N_{i^*}} & , for\ y_{i^*} = 0 \\ (1 - p_{i^*})^{y_{i^*}} \frac{e^{-\lambda_{i^*} t_{i^*}} (\lambda_{i^*} t_{i^*})^{y_{i^*}}}{y_{i^*}! \left(1 - e^{-\lambda_{i^*} t_{i^*}}\right)} & , for\ y_{i^*} > 0 \end{cases} \tag{3-4}$$

where $0 \leq p_{i^*} \leq 1$; $\lambda_{i^*}, t_{i^*} > 0$; $N_{i^*}, y_{i^*} \geq 0$.

The general form of the likelihood function for the hurdle model is

doi:10.6342/NTU201901194

$$L = \prod_{i^* \epsilon i_1} {p_{i^*}}^{N_{i^*}} \prod_{i^* \epsilon i_2} (1 - p_{i^*})^{y_{i^*}} \frac{e^{-\lambda_{i^*} t_{i^*}} (\lambda_{i^*} t_{i^*})^{y_{i^*}}}{y_{i^*}! \left(1 - e^{-\lambda_{i^*} t_{i^*}}\right)} \quad (3\text{-}5)$$

To identify the effect of relevant covariates ($\boldsymbol{x}_{i^*}$), we model p_{i^*} using the logistic regression model

$$logit(p_{i^*}) = \boldsymbol{x}_{i^*}\boldsymbol{\gamma} \quad (3\text{-}6)$$

and λ_{i^*} using the Poisson regression model

$$log(\lambda_{i^*}) = \boldsymbol{x}_{i^*}\boldsymbol{\beta} \quad (3\text{-}7)$$

Thus, the log-likelihood can be written

$$\begin{aligned} lnL &= ln\left\{ \prod_{i^* \epsilon i_0} \left(\frac{e^{\boldsymbol{x}_{i^*}\boldsymbol{\gamma}}}{1 + e^{\boldsymbol{x}_{i^*}\boldsymbol{\gamma}}} \right)^{N_{i^*}} \prod_{i^* \epsilon i_1} \left(1 - \frac{e^{\boldsymbol{x}_{i^*}\boldsymbol{\gamma}}}{1 + e^{\boldsymbol{x}_{i^*}\boldsymbol{\gamma}}} \right)^{y_{i^*}} \frac{e^{-e^{\boldsymbol{x}_{i^*}\boldsymbol{\beta}} t_{i^*}} \left(e^{\boldsymbol{x}_{i^*}\boldsymbol{\beta}} t_{i^*} \right)^{y_{i^*}}}{y_{i^*}! \left(1 - e^{-e^{\boldsymbol{x}_{i^*}\boldsymbol{\beta}} t_{i^*}} \right)} \right\} \\ &= \left\{ \sum_{i^* \epsilon i_0} N_{i^*} [\boldsymbol{x}_{i^*}\boldsymbol{\gamma} - ln(1 + e^{\boldsymbol{x}_{i^*}\boldsymbol{\gamma}})] - \sum_{i^* \epsilon i_1} y_{i^*} \, ln(1 + e^{\boldsymbol{x}_{i^*}\boldsymbol{\gamma}}) \right\} \\ &+ \left\{ \sum_{i^* \epsilon i_1} y_{i^*} [\boldsymbol{x}_{i^*}\boldsymbol{\beta} + ln(t_{i^*})] - \sum_{i^* \epsilon i_1} e^{\boldsymbol{x}_{i^*}\boldsymbol{\beta}} t_{i^*} - \sum_{i^* \epsilon i_1} ln\left(1 - e^{-e^{\boldsymbol{x}_{i^*}\boldsymbol{\beta}} t_{i^*}} \right) - \sum_{i^* \epsilon i_1} ln(y_{i^*}!) \right\} \\ &= l\,n\{L_1(\boldsymbol{\gamma})\} + l\,n\{L_2(\boldsymbol{\beta})\} \end{aligned} \quad (3\text{-}8)$$

3.3 Result

3.3.1 Univariate and multivariate analysis of hurdle model

We used the regression coefficients of the hurdle model to form the scores in association with compliance with and the WT for colonoscopy. In hurdle part, the higher the score, the higher odds of refusing to take colonoscopy. In univariate analysis, the results of hurdle part in

doi:10.6342/NTU201901194

Table 3.2 show male and the elderly people had higher score among individual characteristics; those taking part in the rolling-out period, prevalent cases, central or eastern or offshore islands residents, non-metropolitan people, and local clinic among contextual factors had higher score.

In non-hurdle part, the higher the score, the shorter WT for undergoing colonoscopy was. The results of non-hurdle part indicate both individual characteristics among compliers were not associated with the WT for colonoscopy. Among contextual factors, those participating in screening at the inaugural period, subsequent cases, central Taiwanese, sub-metropolitan residents and public health centers had higher score, which means they had shorter WT for colonoscopy.

In the comparison between multivariate analysis and univariate analysis, the results of individual characteristics indicate female had higher score for non-compliance whereas male had higher risk score with short WT. The elderly had higher score of non-compliance and the youngest had the highest score with short WT. Among contextual factors, the results were similar to those in univariate analysis on calendar periods, round of screening, and type of screening units. In addition, it shows central Taiwanese had the highest score in both parts. People dwelling at less urbanized area had higher score of non-compliance. Sub-metropolitan residents had the highest score with short WT compared with other level of urbanization.

Because the change in the structure of screening program between the inaugural period

doi:10.6342/NTU201901194

and rolling out period, we further evaluated the interaction between factors and periods of screening program. The results of model selection reveal that the effect of the period of screening program was modified by geographic areas and type of screening units in hurdle part. In addition to these two factors, level of urbanization was also included in non-hurdle part.

3.3.2 Risk stratification

Based on the regression coefficients of the multivariate hurdle model, we computed the score based on the trained regression coefficients multiplied by the covariates of interests depending on the risk groups classified by the underlying covariates. The hurdle $\beta\text{-score}_1$ represents the risk score in association with the non-compliance with colonoscopy. The higher $\beta\text{-score}_1$ value was, the higher odds of refusing to take colonoscopy. The WT $\beta\text{-score}_2$, stands for the rate of taking colonoscopy. The higher $\beta\text{-score}_2$ value, the shorter WT for undergoing colonoscopy was.

To predict whether and when screenees comply with the procedure of receiving colonoscopy given contextual factors and individual characteristics, we therefore classified the $\beta\text{-score}_1$ into three groups according to the cutoff of its first quartile value (Q1) and the third quartile value (Q3), and $\beta\text{-score}_2$ into five groups according to the cutoff of its 20^{th}, 40^{th}, 60^{th}, 80^{th} percentile value. In hurdle part, if $\beta\text{-score}_1$ was higher than -0.5320 (Q3), termed as the

doi:10.6342/NTU201901194

intractable group, more than 37% had the probability of not complying with colonoscopy; if β-score$_1$ was lower than -1.0515 (Q1), termed as the compliant group, 26% had the chance of not complying with the uptake of colonoscopy; the middle category was the neutral group. In non-hurdle part, those with β-score$_2$ higher than -3.7246 corresponding to the WT for colonoscopy shorter than 41 days could be considered as high health awareness group; those with β-score$_2$ between -3.7709 and -3.7246 corresponding to the WT for colonoscopy between 41 to 43 days were moderate health awareness group; those with β-score$_2$ between -3.8063 and -3.7709 corresponding to the WT for colonoscopy between 43 and 45 days were modest health awareness group; those with β-score$_2$ between -3.8830 and -3.8063 corresponding to the WT for colonoscopy between 45 and 49 days were low health awareness group; those with β-score$_2$ lower than -3.8830 corresponding to the WT for colonoscopy longer than 49 days were shilly-shally group.

Assuming there are 4,000,000 eligible subjects attending nationwide population-based CRC screening, approximately 2,400,000 of them will be positive FIT tests. In Table 3.3 the results of prediction show around 880,000 attendees are the intractable group, 888,000 are the neutral group, and 624,000 are the compliant group. An example of characteristics representing the intractable group is shown by those female aged 65 to 69, who lived in non-metropolitan and central Taiwan, underwent screening at local clinic during the rolling-out period and were

doi:10.6342/NTU201901194

detected as positive case of prevalent screen. Her β-score$_1$ would be 0.4578 with 61% of probability of non-compliance with colonoscopy. Similar applications to the neutral and the compliant groups are described in Table 3.3. In Table 3.4, we also give an example of different WT groups. Referrals at subsequent rounds of those male aged 50 to 54, living in metropolitan and central Taiwan, undergoing screening at public health center during the inaugural period would be classified as high health awareness (β-score$_2$=-3.3546), and the predicted WT for colonoscopy would be 29 days. Similar applications to other WT groups are described in Table 3.4.

3.4 Discussion

The hurdle Poisson regression model is applied to tackling the issue of both compliance and WT while undergoing confirmatory diagnosis of positive screenees that is the routine procedure of second stage in population-based cancer screening program. To discuss why certain factors predispose to non-compliance and long WT, is particularly meaningful for the contextual factors elucidated in the current model. Firstly, for those prevalent screen cases, their behavior might think they were not positive cases, so they were not willing to undergo the following confirmatory colonoscopy; conversely for those subsequent screen cases, they might have higher health consciousness so that they would undergo screening again and receive

doi:10.6342/NTU201901194

colonoscopy exam after being detected as positive case; Secondly, Taiwan nationwide screening policy has changed since 2010. Before 2010, we offered screening service mainly at the public health center, and during the rolling-out period, hospitals joined to actively invite peoples for FIT. Therefore, we could find that screenees participating in the place of hospital were more willing to undergo colonoscopic exam after entering into the rolling-out period, besides, their WT for colonoscopy also shortened with an increase in manpower in the hospital. Although local clinics joined to actively invite peoples for FIT during the rolling-out period as well, their compliance for colonoscopy was still low and they needed to wait the longest time for colonoscopy compared with other institutions; Thirdly, most hospitals and clinics are concentrated on northern Taiwan, and in eastern or offshore islands, their medical resources are deficient. As a result, eastern or offshore islands residents had higher probability of non-compliance for colonoscopy compared to northern Taiwanese, and their WT for colonoscopy was longer as well. It is an urgent task for the screening organizer to have a remedy for the shortage of endoscopic capacity to meet the demand; Finally, metropolitan people might have more health information, so they had lower probability of non-compliance for colonoscopy compared with non-metropolitan residents. Although metropolitan area had more medical resources and capacity of colonoscopy, its number of positive cases was too many to provide colonoscopic exam immediately, so they still needed to wait longer time.

doi:10.6342/NTU201901194

The proposed model may be worthy of being compared with other previously proposed model (Brown ER et al. (2003), and Cheung YB et al. (2002)). In terms of statistical property, there are several similar statistical models proposed to deal with the issue. In contrast to the cured rate model applied to susceptible and non-susceptible subjects, the proposed hurdle model estimates the parameters governing the hurdle part on the basis of the observed information whereas the cured model estimated the non-susceptible analogous to the hurdle part in the way of latent variable. In addition, there are other models which have the similar concept with the hurdle model called the zero-inflated model that have been dealt with count responses with excess zeros by COM-Poisson and generalized Poisson model (Conway RW et al. (1962), and Consul PC et al. (1992)). The zero-inflated model can deal with zero part (non-complier) as well, but it is not appropriate to apply to screen data because that most of invited subjects are willing to undergo colonoscopy. The essence of this hurdle model is to consider the hurdle part as the non-compliance with the referral using a dichotomous variable (yes/no) and the non-hurdle part as how long it would take to receive confirmatory diagnosis in terms of WT distribution assuming follow the Poisson regression model. Factors associated with the hurdle part can be modelled through logistic regression model whereas factors affecting the WT are modelled by the truncated Poisson regression model as the counts of event are positive. The main advantage of using the proposed hurdle model is to reduce the heterogeneity of the

doi:10.6342/NTU201901194

outcome of interest by introducing a mixture of two outcomes corresponding to the logistic regression model and the truncated Poisson regression model because the former captures the mass function at zero point and the latter part captures the WT of complying with confirmatory diagnosis using the WT distribution. The other advantage is that compared with the use of two independent models, the simple Poisson regression model for WT and the logistic regression model for complier and non-complier, the hurdle model can accommodate the correlation between compliance (hurdle) and WT as indicated earlier. In our example, the heterogeneity (overdispersion) of using the Poisson regression model can be reduced by using the proposed model. The Akaike information criterion (AIC) has been greatly reduced from 839432 to 362771. In addition, it contains two types of people in the zero part: one is non-complier (Subject B), and the other is those who are still waiting for confirmatory examination (Subject A) that is regarded as censoring cases. If we use the traditional Poisson regression model, we cannot identify which zero given time t is derived from and which factors will affect their health behavior of non-compliance.

The current study still leaves something desired. If the data can provide other information such as social-economic status, income or education level, the model may be more explainable and can predict well. Although the current model is better than the Poisson model, overdispersion may still exists, which means variability still cannot be explained due to

doi:10.6342/NTU201901194

correlation and heterogeneity. Nevertheless, the problem of overdispersion dealt with in this study is not entirely due to correlated outcome or the heterogeneity beyond the explained covariates but the mixture of two outcomes, compliance with and WT for colonoscopy that may lead to the identifiability problem without using the hurdle model. This accounts for why we did not use negative binomial regression model in our study as it may suffer from such a problem although the heterogeneity in terms of AIC may be improved with the incorporation of two parameters in negative binomial model rather than one in Poisson model. However, if the truncated Poisson regression cannot be explained by covariates or correlated WT one may attempt hurdle negative binomial regression model under the framework of generalized mixed regression model, which may be the subject of an ongoing research.

In conclusion, we proposed the hurdle Poisson regression model to evaluate the compliance with and the WT for complete confirmatory diagnosis among referrals at first stage of screen. This model was successfully applied to evaluating individual attributes and contextual factors associated with factors related to the compliance and the WT distribution of undergoing colonoscopy, which was further developed into a useful clinical assessment model for quantifying the probability of complying with and the rate of taking colonoscopy given different individual attributes and contextual factors.

doi:10.6342/NTU201901194

Table 3.1 Comparison of referral rate and median waiting time for colonoscopy in inaugural and rolling out period

Characteristics	No. of FIT tests did not receive colonoscopy (%)	No. of FIT tests received colonoscopy (%)	Median waiting time (days)
Individual characteristics			
Gender			
Male	50,447 (32.33)	105,599 (67.67)	46
Female	438,10 (31.65)	94,613 (68.35)	45
Age (years)			
50-54	23,230 (31.75)	49,930 (68.25)	45
55-59	25,130 (31.70)	54,144 (68.30)	45
60-64	24,100 (31.83)	51,625 (68.17)	45
65-69	21,797 (32.87)	44,513 (67.13)	47
Contextual factors			
Calendar period			
2004-2009	7,608 (18.29)	33,983 (81.71)	29
2010-2013	86,649 (34.27)	166,229 (65.73)	51
Screening round			
Prevalent	65,151 (35.02)	120,908 (64.98)	51
Subsequent	29,106 (26.85)	79,304 (73.15)	39
Geographic area			
Northern	39,870 (31.60)	86,304 (68.40)	45
Central	23,172 (33.37)	46,278 (66.63)	43
Southern	27,604 (31.36)	60,419 (68.64)	47
Eastern/offshore islands	3,611 (33.37)	7,211 (66.63)	53

doi:10.6342/NTU201901194

Level of urbanization			
Metropolitan	72,258 (31.72)	155,552 (68.28)	46
Sub-metropolitan	6,743 (31.30)	14,798 (68.70)	41
Non-metropolitan	15,256 (33.81)	29,862 (66.19)	46
Type of screening units			
Hospital	54,435 (32.09)	115,180 (67.91)	47
Public health centers	22,609 (25.53)	65,962 (74.47)	35
Local clinics	17,213 (47.44)	19,070 (52.56)	149
Overall	94,257 (32.01)	200,212 (67.99)	46

doi:10.6342/NTU201901194

Table 3.2 Univariate and multivariate analysis of hurdle part on factors affecting non-compliance for colonoscopy and non-hurdle part on factors affecting waiting time for undergoing colonoscopy

Characteristics	Univariate analysis				Multivariate analysis			
	Hurdle part		Non-hurdle part		Hurdle part		Non-hurdle part	
	Coefficient	Risk Score	Coefficient	Risk Score	Coefficient	Risk Score	Coefficient	Risk Score
Individual characteristics								
Gender								
Male	-2.0920*	0	0.0095	10	0.0312	31	-3.7790*	0
Female	0.0319	32	-4.0402*	0	-0.7699*	0	0.0046	5
Age (years)								
50-54	-2.0920*	0	0.0253	25	0.0024	2	0.0041	4
55-59	0.0989	99	0.0214	21	-0.7676*	0	0.0084	8
60-64	0.1248	125	0.0227	23	0.0058	6	0.0075	8
65-69	0.2232	223	-4.0402*	0	0.0536	54	-3.7820*	0
Contextual factors								
Calendar period								
2004-2009	-2.0920*	0	0.2409	241	-1.4967*	0	0.2655	266
2010-2013	0.7926	793	-4.0402*	0	0.8452	845	-3.8172*	0
Screening round								
Prevalent	0.4722	472	-4.0402*	0	0.3840	384	-3.7824*	0
Subsequent	-2.0920*	0	0.0376	38	-1.0023*	0	0.0142	14
Geographic area								
Northern	0.0181	18	0.0891	89	0.0111	11	0.0630	63
Central	0.1029	103	0.1424	142	0.0916	92	0.1150	115
Southern	-2.0920*	0	0.0284	28	-0.7833*	0	-3.8298*	0
Eastern/offshore islands	0.0379	38	-4.0402*	0	0.0917	92	0.0052	5
Level of urbanization								
Metropolitan	-2.0920*	0	-4.0402*	0	0.0193	19	-3.7864*	0

doi:10.6342/NTU201901194

Sub-metropolitan	0.0680	68	0.0669	67	-0.7860*	0	0.0496	50
Non-metropolitan	0.1894	189	0.0174	17	0.1144	114	0.0411	41
Type of screening units								
Hospital	0.0811	81	0.1051	105	0.3212	321	0.1215	122
Public health centers	-2.0920*	0	0.1311	131	-1.0707*	0	0.2276	228
Local clinic	0.7142	714	-4.0402*	0	0.9683	968	-3.9194*	0

* represents intercept.

doi:10.6342/NTU201901194

Table 3.3 Example of characteristics related to compliance groups

Risk group	Risk score	Number of screenee	Example of Characteristics	Probability of Noncompliance
Intractable	> -0.5320	888,000	A female aged 65 to 69, lived in non-metropolitan and central Taiwan, underwent screening at local clinic during the rolling-out period and was detected as positive case at prevalent screen	61%
Neutral	-1.0515 ~ -0.5320	888,000	A male aged 50 to 54, lived in sub-metropolitan and southern Taiwan, underwent screening at hospital during the rolling-out period and was detected as positive case at prevalent screen	33%
Compliant	< -1.0515	624,000	A male aged 50 to 54, lived in metropolitan and northern Taiwan, underwent screening at public health center during the inaugural period and was detected as positive case at subsequent rounds	7%

doi:10.6342/NTU201901194

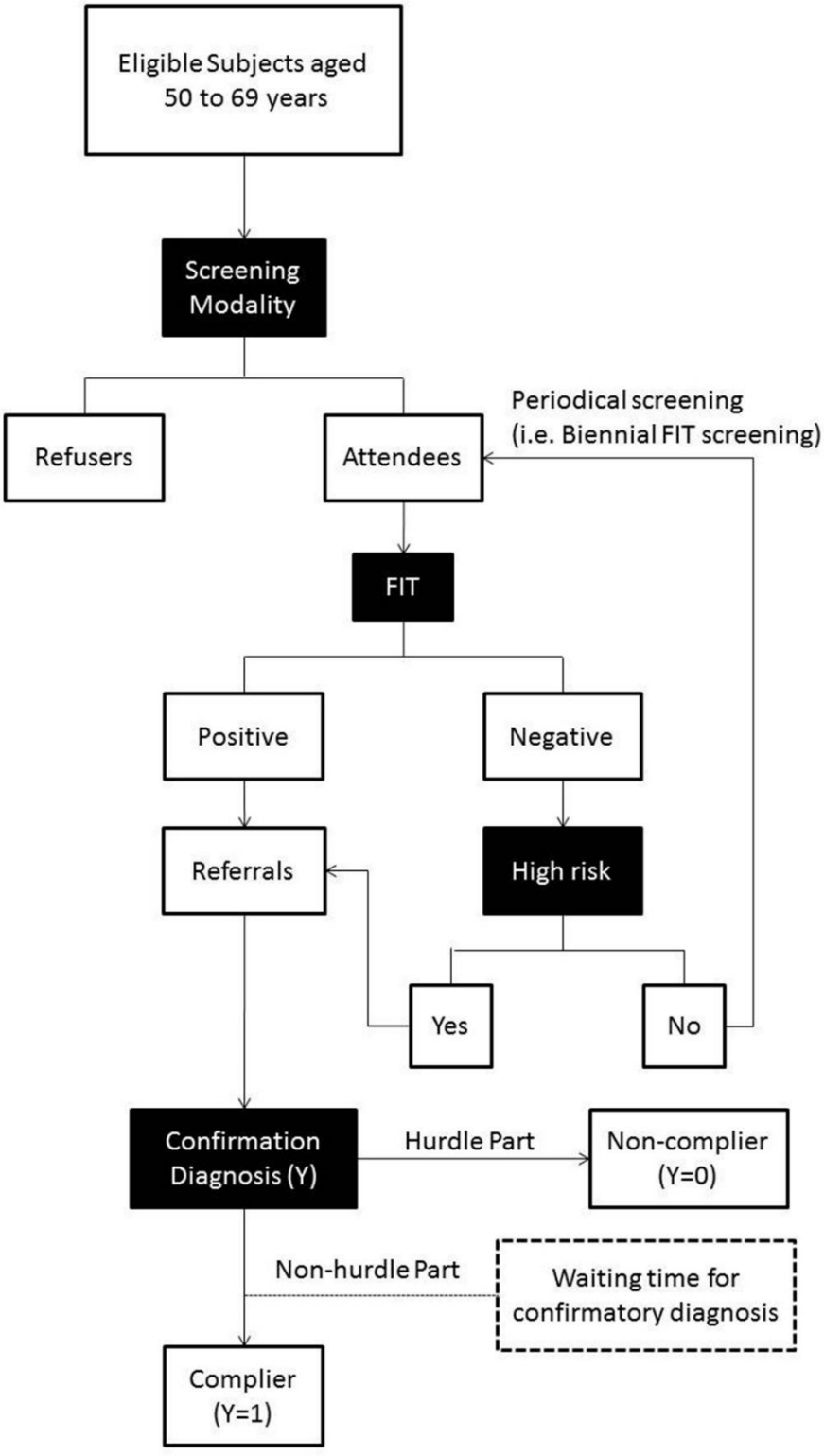
Table 3.4. Example of Characteristics related to waiting-time groups

Awareness group for waiting-time	Risk score (WT days)	Example of Characteristics
High health awareness	> -3.7246 (< 41)	A male aged 50 to 54, lived in metropolitan and central Taiwan, underwent screening at public health center during the inaugural period and was detected as positive case at subsequent rounds (WT=29)
Moderate health awareness	-3.7709 ~ -3.7246 (41 ~ 43)	A male aged 65 to 69, lived in non-metropolitan and central Taiwan, underwent screening at public health center during the rolling-out period and was detected as positive case at subsequent rounds (WT=43)
Modest health awareness	-3.8063 ~ -3.7709 (43 ~ 45)	A female aged 60 to 64, lived in metropolitan and northern Taiwan, underwent screening at hospital during the rolling-out period and was detected as positive case at prevalent screen (WT=44)
Low health awareness	-3.8830 ~ -3.8063 (45 ~ 49)	A female aged 60 to 64, lived in non-metropolitan and central Taiwan, underwent screening at local clinic during the rolling-out period and was detected as positive case at prevalent screen (WT=47)
Shilly-shally	< -3.8830 (> 49)	A female aged 65 to 69, lived in metropolitan and northern Taiwan, underwent screening at local clinic during the inaugural period and was detected as positive case at prevalent screen (WT=66)

Abbreviations: WT, waiting time

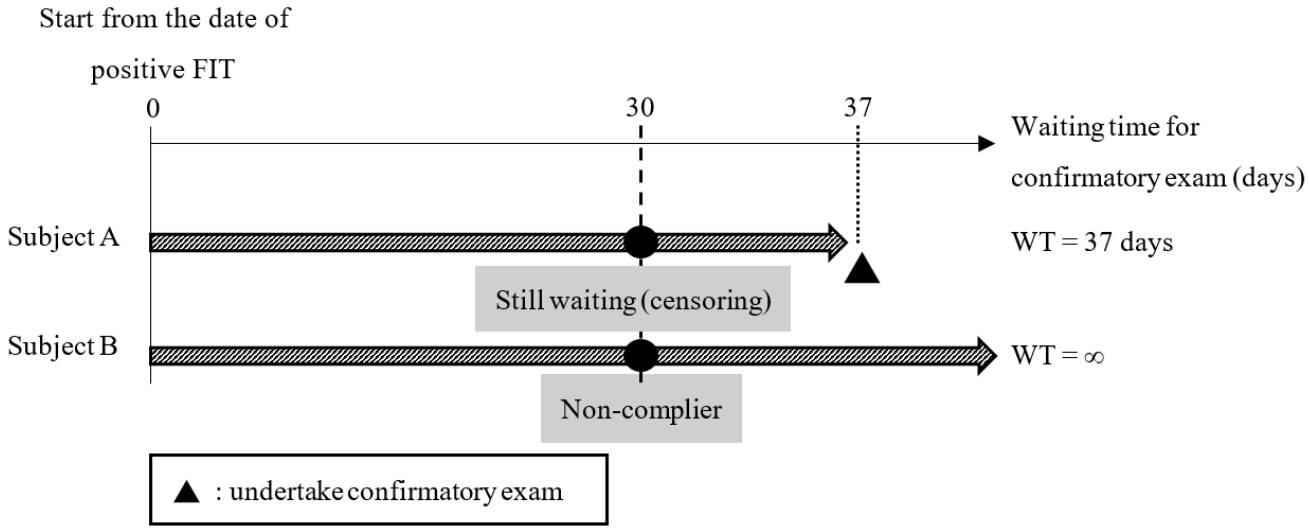
doi:10.6342/NTU201901194

Figure 3.1. Study framework of hurdle model for referral to confirmatory diagnosis in Taiwanese nationwide colorectal cancer screening program



doi:10.6342/NTU201901194

Figure 3.2. Illustration of two type of zero on the waiting time (WT) for confirmatory examination



doi:10.6342/NTU201901194

Chapter 4. Queue Hurdle Coxian Phase-type Model for Waiting Time in Two-stage Population-based Screening

Summary

Introduction Waiting time (WT) for confirmatory diagnosis in those screened as positives plays an important role in the quality assurance of two-stage cancer screening program. It requires a sophisticated statistical model considering three joint effects, the arrival rate (attending screening), positive rate, and the compliance with confirmatory diagnosis.

Methods My thesis proposed a Queue Hurdle Coxian phase-type model and applied it to Taiwanese colorectal cancer screening program. A series of simulations were applied to assess how bias and mean square error behaved.

Results Projected WT by various scenarios was performed to provide a quantitative assessment for the referral policy of WT given clinical capacity. Given 3% positive rate and 60% compliance rate, the mean WT increases from 4 days to 43 days as the annual coverage rate from 5% to 50%. The mean WT is longer by 10-fold as compliance rate changes from 10% to 100%. In addition, the mean WT for those with lower WT score is longer than those with higher WT score under the same coverage rate, positive rate, and compliance rate.

Conclusion Queue Hurdle Coxian phase-type model was proposed to model the elements of structure and process of population-based screening for colorectal cancer with FIT given

doi:10.6342/NTU201901194

limited capacity.

doi:10.6342/NTU201901194

4.1 Introduction

4.1.1 Waiting time in two-stage population-based screening

Population-based mass screening often involves in two-stage confirmatory procedure such as colorectal cancer (CRC) with fecal immunological test (FIT) as first line method and then colonoscopy confirmation, and breast cancer with mammography as first line and biopsy confirmation. During this two-stage process, the waiting time (WT) for confirmatory diagnosis becomes the first concern when clinical capacity of offering confirmatory procedure is limited as long WT may cause delayed diagnosis of disease but shortening WT may impose tremendous burden on such a confirmatory procedure. The change in cutoff of screening test may affect the WT for confirmatory diagnosis. To model the pattern of WT plays an important role in the optimal resource allocation to two-stage population-based screening.

4.1.2 Coxian-phase type process

One of statistical models considered here for modelling WT is the application of Phase-type time distributions introduced by Neuts (1974) and modified by Fackrell (2009) with flexibility to elucidate the underlying dynamic hidden phases to account for progressive multi-phase transitions. One of a subclass that has been extensively utilized in the healthcare industry is the Coxian phase-type (CPH) distribution. The most advantage of using this model is that it

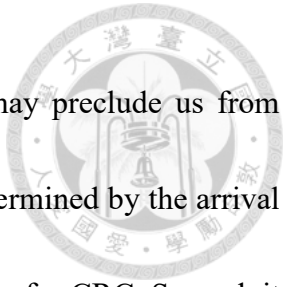
doi:10.6342/NTU201901194

can be interpreted as distinct clinical stages of patients in hospital by the clinical experts for the ease of understanding. A classical CPH distributions applied to heath care has been proposed by Marshall (2007) to classify the potential classifications of hospital length of stay of elderly patients so as to evaluate how to allocate the limited medical resources and costs. For a better understanding of the system, Marshall further incorporated the Bayesian belief networks with CPH distributions to take additional patient information into account (Marshall, & McClean, (2003)), and also modelled the multiple patient movements between hospital and community by the mixture of conditional CPH distributions for providing alternative care in the community and preventing readmission to hospital (Gordon, Marshall, & Cairns, (2016)).

As the CPH has been used in various scenarios as indicated above it is therefore of great interest to apply it to modelling the WT for those who are screened positives with FIT and wait for the referral to undergo confirmatory diagnosis with colonoscopy. In addition to estimating the WT, it is also very interesting to model how the WT are affected by relevant postulated factors such as demographic features, type of institution, geographic areas, and prevalent screen or subsequent screen. Taking such information into account enables one to know the heterogeneity of the WT for confirmatory diagnosis among compliers by dividing those into the short WT group and the long WT group.

While considering the application of the CPH model to such a two-stage population-based

doi:10.6342/NTU201901194



screening data we are faced with two statistical complications that may preclude us from directly using the CPH model. First, the WT distribution is strongly determined by the arrival process among those who have the uptake of population-based screening for CRC. Second, it may also be affected by the compliance with the referral to confirmatory diagnosis. To solve the first complication, the Poisson Queue process and its regression model are proposed to embed the arrival process of attending CRC screening into the CPH model by estimating the arrival rate and its relevant factors. The second is related to the type of the Queue process. The CPH model is a specialized case of hyper-exponential queue model. It is natural to consider whether it can be used for hypo-exponential as the referral of participants with positive test may suffer from the problem of non-compliance. From the methodological viewpoint, how to modify the CPH model to accommodate both hyper-exponential and hypo-exponential Queue models, which may be adequate for modelling data on WT, making allowance for the non-compliance of undergoing confirmatory diagnosis. To solve this issue, the Poisson Hurdle regression model (Mullahy, 1986) is therefore embedded into the CPH model.

Therefore, we develop a unifying model to incorporate three processes, the Poisson Queue process to model the arrival of having the uptake of screening, the hurdle process to model non-compliance, and the CPH process to model WT among compliers, named as Queue Hurdle Coxian phase-type model (abbreviated as QH-CPH) to estimate the arrival rate, the rate of

doi:10.6342/NTU201901194

complying with confirmatory diagnosis, and the WT and to assess how relevant factors affect three processes. The novel model not only can take the arrival rate of eligible screenee into account, but also determine factors affecting non-compliance of a follow-up colonoscopy and further understand the underlying dynamic hidden phases of the WT according to relevant postulated factors.

4.2 Application to confirmatory diagnosis of population-based colorectal cancer screening

Data

The data considered here is from Taiwanese nationwide CRC screening program between 2004 and 2009, which offer FIT for subjects aged 50 to 69 years biennially. Those positive FIT tests or negative FIT tests with family history were considered as referrals and all of them were referred for confirmatory colonoscopic diagnosis. Details about Taiwanese nationwide CRC screening program have been described in our previous study (Chiu et al., (2015)). In brief, this nationwide program was served in 25 municipalities with a total of 5,417,699 eligible residents, and 1,160,895 of them attended this program (21.4% coverage of population). They undertook screening services mainly at public health center during this period. The cutoff value for a positive test was equivalent to 20 μg hemoglobin/g feces (Chen et al., (2007)). If those who screened with unspecified FIT brands or at nonqualified

doi:10.6342/NTU201901194

laboratories were excluded for our analysis, therefore there were 1,013,183 subjects in our study, which yielded 1,232,145 tests of FIT. The eligible FIT among this period was 26,720,380 visits, so the mean arrival rate per year was 4.61%. Among these FIT tests, 45,305 of them were positive results (3.7% of positive rate) and 33,983 had undergone the following confirmatory diagnosis with colonoscopy (75% of compliance rate).

In our data, it demonstrates that if the cutoff FIT value changes to 10 μg hemoglobin/g feces, the positive rate will be 7% approximately, which might yield 86,250 positive tests; however, after the cutoff value turns to 30 μg hemoglobin/g feces, the number of positive results decline to 36,963 tests and the positive rate is around 3%. It is apparent that the number of positive tests will be influenced by the cutoff value and further affect the following procedure of confirmatory diagnosis such as compliance and the WT given the equivalent clinical capacity.

In order to analyze this kind of complicated structure data, we proposed an idea to combine queue process, Hurdle model, and CPH. We are faced with high demand for around over one million participants eligible for the uptake of CRC screening with FIT, yielding a high demand for the referral of positive FIT to undergo colonoscopy. In contrast to the conventional Queue process that evaluates the arrival rate as opposed to departure rate relating to service time distribution, the non-compliance (non-susceptibility) problem for the referral of positive FIT made the traditional Queue process infeasible and may resort to the use of Hurdle model. In

doi:10.6342/NTU201901194

addition, those who were willing to consent to undergo colonoscopy may be classified into different types according to WT for colonoscopy, for example, those who had higher health awareness or were prone to anxiety might have shorter WT. This raised the rationale for using the CPH model for detecting whether it can identify hidden phase during the WT so as to provide a new insight into information used for health promotion for enhancing the referral rate. However, the real scenario of WT distribution also include non-response data (time=0) and Queue process that render the conventional CPH model inadequate. To solve these issues, we therefore developed the Hurdle model in combination with the CPH. The details of this unifying QH-CPH model are described in the following section.

4.3 The Queue Hurdle Coxian Phase-type model for CRC screening data

4.3.1 Model Description

For the consideration of the arrival rate for eligible screenees, non-compliance with colonoscopy (non-complier), and the WT for undergoing colonoscopy (complier) simultaneously, we used the Poisson distribution to model arrival rate, the Hurdle model to identify factors affecting non-compliance, and the CPH distribution to understand the underlying dynamic hidden phases of the WT among compliers. We developed a new Queue Hurdle Coxian Phase-type model (QH-CPH), which is the combination of three processes

doi:10.6342/NTU201901194

(Figure 4.1):

(1) Poisson Queue process:

- $Z_i = 1$ is participant of CRC screening and $Z_i = 0$ is non-participant *i=1,...N* (=26,720,380) denoted as individuals eligible for FIT tests during six years
- ν is the mean arrival rate of screenees per person-years

The Poisson distribution can be expressed as

$$f(Z_i; \nu) = \frac{\nu^{Z_i} exp(-\nu)}{Z_i!},\ \nu > 0 \quad (4\text{-}1)$$

Most of people would participate in CRC screening during the period of March to September but some people may join with this screening program between October and February. Moreover, the arrival rate is affected by sex and age as well. Then, the Poisson regression model was proposed to take these covariates (age, sex, and season) into consideration:

$$log(\nu) = \boldsymbol{\alpha}'\boldsymbol{X} \quad (4\text{-}2)$$

(2) Hurdle model:

To accommodate the non-complier of undergoing colonoscopy and also WT for undergoing colonoscopy among the compliers, the hurdle model was applied. The Hurdle model (Mullahy, (1986)) has two components: a binary part (Hurdle part) to measure the

doi:10.6342/NTU201901194

outcome of compliance, which is the application of logistic regression model, and a truncated Poisson part (non-hurdle part) to explain the response among the compliers, which is modelled by the truncated Poisson regression model given the count greater than one. Covariates included in the regression model consist of sex, age, geographic area, type of screening units, level of urbanization, and screening round.

- N_j is the total number of tests in the *j*th (*j*=1,2…,m) category of covariates
- Y_{ij} represents whether the *i*th individual who had been screened positive underwent colonoscopy exam (1=yes, 0=no). $Y_j = \sum_{i=1}^{N_j} Y_{ij}$ is the total number of screenee positive undergoing colonoscopy in the *j*th category
- T_{ij}^w is the WT for colonoscopy. $T_j^w = \sum_{i=1}^{N_j} T_{ij}^w$ represents the total WT for colonoscopy in the *j*th category (WT was defined from the date of reporting as positive FIT to the date of receiving colonoscopy)
- p_j is the probability of refusing to undergo colonoscopy (non-complier)
- τ_j is the mean rate of undergoing colonoscopy

The probability mass function (p.m.f) of truncated Poisson process (given the count greater than one) is

$$\mathrm{P}\left(Y_j = y_j \middle| Y_j > 0\right) = \frac{\left(\tau_j t_j^w\right)^{y_j} exp(-\tau_j t_j^w)}{y_j!\left(1-exp(-\tau_j t_j^w)\right)} \quad (4\text{-}3)$$

doi:10.6342/NTU201901194

For $y_j = 0$, the p.m.f of Hurdle model is expressed by $\mathrm{P}(Y_j = 0) = {p_j}^{N_j}$; for $y_j > 0$, $\mathrm{P}(Y_j = y_j) = (1 - p_j)^{y_j} \frac{1}{(1-e^{-\eta_j})} \times \frac{{\eta_j}^{y_j} e^{-\eta_j}}{y_j!}$ where $\eta_j = \tau_j \times t_j^w$.

To identify the effect of relevant covariates ($\boldsymbol{x}_j$) such as sex, age at screening, geographic areas, type of screening units, level of urbanization, and round of screening, p_j can be modelled by the logistic regression model

$$p_j = \frac{exp(\boldsymbol{x}_j\boldsymbol{\gamma})}{1+exp(\boldsymbol{x}_j\boldsymbol{\gamma})} \tag{4-4}$$

and η_j , the estimate of mean count in the *j*th category, by the Poisson regression model

$$log(\eta_j) = offset(log\, t_{wj}) + \boldsymbol{x}_j\boldsymbol{\beta} \tag{4-5}$$

Thus, the log-likelihood can be expressed as

$$lnL = ln\left\{\prod (p_j)^{N_j}(1-p_j)^{y_j} \frac{1}{(1-e^{-\eta_j})} \times \frac{{\eta_j}^{y_j} e^{-\eta_j}}{y_j!}\right\}$$

$$= l\,n\{L_1(\boldsymbol{\gamma})\} + l\,n\{L_2(\boldsymbol{\beta})\} \tag{4-6}$$

(3) Coxian Phase-type distribution (CPH):

Following the notation of Marshall (2007), the CPH model describes the time to absorption of a finite Markov chain in continuous time over the phases {*1,2,...,k,k+1*}. This Markov chain has one absorbing phase (*k+1*th), and *k* transient phases (*1,...,k*), and the process only starts in the first transient phase (Figure 4.2). While analyzing the WT data, transient phases can represent the hidden transition and absorbing phase as

doi:10.6342/NTU201901194

compliance for colonoscopy. In Figure 4.2, the parameters are defined as follows:

- λ_s : transition rate from transient phase (s) to transient phase (s+1)

$$P(X(t+\Delta t)=s+1\,|X(t)=s)=\lambda_s\Delta t+o(\Delta t) \text{ for } s=1,\ldots,k\text{-}1 \quad (4\text{-}7)$$

- μ_s : absorbing rate from transient phase (s) to absorbing phase (k+1), we express 'referral rate' here.

$$P(X(t+\Delta t)=k+1\,|X(t)=s)=\mu_s\Delta t+o(\Delta t) \text{ for } s=1,\ldots,k. \quad (4\text{-}8)$$

The random variable T defined as the time to compliance (WT) is said to have a CPH distribution. The infinitesimal generator for the Markov chain can be written in block-matrix form as

$$\boldsymbol{R}=\begin{pmatrix}\boldsymbol{Q} & \boldsymbol{q}\\ \boldsymbol{0} & 0\end{pmatrix}$$

$$\boldsymbol{Q}=\begin{bmatrix}-(\lambda_1+\mu_1) & \lambda_1 & 0 & \cdots & 0 & 0\\ 0 & -(\lambda_2+\mu_2) & \lambda_2 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & 0 & \cdots & 0 & -\mu_k\end{bmatrix}$$

$$\boldsymbol{q}=(\mu_1\ \mu_2\ \cdots\ \mu_k)^T$$

where $\boldsymbol{Q}$ is a matrix of transition rates restricted to the transient phases and $\boldsymbol{q}$ is a vector of transition rates from transient phases to the absorbing phase.

To ensure absorption in a finite time with probability one, it requires that every non-absorbing state is transient, so the matrix $\boldsymbol{R}$ can be blocked and let the matrix $\boldsymbol{Q}$ do not

doi:10.6342/NTU201901194

consider the absorbing state. Due to the absorption in a finite time with probability one, the process with $\boldsymbol{Q}$ is an honest Markov process. As getting solution of the differential equations by forward and backward Kolmogorov equations, both sets of equations have the same unique solution to an honest Markov process (Cox, & Miller, (1965)). Therefore, the formal solution of transition probability matrix is

$$\boldsymbol{P}(\boldsymbol{t}) = exp(\boldsymbol{R}t) \tag{4-9}$$

$$= \sum_{n=0}^{\infty} \boldsymbol{R}^n \frac{t^n}{n!} \tag{4-10}$$

When $\boldsymbol{R}$ is a finite matrix, that is when the number of states of the process is finite, the series (4-10) is convergent and (4-9) is the unique solution of both forward and backward equations. As a result, the probability density function (p.d.f) of CPH is the derivative of the transition probabilities derived from (4-9) as follows:

$$f_c(t) = \frac{d}{dt} P_{1k}(t) \tag{4-11}$$

where $P_{1k}(t)$ is the (1,*k*)th element of matrix *P*(*t*), represents the probability from the I^{st} phase to absorbing phase at time *t*. The WT in each phase and the marginal WT are

- T_s represents the length of WT in phase *s*, where $T_s \sim exp(\lambda_s + \mu_s)$, the moment-generating function (MGF) is given by

$$M_{T_s}(\theta) = \frac{\lambda_s + \mu_s}{\lambda_s + \mu_s - \theta}$$

doi:10.6342/NTU201901194

$$E[T_s] = \frac{1}{\lambda_s + \mu_s}$$

- T is the marginal mean WT in the system, it can also be calculated by MGF

$$M_T(\theta) = \int_{-\infty}^{\infty} e^{t\theta}\, dP_{1k}(t)$$

$$E[T] = M_T'(0)$$

4.3.2 The Queue Hurdle Coxian Phase-type (QH-CPH) model

With the integration of these three models, the QH-CPH distribution can be expressed as

$$f(z, x, y, t; \nu, \pi, p, \lambda, \mu) = e^{-\nu}\{\nu(1-\pi)^{1-x}[\pi p^{1-y}[(1-p)f_C(t)]^y]^x\}^z \quad (4\text{-}12)$$

- ν is the mean arrival rate of screenees per person-years, which comes from the Poisson Queue model ($Z = 1$ is participant and $Z = 0$ is non-participant). The Poisson regression form was used to take covariates into account as in (2)
- π is the FIT positive rate ($X = 1$ is positive result and $X = 0$ is negative result) determined by the pre-determined cutoff
- p is the probability of refusing to undergo colonoscopy, which can be estimated by the Hurdle model ($Y = 1$ is complier for colonoscopy and $Y = 0$ is non-complier). The logistic regression form was adopted to distinguish those with specific characteristics by transforming coefficients estimated from the binary part of the

doi:10.6342/NTU201901194

Hurdle regression model into a new continuous covariate (Noncompliance score; $score_1$) then dividing it into a binary outcome according to the cutoff of median value:

$$logit(p) = \beta_0 + \beta_1 \times G_1$$

$$G_1 = \begin{cases} 1\,, if\ score_1 \geq -1.6711 \\ 0\,, if\ score_1 < -1.6711 \end{cases}$$

- T is the marginal WT for undergoing colonoscopic exam, the p.d.f is $f_C(t)$ of CPH distribution in (4-11). The proportional hazards regression form was applied to transition/referral rate by the similar method mentioned above except using coefficients estimated from the non-hurdle part as $score_2$ (WT score):

$$\lambda_s = \lambda_{0s}\exp(\gamma \times G_2)$$

$$\mu_s = \mu_{0s}\exp(\gamma \times G_2)$$

$$G_2 = \begin{cases} 1\,, if\ score_2 \geq -3.4869 \\ 0\,, if\ score_2 < -3.4869 \end{cases}$$

4.3.3 Criteria for Model Selection

After constructing the likelihood function of QH-CPH distribution in (12), we decided to use the Bayesian information criterion (BIC) (Schwarz, 1978), which is widely applied to likelihood-based model, for model selection. In our analyses, we have to determine how many phases will be fitted on WT in the beginning, then taking covariates into account on arrival rate, positive rate, as well as compliance rate. As parameters need to be estimated increase, the

doi:10.6342/NTU201901194

overfitting might occur. As a result, the penalty term, which is greater in BIC than in Akaike information criterion, was introduced to deal with this problem. After testing in each stage, the most appropriate model would be determined with minimum BIC score.

4.4 Simulation Study

We set up a scenario similar to Taiwanese population-based two-stage screening for CRC between 2004 and 2009, which is a six-year biennual screening program with 4.5% of arrival rate (per year), 3.7% of positive rate and 75% of compliance rate. The 2-phase QH-CPH models (the optimal model identified with BIC and presented in the results of the following section) to assess whether the number of eligible subject would affect the performance of the QH-CPH model or not. In Table 4.1, we assume that the manpower of colonoscopists would be increased along with the expanded compliers as well as decreased along with the shrinked compliers so that it could be affordable to receive colonoscopy exam in 37 waiting days. The result demonstrates the marginal expected WTs are very close to our true value (37 days) by using the 2-phase QH-CPH model and declined from 39.1 days to 38.3 days as the number of eligible subjects increased. With 100 replications, it has a great variation if only 4500 eligible subjects in the program; however, after reaching to 4500000 eligible subjects, the variance drops substantially. Both of biases and variances decline with an increase of the number of eligible

doi:10.6342/NTU201901194

subjects, despite the biases of WT are barely changed after the number of eligible subjects exceeded up to 10000. As the sample sizes increase, it results in a tremendous decrease in variance from 44.49 to 0.03 as well as the mean square error (MSE) by 47.18 (from 48.9 to 1.72). Until it reaches up to 80000 eligible subjects, the MSE becomes stable gradually. Nevertheless, it is incompatible with a real situation that the mean WT for colonoscopy remains on 37 days as the clinical capacity can always suffice such few eligible subjects. Therefore, there is a large of gap on MSE between 4500 and 10000 of eligible subjects but if the number of eligible subjects equals 4500000, which is pretty close to Taiwanese CRC screening program, then the performance of the 2-phase QH-CPH model is stable. Accordingly, if the number of eligible subjects is less than 80,000, the 2-phase in QH-CPH model may lead to large variation.

4.5 Results

4.5.1 Empirical results on Taiwanese two-stage population-based screening for CRC

According to data on Taiwanese CRC screening program, the empirical results show that the continuous WT data are positively skewed with a long tail, representing those few positive FIT cases who had not received colonoscopic exam with an extremely long WT (Figure 4.3). This suggests the WT had better be modelled by the CPH distribution. The result shows the 2-phase QH-CPH model is the most appropriate one to capture the WT by the minimum BIC

doi:10.6342/NTU201901194

value (Table 4.2). Compared with the 1-phase QH-CPH model, the BIC was reduced by 1324; however, it was increased by 15 in comparison to the 3-phase QH-CPH model. Furthermore, in the 3-phase QH-CPH model, the estimated referral rate from the moderate waiting phase (μ_2) equals 0, because the identifiability problem may exist between the referral rate resulting from the moderate waiting phase (μ_2) and the transition rate from the short waiting phase to moderate waiting phase (λ_1), which might suggest the 3-phase QH-CPH model was inappropriate.

The 2-phase QH-CPH model can be classified as short waiting phase and long waiting phase (Figure 4.4). The corresponding mean WT was 33 days and 152 days, respectively, and the marginal mean WT was 37 days. The mean arrival rate was 4.01% (per person-year). We found that most screenees participated in the FIT screening during the period between March and September and the rate was pretty low rate from October to February. In addition to season effect, the arrival rate might vary with sex and age as well, so that we incorporated these covariates into the 2-phase QH-CPH model using the proportional hazards regression form. The results suggest the model including season, sex, and age (labelled 'Model B' in Table 4.3) was better than the original one only with season (BIC was reduced by 93137) as well as the model not including age (labelled 'Model A' in Table 4.3; BIC was reduced by 70986) and show those who were male or young age group had lower arrival rate despite the similar findings on the positive rate, the probability of non-compliance with colonoscopy and the WT.

doi:10.6342/NTU201901194

The estimated positive rate was 3.7%. The estimated non-compliance rate was 26.3%. It is very interesting to note that the referral rate was five times greater in short waiting phase than long waiting phase. Around 15% subjects were in a dilemma to be referred to undergo colonoscopy so as to be trapped in long waiting phase.

Apart from the variation on arrival rates, compliance rates with colonoscopy and the median WT varied with these specific characteristics as well. The results of the QH-CPH model in Table 4.3 show if we considered the variation of the non-compliance rate as shown in Model C, the BIC score was smaller than Model B by 581. To further accommodate the variation of WT, we extended Model C to Model D that assessed whether $score_2$ has more impact on the transition rate of undergoing colonoscopy (μ_1) with short waiting phase, or on the transition rate of undergoing colonoscopy (μ_2) with long waiting phase or on the transition rate from short waiting phase to long waiting phase (λ_1). The results of model selection show the model with $score_2$ related to μ_1 was the most appropriate model (labelled 'Model D' in Table 4.3), the differences of BIC are 289 and 296 compared to both models with $score_2$ related to λ_1 and μ_2, separately. Table 4.3 shows that this final model (Model D) was much more interpretable than the previous one (Model C) due to the reduction of BIC by 287. The probability of not complying with colonoscopy was 21% in low $score_1$ group and 31% in high $score_1$ group. Among compliers, the mean WT in short waiting phase was 36 days and 29 days in low $score_2$

doi:10.6342/NTU201901194

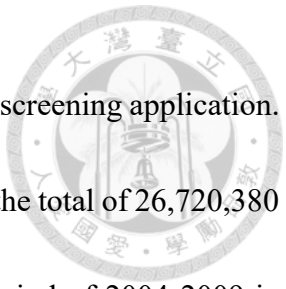
group and high score$_2$ group, separately, and both of the mean WT in longer waiting phase were 151 days. The corresponding marginal mean WT were 40 days and 33 days for low score and high score group, respectively.

According to the Model D as shown in Table 4.3, we could predict its transition probabilities by different WTs among compliers. In Figure 4.5, the probability of staying in short waiting phase (P_{11}) declined over time and those with lower score$_2$ had longer WT in short waiting phase than higher score$_2$ given the same probability of staying in short waiting phase. The transition probability from short waiting phase to long waiting phase (P_{12}) was pretty small and no difference was noted between these two groups. The transition probability of undergoing colonoscopy (P_{13}) increased over time, because attendees would receive colonoscopic exam eventually. Under the same transition probability to undergo colonoscopy, low score$_2$ group had longer WT than high score$_2$ group.

4.5.2 Projection of WT by different scenarios

Except the number of eligible subjects, there are still other factors which might affect the results. Based on our model, the number of complier with colonoscopy would be influenced by three parameters (arrival rate, positive rate, and compliance rate). It is worthwhile to examine how the WT will be changed by these parameters by doing sensitivity analysis. Here,

doi:10.6342/NTU201901194



we turn arrival rate into annual coverage rate for better understanding of screening application. In Figure 4.6, we set 3% of positive rate and 60% of compliance rate for the total of 26,720,380 eligible FIT tests and assume the equivalent clinical capacity as the period of 2004-2009 in Taiwan, the mean WT for colonoscopy prolongs from 4 days to 43 days with increased annual coverage rate (dotted line). While positive rate raises to 7% and compliance rate to 90%, the mean WT ranges from 15 days to 151 days (solid line). In Figure 4.7, given 20% of annual coverage rate and 3% of positive rate, the mean WT increases with enlarged compliance rate (dotted line), while annual coverage rate rises to 30% and positive rate to 7%, the mean WT lengthens as well (solid line). We can also find that if compliance rate is expanded from 30% to 60%, the expected WT doubles in these four situation. Furthermore, we predict the WTs taking WT $score_2$ into account given 30% of annual coverage rate and 3% of positive rate (see Figure 4.8). If positive cases are in different WT $score_2$ groups, the increasing rates of WT between low WT score group and high WT score group are differential. The larger the compliance rate, the more incremental WT in low WT score group. For example, it would increase by just 1 waiting day for low WT score group compared with high score group under 10% of compliance rate, however, after reaching to 100% of compliance rate, the corresponding figure is increased to 8 days.

doi:10.6342/NTU201901194

4.5.3 The Referral Policy of WT for Colonoscopy

Based on the results of projections as above, the QH-CPH model provides a quantitative assessment of various scenarios on the referral policy of WT given clinical capacity in our stimulation study. For example, the Taiwanese national cancer screening policy has changed since 2010, the volume of screenees expand rapidly and positive rate ascended (7.3%) so that the compliance rate decline (66%) and the WT prolongs (Jen, Hsu, & Chen, (2017)). If the mean WT of two months is allowed, we have to recruit screenees in terms of the scale of Taiwanese population with a three-year program to reach 90% coverage rate given the positive rate is 7% and the compliance rate is 60%. The similar expected WT can be estimated by changing the cutoff of screening test and also the compliance rate with confirmatory diagnosis. For instance, if we change in cutoff value from 10 μg hemoglobin/g feces to 30 μg hemoglobin/g feces, which might result in increased positive rate from 3% to 7%, the estimated WT will be prolonged by 2.35-fold under 20% of annual coverage rate and 60% of compliance rate.

4.6 Discussion

The proposed QH-CPH model here has succeeded in linking all the aspects of determinants from arrival rate of attending screen, referral rate due to the positivity of screening

doi:10.6342/NTU201901194

test, and the compliance with colonoscopy in relation to WT for colonoscopy. The merits of this study not only make contribution to developing a new statistical method for modeling joint effects of three crucial parameters but also provide a series of projections of WT by different scenarios. The further incorporation of relevant covariates is to elucidate the heterogeneity of the proposed QH-CPH model, which added to another novelty in the study. When applying the QH-CPH model to population-based screening program with the problems of queue and non-response to colonoscopy the findings gave a clue to explore the reasons dominating such differences including provider factors such as the implementation of screening program and medical resources and population factors such as the knowledge and attitude toward CRC screening and medical interventions. They also provide more insight on the promotion of the referral of positive FIT identified from the participants with the uptake of screening program. This model can also be extended to estimate the expected costs or utility of screening, because each phase might have various costs so that it can evaluate how to allocate the limited medical resources and costs.

One of limitations of the proposed QH-CPH model is related to whether the homogeneous processes used in arrival rate, the hurdle part, and the Coxian phase-type process can account for the heterogeneity of empirical scenario. To reduce this concern, relevant covariates have been incorporated into each process by using proportional hazards regression

doi:10.6342/NTU201901194

forms. However, the variation of such a heterogeneity may be explained beyond these covariates. The introduction of random effect parameters to each process may provide a solution but the likelihood function would become computationally intractable. This becomes the subject of ongoing research.

In conclusion, we developed a new QH-CPH model to solve the compliance with the uptake of screening using the Queue process, the problem of non-compliance with the referral of positive results of screenees to have confirmatory diagnosis using the Hurdle model in combination with the CPH model to identify hidden phases during the WT for undergoing colonoscopy for the referrals. With the limited clinical resources, the development of this new model not only provides a new insight into the underlying mechanism of WT for early detection of CRC, but also can help clinicians or hospital managers improve the quality of service and provide some useful information for making decisions.

doi:10.6342/NTU201901194

Table 4.1 Simulated results of the 2-phase QH-CPH models based on 100 replications.

Number of eligible subjects per year	Marginal True WT (days)	Marginal Expected WT (days)	Bias	Variance	Mean Square Error (MSE)
4500	37	39.1	2.1	44.492	48.902
10000	37	38.6	1.6	16.449	19.009
20000	37	38.7	1.7	10.233	13.123
30000	37	38.8	1.8	6.271	9.511
45000	37	38.5	1.5	4.911	7.161
80000	37	38.4	1.4	2.179	4.139
100000	37	38.5	1.5	1.551	3.801
200000	37	38.5	1.5	0.613	2.863
450000	37	38.3	1.3	0.296	1.986
4500000	37	38.3	1.3	0.026	1.716

***arrival rate=4.5%, positive rate=3.7%, compliance rate=75%**

doi:10.6342/NTU201901194

Table 4.2 The estimated results of 2-phase Queue Hurdle Coxian phase-type (QH-CPH) models

Parameters	**Estimate (SD$\times 10^{-3}$)**
Arrival rate $(\hat{v})$	
Intercept $(\hat{v}_0)$	-6.8521 (6.64)
Month (vs. Period 1 (Jan-Feb))	
Period 2 (Mar.-Sep.) $(\hat{\alpha}_1)$	3.6349 (6.71)
Period 3 (Oct.-Dec.) $(\hat{\alpha}_2)$	1.6403 (7.16)
Positive rate $(\hat{\pi})$	0.0368 (0.17)
Probability of non-compliance $(\hat{p})$	0.2632 (2.07)
Referral rate $(\hat{\mu}_1)$	0.0299 (0.19)
Referral rate $(\hat{\mu}_2)$	0.0066 (0.38)
Transition rate $(\hat{\lambda}_1)$	0.0008 (0.08)

$\hat{\mu}_s$: referral rate from sth phase to absorbing phase

$\hat{\lambda}_s$: transition rate from sth phase to s+1th phase

doi:10.6342/NTU201901194

Table 4.3 The comparison of five 2-phase QH-CPH models

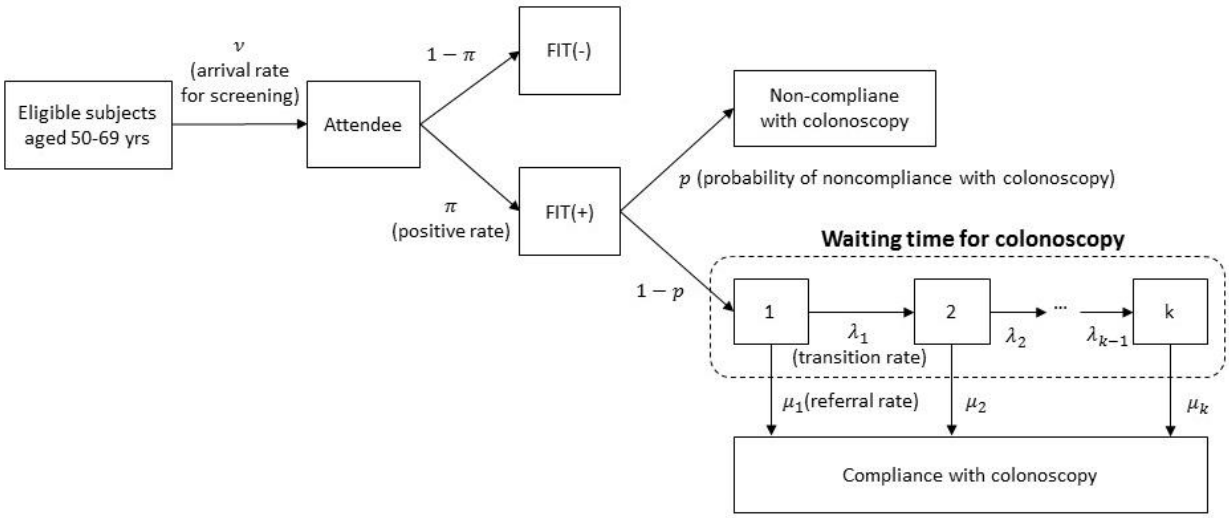
Parameters	Estimate (SD× 10^{-3})			
	Model A	**Model B**	**Model C**	**Model D**
Arrival rate ($\hat{v}$)				
Intercept ($\hat{v}_0$)	-6.6478 (6.67)	-6.4483 (6.92)	-6.4483 (6.90)	-6.4484 (6.92)
Month (vs. Period 1 (Jan-Feb))				
Period 2 (Mar.-Sep.) ($\hat{\alpha}_1$)	3.6412 (6.70)	3.6357 (6.72)	3.6357 (6.69)	3.6357 (6.72)
Period 3 (Oct.-Dec.) ($\hat{\alpha}_2$)	1.6491 (7.15)	1.6442 (7.17)	1.6442 (7.15)	1.6442 (7.17)
Sex				
Male ($\hat{\alpha}_3$)	-0.4886 (1.86)	-0.4844 (1.86)	-0.4844 (1.86)	-0.4844 (1.86)
Age (vs. 65-69 yr)				
50-54 yr ($\hat{\alpha}_4$)	-	-0.3640 (2.57)	-0.3640 (2.57)	-0.3640 (2.57)
55-59 yr ($\hat{\alpha}_5$)	-	-0.1719 (2.62)	-0.1719 (2.62)	-0.1719 (2.61)
60-64 yr ($\hat{\alpha}_6$)	-	-0.1178 (2.85)	-0.1178 (2.85)	-0.1178 (2.85)
Positive rate ($\hat{\pi}$)	0.0368 (0.17)	0.0368 (0.17)	0.0368 (0.17)	0.0368 (0.17)
Probability of non-compliance ($\hat{p}$)	0.2632 (2.07)	0.2632 (2.07)		
Intercept ($\hat{\beta}_0$)	-	-	-1.3300 (17.04)	-1.3299 (18.54)
Noncompliance Score ($\hat{\beta}_1$)	-	-	0.5703 (22.07)	0.5267 (24.84)
Referral rate ($\hat{\mu}_1$)	0.0299 (0.19)	0.0299 (0.19)	0.0299 (0.19)	
Baseline of referral rate ($\hat{\mu}_{01}$)	-	-	-	0.0272 (0.23)
WT Score ($\hat{\gamma}$)	-	-	-	0.2013 (11.78)
Referral rate ($\hat{\mu}_2$)	0.0066 (0.38)	0.0066 (0.38)	0.0066 (0.38)	0.0066 (0.39)
Transition rate ($\hat{\lambda}_1$)	0.0008 (0.08)	0.0008 (0.08)	0.0008 (0.08)	0.0008 (0.08)

$\hat{\mu}_i$: referral rate from ith phase to absorbing phase

$\hat{\lambda}_i$: transition rate from ith phase to i+1th phase

doi:10.6342/NTU201901194

Figure 4.1 Queue Hurdle Coxian Phase-type model



FIT(+) consists of positive fecal immunochemical tests (FIT) and negative FIT tests with family history

doi:10.6342/NTU201901194

Figure 4.2 Coxian Phase-type distribution

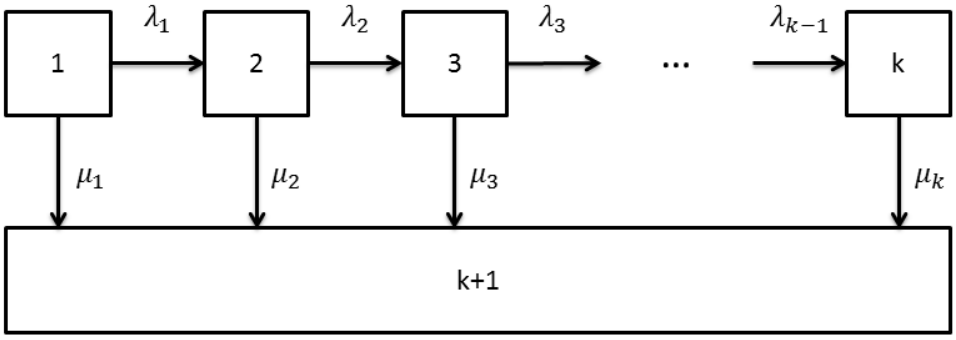
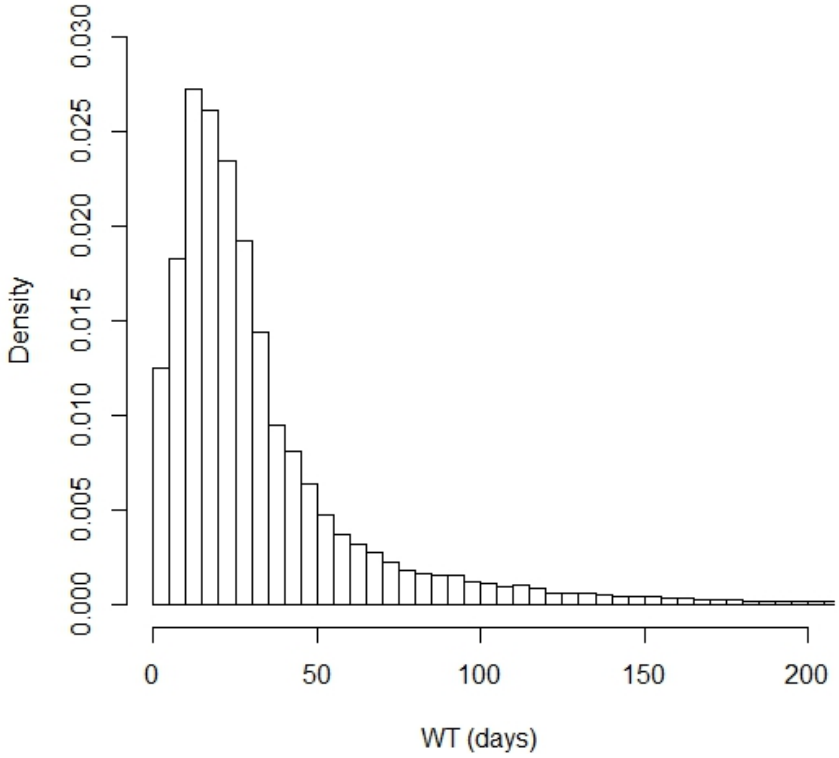


Figure 4.3 Empirical data on WT for colonoscopy



doi:10.6342/NTU201901194

Figure 4.4 Queue Hurdle 2-phase Coxian phase-type model

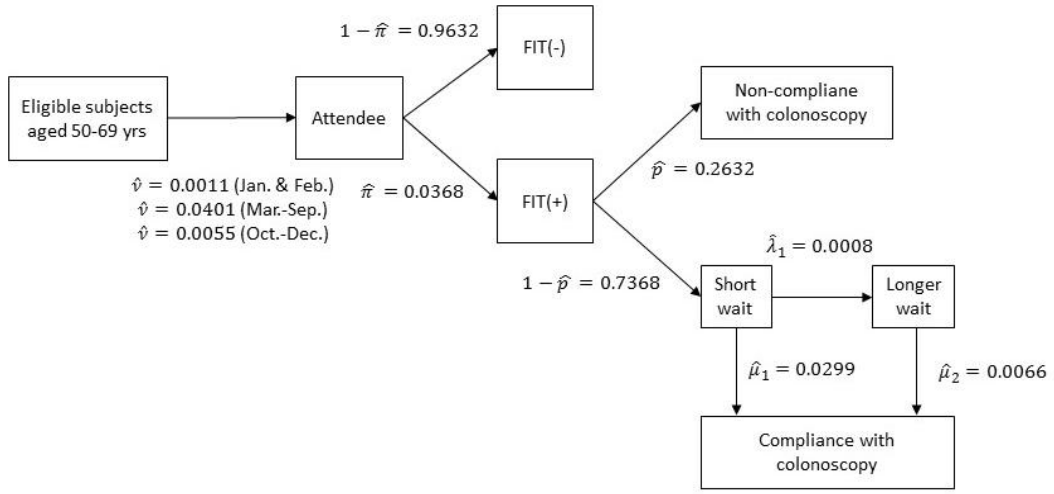
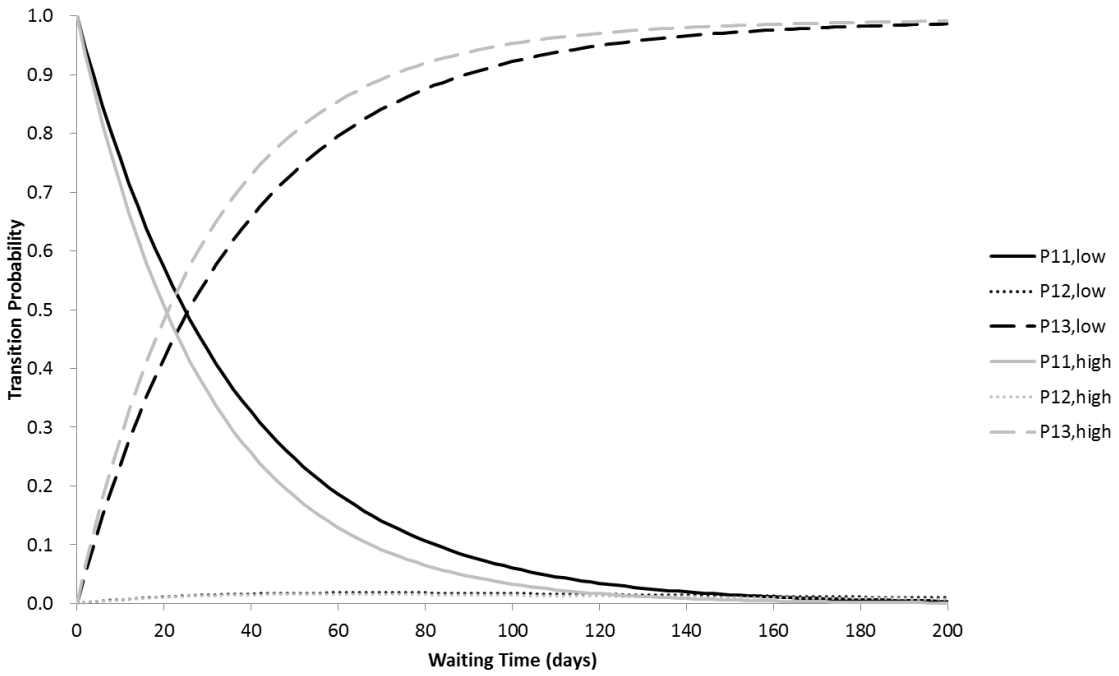
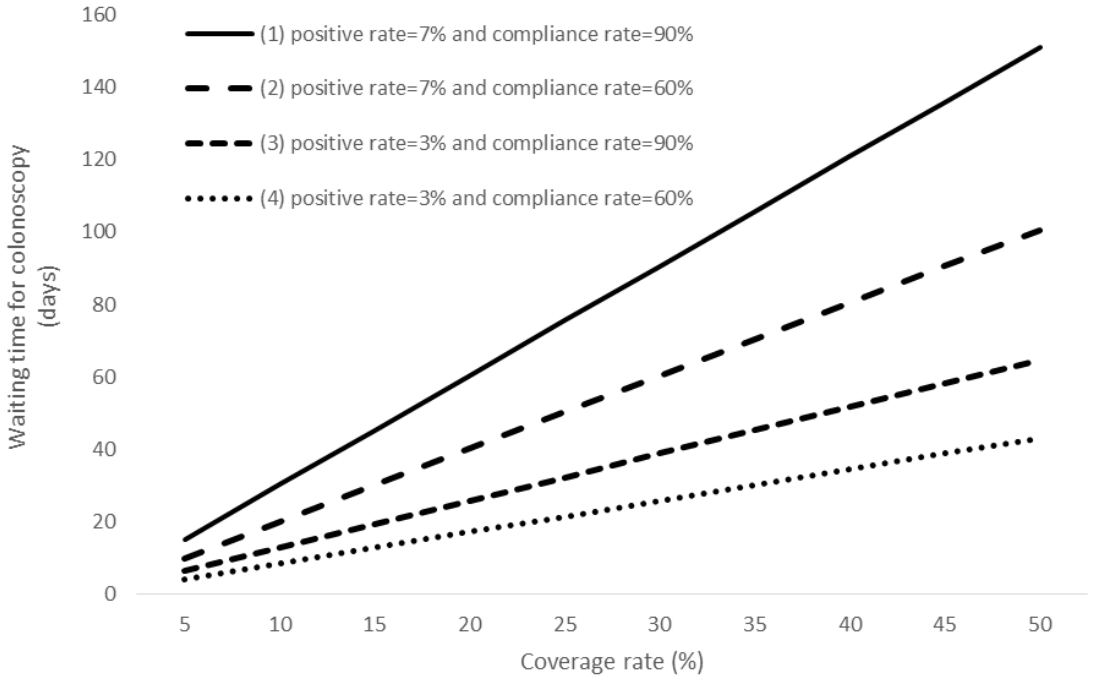


Figure 4.5 Transition probabilities of Coxian two-phase model by risk score



doi:10.6342/NTU201901194

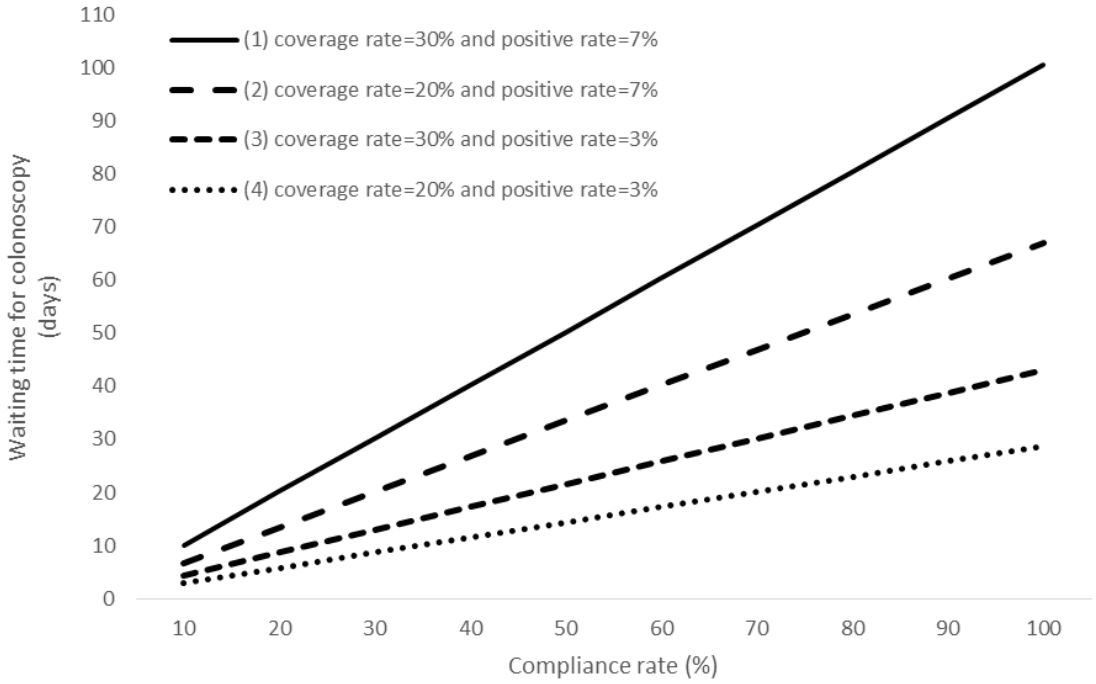
Figure 4.6 Predicted waiting time for colonoscopy by different coverage rates as given positive rate and compliance rate.



Scenarios	Waiting days									
(1)	15	30	45	60	75	91	106	121	136	151
(2)	10	20	30	40	50	60	70	80	91	101
(3)	6	13	19	26	32	39	45	52	58	65
(4)	4	9	13	17	22	26	30	34	39	43

doi:10.6342/NTU201901194

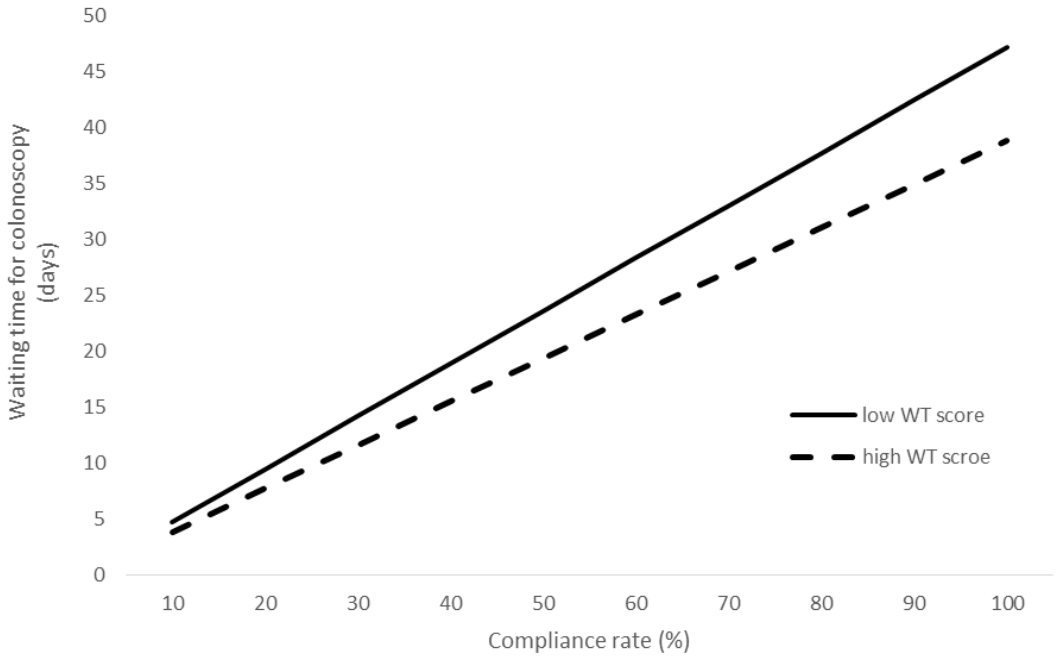
Figure 4.7 Predicted waiting time for colonoscopy by different compliance rates as given arrival rate and positive rate.



Scenarios	Waiting days									
(1)	10	20	30	40	50	60	70	80	91	101
(2)	7	13	20	27	34	40	47	54	60	67
(3)	4	9	13	17	22	26	30	34	39	43
(4)	3	6	9	11	14	17	20	23	26	29

doi:10.6342/NTU201901194

Figure 4.8 Predicted waiting time for colonoscopy by different compliance rates with different WT score groups as given 30% of annual coverage rate and 3% of positive rate.



WT Score	Waiting days									
Low	5	9	14	19	24	28	33	38	43	47
High	4	8	12	16	19	23	27	31	35	39

doi:10.6342/NTU201901194

Chapter 5 Generalized Coxian phase-type Markov Process for Disease Natural History Model of CRC

Summary

Introduction While stochastic process is useful for evaluation of effectiveness of population-based service screening with the creation of the pseudo control group**,** data with hidden states within the pre-clinical detectable phase (PCDP) feature with non-Markov stochastic are often encountered. The device of stage method with a mixture of serial and parallel form is therefore proposed to tackle this issue.

Aim The Coxian phase-type Markov process was therefore created to evaluate the effectiveness of population-based cancer screening program

Data Taiwanese population-based screening for CRC with FIT with the incorporation of AJCC tumour stage.

Method The Coxian phase-type Markov process was used to model hidden transient states within PCDP and the corresponding absorbing states reaching to clinical phase. The proposed Coxian phase-type Markov process constituted the likelihood functions with different detection modes, making allowance for the sensitivity of the screening method. Bayesian DAG model was used to estimate the parameters of interest. Simulation studies with asymptotic analysis and perturbation analysis were performed to elucidate statistical inference and power function of each parameter.

doi:10.6342/NTU201901194

Results The transition rates of early and late CRC in the PCDP and both absorbing rates to CP using the Coxian phase-type Markov process were estimated, making allowance for the sensitivity of the screening method, to represent the force of upstaging and downstaging. As considering the measurement errors (sensitivity), the estimated CRC incidence is 1.28 per 1000. The transition rate from early PCDP to late PCDP compared with that from early PCDP to early CP is approximately 3-fold. Based on the five-state model, the projected advanced cancer reduction is 15.44% under 20% coverage rate. After coverage rate reaches to 80%, the effectiveness of advanced cancer reduction is 29.30%. The effectiveness of FIT screening with different inter-screening interval and coverage rate was evaluated with the application of the proposed five-state Coxian phase-type model.

Conclusion The Coxian phase-type Markov process with the device stage method was proposed to evaluate the effectiveness of population-based screening with non-Markovian property.

doi:10.6342/NTU201901194

5.1 Introduction

5.1.1 The pseudo control group created from stochastic process

With the generalized Coxian phase-type Markov process for disease natural history model of CRC, one is able to evaluate the long-term effectiveness in organized service screening program with escape from the following obstacles. Firstly, there is a lacking of the comparator, the control group in the RCT and the unscreened in the observational follow-up cohort or the pre-screened group in non-RCT (Chen LS et al.(2010)). Second, it may require long-term follow-up data that involve the difficulty of logistics of follow-up in both cost and time aspects. The first difficulty may be solved by the adjustment for self-selection bias with the attendance rate while the unscreened group is taken as the comparator and also the allowance made for the maturation of incidence in colorectal cancer when the pre-screened group is used as the comparator. The solution to the second issue is the better use of surrogate endpoints. Previous studies have already shown that the use of tumour staging as surrogate endpoints may not only dispense with the logistics of follow-up but also enhance statistical power (Chen THH et al. (1999), and Chiu SY et al. (2011)) compared with the use of primary endpoint like CRC death on the basis of the randomized controlled trial. More importantly, evaluation of organized screening program with quasi-experimental design may answer whether it is effective in reducing CRC mortality and advanced cancer but

doi:10.6342/NTU201901194

it may not throw light on why and how it works, which is related to the optimal screening interval and age to begin with screen.

To this end, one may consider a modelling approach like the classical example of the MISCAN model and the stochastic process used in the previous studies (Buskermolen M et al. (2018), Chen THH et al.(1999), Chen LS et al. (2007), and Wu HM et al. (2006)). It builds up the disease natural history of colorectal neoplasia by AJCC tumour stage for estimating the expected advanced CRC and death to form the pseudo-control group. The expected results from this comparator were compared with the corresponding findings given the different screening policies of FIT. So doing provides an opportunity to evaluate long-term effectiveness of organized colorectal cancer screening in reducing advanced CRC and also subsequent deaths from CRC dispensing with the requirement of the comparator and also the necessity of waiting for longitudinal follow-up of the entire screened cohort.

5.1.2 Device of stage method for non-Markovian process

While Markov process has been extensively applied to multistate disease process, which is particularly important for evaluation of population-based periodical screening data for example. The most concern over its application is often raised as to whether the empirical data are amenable to the assumption of Markov property, that is, the

doi:10.6342/NTU201901194

probability of the next state transition is only dependent on the current state rather than the previous state.

In modelling data on population-based screening for cancer and chronic disease, we are often faced with three main phenotypes on the status of disease, free of disease (FOD), pre-clinical detectable phenotype (PCDP), and clinical phenotype (CP). In the case of colorectal cancer screening, three corresponding states are, FOD on colorectal cancer (CRC), PCDP CRC, CP CRC. There are several literatures on the application of three-state Markov process to modelling the disease natural history of CRC from FOD to CP, yielding two important parameters, annual incidence rate of PCDP CRC and the dwelling time of lingering with PCDP without progression to CP.

It is often criticized that three-state Markov process is not sufficient for modelling the sample path of empirical data on the disease natural history of CRC from FOD to CP when Markov property is applied. The reason is that the transition to CP is highly dependent on how long PCDP CRCs stay or what kind of attributes they carry with. The Markov property is often problematic when one encounters the empirical data on organized service screening with imbalanced design on irregular inter-screening interval. If there are measurement errors such as the sensitivity of the screening method the Markov property is even worse and should be dealt with. According to the classification of AJCC staging classification, low grade of tumour stage of PCDP CRC

doi:10.6342/NTU201901194

often stay longer than high grade of tumour stage. In the language of stochastic process, this means the sample path of data may be amenable to non-Markovian process.

To accommodate non-Markovian process, three approaches are considered, including device of stage method, supplementary variable, and embedded Markov chain, in the development of stochastic process. The two latter are evolved into semi-Markov process and the former is characterized by phase-type Markov process if the assumption of progressive property is made. The device stage of method is the main subject of this thesis.

5.1.3 Evaluation of population-based service screening program

In Taiwan, a nationwide organized colorectal cancer screening with biennial FIT has been conducted from 2004 until so far. We are faced with the challenge of whether this nationwide organized service screening program is effective in reducing CRC mortality and advanced cancer. Is the policy changed into annual regime or triennial regime? Do we need to start the screening earlier as the young age has rising incidence? It is very interesting to assess these subsidiary issues of screening policy by using a modelling approach.

The aim of this chapter is to estimate the rate of incidence and progression of CRC between tumour staging classified by AJCC based on a Coxian phase-type Markov

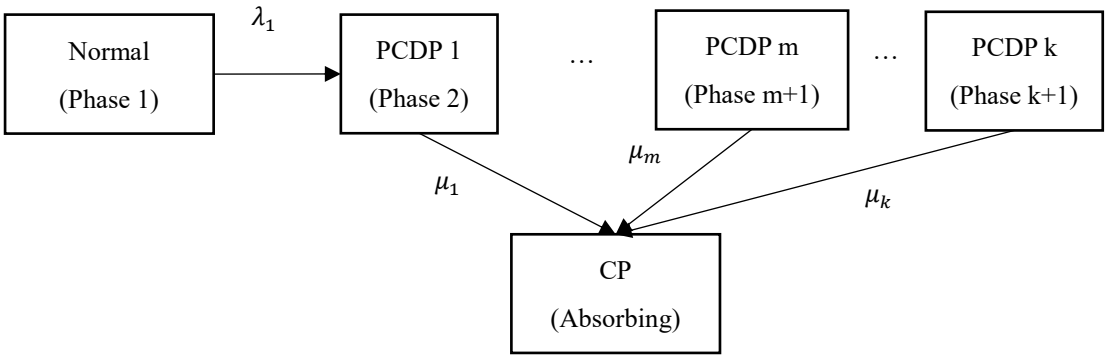
doi:10.6342/NTU201901194

process and then to evaluate long-term effectiveness of Taiwanese nationwide biennial FIT screening program.

5.2 Coxian phase-type with devices of stage for non-Markovian process

5.2.1 Coxian Phase-type Model

A generalized Coxian phase-type with devices of stage, *k* phases in PCDP, for non-Markovian process has been depicted below. The Coxian phase-type model is a special case of multi-state Markov process. The continuous-time Markov process has k+1 transient phases, and one absorbing phase.



The disease progression from phase i to phase (i+1) in a short time interval can be denoted by λ_i and the absorbing rate from phase i to absorbing phase denoted by μ_i. The probability density function (p.d.f) of Coxian phase-type distribution is

$$f(t) = \boldsymbol{\pi}\exp(\mathbf{Q}t)\mathbf{q}$$

doi:10.6342/NTU201901194

$$\boldsymbol{\pi} = (1,0,\ldots,0)$$

$$\mathbf{Q} = \begin{bmatrix} -\lambda_1 & \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & -(\lambda_2+\mu_1) & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -\mu_k \end{bmatrix}$$

$$\mathbf{q} = (0, \mu_1, \mu_2, \cdots, \mu_k)^T$$

where $\boldsymbol{\pi}$ is initial probability vector (1×(k+1)), $\mathbf{Q}$ is intensity matrix restricted to the transient phases ((k+1)×(k+1)), and $\mathbf{q}$ is a vector transition rates from transient phases to the absorbing phase ((k+1)×1).

To ensure absorption with probability one in a finite time, we rewrite the infinitesimal matrix in block-matrix,

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{q} \\ \mathbf{0} & 0 \end{bmatrix}.$$

By using the Kolmogorov differential equations, it has the same unique solution with forward equations:

$$\boldsymbol{P'(t)} = \boldsymbol{P(t)R}$$

and backward equations (Cox, & Miller, 1965):

$$\boldsymbol{P'(t)} = \boldsymbol{RP(t)}$$

where the matrix $\boldsymbol{P(t)}$ is the transition probabilities. If the matrix $\mathbf{R}$ is finite with initial condition $\boldsymbol{P(0)} = \boldsymbol{I}$, the formal solution can be derived as

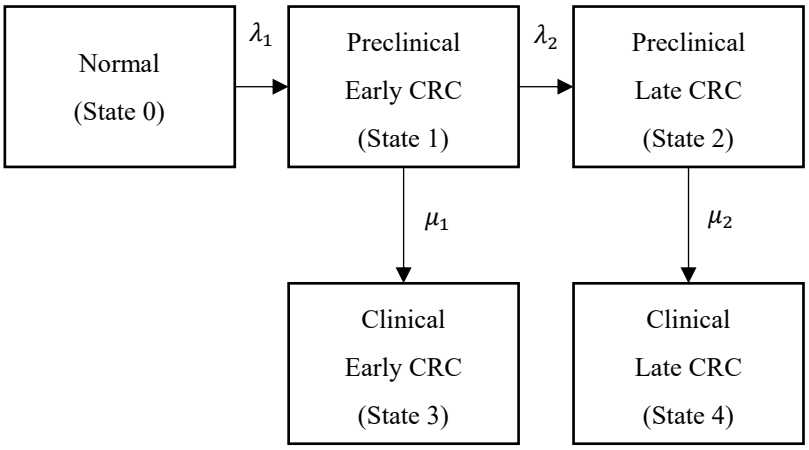
$$\boldsymbol{P(t)} = exp(\boldsymbol{R}t)$$

doi:10.6342/NTU201901194

5.2.2 Consideration of measurement errors

Free of measurement errors

We take a five-state Markov model for example



, the state space is X_t={0: Normal, 1: Preclinical Early CRC (Stage 0/I), 2: Preclinical Late CRC (Stage II/III/IV), 3: Clinical Early CRC (Stage 0/I), 4: Clinical Late CRC (Stage II/III/IV)}. State 3 and State 4 are absorbing states. The p.d.f is

$$f(t) = \boldsymbol{\pi}\exp(\mathbf{Q}t)\mathbf{q}$$

$$\boldsymbol{\pi} = (1,0,0,0,0)$$

$$\mathbf{Q} = \begin{bmatrix} -\lambda_1 & \lambda_1 & 0 \\ 0 & -(\lambda_2+\mu_1) & \lambda_2 \\ 0 & 0 & -\mu_2 \end{bmatrix}$$

$$\mathbf{q} = \begin{bmatrix} 0 & 0 \\ \mu_1 & 0 \\ 0 & \mu_2 \end{bmatrix}$$

$\mathbf{q}$ is a 3×2 matrix due to two absorbing phases.

By using the Kolmogorov differential equations, the transition probability can be given by

doi:10.6342/NTU201901194

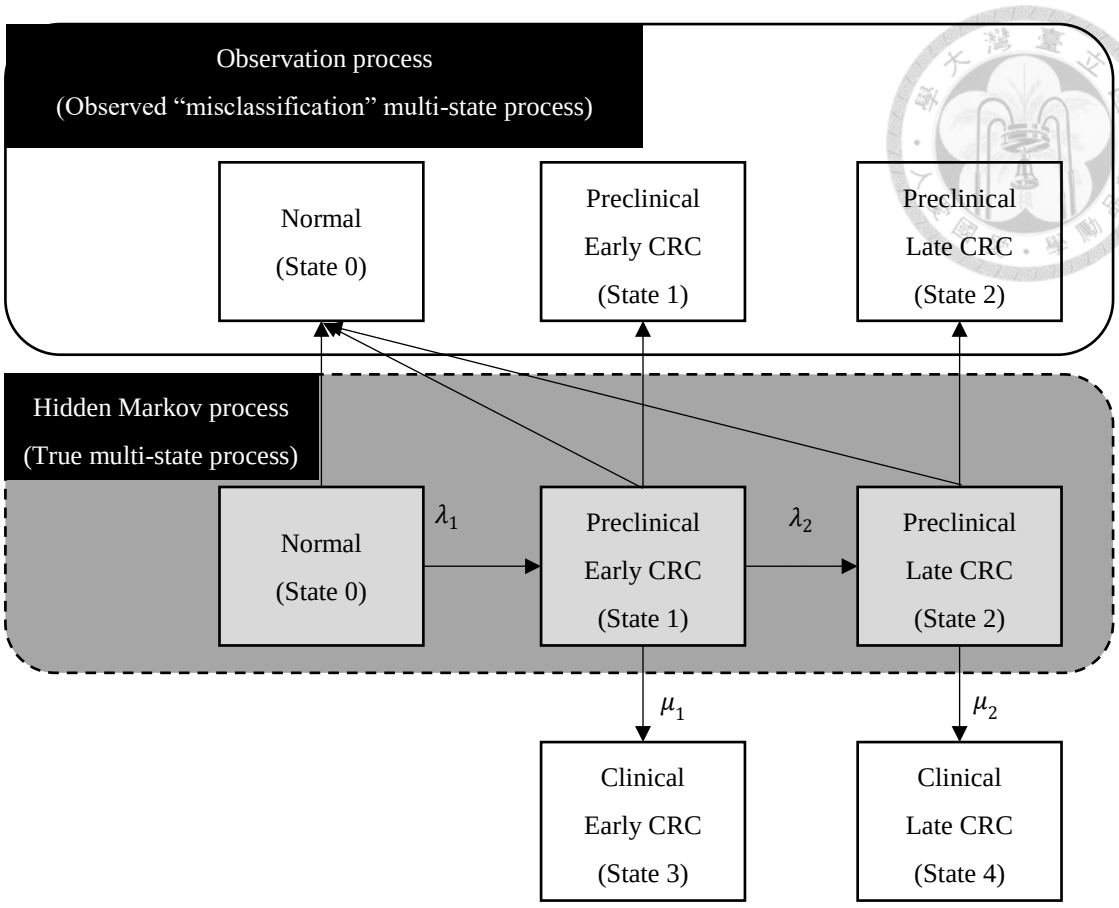
$$\mathbf{P}(\mathrm{t}) = \mathbf{exp}(\mathbf{R}\mathrm{t}) = \begin{bmatrix} P_{00}(t) & P_{01}(t) & P_{02}(t) & P_{03}(t) & P_{04}(t) \\ 0 & P_{11}(t) & P_{12}(t) & P_{13}(t) & P_{14}(t) \\ 0 & 0 & P_{22}(t) & 0 & P_{24}(t) \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{q} \\ \mathbf{0} & 0 \end{bmatrix} = \begin{bmatrix} -\lambda_1 & \lambda_1 & 0 & 0 & 0 \\ 0 & -(\lambda_2 + \mu_1) & \lambda_2 & \mu_1 & 0 \\ 0 & 0 & -\mu_2 & 0 & \mu_2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Model with measurement errors

In clinical reality, the state space X_t is usually measured by a biomarker of screening test (ex: FIT for CRC screening) which may not be a perfect detective modality so that it has a measurement error misclassifying outcome results (sensitivity and specificity). The Hidden Markov model (HMM) often used to deal with the issue of misclassification, which contains initial probability (π), transition probability (P), and emitted probability (Φ). In the HMM, there are two process: one is the Hidden Markov process (X_t), and the other is the observation process (O_t). Taking a five-state Markov model of CRC disease for example

doi:10.6342/NTU201901194



In the screening setting, if the sensitivity of screening test (i.g. FIT) is not 100%, it will yield the false negative results, which means those asymptomatic subject might have stayed in the preclinical early or late CRC, but they were regarded as a normal subject due to the measurement errors. Because we do not know whether he/she had preclinical early/late CRC or normal status, it is a Hidden Markov process (gray area with dashed line) representing underlying true states.

As using the HMM, it has to follow

(1) the Markov assumption,

$$\mathrm{P}(x_{t+1} = \mathrm{j} | x_t = \mathrm{i}, x_{t-1} = \mathrm{m}, \dots, x_0 = \mathrm{n}) = \mathrm{P}(x_{t+1} = \mathrm{j} | x_t = \mathrm{i}) = P_{ij}$$

doi:10.6342/NTU201901194

(2) the stationary assumption,

$$P_{ij} = \mathrm{P}\left(x_{t_1+1} = \mathrm{j} \middle| x_{t_1} = \mathrm{i}\right) = \mathrm{P}\left(x_{t_2+1} = \mathrm{j} \middle| x_{t_2} = \mathrm{i}\right)$$

(3) the observation independence

$$\mathrm{P}(\mathrm{O}|x_1, x_2, \dots, x_T, \pi, \mathrm{P}, \Phi) = \prod_{t=1}^{T} P(o_t|x, \pi, \mathrm{P}, \Phi).$$

So the joint probability distribution can be derived as

$$\mathrm{P}(\mathrm{X}, \mathrm{O}) = \prod_{t=1}^{T} P(x_t|x_{t-1}, \pi, \mathrm{P}, \Phi) P(o_t|x_t)$$

Here, we define X_t as the hidden process with five states (0: Normal, 1: Preclinical Early CRC (Stage 0/I), 2: Preclinical Late CRC (Stage II/III/IV), 3: Clinical Early CRC (Stage 0/I), 4: Clinical Late CRC (Stage II/III/IV)). State 3 and State 4 are absorbing states. The HMM cannot be observed directly but yielded a correlated state with a presumed probability distribution (emission probability), and the observed process was defined as (O_t). The hidden process follows the Markov assumption:

$$\mathrm{P}(X_t = x_t|X_1, \dots, X_{t-1}, O_1, \dots, O_{t-1}) = \mathrm{P}(X_t = x_t|X_{t-1} = x_{t-1}).$$

The transition probability from state i to state j is

$$P_{ij} = \mathrm{P}(X_t = \mathrm{j}|X_{t-1} = \mathrm{i}),$$

and the initial probability defined as

$$\boldsymbol{\pi} = (\pi_1, \dots, \pi_m) = \big(\mathrm{P}(X_1 = 0), \dots, \mathrm{P}(X_1 = 4)\big) = (1,0,0,0,0)$$

with $\sum_{i=1}^{m} \pi_i = 1$. On the other hand, the observed process has the conditional

doi:10.6342/NTU201901194

independence assumption:

$$\mathrm{P}(O_t = o_t | X_1, \dots, X_t, O_1, \dots, O_{t-1}) = \mathrm{P}(O_t = o_t | X_t = x_t) = e_{ox}$$

which is the definition of emitted probability. Consider the emitted probability in the CPH model to adjust sensitivities of FIT test, which can be defined as

Sensitivity of early CRC: $\mathrm{P}(O_t = 1 | X_t = 1) = e_{11}$ and

Sensitivity of late CRC: $\mathrm{P}(O_t = 2 | X_t = 2) = e_{22}$.

5.2.3 Likelihood functions

In the screening program, there are four types of detection mode, the corresponding likelihood are given as

Free of measurement errors

Prevalent screen

Because screening program is provided to those asymptomatic subjects, the likelihood for participant at the first screen is the conditional probability, which is conditional on no previous clinical cancer. If the subject is diagnosed as prevalent screen-detected cancer then the likelihood is given by

$$L_{PSD,Early}(.) = \frac{P_{01}(age_s)}{P_{00}(age) + P_{01}(age) + P_{02}(age)}$$

$$L_{PSD,Late}(.) = \frac{P_{02}(age_s)}{P_{00}(age) + P_{01}(age) + P_{02}(age)}$$

If the subject is free of CRC, the likelihood is

doi:10.6342/NTU201901194

$$L_{N1}(.) = \frac{P_{00}(age)}{P_{00}(age_e) + P_{01}(age_e) + P_{02}(age_e)}$$

where age_s is age at first screen for ith individual, and age is the eligible age for screening in the first year.

Subsequent subjects

The likelihood of subject diagnosed as subsequent screen-detected cancer is given by

$$L_{SSD,Early}(.) = P_{01}(t_j - t_{j-1})$$

$$L_{SSD,Late}(.) = P_{02}(t_j - t_{j-1})$$

If the subject is free of CRC, the likelihood is

$$L_{N2}(.) = P_{00}(t_j - t_{j-1})$$

where $t_j - t_{j-1}$ is the interscreening interval between the time of last screen (t_{j-1}) and the current screen (t_j)

If subject had a normal results but was diagnosed as clinical CRC before the next screen (Interval cancer), the likelihood is

$$L_{IC,Early}(.) = P_{01}(D - t_{j-1})\mu_1$$

$$L_{IC,Late}(.) = P_{02}(D - t_{j-1})\mu_2$$

where D is the diagnosis time of CRC (the date of being reported in national cancer registry), and t_{j-1} is the last screen date. Because we can know the exact diagnosis time, it can be expressed by the p.d.f.

doi:10.6342/NTU201901194

If those ever screened subjects had a normal results at the last time of screening, they would be followed until the end of study or the upper limit of screening age, the likelihood is

$$L_{nonattendee}(.) = 1 - P_{03}(E - t_{j-1}) - P_{04}(E - t_{j-1})$$

where E is the end date of study or upper limit screening age, and t_{j-1} is the last screen date.

Refuser

As the prevalent subjects, the likelihood of refuser is also conditional on no previous cancer, but we know no if he/she had PCDP early cancer or PCDP late cancer because he/she did not participant the screening program during the period, so if refuser was diagnosed as clinical CRC, the likelihood can be given by

$$L_{Refuser,Early}(.) = \frac{P_{00}(age)P_{01}(D - age)\mu_1 + P_{01}(age)P_{11}(D - age)\mu_1}{P_{00}(age) + P_{01}(age) + P_{02}(age)}$$

$$L_{Refuser,Late}(.)$$

$$= \frac{P_{00}(age)P_{02}(D - age)\mu_2 + P_{01}(age)P_{12}(D - age)\mu_2 + P_{02}(age)P_{22}(D - age)\mu_2}{P_{00}(age) + P_{01}(age) + P_{02}(age)}$$

doi:10.6342/NTU201901194

If non-attendee is free of CRC, the likelihood is

$$L_{N3}(.) = 1 - \frac{P_{00}(age)[P_{03}(E-age)+P_{04}(E-age)]}{P_{00}(age)+P_{01}(age)+P_{02}(age)} - \frac{P_{01}(age)[P_{13}(E-age)+P_{14}(E-age)]}{P_{00}(age)+P_{01}(age)+P_{02}(age)} - \frac{P_{02}(age)P_{24}(E-age)}{P_{00}(age)+P_{01}(age)+P_{02}(age)}$$

where D is the diagnosis time of CRC (the date of being reported in national cancer registry), E is the end date of study or upper limit screening age, and age is the eligible age for screen in the first year.

With Measurement errors

To take sensitivities of FIT test into account, we firstly define the important indicators as follows,

$$\text{Sensitivity of early CRC: } \text{sen1} = e_{11} = \text{P}(O_t = 1 | X_t = 1),$$

$$\text{Sensitivity of late CRC: } \text{sen2} = e_{22} = \text{P}(O_t = 2 | X_t = 2),$$

$$\text{True negative of normal result: TN} = \frac{P_{00}(t)}{P_{00}(t)+P_{01}(t)(1-sen1)+P_{02}(t)(1-sen2)}$$

$$\text{False negative of early CRC: FN1} = \frac{P_{01}(t)(1-sen1)}{P_{00}(t)+P_{01}(t)(1-sen1)+P_{02}(t)(1-sen2)}$$

$$\text{False negative of late CRC: FN2} = \frac{P_{02}(t)(1-sen2)}{P_{00}(t)+P_{01}(t)(1-sen1)+P_{02}(t)(1-sen2)}$$

doi:10.6342/NTU201901194

The likelihoods of different detection mode are given as

Prevalent screen

$$L_{PSD,Early}(.) = \frac{P_{01}(age_s)sen1}{P_{00}(age) + P_{01}(age) + P_{02}(age)}$$

$$L_{PSD,Late}(.) = \frac{P_{02}(age_s)sen2}{P_{00}(age) + P_{01}(age) + P_{02}(age)}$$

$$L_{N1}(.) = \frac{P_{00}(age_s) + P_{01}(age_s)(1 - sen1) + P_{02}(age_s)(1 - sen2)}{P_{00}(age) + P_{01}(age) + P_{02}(age)}$$

where age_s is age at first screen for ith individual, and age is the eligible age for screen in the first year.

Subsequent subjects

$$L_{SSD,Early}(.) = TN \times P_{01}(t_j - t_{j-1})sen1 + FN1 \times P_{11}(t_j - t_{j-1})$$

$$L_{SSD,Late}(.) = TN \times P_{02}(t_j - t_{j-1})sen2 + FN1 \times P_{12}(t_j - t_{j-1})$$

$$+FN2 \times P_{22}(t_j - t_{j-1})$$

$$L_{N2}(.) = TN \times \Big(P_{00}(t_j - t_{j-1}) + P_{01}(t_j - t_{j-1})(1 - sen1)$$

$$+ P_{02}(t_j - t_{j-1})(1 - sen2)\Big)$$

$$L_{IC,Early}(.) = TN \times P_{01}(D - t_{j-1})\mu_1 + FN1 \times P_{11}(D - t_{j-1})\mu_1$$

$$L_{IC,Late}(.) = TN \times P_{02}(D - t_{j-1})\mu_2 + FN1 \times P_{12}(D - t_{j-1})\mu_2$$

$$+FN2 \times P_{22}(D - t_{j-1})\mu_2$$

$$L_{nonattendee}(.) = TN \times \Big(1 - P_{03}(E - t_{j-1}) - P_{04}(E - t_{j-1})\Big)$$

$$+FN1 \times \Big(1 - P_{13}(E - t_{j-1}) - P_{14}(E - t_{j-1})\Big) + FN2 \times \Big(1 - P_{24}(E - t_{j-1})\Big)$$

doi:10.6342/NTU201901194

where $t_j - t_{j-1}$ is the interscreening interval between the time of last screen (t_{j-1}) and the current screen (t_j), D is the diagnosis time of CRC (the date of being reported in national cancer registry), and E is the end date of study or upper limit screening age.

Because the false negative case was missed at last screen, the underlying population for the present screen comprise true negative (TN), and false negative ($FN1, FN2$), which are the function of t that depend on the round of screening,

$$\mathrm{t} = \begin{cases} age_s & , for\ the\ 2nd\ round \\ t_j - t_{j-1} & , for\ the\ subsequent \end{cases}$$

Refuser

Because subjects have never participated in the FIT screening program, the result will not be affected by the sensitivities of FIT. Therefore, the likelihood functions are the same as the approach of free of measurement errors.

5.2.4 Bayesian DAG (directed acyclic graphic model)

We constructed a Bayesian directed acyclic graphic (DAG) model for the estimation of disease natural history of CRC. Figure 5.1 depicts the Bayesian DAG model for a five-state model. Let Ω={0,1,2,3,4} denote the state space, where 0 is for CRC free, 1 and 2 for early and late CRC in PCDP, respectively, and 3 and 4 for their counterparts in CP. The numbers of participants classified as state 0, 1, 2 at the first

doi:10.6342/NTU201901194

screen, denoted by $\boldsymbol{X_F} = \{O_0, O_1, O_2\}$ following a multinomial distribution with $\boldsymbol{P_F}$, which are the same as the likelihood of prevalent subjects and can be expressed as

$$\boldsymbol{X_F} \sim Multinomial\ (\boldsymbol{P_F}, \boldsymbol{N_F}),$$

the numbers of attendees in state 0, 1, 2, 3, and 4 at the subsequent screen (*k*-th) given the state m (m=0, 1, 2) classified in the last screen (*(k-1)*-th), denoted by $\boldsymbol{X_{S_m}} = \{O_0^k | O_m^{k-1}, O_1^k | O_m^{k-1}, O_2^k | O_m^{k-1}, O_3^k | O_m^{k-1}, O_4^k | O_m^{k-1}\}$ following multinomial distributions with $\boldsymbol{P_{S_m}}$, which are the same as the likelihood of subsequent subjects and can be expressed as

$$\boldsymbol{X_{S_m}} \sim Multinomial\ (\boldsymbol{P_{S_m}}, \boldsymbol{N_{S_m}}).$$

Because FIT screening is only provided to those asymptom subjects, we assume $\boldsymbol{X_F}$ in state 0/1/2 and $\boldsymbol{X_{S_m}}$ with m=0, 1, 2.

For those who never screen, the numbers of refusers classified as state 0, 1, 2, 3, and 4, denoted by $\boldsymbol{X_C} = \{O_0, O_1, O_2, O_3, O_4\}$ following a multinomial distribution with $\boldsymbol{P_C}$, which are the same as the likelihood of refuser and can be expressed as

$$\boldsymbol{X_C} \sim Multinomial\ (\boldsymbol{P_C}, \boldsymbol{N_C}).$$

Bayesian MCMC estimation

The parameters of the five-state model of natural history were estimated by using the Markov Chain Monte Carlo simulation with Bayesian approach. We assigned a non-

doi:10.6342/NTU201901194

informative prior distribution for parameters $\boldsymbol{\Theta} = (\lambda_1, \lambda_2, \mu_1, \mu_2)$ in the transition probabilities including incidence rate and three transition rates as gamma distribution;

$$\boldsymbol{\Theta} \sim \text{Gamma}(0.001, 0.001),$$

and the prior of parameters of sensitivities $\boldsymbol{\Phi} = (e_{11}, e_{22})$ are assigned as beta distribution;

$$\boldsymbol{\Phi} \sim Beta(10, 10).$$

Then, the full joint probability distribution can be given as

$$\text{P}\left(\boldsymbol{X_F}, \boldsymbol{X_{S_m}}, \boldsymbol{X_C}, \boldsymbol{\Theta}, \boldsymbol{\Phi}\right) \propto \text{P}(\boldsymbol{X_F}|\boldsymbol{\Theta}, \boldsymbol{\Phi}) \times \text{P}\left(\boldsymbol{X_{S_m}}\middle|\boldsymbol{\Theta}, \boldsymbol{\Phi}\right) \times \text{P}(\boldsymbol{X_C}|\boldsymbol{\Theta}) \times \text{P}(\boldsymbol{\Theta}) \times \text{P}(\boldsymbol{\Phi}).$$

We set 30 of tuning iterations with 50 of burn-in iterations for reducing the correlation, and run 150,000 of MCMC iterations excluding the burn-in samples. The total samples of the posterior distribution are 5,000. After setting the initial value of parameters $\boldsymbol{\Theta} = (0.001, 0.1, 0.1, 0.1)$ and $\boldsymbol{\Phi} = (e_{11}, e_{22}) = (0.7, 0.7)$, the priors would be updated by data (likelihood), and we can get the estimated parameters and the 95% highest posterior density (HPD) region from samples of the posterior distribution by using Gibbs Sampling algorithm.

Gibbs Sampling algorithm

Gibbs sampler samples from the conditional distributions which was decomposed by the full joint posterior distribution. Let $\boldsymbol{\Omega} = (\boldsymbol{\Theta}, \boldsymbol{\Phi}) =$

doi:10.6342/NTU201901194

$(\lambda_1, \lambda_2, \mu_1, \mu_2, e_{11}, e_{22}) = (\theta_1, \dots, \theta_6)$ be the parameter space and the full posterior conditional distribution be $\mathrm{P}\left(\theta_i \middle| \theta_j, i \neq j, \boldsymbol{X_F}, \boldsymbol{X_{S_m}}, \boldsymbol{X_C}\right) \propto \mathrm{P}\left(\boldsymbol{X_F}, \boldsymbol{X_{S_m}}, \boldsymbol{X_C} | \boldsymbol{\Omega}\right) \times \mathrm{P}(\boldsymbol{\Omega})$, then the conditional distribution of θ_1 (λ_1) given $\theta_j = \theta_j'$, where $2 \leqq \mathrm{j} \leqq 6$, is

$$\mathrm{P}\left(\theta_1 \middle| \theta_j = \theta_j', 2 \leq \mathrm{j} \leq 6, \boldsymbol{X_F}, \boldsymbol{X_{S_m}}, \boldsymbol{X_C}\right)$$

$$\propto \mathrm{P}\left(\boldsymbol{X_F}, \boldsymbol{X_{S_m}}, \boldsymbol{X_C} | \boldsymbol{\Omega} = (\theta_1, \theta_2', \dots, \theta_6')\right) \times \mathrm{P}\left(\boldsymbol{\Omega} = (\theta_1, \theta_2', \dots, \theta_6')\right).$$

The steps of Gibbs sampling algorithm:

(1) Start from the initial value of $\boldsymbol{\Omega^0} = (\theta_1^0, \theta_2^0, \dots, \theta_6^0) = (0.001, 0.1, 0.1, 0.1, 0.7, 0.7)$.

(2) Update parameters at first time:

- Sample θ_1^1 from $\mathrm{P}\left(\theta_1 \middle| \theta_2^0, \dots, \theta_6^0, \boldsymbol{X_F}, \boldsymbol{X_{S_m}}, \boldsymbol{X_C}\right)$
- Sample θ_2^1 from $\mathrm{P}\left(\theta_2 \middle| \theta_1^1, \dots, \theta_6^0, \boldsymbol{X_F}, \boldsymbol{X_{S_m}}, \boldsymbol{X_C}\right)$
- Sample θ_3^1 from $\mathrm{P}\left(\theta_3 \middle| \theta_1^1, \theta_2^1, \dots, \theta_6^0, \boldsymbol{X_F}, \boldsymbol{X_{S_m}}, \boldsymbol{X_C}\right)$

⋮

- Sample θ_6^1 from $\mathrm{P}\left(\theta_6 \middle| \theta_1^1, \theta_2^1, \dots, \theta_5^1, \boldsymbol{X_F}, \boldsymbol{X_{S_m}}, \boldsymbol{X_C}\right)$

(3) Update parameters at (t+1) times:

- Sample θ_1^{t+1} from $\mathrm{P}\left(\theta_1 \middle| \theta_2^t, \dots, \theta_6^t, \boldsymbol{X_F}, \boldsymbol{X_{S_m}}, \boldsymbol{X_C}\right)$
- Sample θ_2^{t+1} from $\mathrm{P}\left(\theta_2 \middle| \theta_1^{t+1}, \dots, \theta_6^t, \boldsymbol{X_F}, \boldsymbol{X_{S_m}}, \boldsymbol{X_C}\right)$

⋮

- Sample θ_6^{t+1} from $\mathrm{P}\left(\theta_6 \middle| \theta_1^{t+1}, \theta_2^{t+1}, \dots, \theta_5^{t+1}, \boldsymbol{X_F}, \boldsymbol{X_{S_m}}, \boldsymbol{X_C}\right)$

until t=5000, then stop.

doi:10.6342/NTU201901194

5.3 Illustration with data from service screening program

5.3.1 Taiwanese population-based service CRC screening program

Data used in this study were derived from the Taiwanese nationwide CRC screening program which was organized by the Health Promotion Administration and has been launched since 2004. The details of the Taiwanese CRC program have been provided elsewhere (Chiu et al. (2015)). In brief, the program provided FIT screening biennially for residents aged 50-69 years since 2004. Subjects with positive FIT would be invited to take the follow-up confirmatory exam by public health nurses from community or medical staff from medical units. A computerized system for reporting screening finding, referral outcomes, and diagnosis has been built at the inception of the program. All records of adenomas lesion and CRC (screen-detected cases) identified at screening were kept in the system. The screening dataset regularly is compared with the National Cancer Registry to systematically identified CRCs outside the screening visits, such as interval cancers and refusers. The screening dataset is also regularly linked to the National Death Registry to ascertain the vital status. In this study, we enrolled all eligible population (aged 50-69 years) in 2004-2009 as the basis of my analysis cohort, which involved 5,417,699 eligible population. All screening records were retrieved until the end of 2014.

The screening findings in first and subsequent rounds of screening is shown in

doi:10.6342/NTU201901194

Table 5.1. Among the eligible population, there were 3,074,538 (56.7%) subjects attended screens at least once. The proportion of female attendees (55% and 63% in the first and later screens, respectively) is larger than male. Positive FIT findings were found in 7% attendees in the first screening round, and 6% in later rounds. The referral rates for confirmatory colonoscopy among those with positive FITs were 67% and 75% in the first and later rounds of screening. Until the end of 2014, the mean follow-up time is 4.86 years in screened group, and 9.02 years in unscreened group, and 9168 CRCs were detected at the first screen, 4653 were found in the later screens, 11,904 cancers diagnosed due to symptoms after their negative screening finding (interval cancer), and 45,818 CRCs were identified in the Cancer Registry among those refused to attend screening. The proportion of Stage 0/I CRCs was highest in cancers detected in screen-detected cancers (48% in prevalence screening, and 58% in subsequent screens), followed by interval cancer (39%), and refuser CRCs (22%) (Figure 5.2). On the other hand, the distribution of Stage IV cancers was highest in refuser CRCs (26%), followed by interval cancer (17%), and screen-detected cases (10% in prevalent screen, and 6% in subsequent screens). Table 5.2 lists number of subjects in different detection modes as defined in the above section.

doi:10.6342/NTU201901194

5.3.2 Estimated results for the multi-state evolution of colorectal neoplasm

Table 5.3 shows the estimated results of Coxian phase-type Markov model for the disease natural history of colorectal cancer. We characterized two phases in light of the AJCC staging in both PCDP and CP, but with different definitions for the two phases:

- Model 1: Phase I—stage 0/I; Phase II—stage II/III/IV
- Model 2: Phase I—stage 0/I/II; Phase II—stage III/IV
- Model 3: Phase I—stage 0/I/II/III; Phase II—stage IV

The estimated results for the abovementioned three models are shown in Table 5.3. The incidence of preclinical cancer were estimated as 1.28 (95% CI: 1.27, 1.29) per 1000 in all the three models. In Model 1, the instantaneous transition rate from stage 0/I to stage II/III/IV in PCDP before developing clinical symptom was estimated as 0.4624. The transition from PCDP to CP was faster in late stage (stage II/III/IV) (0.5231, 95% CI: 0.5008, 0.5462) than in the early stage (stage 0/I) (0.1448, 95% CI: 0.1364, 0.1543). The sensitivity for stage 0/I and stage II+ were estimated as 61% (95% CI: 57%, 64%) and 78% (95% CI: 75%, 81%), respectively.

If stage II was classified as early stage (Model 2) rather than late stage (Model 1), the transition within PCDP become more difficult with slower transition rate (0.2648, 95% CI: 0.2525,0.2780), but the transition from PCDP to CP for both early and late

doi:10.6342/NTU201901194

stage became faster. Because both the two phases in Model 2 were more severe than their counterparts in Model 1, the two estimates of sensitivity became larger (69% for early cancer of stage 0/I/II, and 74% for late cancer of stage III/IV). The similar phenomena were observed in Model 3. The DIC statistics for Model 1-3 were 1162853, 1174327, and 1152114, respectively. Although Model 3 defining Stage IV has a smallest DIC, defining stage IV as advanced CRC is contradict to the concept of early detection in screening, we therefore chose Model 1 as the best model for further investigation.

Table 5.4 shows the estimated results of Coxian phase-type for the natural history of CRC with different number of phases in PCDP and CP, from 2-phase (five-state model), 3-phase (seven-state model), to 4-phase (nine-state) model. Regardless of number of phases in PCDP and CP, the estimated incidence of CRC was consistent across models. The further split of stage IV from late CRC in the seven-state model (DIC=1,217,330), and further division of stage II and III in the nine-state model (DIC=1,260,090) showed poor performance in terms of higher DIC than that in the five-state model (DIC=1,164,440). Following results in Table 5.3 and 5.4, the five-state model defining stage 0/I as early cancer and stage II/III/IV as advanced cancer is the best model structure for the disease natural history for CRC. The final model would be used in further analysis.

We further explored whether the estimated results would be affected by exclusion

doi:10.6342/NTU201901194

of refusers. The comparators would be the data with inclusion of refusers without (also the five-state model in Table 5.4) and with (also Model 1 in Table 5.3) considering sensitivity. Note that the former assumes sensitivity as 100%. The last column lists their corresponding estimates without non-attenders. Table 5.5 shows that if we do not take non-attendee into consideration, the estimated incidence rate (λ_1) would be inflated to 0.2 per 1000 regardless of with or without measurement error. Taking sensitivities of FIT into account, all estimated transition rates in the model without non-attendee were greater than that in the model with non-attendee. After tackling the problem of selection-bias and measurement error, the estimated parameters of incidence rate, transition rate from early PCDP to advanced PCDP, transition rate from early PCDP to early CP, and transition rate from advanced PCDP to advanced CP are 12.8 per 10,000, 0.4624, 0.1448, and 0.5231, respectively. The sensitivities are 60.68% for early CRC and 77.98% for advanced CRC. The estimated results of this final model were used as the base case of further simulation study. Figure 5.3 shows the dynamic curves of cumulative risk of developing early PCDP, late PCDP, early CP, and late CP from normal, and early PCDP based on the final model.

5.4 Simulation Study of Asymptotic Analysis

In order to elucidate how the screening characteristics, including screening rate,

doi:10.6342/NTU201901194

inter-screening interval, and sensitivity of screening tool (FIT) influence the efficacy of screening in terms of the reduction of advanced CRC and mortality from CRC, we conducted a series of simulations with varying values of abovementioned parameters compared with no screening as a control group.

Each scenario was simulated with different cohort size from 10,000 to 1,000,000 to check the asymptotic property for the simulated effectiveness, and yield the asymptotic estimate of number needed to screen (NNS) to reduce one adverse event (advanced CRC or CRC death). We also reported the bias, coverage rate, and power to its reference results of effectiveness given the simulation with a true cohort size (n=5,417,699) given 100 times trials for each scenario.

5.4.1 Design for the simulation of asymptotic analysis

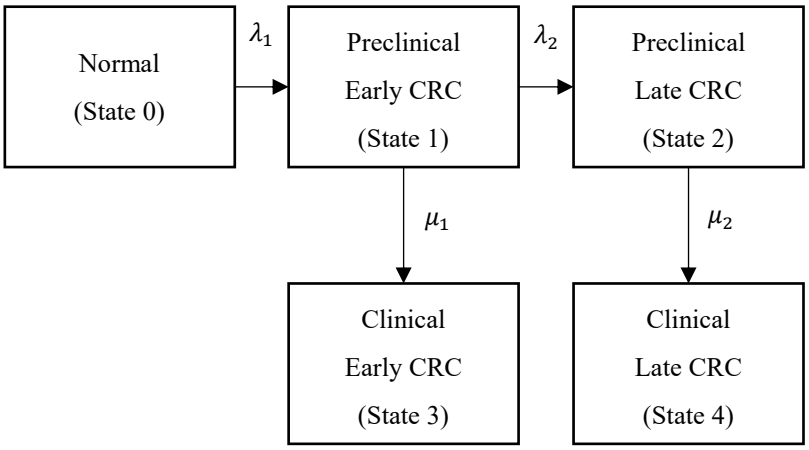
We set the base case of screening program was 60% screening rate, 2-year inter-screening interval with sensitivity of early CRC and advanced CRC as 60.68% and 77.98%, respectively. The three scenarios were like the follows.

- Scenario I: changing screening rate with 40%, 60%, and 80%
- Scenario II: changing inter-screening interval with 1-, 2-, and 3-yearly
- Scenario III: with improved sensitivity for early CRC and advanced CRC as 75% and 85%, respectively

doi:10.6342/NTU201901194

5.4.2 Algorithm for the computer simulation

Given the best model we had in Section 5.3 (see below), we developed the computer micro-simulation algorithm to generate the underlying distribution of sojourn times, also called as holding time in different states, and the embedded discrete-time Markov chain. Secondly, we superimposed the assigned screening characteristics to yield cases of different detection modes with different scenarios. The detailed procedure was described below.



Step 1. Simulate the underlying distribution of sojourn time

Let random variables T_0, T_1, and T_2 denote the sojourn time of staying in state 0 (normal), state 1 (early PCDP), and state 2 (late PCDP) before further departure. Each sojourn time would independently follow an exponential distribution with parameters of $\lambda_1 = 0.00128$, $\lambda_2 + \mu_1 = 0.6072$, and $\mu_2 = 0.5231$ (see Table 5.4). The transition probability matrix of the embedded Markov Chain matrix is

doi:10.6342/NTU201901194

$$
\begin{array}{c}
Normal \\
Early\ PCDP \\
Late\ PCDP \\
Early\ CP \\
Late\ CP
\end{array}
\begin{pmatrix}
0 & 1 & 0 & 0 & 0 \\
0 & 0 & \frac{\lambda_2}{\lambda_2+\mu_1} & \frac{\mu_1}{\lambda_2+\mu_1} & 0 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}
$$

Accordingly, the pathway for departing from State 1 can be either pathway (I) 0→1→2→4, or pathway (II) 0→1→3. For any subject arrives at state 1, the departing pathway will be randomly allocate with a randomly number following uniform distribution between 0 and 1. The detailed procedures are as follow.

(1) Independently generate T_0, T_1, and T_2 with distribution as follow:

$T_0 \sim \exp(\lambda_1)$

$T_1 \sim \exp(\lambda_2 + \mu_1)$

$T_2 \sim \exp(\mu_2)$

(2) Generate $U \sim$Uniform(0,1)

(3) If U lies between 0 and $\frac{\lambda_2}{\lambda_2+\mu_1}$, then the underlying state transition will follow the pathway (I) 0→1→2→4. Otherwise, the transition will follow pathway (II) 0→1→3.

(4) The total time (T) from birth to absorbing state (state 3 or state 4) is $T_0+T_1+T_2$ for pathway (I) or T_0+T_1 for pathway (II).

doi:10.6342/NTU201901194

Step 2. Scheduled screen and consequent data on detection mode

The following algorithms are used to get the data in screen program.

(1) Given *r* times scheduled screening at ages, $C_1,...,C_r$.

(2) For pathway (I) 0→1→2→4, if age at any specific schedule time point $C_m < T_0$ then the occupied state at C_m denoted by $Y(C_m)$ is assigned as 0. If C_m lies between T_0 and $T_0+ T_1$ then $Y(C_m)$ is assigned as 1. If C_m lies between $T_0+ T_1$ and the total time (T), then $Y(C_m)$ is assigned as 2. If C_m is beyond T, then $Y(C_m)$ is assigned as 4.

(3) For pathway (II) 0→1→3, $C_m < T_0$ then $Y(C_m)$ is assigned as 0. If C_m lies between T_0 and the total time ($T= T_0+ T_1$), then $Y(C_m)$ is assigned as 1. If C_m is beyond T, then $Y(C_m)$ is assigned as 3.

(4) Repeat m from 1 to r to get individual's screening history and decide the data realization of the r time points. The number of detected cases in each screen can be calculated. For example, in the first round of screen, the number of early screen-detected cases is $\sum_{n_1} I\left(Y(C_1)=1\right)$, and the number of late screen-detected cases is $\sum_{n_1} I\left(Y(C_1)=2\right)$.

5.4.3 Simulation results

Table 5.6 shows the results of asymptotic analysis for the effectiveness in terms of advanced cancer reduction with varying simulated cohort size under different

doi:10.6342/NTU201901194

scenarios. Compared to the reference value of 24.86% reduction, the bias across three scenarios and different sample size yielded a bias less than 1%, and all the coverage rates were higher than 92%. The power to detect a significant reduction would be affected by sample size (Figure 5.4). Almost all scenarios with sample size larger than 30,000 can reach 90% or higher power. The power function for annual screening was better than its counterpart of two-yearly and three-yearly. The asymptotic NNS to reduce one advanced CRC with 40%, 60%, and 80% screening rate were 380, 309, and 260, respectively.

Compared to the base case: 2-yearly screening of 60% screening rate (NNS=309), a shorter inter-screening interval had smaller NNS, 220 for annual screening. Similar phenomenon was observed with better performed screening tool, 265.

Table 5.7 shows the results of asymptotic analysis for the effectiveness in terms of CRC mortality reduction with varying simulated cohort size under different scenarios. Compared to the reference value of 20.93%, the bias for simulation size larger than 30,000 can be restrained with 1% across all the three scenarios. The coverage rate of mortality reduction was larger than 92%. For the power to provide a significant reduction in CRC mortality, we would need a cohort size no smaller than 100,000 (Figure 5.5). Almost all scenarios with sample size larger than 100,000 can reach 90% or higher power. The power function for annual screening was better than its counterpart

doi:10.6342/NTU201901194

of two-yearly and three-yearly.. The asymptotic NNS to reduce one CRC death with 40%, 60%, and 80% screening rate were 882, 750, and 656, respectively.

Compared to the basecase: 2-yearly screening of 60% screening rate (NNS=750), a shorter inter-screening interval had smaller NNS, 544 for annual screening. Similar phenomenon was observed with better performed screening tool of enhanced sensitivity for early and late CRCs, 638.

5.5 Perturbation analysis on effectiveness of FIT screening

In order to test the influence of screening characteristics on the effectiveness of FIT screening—screening rate, inter-screening interval, and sensitivity of FIT, we applied another series of screening, with the following three scenarios:

- Scenario I: changing screening rate from 10% to 100% with 10% increments
- Scenario II: changing inter-screening interval with 1-, 2-, and 3-yearly
- Scenario III: with improved sensitivity for early CRC and advanced CRC from 60% to 80% and 80% to 100%, respectively

The parameter of the base case were the same as the ones used in previous section.

Table 5.8 shows the effectiveness on advanced cancer reduction. With increasing coverage rate, the advanced cancer reduction increased from 13.18% for 10% screen rate to 34.10% with 100% screen rate, and the number needed to be screened decreased

doi:10.6342/NTU201901194

from 599 to 225 (Table 5.8, Figure 5.6). With lengthen inter-screening intervals, the efficacy declines and the number need to screen increased (Table 5.9, Figure 5.7). The advanced cancer reduction increased with improved sensitivities. The higher sensitivities (no matter in early stage of advanced stage), the less advanced cancers at diagnosis, which yielded more efficient screening (smaller number needed to be screened). The similar findings were seen in reducing clinically-detected cancers (Table 5.9), and mortality reduction (Table 5.10).

5.6 Discussion

5.6.1 Device of stage method with Coxian phase-type Markov process

This chapter demonstrate the use of device of stage method with the mixture of serial form and the parallel form, named as Coxian phase-type Markov process to tackle the problem of data with non-Markovian structure. The concept of device of stage method incorporated tumour stage into transient states in the PCDP built in the proposed Coxian phase-type Markov process provide an opportunity to evaluate the effectiveness of reducing advanced CRC as a result from screening with perturbation analysis on three parameters, including coverage rate, sensitivity, and inter-screening interval.

doi:10.6342/NTU201901194

5.6.2 The comparison of transition rate between g-FOBT and FIT

In the previous study of Nottingham and Denmark RCTs for CRC screening which provided the biennial fecal occult blood (FOB) test to subjects aged 45-74 years (Chiu SYH et al. (2011)), the estimated results of five-state CRC natural history model by Dukes' stage demonstrated that the CRC incidence was 1.48 per 1000, which was higher than our results in Taiwan, and the annual progression rate from PCDP localized CRC (Dukes' A, B) to PCDP non-localized CRC (Dukes' C, D), the rate from PCDP localized CRC to clinical localized CRC, and the rate from PCDP non-localized CRC to clinical non-localized CRC were 0.28, 0.22, and 0.72, respectively. Despite the definition of advanced CRC are different in comparison with our study, we had a greater rate from early PCDP to late PCDP (0.4624).

The Coxian phase-type stochastic process was used to estimate the tumour-stage-based natural history of CRC and project the effectiveness of population-based screening program. Under 60% coverage rate, biennial FIT screening led to 25% advanced CRC reduction, and 311 number need to be screened (NNS). As the inter-screening interval shortened to annual program, the reduction increased by 10% and NNS was reduced to 222. Asymptotic and perturbation analysis also revealed the required sample size to reach statistical power and also statistical impacts regarding different effect sizes as a result of various screening policies. Such a modelling

doi:10.6342/NTU201901194

approach also provides an alternative method in contrast to the traditional study design based on population-based experimental and quasi-experimental design requiring the control group.

5.6.3 Limitations

The device stage method may not be appropriate for non-Markov structure in some occasions particularly when number of stages increase, which may lead to too many parameters in contrast to sparse data. To relieve this concern, the semi-Markov model with the supplementary approach or the embedded Markov chain method can be considered.

In conclusion, we propose Coxian phase-type Markov process to deal with non-Markovian stochastic process. The model form with the incorporation of tumour stage, making allowance for measurement error, was applied to evaluating the effectiveness of population-based screening for CRC with FIT.

doi:10.6342/NTU201901194

Table 5.1 Descriptive results of demographic features, and basic screening characteristics.

	First Round	Subsequent Round
Number of FIT test	**3,074,538**	**2,552,925**[#]
Gender		
Male	1,376,955 (44.8%)	949,204 (37.2%)
Female	1,697,534 (55.2%)	1,603,721 (62.8%)
Age group		
50-54	808,767 (26.3%)	215,411 (8.4%)
55-59	991,314 (32.2%)	878,488 (34.4%)
60-64	720,564 (23.4%)	819,006 (32.1%)
65-69	553,893 (18.1%)	640,020 (25.1%)
Number of Positive test	**211,888 (6.9%)**	**159,382 (6.2%)**
Number of referral	**142,800 (67%)**	**119,648 (75%)**
Adenoma	**65,557**	**54,971**
PPV[+]	46%	46%
Detection rate	21‰	22‰
Advanced adenoma	**18,346**	**11,290**
PPV[+]	13%	9%
Detection rate	6‰	4‰
Colorectal Cancer	**9,168**	**4,653**
PPV[+]	6%	4%
Detection rate	3‰	2‰

[#] Number of repeated subject is 1,583,805 (repeated rate=52%)

[+] PPV (Positive Predictive Value) = number of disease / number of compliers

doi:10.6342/NTU201901194

Table 5.2 Frequencies of different types of CRC disease progression from normal, through PCDP, until CP.

Screening finding by round	Status	Frequency
Prevalent screening	Normal	3065370
	Stage 1 CRC	3899
	Stage 2 CRC	1445
	Stage 3 CRC	2010
	Stage 4 CRC	777
	Stage missing CRC	1037
Subsequent screening	Normal	2548272
	Stage 1 CRC	2439
	Stage 2 CRC	599
	Stage 3 CRC	925
	Stage 4 CRC	272
	Stage missing CRC	418
Interval cancer	Stage 1 CRC	4111
	Stage 2 CRC	2048
	Stage 3 CRC	2635
	Stage 4 CRC	1851
	Stage missing CRC	1259
Refuser	Normal	2540619
	Stage 1 CRC	8539
	Stage 2 CRC	9286
	Stage 3 CRC	11165
	Stage 4 CRC	10152
	Stage missing CRC	6676
Censoring[+]		2927405

[+] Censoring:normal state until the end of follow-up

doi:10.6342/NTU201901194

Table 5.3 Estimated results of dwell times of staying in non-advanced and advanced CRC and the corresponding sensitivity.

	Model 1 (Stage 2 as cutoff)	Model 2 (Stage 3 as cutoff)	Model 3 (Stage 4 as cutoff)
	Estimate (HPD interval)	Estimate (HPD interval)	Estimate (HPD interval)
Incidence rate	0.00128 (0.00127,0.00129)	0.00128 (0.00127,0.00129)	0.00128 (0.00126,0.00128)
Transition rate			
Early PCDP → Late PCDP	0.4624 (0.4354,0.4901)	0.2648 (0.2525,0.2780)	0.0900 (0.0865,0.0934)
Early PCDP → Early CP	0.1448 (0.1364,0.1543)	0.2283 (0.2173,0.2397)	0.2708 (0.2608,0.2804)
Late PCDP → Late CP	0.5231 (0.5008,0.5462)	0.5344 (0.5066,0.5620)	0.9748 (0.8880,1.0771)
Sensitivity			
Early PCDP	0.6068 (0.5717,0.6432)	0.6873 (0.6559,0.7167)	0.7145 (0.6927, 0.7374)
Late PCDP	0.7798 (0.7535,0.8066)	0.7417 (0.7121,0.7706)	0.8028 (0.7484,0.8579)

PCDP: pre-clinical detectable phase; CP: clinical phase; MST: mean sojourn time

doi:10.6342/NTU201901194

Table 5.4 Estimated results of transition rates of CRC natural history by using multi-state models without measurement errors

Transition rates	Estimate (Mean)	HPD Interval Lower	HPD Interval Upper	DIC
Five-state				
Incidence rate (λ_1)	0.00131	0.00130	0.00132	1164440
PCDP (Stage 0/I) → PCDP (Stage II+) (λ_2)	0.8917	0.8689	0.9134	
PCDP (Stage 0/I) → CP (Stage 0/I) (λ_3)	0.2748	0.2672	0.283	
PCDP (Stage II+) → CP (Stage II+) (λ_4)	0.7485	0.7285	0.7685	
Seven-state				
Incidence rate (λ_1)	0.00132	0.00131	0.00133	1217330
PCDP (Stage 0/I) → PCDP (Stage II/III) (λ_2)	0.8931	0.8710	0.9144	
PCDP (Stage II/III) → PCDP (Stage IV) (λ_3)	0.3226	0.3127	0.3328	
PCDP (Stage 0/I) → CP (Stage 0/I) (λ_4)	0.2739	0.2662	0.2814	
PCDP (Stage II/III) → CP (Stage II/III) (λ_5)	0.6371	0.6174	0.6553	
PCDP (Stage IV) → CP (Stage IV) (λ_6)	1.6644	1.5593	1.7701	
Nine-state				
Incidence rate (λ_1)	0.00132	0.00131	0.00133	1260090
PCDP (Stage 0/I) → PCDP (Stage II) (λ_2)	0.8832	0.8607	0.9043	
PCDP (Stage II) → PCDP (Stage III) (λ_3)	2.1830	2.0752	2.2909	
PCDP (Stage III) → PCDP (Stage IV) (λ_4)	0.6297	0.6057	0.6567	
PCDP (Stage 0/I) → CP (Stage 0/I) (λ_5)	0.2685	0.2608	0.2763	
PCDP (Stage II) → CP (Stage II) (λ_6)	0.8624	0.8188	0.9059	
PCDP (Stage III) → CP (Stage III) (λ_7)	0.6812	0.6549	0.7097	
PCDP (Stage IV) → CP (Stage IV) (λ_8)	1.7889	1.6550	1.9084	

doi:10.6342/NTU201901194

Table 5.5 Estimated results of transition rates of CRC natural history by using five-state models considering with and without those non-attendee of FIT screening

	With non-attendee	Without non-attendee
	Estimate (HPD interval)	Estimate (HPD interval)
Without measurement error		
Incidence rate (λ_1)	0.00131 (0.00130,0.00132)	0.00152 (0.00150,0.00154)
PCDP (Stage 0/I) → PCDP (Stage II+) (λ_2)	0.8917 (0.8689,0.9134)	0.8650 (0.8437,0.8885)
PCDP (Stage 0/I) → CP (Stage 0/I) (λ_3)	0.2748 (0.2672,0.2830)	0.3380 (0.3262,0.3491)
PCDP (Stage II+) → CP (Stage II+) (λ_4)	0.7485 (0.7285,0.7685)	0.7727 (0.7534,0.7928)
With measurement error		
Incidence rate (λ_1)	0.00128 (0.00127,0.00129)	0.00142 (0.00140,0.00144)
PCDP (Stage 0/I) → PCDP (Stage II+) (λ_2)	0.4624 (0.4354,0.4901)	0.4934 (0.4578,0.5247)
PCDP (Stage 0/I) → CP (Stage 0/I) (λ_3)	0.1448 (0.1364,0.1543)	0.2048 (0.1895,0.2198)
PCDP (Stage II+) → CP (Stage II+) (λ_4)	0.5231 (0.5008,0.5462)	0.5294 (0.5046,0.5539)
Sensitivity of Early PCDP	0.6068 (0.5717,0.6432)	0.6722 (0.6296,0.7150)
Sensitivity of Late PCDP	0.7798 (0.7535,0.8066)	0.7790 (0.7530,0.8066)

doi:10.6342/NTU201901194

Table 5.6 Simulated results of asymptotic analysis on the effectiveness of advanced cancer reduction

Sample size	Bias	Coverage rate	Power	NNS (95% C.I.)
Scenario I: Changing Screening rate				
40% Screening rate				
10,000	0.0107	95.2%	54%	582 (-2289, 3454)
30,000	0.0050	94.2%	91%	473 (-519, 1465)
50,000	-0.0002	95.6%	99%	405 (153, 657)
100,000	-0.0002	92.8%	100%	392 (235, 550)
300,000	0.00002	93.8%	100%	382 (301, 464)
500,000	0.0005	95.6%	100%	382 (320, 444)
1,000,000	0.0001	96.2%	100%	380 (337, 423)
60% Screening rate				
10,000	0.0045	96.2%	69%	369 (-891, 1629)
30,000	-0.0024	96.2%	98%	330 (50, 610)
50,000	-0.0006	94.6%	100%	321 (169, 473)
100,000	-0.0004	95.4%	100%	314 (220, 407)
300,000	0.0010	93.0%	100%	313 (254, 371)
500,000	0.0013	94.2%	100%	311 (270, 353)
1,000,000	-0.0001	96.8%	100%	309 (281, 336)
80% Screening rate				
10,000	0.0019	95.2%	83%	318 (-260, 896)
30,000	0.0059	95.0%	100%	280 (129, 432)
50,000	0.0002	95.6%	100%	267 (173, 360)
100,000	0.0003	94.2%	100%	264 (195, 333)
300,000	0.0010	94.8%	100%	266 (226, 298)
500,000	0.0004	95.2%	100%	261 (232, 290)
1,000,000	-0.0002	94.6%	100%	260 (240, 280)

NNS: Number needed to be screened

doi:10.6342/NTU201901194

Table 5.6 (Continued)

Sample size	Bias	Coverage rate	Power	NNS (95% C.I.)
Scenario II: Changing Inter-screening interval				
Annual screening interval				
10,000	0.0044	93.8%	93%	252 (20, 483)
30,000	-0.0024	93.4%	100%	226 (135, 316)
50,000	0.0013	95.2%	100%	226(160, 291)
100,000	0.0021	94.6%	100%	224 (177, 271)
300,000	0.0010	95.8%	100%	222 (196, 247)
500,000	0.0011	95.0%	100%	221 (201, 241)
1,000,000	-0.0004	95.8%	100%	220 (207, 233)
Biennial screening interval				
10,000	0.0028	94.6%	70%	361 (-1225, 1947)
30,000	0.0017	95.4%	98%	335 (109, 562)
50,000	-0.0018	96.2%	100%	318 (176, 461)
100,000	0.0003	95.2%	100%	315 (218, 413)
300,000	-0.0006	94.8%	100%	309 (257, 361)
500,000	0.0012	94.0%	100%	311 (270, 353)
1,000,000	0.0002	94.8%	100%	309 (280, 338)
Triennial screening interval				
10,000	-0.0073	94.8%	51%	403 (-1951, 2756)
30,000	0.0032	94.0%	89%	491 (-378, 1359)
50,000	0.0018	96.2%	98%	430 (115, 745)
100,000	0.0016	96.2%	100%	411 (236, 587)
300,000	0.0009	94.2%	100%	399 (310, 488)
500,000	0.0003	95.6%	100%	395 (329, 462)
1,000,000	0.0007	94.8%	100%	395 (350, 440)

NNS: Number needed to be screened

doi:10.6342/NTU201901194

Table 5.6 (Continued)

Sample size	Bias	Coverage rate	Power	NNS (95% C.I.)
Scenario III: Changing sensitivities of FIT: 75% for early CRC and 85% for advanced CRC				
10,000	0.0061	96.4%	81%	339 (-330, 1008)
30,000	-0.0026	94.6%	100%	274 (137, 411)
50,000	0.0047	95.4%	100%	278 (170, 387)
100,000	0.0019	96.6%	100%	270 (204, 336)
300,000	0.0002	94.6%	100%	266 (228, 304)
500,000	0.0005	94.6%	100%	266 (236, 295)
1,000,000	0.0004	95.2%	100%	265 (244, 286)

NNS: Number needed to be screened

doi:10.6342/NTU201901194

Table 5.7 Simulated results of asymptotic analysis on the effectiveness of CRC mortality reduction

Sample size	Bias	Coverage rate	Power	NNS (95% C.I.)
Scenario I: Changing Screening rate				
40% Screening rate				
5,000	0.0424	95.4%	19%	328 (-3065, 3721)
10,000	0.0135	95.6%	28%	592 (-3878, 5063)
30,000	0.0011	95.0%	58%	1092 (-4684, 6742)
50,000	0.0035	92.0%	77%	1399 (-5521, 8318)
100,000	0.0024	95.0%	96%	1031 (-107, 2132)
300,000	0.0023	94.6%	100%	915 (585, 1246)
500,000	0.0003	95.4%	100%	893 (663, 1122)
1,000,000	-0.0007	96.2%	100%	880 (721, 1039)
3,000,000	0.0003	95.4%	100%	882 (790, 973)
60% Screening rate				
5,000	0.0347	94.2%	22%	436 (-2969, 3842)
10,000	0.0153	94.8%	35%	845 (-3743, 5433)
30,000	0.0020	95.4%	70%	1018 (-2328, 4364)
50,000	0.00002	96.4%	88%	840 (-252, 1932)
100,000	0.0005	95.2%	99%	803 (298, 1308)
300,000	0.0014	96.6%	100%	770 (541, 1000)
500,000	0.0013	95.6%	100%	763 (592, 935)
1,000,000	0.0006	95.0%	100%	756 (635, 877)
3,000,000	-0.00009	94.8%	100%	750 (682, 819)
80% Screening rate				
5,000	0.0181	94.6%	27%	453 (-2523, 3429)
10,000	0.0193	96.6%	42%	837 (-3168, 4842)
30,000	0.0045	94.4%	81%	856 (-2680, 4392)
50,000	0.0079	95.4%	95%	778 (-101, 1657)
100,000	0.0020	97.2%	100%	690 (353, 1026)
300,000	0.0029	95.4%	100%	672 (499, 846)
500,000	0.0013	96.4%	100%	662 (536, 789)
1,000,000	-0.0003	96.0%	100%	655 (568, 742)
3,000,000	0.0005	95.8%	100%	656 (605, 707)

NNS: Number needed to be screened

doi:10.6342/NTU201901194

Table 5.7 (Continued)

Sample size	Bias	Coverage rate	Power	NNS (95% C.I.)
Scenario II: Changing Inter-screening interval				
Annual screening interval				
5,000	0.0221	96.0%	35%	538 (-1874, 2951)
10,000	0.0055	96.2%	55%	624 (-2197, 3445)
30,000	0.0085	95.2%	92%	666 (-2313, 3644)
50,000	0.0002	95.4%	99%	582 (198, 965)
100,000	-0.0011	93.8%	100%	560 (333, 786)
300,000	-0.0001	95.6%	100%	548 (437, 660)
500,000	0.0017	94.6%	100%	551 (460, 641)
1,000,000	0.0010	95.2%	100%	547 (485, 610)
3,000,000	-0.0003	95.4%	100%	544 (509, 579)
Biennial screening interval				
5,000	0.0347	94.2%	22%	436 (-2969, 3842)
10,000	0.0153	94.8%	35%	845 (-3743, 5433)
30,000	0.0020	95.4%	70%	1018 (-2328, 4364)
50,000	0.00002	96.4%	88%	840 (-252, 1932)
100,000	0.0005	95.2%	99%	803 (298, 1308)
300,000	0.0014	96.6%	100%	770 (541, 1000)
500,000	0.0013	95.6%	100%	763 (592, 935)
1,000,000	0.0006	95.0%	100%	756 (635, 877)
3,000,000	-0.00009	94.8%	100%	750 (682, 819)
Triennial screening interval				
5,000	0.0248	95.0%	17%	373 (-3236, 3982)
10,000	0.0015	95.4%	26%	702 (-3486, 4890)
30,000	0.0077	94.0%	53%	976 (-6638, 8589)
50,000	0.0025	94.8%	72%	1263 (-3334, 5859)
100,000	0.0020	96.8%	93%	1049 (164, 1934)
300,000	0.0002	96.8%	100%	970 (579, 1360)
500,000	-0.0009	95.8%	100%	948 (662, 1235)
1,000,000	-0.0004	95.0%	100%	942 (756, 1128)
3,000,000	0.0003	95.8%	100%	941 (835, 1047)

NNS: Number needed to be screened

doi:10.6342/NTU201901194

Table 5.7 (Continued)

Sample size	Bias	Coverage rate	Power	NNS (95% C.I.)
Scenario III: Changing sensitivities of FIT: 75% for early CRC and 85% for advanced CRC				
5,000	0.0374	95.6%	26%	312 (-2549, 3172)
10,000	0.0091	96.4%	41%	784 (-3464, 5033)
30,000	-0.0003	94.4%	79%	735 (-1488, 2958)
50,000	-0.0003	94.6%	93%	814 (-3578, 5206)
100,000	0.0011	95.2%	100%	670 (340, 1001)
300,000	0.0003	95.6%	100%	649 (477, 821)
500,000	0.0004	94.8%	100%	645 (520, 770)
1,000,000	0.0006	94.4%	100%	642 (558, 727)
3,000,000	-0.0003	95.0%	100%	638 (589, 687)

NNS: Number needed to be screened

doi:10.6342/NTU201901194

Table 5.8 Relative risk, absolute risk, and NNS for assessing FIT screening policy on advanced cancer reduction

Coverage rate (%)	Number of advanced cancer		Advanced cancer reduction (95% CI; %)	Number needed to be screened
	Control	Invited		
Scenario I: Change coverage rates				
10%	3913	3396	13.18 (9.12-17.25)	599 (372-826)
20%	3914	3309	15.44 (11.47-19.41)	507 (352-662)
30%	3906	3212	17.73 (13.91-21.55)	439 (329-549)
40%	3916	3121	20.27 (16.50-24.04)	382 (300-464)
50%	3910	3031	22.48 (18.84-26.11)	344 (279-409)
60%	3912	2940	24.83 (21.21-28.46)	311 (257-364)
70%	3916	2849	27.22 (23.70-30.75)	283 (238-328)
80%	3909	2763	29.30 (25.83-32.77)	263 (225-301)
90%	3914	2667	31.84 (28.70-34.99)	241 (212-271)
100%	3914	2579	34.10 (30.90-37.30)	225 (199-252)
Scenario II: Change inter-screening interval				
1-yearly	5379	2553	34.72 (31.31-38.14)	222 (195-249)
2-yearly	5379	2937	24.84 (21.43-28.25)	311 (259-362)
3-yearly	5376	3150	19.45 (15.65-23.25)	399 (309-489)
Scenario III: sensitivity of early and advanced CRC				
(Early	**Late)**			
60%	80%	3922	2947	24.85 (21.26-28.44)	310 (257-363)
65%	80%	3918	2902	25.91 (22.42-29.41)	297 (250-345)
70%	85%	3950	2863	27.51 (24.09-30.94)	277 (236-319)
75%	85%	3953	2819	28.65 (24.99-32.31)	266 (225-307)
75%	90%	3975	2825	28.92 (25.50-32.33)	262 (225-299)
80%	95%	4010	2782	30.60 (27.38-33.83)	245 (214-277)
80%	100%	4039	2787	30.97 (27.57-34.36)	241 (209-273)

doi:10.6342/NTU201901194

Table 5.9 Relative risk, absolute risk, and NNS for assessing FIT screening policy on clinically detected CRCs

Coverage rate (%)	Number of clinically-detected CRC		Number of clinically-detected CRC reduction	RR (95% CI)
	Control	Invited		
Scenario I: Change coverage rates				
10%	4575	4270	305	6.64 (2.77-10.52)
20%	4575	3966	609	13.28 (9.66-16.90)
30%	4570	3659	916	19.92 (16.35-23.49)
40%	4570	3359	1216	26.49 (23.15-29.82)
50%	4574	3056	1518	33.16 (30.09-36.23)
60%	4568	2755	1820	39.69 (36.82-42.56)
70%	4577	2446	2128	46.53 (43.91-49.16)
80%	4567	2148	2426	52.95 (50.59-55.32)
90%	4569	1841	2734	59.70 (57.49-61.92)
100%	4572	1544	3030	66.22 (64.23-68.20)
Scenario II: Change inter-screening interval				
1-yearly	4566	2315	2251	49.29 (46.66-51.93)
2-yearly	4566	2755	1811	39.65 (36.79-42.52)
3-yearly	4565	3087	1479	32.35 (29.30-35.41)
Scenario III: sensitivity of early and advanced CRC				
(Early	**Late)**			
60%	80%	4573	2747	1826	39.92 (36.98-42.86)
65%	80%	4568	2713	1855	40.59 (37.68-43.49)
70%	85%	4571	2652	1919	41.97 (39.37-44.57)
75%	85%	4569	2617	1951	42.69 (39.91-45.48)
75%	90%	4570	2590	1980	43.32 (40.63-46.01)
80%	95%	4573	2528	2045	44.72 (42.01-47.42)
80%	100%	4575	2500	2075	45.34 (42.76-47.91)

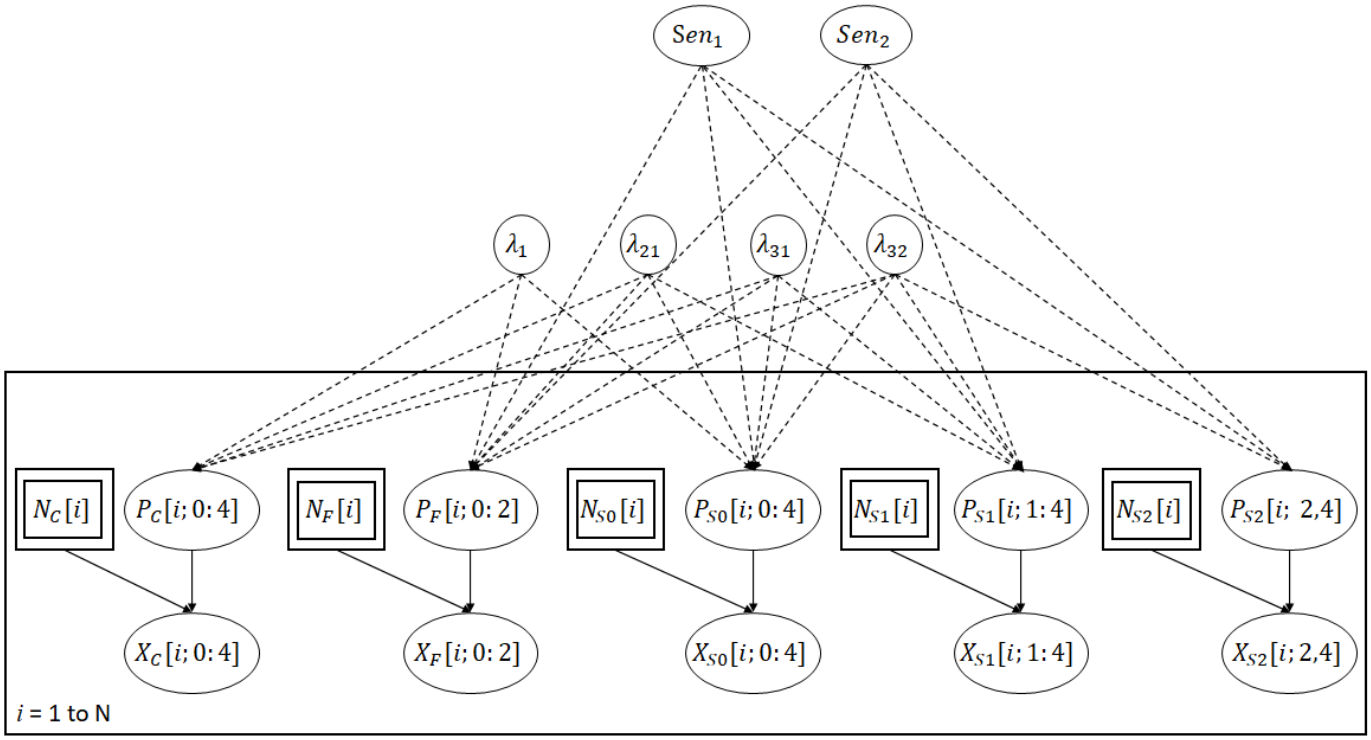
doi:10.6342/NTU201901194

Table 5.10 Relative risk, absolute risk, and NNS for assessing FIT screening policy on CRC mortality reduction

Coverage rate (%)	Number of CRC death		CRC mortality reduction (95% CI; %)	Number needed to be screened
	Control	Invited		
Scenario I: Change coverage rates				
10%	1914	1661	13.18 (7.32-19.04)	1384 (-3822-6590)
20%	1914	1632	14.69 (9.02-20.36)	1122 (543-1701)
30%	1915	1601	16.34 (10.53-22.15)	1000 (538-1462)
40%	1911	1572	17.70 (12.39-23.01)	913 (578-1248)
50%	1913	1544	19.24 (13.67-24.81)	837 (539-1136)
60%	1913	1512	20.93 (15.56-26.29)	765 (534-995)
70%	1914	1484	22.44 (17.20-27.69)	711 (502-920)
80%	1915	1457	23.88 (18.63-29.13)	666 (491-841)
90%	1915	1426	25.53 (20.47-30.59)	621 (475-767)
100%	1913	1397	26.92 (22.13-31.72)	588 (465-712)
Scenario II: Change inter-screening interval				
1-yearly	1912	1362	28.71 (23.68-33.75)	552 (433-671)
2-yearly	1209	1514	20.63 (15.10-26.17)	780 (523-1036)
3-yearly	1914	1595	16.65 (11.39-21.91)	972 (583-1361)
Scenario III: sensitivity of early and advanced CRC				
(Early	**Late)**			
60%	80%	1923	1517	21.08 (15.59-26.57)	756 (522-989)
65%	80%	1921	1498	21.97 (16.80-27.14)	723 (522-989)
70%	85%	1934	1485	23.17 (18.19-28.15)	679 (520-857)
75%	85%	1936	1466	24.28 (19.18-29.38)	647 (484-810)
75%	90%	1948	1473	24.35 (19.28-29.42)	641 (480-802)
80%	95%	1967	1456	25.93 (20.79-31.08)	595 (454-737)
80%	100%	1978	1458	26.27 (21.43-31.11)	584 (450-717)

doi:10.6342/NTU201901194

Figure 5.1 The Bayesian directed acyclic graph (DAG) model for the five-state model with measurement errors



Annotation:

(a) N_c, N_F, N_{S0}, N_{S1}, N_{S2}: total number of subjects in the control group (refuser), attending first screen, and in State 0-2 from previous time point among those attending subsequent screens.

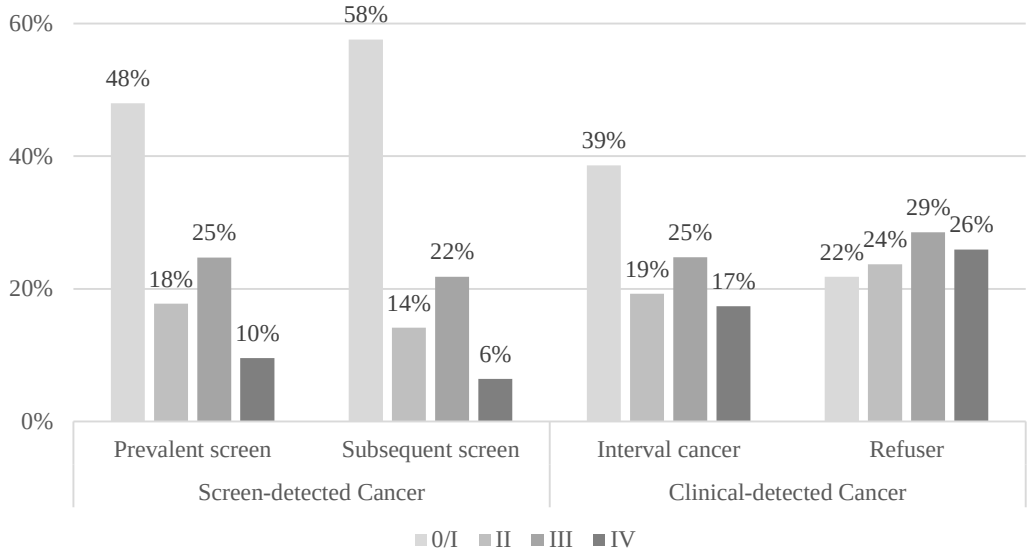
(b) $X_c[m]$, $X_F[m]$, $X_{S0}[m]$, $X_{S1}[m]$, $X_{S2}[m]$: number of subjects in state *m* in the control group (refuser), attending first screen, and in State 0-2 from previous time point among those attending subsequent screens.

(c) λ_i: transition rates

(d) Sen_1 and Sen_2: sensitivity of FIT for early and late CRC

doi:10.6342/NTU201901194

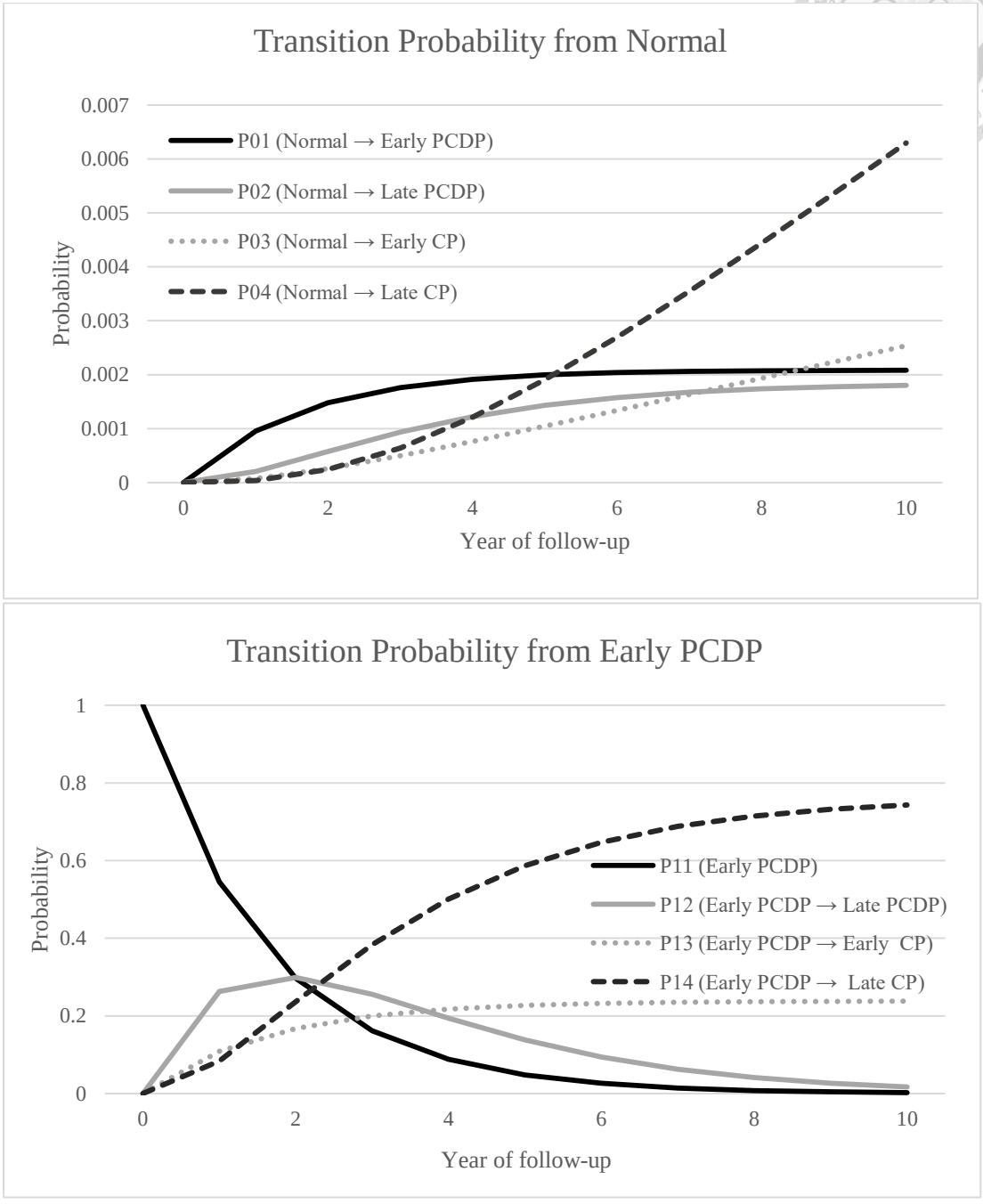
Figure 5.2 Distribution of AJCC tumour staging of CRC



	Screen-detected Cancer		Clinical-detected Cancer	
	Prevalent	Subsequent	Interval cancer	Refuser
0/I	3899	2439	4111	8539
II	1445	599	2048	9286
III	2010	925	2635	11165
IV	777	272	1851	10152
NA	1037	418	1259	6676
Overall	9168	4653	11904	45818

doi:10.6342/NTU201901194

Figure 5.3 Dynamic curves of early and late CRC progressions



doi:10.6342/NTU201901194

Figure 5.4 Power function of asymptotic analysis on the effectiveness of advanced cancer reduction by inter-screening interval

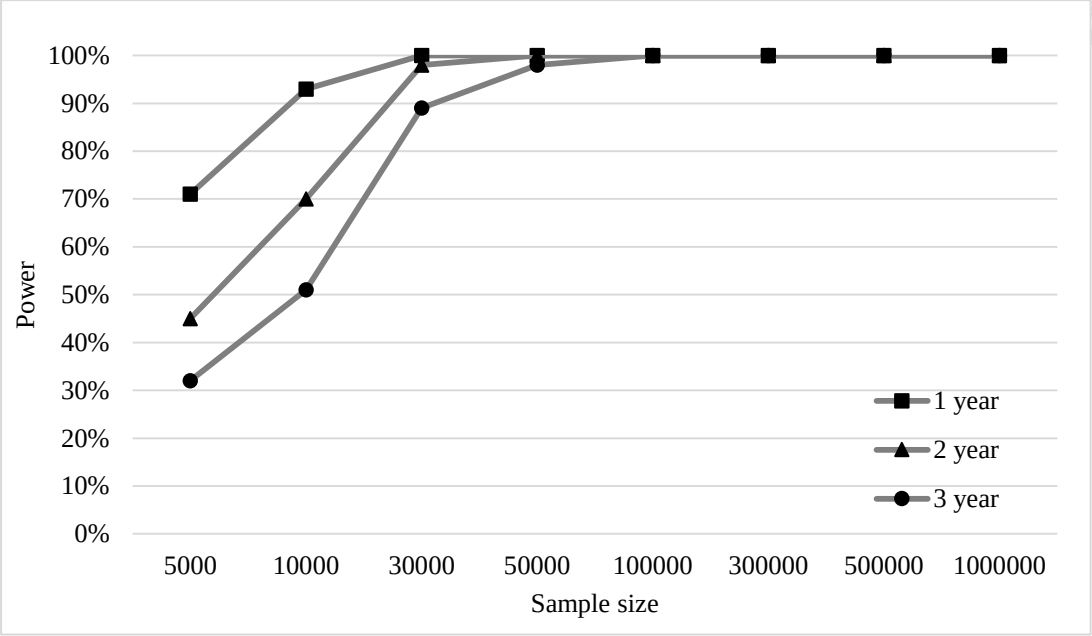
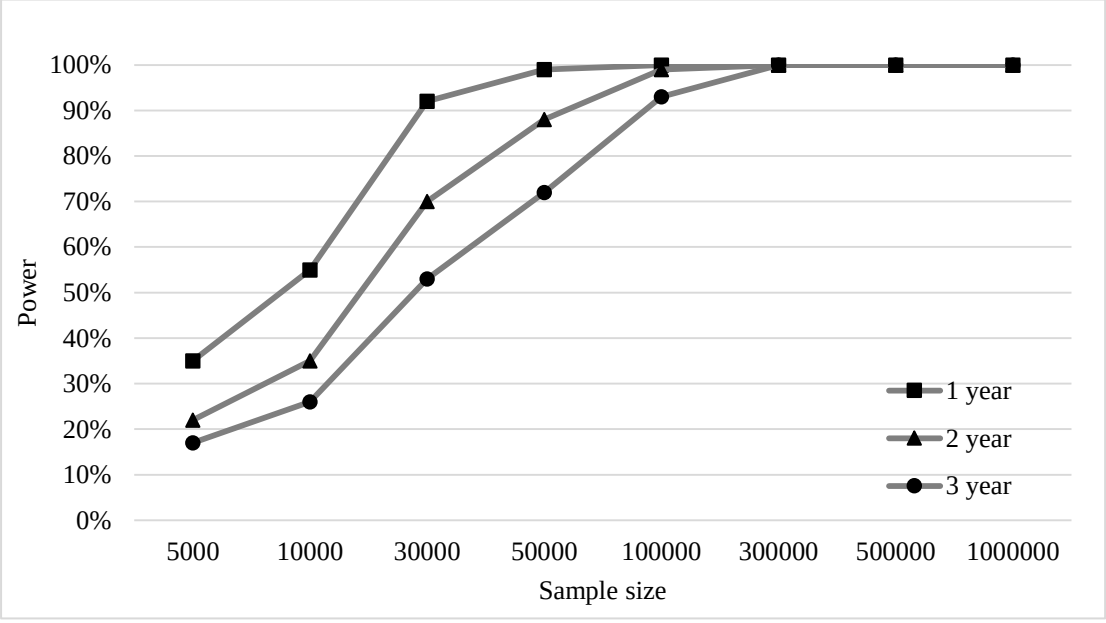


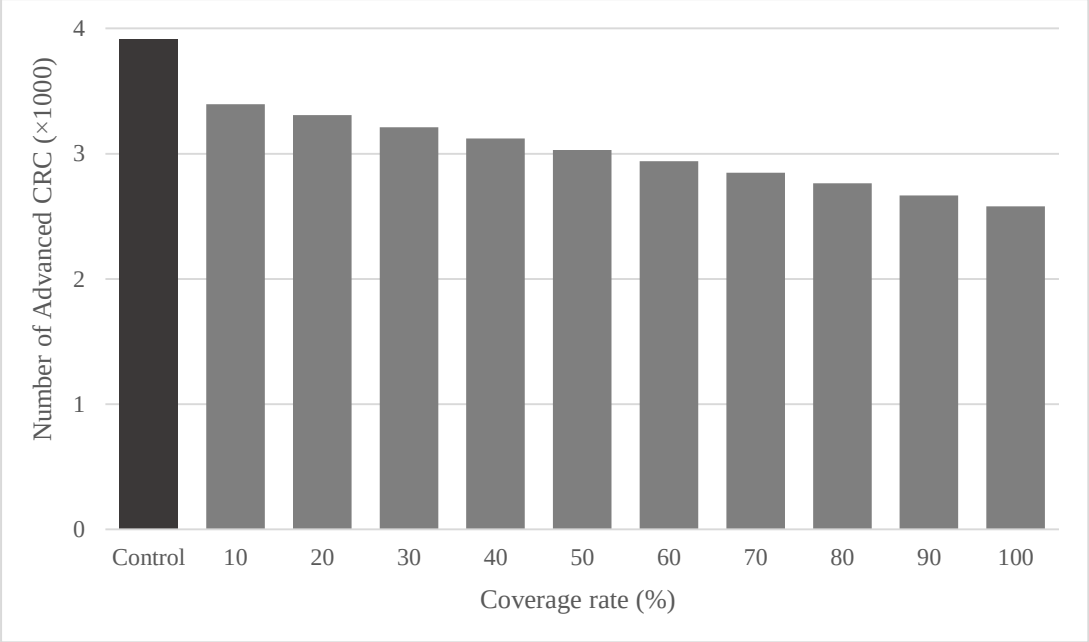
Figure 5.5 Power function of asymptotic analysis on the effectiveness of CRC mortality reduction by inter-screening interval



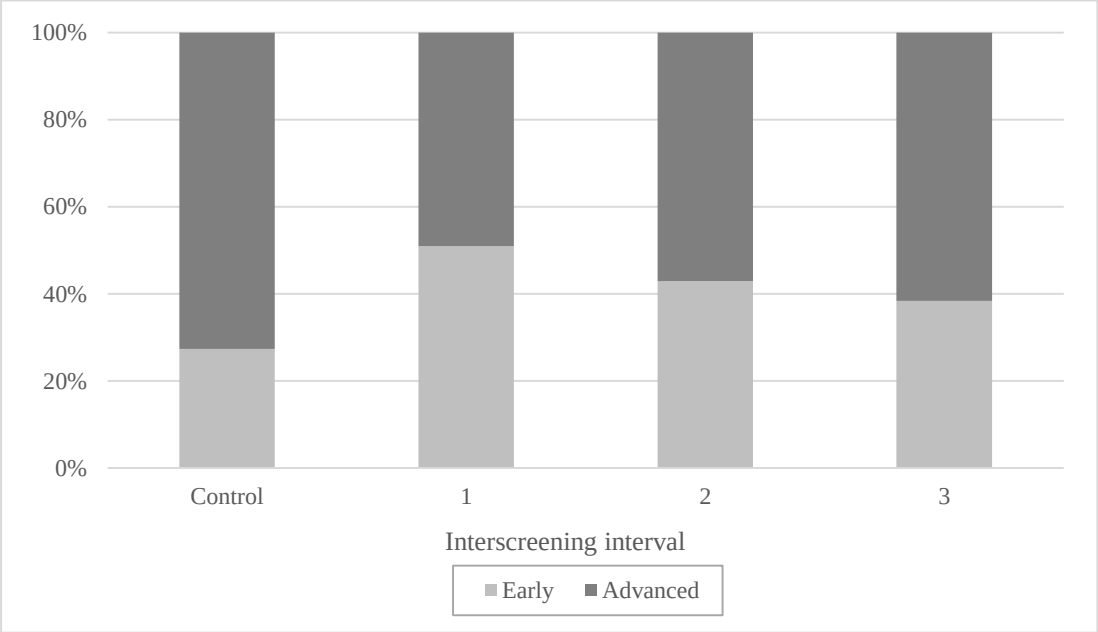
doi:10.6342/NTU201901194

Figure 5.6 Number and proportion of advanced cancer by different scenarios

(a) Change coverage rates



(b) Change inter-screening intervals



doi:10.6342/NTU201901194

Chapter 6 Statistical Indicators for Up-staging of CRC Assessed by Coxian Phase-type Markov Process with Device Stage Method

Summary

Introduction Despite the fact that non-Markov process has been developed for modelling the disease natural history of colorectal cancer with three phenotypes (free of colorectal cancer (CRC) (state 1), pre-clinical detectable phase (PCDP) (state 2), and clinical phase (CP)) (state 3) by tumour stage, how to assess the force of upstaging before surfacing to CP has been never quantified.

Aims This chapter is to develop statistical indicators underpinning with Coxian phase-type Markov process as proposed in Chapter 5 for evaluation of the rate of upstaging between transient states within PCDP relative to that of downstaging to absorbing state of CP

Statistical methods The indicator was developed with the relative rate between upstaging within the latent state of PCDP (λ_{2m}, m=1,…..,k sub-state of state 2) and surfacing to CP with downstaging stage ((λ_{3m}, m=1,…..,k), sub-state of state 3) to assess whether the poorer performance of proximal CRC as compared with distal CRC was due to poor sensitivity or high potential of upstaging ad to quantify the force of upstaging by location. The other indicator characterized by multi-nominal odds ratios was also proposed to assess how the effect of inter-screening interval on the reduction

doi:10.6342/NTU201901194

of advanced CRC varies with the location (proximal and distal site) of occurrence of CRC.

Analytical formula and simulation with stochastic integration were developed to prove the use of relative rate ($RR=\lambda_{2m}/\lambda_{3m}$) for assessing the force of upstaging as opposed to downstaging. A series of simulation studies with perturbation analysis by the order of λ_{2m} and λ_{3m} were conducted to assess the influence of three scenarios ($\lambda_{2m} > \lambda_{3m}$, $\lambda_{2m} < \lambda_{3m}$, and $\lambda_{2m} = \lambda_{3m}$) associated with the effect of inter-screening interval the reduction of advanced CRC.

Applications These indicators were applied to five-state (k=2) Coxian phase-type Markov process to evaluating the force of upstaging from early CRC to late CRC within PCDP by the location (proximal and distal site) of CRC using data on Taiwanese national colorectal cancer screening program.

Results Estimated results show the incidence of distal CRC is three-fold of the proximal CRC, but all of transition rates after entering PCDP for proximal CRC is faster than distal CRC. The proximal CRC (RR=4.46) had statistically significantly higher potential of upstaging from early to late CRC in the PCDP compared with the distal one (RR=2.87) but the estimates of the sensitivity of FIT test in early and later stage of CRC were identical between the proximal CRC and the distal one. The effect of shorting inter-screening interval from triennial to annual regime on the reduction of advanced

doi:10.6342/NTU201901194

CRC was therefore more remarkable in the proximal site in comparison with the distal one. In addition, the higher folds of λ_{2m} / λ_{3m}, the greater effectiveness of advanced CRC reduction as well.

Conclusion The proposed statistical indicators under the context of Coxian phase-type Markov process were not only theoretically sound but had high potential of practical use in quantifying the force of upstaging within latent PCDP. The application of these indicators to evaluation of upstaging in population-based organized service screening provides a new insight into a higher potential of upstaging for proximal CRCs compared with distal CRCs, which also accounts for the performance of FIT associated with proximal CRC attributed to upstaging rather than poor sensitivity.

doi:10.6342/NTU201901194

6.1 Introduction

While the stage device method underpinning Coxian phase-type Markov process is proposed to turn non-Markovian process into Markov process statistical application to elucidating the disease progression, particularly latent state transition, has been not well studies yet. In the real world, there are many classifications of states defined by various phenotypes of disease from different aspects. These include symptoms and signs of disease from patients' viewpoint, clinical attributes diagnosed from medical devices or physician, psychosocial contextual factors derived from social environments, pathological diagnosis of stage and class of disease with or without aided by molecular markers originated from laboratory science. In the domain of disease screening for instance, the administration of new screening method has identified a series of asymptomatic states in parallel with symptomatic states defined by pathological stages.

In the previous literatures, there are numerous mathematical modelling approaches that have been developed for quantifying the disease progression but few studies have been devoted to the development of statistical indicators for assessing the relative force of transition to severe disease status (up-staging) with different aspects. If the disease can be defined by three main classes according to patient symptoms and signs, apparently healthy, asymptomatic, and symptomatic, how can information derived from other aspects be incorporated into three main classes for assessing the

doi:10.6342/NTU201901194

force of transition to different pathways?

In the context of Coxian phase-type Markov process applied to the progression of colorectal cancer (CRC), asymptomatic state is defined as pre-clinical detectable phase (PCDP) and symptomatic state is defined as clinical phase (CP). Pathological staging according to AJCC system can be factored into PCDP and CP. The transient states between the stages of disease within the PCDP can be modelled by phase-type process and the clinical phase become absorbing state characterized by the corresponding pathological stage.

It is therefore interesting to assess whether the force of the transition to adjacent state with severe stage (up-staging) within PCDP is greater than that of the absorbing into CP with the same stage. For instance, for two transient states, early and late CRC, and two corresponding absorbing states it is very interesting to ask whether the transition rate form early PCDP to late PCDP is greater than the transition from early PCDP to early CP. There is lacking of a statistical indicator for evaluating such a relative contribution particularly when there are numerous states in the PCDP. It should be noted that the sensitivity of the screening test may further complicate such an assessment.

The usefulness of this indicator is two-fold. It is useful for evaluating the difference between the proximal CRC and the distal CRC with adjustment for the sensitivity of the screening test. The indicator can also aid in the evaluation of the effect of inter-

doi:10.6342/NTU201901194

screening interval on advanced cancer when the location of CRC is taken into account.

The specific aim of this study is to develop a useful statistical indicator underpinning Coxian phase-type Markov process on which we are based to evaluate the relative force of the transition to severe PCDP in comparison with the transition to the absorbing state of early CRC. The impact of this relative force on inter-screening interval can be assessed.

6.2 Data

Data used in this chapter is the same as those before. The chapter only adds information on CRC by location. The frequencies, demographic characteristics and tumour attributes of distal and proximal cancers are listed in Table 6.1. The higher male proportion (62%) was observed in distal colon compared with proximal colon (51%). There were 25602 CRCs detected in this screening program. 45502 CRCs were classified as cancers from refuser. Of these, there were 53179 cases (75%) of distal cancer and 17925 cases (25%) of proximal cancer. There were 18914 cases (31%) of carcinoma in situ and stage I (35.0%), 13260 stage II (21%), 16649 stage III (27%), and 13007 stage IV (12%).

doi:10.6342/NTU201901194

6.3 Methodology

6.3.1 Model specification

Figure 6.1 is a five-state Coxian phase-type Markov model that includes free of CRC (state 1), early and late PCDP (state 22 and state 23), and two corresponding states in the CP making allowance for the sensitivity of FIT test captured by false negative cases. The main focus of this chapter here is to compare relative rate of λ_{21} and λ_{31} in order to assess the force of upstaging within latent PCDP as indicated below. Information on screen-detected CRC and the counterpart of those free of CRC was used to capture the transition of upstaging from early to late CRC within the transient state of PCDP. Information accrued from interval cancers and refuser was used to capture downstaging to CP taking into account false negative CRC.

6.3.2 State-specific regression model and likelihood function

Coxian phase-type Markov regression model was developed to model each transition rate as a function of age, gender, and location of CRC (proximal and distal site). That means those are treated as state-specific covariates incorporated into the transition rate in proportional hazard regression form as follows.

$$\lambda_{ij} = \lambda_{0,ij} \exp(\boldsymbol{\beta X}), \mathrm{i,j} = 1, \ldots, \mathrm{m} \qquad (6.1)$$

Note that these covariates are also modelled as a function of stage-specific sensitivity

doi:10.6342/NTU201901194

by using logistic regression model.

The likelihood functions based on the Coxian phase-type Markov regression model were formed in a similar manner as developed in Chapter 5.

6.3.3 Statistical Indicator for quantifying latent potential of progression to advanced CRC

When the proposed Coxian phase-type Markov process with (K+1) state is applied to K phase-type in the PCDP (state 21,…..2K) and one absorbing CP but with different states (31,…..,3K) defined by k state of tumour stage, the transition rate of state 2m to the absorbing CP with state 3m denoted by λ_{3m} always competes with that of state 2m to the adjacent progressive 2m+1 state denoted by λ_{2m}. There is a premise that stage 1 to k has ordinal property. It means the higher the value the severe tumour stage is represented. To measure the potential of up-staging, following the incidence rate of entering first state of PCDP denoted by state 21, the relative rate of up-staging between $\lambda_{2m}/\lambda_{3m}$ and has the following property

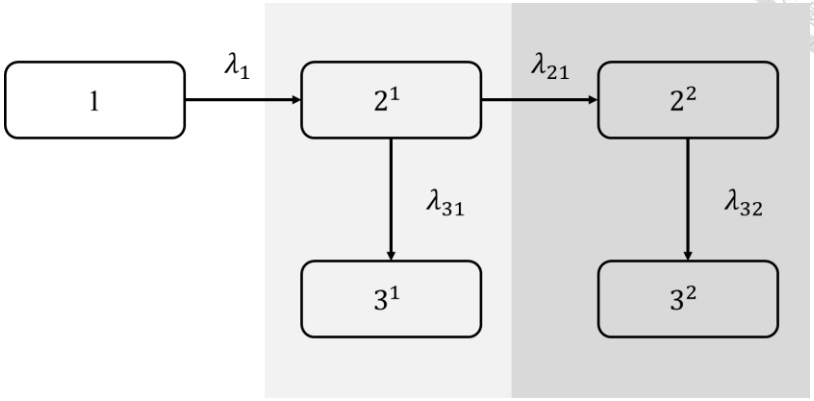
$$\lambda_{2m}/\lambda_{3m} = P_{2m,\,2m+1}(t) + \cdots\cdots + P_{2m,\,2K}(t)/P_{2m,3m}(t), \qquad (6.2)$$

where m=1,2…….,k.

doi:10.6342/NTU201901194

Proof:

When m=2

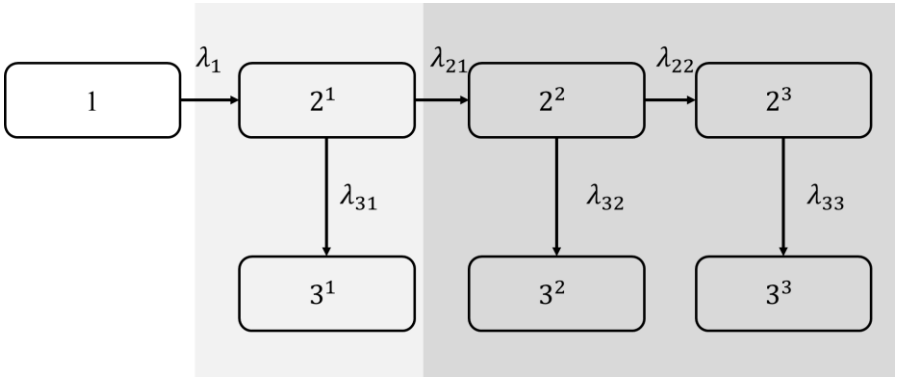


$$P_{21,22}(t) + P_{21,32}(t) \Big/ P_{21,31}(t)$$

$$= \frac{\int_0^t \lambda_{21} \cdot e^{-(\lambda_{21}+\lambda_{31})t_1} \cdot e^{-\lambda_{32}(t-t_1)} dt_1 + \int_0^t \lambda_{21} \cdot e^{-(\lambda_{21}+\lambda_{31})t_1} \cdot \left(1 - e^{-\lambda_{32}(t-t_1)}\right) dt_1}{\int_0^t \lambda_{31} \cdot e^{-(\lambda_{21}+\lambda_{31})t_1} dt_1}$$

$$= \frac{\int_0^t \lambda_{21} \cdot e^{-(\lambda_{21}+\lambda_{31})t_1} dt_1}{\int_0^t \lambda_{31} \cdot e^{-(\lambda_{21}+\lambda_{31})t_1} dt_1} = \frac{\lambda_{21}}{\lambda_{31}} \tag{6.3}$$

When m=3



$$P_{21,22}(t) + P_{21,32}(t) + P_{21,23}(t) + P_{21,33}(t) \Big/ P_{21,31}(t)$$

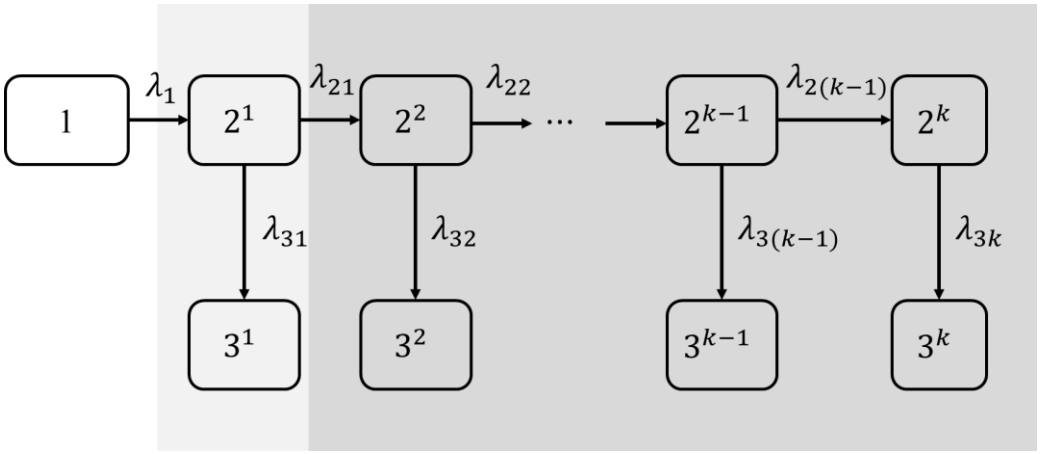
doi:10.6342/NTU201901194

$$= \frac{\left\{\begin{array}{c} \int_0^t \lambda_{21} \cdot e^{-(\lambda_{21}+\lambda_{31})\cdot t_1} \cdot e^{-(\lambda_{22}+\lambda_{32})\cdot(t-t_1)} dt_1 \\ + \int_0^t \lambda_{21} \cdot e^{-(\lambda_{21}+\lambda_{31})\cdot t_1} \cdot \int_0^{t-t_1} \lambda_{32} \cdot e^{-(\lambda_{32}+\lambda_{22})\cdot t_2} dt_2 dt_1 \\ + \int_0^t \lambda_{21} \cdot e^{-(\lambda_{21}+\lambda_{31})\cdot t_1} \cdot \int_0^{t-t_1} \lambda_{22} \cdot e^{-(\lambda_{32}+\lambda_{22})\cdot t_2} \cdot e^{-\lambda_{33}\cdot(t-t_1-t_2)} dt_2 dt_1 \\ + \int_0^t \lambda_{21} \cdot e^{-(\lambda_{21}+\lambda_{31})\cdot t_1} \cdot \int_0^{t-t_1} \lambda_{22} \cdot e^{-(\lambda_{23}+\lambda_{32})\cdot t_2} \cdot \left(1-e^{-\lambda_{33}\cdot(t-t_1-t_2)}\right) dt_2 dt_1 \end{array}\right\}}{\int_0^t \lambda_{31} \cdot e^{-(\lambda_{21}+\lambda_{31})\cdot t_1} dt_1}$$

$$= \frac{\left\{\begin{array}{c} \int_0^t \lambda_{21} \cdot e^{-(\lambda_{21}+\lambda_{31})\cdot t_1} \cdot e^{-(\lambda_{22}+\lambda_{32})\cdot(t-t_1)} dt_1 \\ + \int_0^t \lambda_{21} \cdot e^{-(\lambda_{21}+\lambda_{31})\cdot t_1} \cdot \int_0^{t-t_1} \lambda_{32} \cdot e^{-(\lambda_{32}+\lambda_{22})\cdot t_2} dt_2 dt_1 \\ + \int_0^t \lambda_{21} \cdot e^{-(\lambda_{21}+\lambda_{31})\cdot t_1} \cdot \int_0^{t-t_1} \lambda_{22} \cdot e^{-(\lambda_{32}+\lambda_{22})\cdot t_2} dt_2 dt_1 \end{array}\right\}}{\int_0^t \lambda_{31} \cdot e^{-(\lambda_{21}+\lambda_{31})\cdot t_1} dt_1}$$

$$= \frac{\int_0^t \lambda_{21} \cdot e^{-(\lambda_{21}+\lambda_{31})t_1} dt_1}{\int_0^t \lambda_{31} \cdot e^{-(\lambda_{21}+\lambda_{31})t_1} dt_1} = \frac{\lambda_{21}}{\lambda_{31}} \tag{6.4}$$

When m=k



$$\frac{P_{2_1 2_2}(t)+P_{2_1 2_3}(t)+\cdots+P_{2_1 2_k}(t)+P_{2_1 3_k}(t)}{P_{2_1 3_1}(t)+P_{2_1 3_2}(t)+\cdots+P_{2_1 3_{k-1}}(t)} \tag{6.5}$$

where $P_{2_1 2_2}(t) + P_{2_1 2_3}(t) + \cdots + P_{2_1 2_k}(t) + P_{2_1 3_k}(t)$

$$= \int_0^t \ldots \int_0^{t-t_1-\cdots-t_{k-1}} \lambda_{2_{(k-1)}} e^{-\left(\lambda_{2_{(k-1)}}+\lambda_{3_{(k-1)}}\right)t_{k-1}} e^{-\lambda_{3_k}(t-t_1-t_{k-1})} dt_1 \ldots dt_{k-1}$$

doi:10.6342/NTU201901194

$$+\int_0^t \cdots \int_0^{t-t_1-t_{k-2}} \lambda_{2_{(k-1)}}\, e^{-\left(\lambda_{2_{(k-1)}}+\lambda_{3_{(k-1)}}\right)t_{k-1}}(1$$

$$- e^{-\lambda_{3_k}^{(t-t_1-t_{k-1})}})dt_1 \ldots dt_{k-1}$$

$$= \int_0^t \lambda_{2_1}\, e^{-(\lambda_{2_1}+\lambda_{3_1})t_1} e^{-(\lambda_{2_2}+\lambda_{3_2})(t-t_1)} dt_1$$

$$+ \int_0^t \lambda_{2_1}\, e^{-(\lambda_{2_1}+\lambda_{3_1})t_1}\left(1 - e^{-(\lambda_{2_2}+\lambda_{3_2})(t-t_1)}\right)dt_1$$

$$+ \cdots = \int_0^t \lambda_{2_1} e^{-(\lambda_{2_1}+\lambda_{3_1})t_1}\, e^{-(\lambda_{2_2}+\lambda_{3_2})(t-t_1)} dt_1 +$$

$$\int_0^t \lambda_{2_1} e^{-(\lambda_{2_1}+\lambda_{3_1})t_1} \int_0^{t-t_1} \lambda_{2_2} e^{-(\lambda_{2_2}+\lambda_{3_2})t_2}\, e^{-(\lambda_{2_3}+\lambda_{3_3})(t-t_1-t_2)} dt_2 + \cdots +$$

$$\int_0^t \cdots \int_0^{t-t_1-t_{k-2}} \lambda_{2_{(k-1)}}\, e^{-\left(\lambda_{2_{(k-1)}}+\lambda_{3_{(k-1)}}\right)t_{k-1}} dt_{k-1} \qquad (6.6)$$

It is hypothesized that when λ_{2m} is greater than λ_{3m}, tumour has a stronger tendency to deteriorate into severe stage in the absence of intervention such as screening and otherwise down-staging is noted.

This indicator is used for assessing whether the potential of progression varies with the location of CRC. It is postulated that the potential of progression for CRCs occurring from the proximal site is higher than those originated from the distal site. However, this indicator may be confounded by the sensitivity of screening method that is used for early detection of CRC because the detectability of each screening tool may be also subject to the location of CRC. The proposed Coxian phase-type Markov process in Chapter 3 has taken into account the sensitivity of measurement error.

doi:10.6342/NTU201901194

6.3.4 Evaluation of the efficacy of screening policy by the location of CRC

To evaluate whether and how the efficacy of reducing advanced CRC as a result of the change in inter-screening interval varies with the location of CRC. Another indicator is proposed

$$OR=[P_{2m,\,2m+1}(t)+ \ldots\ldots\ldots+P_{2m,\,2K}(t)]/[P_{2m,2m}(t)+ P_{2m,3m}(t)]=$$

$$=[(\Omega^{P}_{2m+1}+\ldots\ldots\ldots\ldots \Omega^{P}_{2k})/(1+ \Omega^{P}_{3m})]/[(\Omega^{D}_{2m+1}+\ldots\ldots\ldots\ldots \Omega^{D}_{2k})/(1+ \Omega^{D}_{3m})] \quad (6.7)$$

6.4 Estimated results

6.4.1 Transition rates by location of CRC

Figure 6.2 shows the pre-clinical incidence rates of CRC by location of CRC. The incidence rates per 100,000 of CRC for distal and proximal colon were estimated as 95 and 33, respectively, with the approximate ratio of 3:1. To test the potential of progression to up-staging CRC, two sets of transition parameters by location of CRC are compared. The higher transition rate (0.6591) from early PCDP to early PCDP was observed for proximal cancer compared with distal cancer (0.4085). Proximal cancer had also faster transition rate (0.5505) than distal cancer (0.5085). However, compared with other transition parameters, the transition rates from early PCDP to early CP (around 0.14) were lower and identical for both of distal and proximal cancer. The

doi:10.6342/NTU201901194

estimated results are also showed in Table 6.2.

For detecting the early pre-clinical detected CRC, 0.60 of sensitivity was estimated for both of distal and proximal cancer whereas the sensitivity for distal and proximal cancer were 0.79 and 0.75, respectively, for identifying the late pre-clinical detected CRC. There is not much difference in the test sensitivity estimates of early and late CRC between proximal and distal CRCs.

The findings of transition rates of up-staging within PCDP and from PCDP to CP with the same stage together with the test sensitivity suggest that the later tumour stage found in proximal CRC as opposed to that in distal CRC is largely due to a high potential of up-staging before surfacing to CP rather than the poor test sensitivities.

6.4.2 Gender-age-specific transition by location of CRC

For distal colon, the incidence rates of entering PCDP for men increased with advancing age given the finding that aged 50-54, 55-59, 60-64 and 65-69 were estimated as 78, 111, 159 and 226 per 100,000, respectively (Table 6.3). Those incidence rates for each age group were lower in women than in men. It is very interesting to note that the annual transition rates from early PCDP to early CP, from late PCDP to late CP, and from late PCDP to late CP not only increased with advancing age but also faster in women than in men. The higher sensitivity for detection of both

doi:10.6342/NTU201901194

early and late pre-clinical detected cancer for each age group were found in women compared with men. The better sensitivity for cancer detection was observed in late PCDP compared with detection on early PCDP. Similar findings were found in proximal colon. However, again, the higher transition rates for each age group from early PCDP to late PCDP and from late PCDP to late CP as found in the overall group was observed for proximal cancer compared with distal cancer.

Table 6.4 shows the potential of up-staging according to the equation (6.3) indicated in the methodological section. The relative rate of up-staging was 3.19 (95% CI:3.13-3.26) for the overall CRC, 2.87 (95% CI:2.80-2.93) for the distal CRC, and 4.46 (95% CI:4.27-4.65) for the proximal CRC, which indicates the proximal CRC had higher potential of up-staging than the distal CRC.

Table 6.5 shows the corresponding findings by gender and age groups. The similar finding was still noted that the proximal CRC was more likely to have up-staging than the distal one. The additional interesting findings are that a higher potential of up-staging of CRC was found in females and young age group.

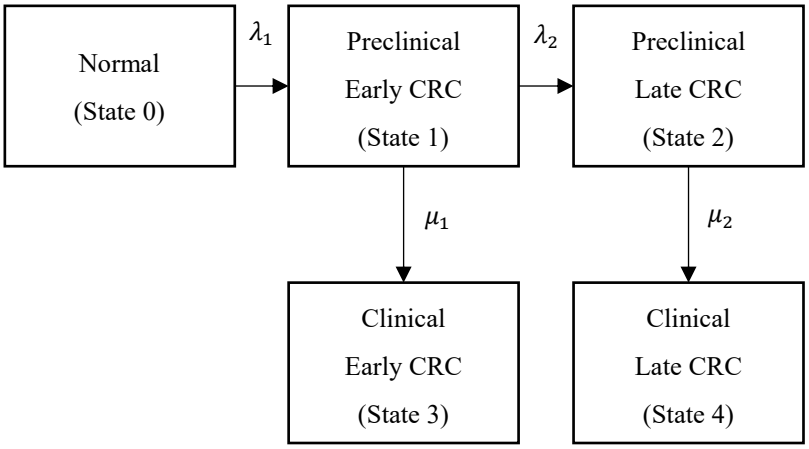
Table 6.6 shows the effect of inter-screening interval on the proportion of advanced CRC by the location of CRC. Generally speaking, the proportion of advanced CRC increased with the length of inter-screening interval. The difference in the proportion of advanced CRC between proximal site and distal site widened when inter-screening

doi:10.6342/NTU201901194

interval was lengthened. The similar findings by age and gender were noted (Table 6.7). Using the micro-simulation method yielded the similar finding as shown in Table 6.8 compared with those shown in Table 6.6

6.5 Simulation Study

In this chapter, we used a simulation study to study how the ratio of transition in the phase-type distribution in PCDP to that entering into absorbing state affects the effectiveness of FIT screening. Taking a five-state model as an example (see below), will be the ratio of transition rate travelling in the PCDP (λ_2) to that entering into CP (μ_1) for early PCDP larger than 1?



To do so, our simulation has to consider the correlation between the estimated transition rates as the empirical data shown. In the simulation algorithm we developed in Chapter

doi:10.6342/NTU201901194

5, we independently generated three sojourn times for state 0, 1, and 2, but confined the total time to the screening schedule. The question is, will the estimates based on data simulated in this way yield correlated estimated results? Or a simulation considering correlated sojourn time, such as Copula algorithm, is the only way to create data for correlated estimated parameters?

Figure 6.3 shows the correlation matrix of estimates based on empirical data, simulation algorithm in Chapter 5, and Copular algorithm. With the empirical data, λ_{21} and λ_{31} are highly positively corrected (r=0.75), and λ_{32} is negatively associated with λ_{21} (r = -0.1871) and λ_{31} (r = -0.2137). Interestingly, although we simulated data with independent sojourn times, with confined total time the estimates of parameters were correlated, but with weaker correlation compared to that with empirical data. For example, correlation between λ_{21} and λ_{31} (r = 0.4693) became weaker than its counterpart with empirical data (r = 0.7437). So are correlation between λ_{32} and λ_{21}, and between λ_{32} and λ_{31}. Nevertheless, the positivity of correlation remains the same. When we used Copula algorithm to create the three distributions for sojourn time, the correlation between λ_{21} and λ_{31} became slightly stronger (r = 0.5924), but λ_{21} and λ_{31} would be over-estimated, whereas λ_{32} been underestimated.

By simulating the data, Tables 6.9-6.10 give the results two scenarios between λ_{21} and $\lambda_{31}$to quantify the benefit of screening. A larger benefit is noted for a higher rate of

doi:10.6342/NTU201901194

λ_{21} /$\lambda_{31}$when the screening is administered. Note that a smaller difference was noted for shortening the length of inter-screening interval given $\lambda_{21} < \lambda_{31}$ in contrast to the opposite $\lambda_{21} > \lambda_{31}$.

6.6 Discussion

6.6.1 Theoretical property

From the statistical viewpoint, the novel advance of this chapter is to develop statistical indicators for quantifying the force of upstaging within PCDP and its application to evaluation the potential of upstaging varying age, gender, and location of CRC. It is very interesting to note that the indicator of upstaging is simply modelled by relative rate of $\lambda_{2m}/\lambda_{3m}$ independent of follow-up time t as the model form after the transition entering the PCDP based on such a kind of Coxian phase-type model has a very good mathematical property because of symmetric structure between transient states within PCDP and absorbing states of CP reflecting the mirror of tumour staging between two opposite states. This has been proven by several mathematical algebra of stochastic integration of multiple integral from time t_1 to t_{k-1} with the general formula for the transient states from 1 to k as seen in the equation (6.5). The similar logic is also applied to another indicator with cumulative logistic regression form with various types of odds ratio as seen in the equation (6.7) for evaluation of the effect of inter-screening

doi:10.6342/NTU201901194

interval on the reduction of advanced CRC.

The former indicator is very useful for different sites of cancer screening by examining whether an intensive screening is required and also whether screening is justified. When there is a situation: $\lambda_{2m} > \lambda_{3m}$, the higher the relative rate, the higher the odds of applying intensive inter-screening interval. In contrast, the opposite ($\lambda_{2m} < \lambda_{3m}$) may render the administration of population-based screening unjustified.

The chapter also did various scenarios on the order of λ_{2m} and λ_{3m}. It can be easily seen that the effect of inter-screening interval on the reduction of advanced cancer is more remarkable in the scenario when $\lambda_{2m} > \lambda_{3m}$ than that in the scenario when $\lambda_{2m} < \lambda_{3m}$.

6.6.2 Statistical evaluation of the disease natural history of CRC by location

These two indicators were applied to evaluating the contrast of the force of upstaging between the proximal colon and the distal colon and assessing whether the impact of inter-screening on the reduction of advanced CRC varied with location. The results found there is strong justification for adopting population-based screening as $\lambda_{2m} > \lambda_{3m}$ and the proximal colon requires an intensive screening policy compared with the distal colon as relative rate of $\lambda_{2m}/\lambda_{3m}$ was higher in the proximal colon than the distal colon. When different inter-screening interval was applied the impact of

doi:10.6342/NTU201901194

shortening inter-screening interval in the proximal colon has a more remarkable effect than that in the distal colon.

6.6.3 Clinical applications

From the practical aspect, this is also the first study to elucidate the disease progression of distal and proximal colorectal cancer and its associated factors based on empirical longitudinal screening data. The major findings accrued from transition parameters may have significant implications for revealing different patterns of disease progression in distal and proximal colon. The incidence rate derived from annual transition rate from free of CRC to pre-clinical cancer for distal colon which is 3-fold higher than proximal colon. The higher incidence populations arise in the distal colon which is consistent with a previous epidemiological study (Haenszel et al., (1973)). Although the higher incidence of CRC indicates the more disease burden on distal colon, the transition rates from asymptomatic PCDP phase to symptomatic CP phase were lower in the distal colon compared with the proximal colon. The comparison between distal and proximal colon for the transition rates from asymptomatic phase to symptomatic phase may reveal that the distal colon is more likely to prevent. Due to the slower progression, subjects with negative finding from FIT are less likely to be missed as interval cancer. On the other hand, the progression rate from asymptomatic

doi:10.6342/NTU201901194

phase to symptomatic phase for the proximal colon is much faster. This reveals that proximal cancer may present at a more advanced stage as shown in the previous study (Slattery et al., (1996)). A review study demonstrated that interval CRCs are 2.4 times more likely to be proximal CRCs than the detected CRCs and were less likely to present at an advanced stage (Singh et al., (2014)). In UK FOB-based bowel cancer screening data, such a tendency to arise in the proximal colon was also observed (Blanks et al., (2019)). However, the estimated FIT sensitivity for late CRC did not show much different between proximal colon (75%) and distal colon (79%). Both of the sensitivity for early CRC in distal and proximal colon were also identical as 60%. Our finding suggests that the higher proportion of proximal colon in interval cancers is not because the sensitivity of FIT but that is the faster progression rate of proximal colon.

For the natural history of distal colon, the transition rates from early PCDP to early CP (0.4085) and from late PCDP to late CP (0.5080) are comparable with those from other studies addressing mass screening for CRC. The meta-analysis from two randomized control trial, the Nottingham and Funen studies, the transition rates also estimated from the same five-state Markov model of 0.2754 for pre-clinical Dukes' stage A or B to pre-clinical Dukes' stage C or D and 0.7627 for pre-clinical Dukes' stage C or D to clinical Dukes' stage C or D (Chiu et al., (2011)). This current estimate for early PCDP to early CP is also slightly higher than the estimate of 0.30 but lower than

doi:10.6342/NTU201901194

the estimate of 0.92 for late PCDP to late CP given for Taiwanese subjects with the pilot FIT screening (Yang et al., (2006)). The different results could be possible due to the different classification of early and late tumour stages. It should be noted that the transition rates for two directions are the opposite for the proximal cancer. The faster progression at early cancer phase but lower progression at late cancer phase were observed. This indicated that biennial FIT may not be sufficient to detect proximal cancer at early cancer phase.

We found that the estimated incidence rates increased by age and males have the higher estimated incidence rate than women. The basic epidemiological profiles from these findings are coherent with our current knowledge. However, the faster transition rates from early PCDP to early CP, from late PCDP to late CP, and from late PCDP to late CP were found in women. The higher sensitivity for detection of both early and late pre-clinical detected cancer for each age group were also found in women. The poorer accuracy in in detecting proximal serrated polyps from fecal immunochemical test particularly in women has been demonstrated in the previous studies (Larsen et al., (2010), Caldarella et al., (2013), Carot et al., (2018)). Again, the inconsistent finding can be explained as the faster progression rate of proximal colon rather than poor sensitivity in women.

This thesis also found the disease progression of CRC varied with age and gender.

doi:10.6342/NTU201901194

We found that the estimated incidence rates increased by age and males have the higher estimated incidence rate than women. The basic epidemiological profiles from these findings are close to our current knowledge. However, the faster transition rates from early PCDP to early CP, from late PCDP to late CP, and from late PCDP to late CP were found in women. The higher sensitivity for detection of both early and late pre-clinical detected cancer for each age group were also found in women. In addition to revealing differences between distal and proximal segments, these findings may also reveal the possibility of differences in gene expression patterns related to age, gender and ethnicity. Such findings could prove relevant not only to the risk of developing CRC but also to the type of cancer that develops. The further study on the mechanism for disease progression related to age and gender should be further conducted.

In conclusion, the multi-state disease progression in distal and proximal cancer was quantified in order to answer the questions for the decent screening interval with individualized characteristics and how the performance of screening program is. Results of transition parameters from models by looking at the two segments can be very useful to evaluate the effectiveness of colorectal cancer screening regime.

doi:10.6342/NTU201901194

Table 6.1 Distribution of selected characteristics associated with distal and proximal cancer

Characteristics	N	%	Distal		Proximal	
			N	%	N	%
Male	41843	59%	32738	62%	9105	51%
Female	29261	41%	20441	38%	8820	49%
Age at diagnosis						
50-54	6954	10%	5437	10%	1517	8%
55-59	13772	19%	10542	20%	3230	18%
60-64	14746	21%	11255	21%	3491	20%
65-69	15884	22%	11803	22%	4081	23%
>=70	19748	28%	14142	27%	5606	31%
Cancer Detection Mode						
Screen-detected at prevalent screen	9151	13%	7170	13%	1981	11%
Screen-detected at subsequent screen	4645	7%	3386	6%	1259	7%
Interval Cancer	11806	17%	8415	16%	3391	19%
Refuser	45502	64%	34208	64%	11294	63%
AJCC Stage						
Stage 0/I	18914	31%	15021	33%	3893	24%
Stage II	13260	21%	9119	20%	4141	26%
Stage III	16649	27%	12514	27%	4135	26%
Stage IV	13007	21%	9197	20%	3810	24%
Not known	9274		7328		1946	
Total	71104		53179	75%	17925	25%

doi:10.6342/NTU201901194

Table 6.2 Estimated results of five-state natural history of CRC by location of CRC

	Incidence rate	Transition rate			Sensitivity	
		Early PCDP → Late PCDP	Early PCDP → Early CP	Late PCDP → Late CP	Early PCDP	Late PCDP
Location of CRC						
Distal	0.00095 (0.00094,0.00096)	0.4085 (0.3818,0.4327)	0.1424 (0.1332,0.1515)	0.5080 (0.4836,0.5335)	0.6068 (0.5701,0.6425)	0.7913 (0.7597,0.8199)
Proximal	0.00033 (0.00032,0.00033)	0.6591 (0.5602,0.7864)	0.1476 (0.1244,0.1758)	0.5505 (0.4996,0.5995)	0.6048 (0.5016,0.7075)	0.7481 (0.6943,0.8064)
Overall	0.00128 (0.00127,0.00129)	0.4624 (0.4354,0.4901)	0.1448 (0.1364,0.1543)	0.5231 (0.5008,0.5462)	0.6068 (0.5717,0.6432)	0.7798 (0.7535,0.8066)

doi:10.6342/NTU201901194

Table 6.3 Estimated results of five-state model of CRC natural history by age, gender, and location of CRC

Gender/Age groups	Incidence rate	Transition rate			Sensitivity*	
		Early PCDP → Late PCDP	Early PCDP → Early CP	Late PCDP → Late CP	Early PCDP	Late PCDP
Distal						
Male						
50-54 yr	0.00078	0.2324	0.0797	0.3037		
55-59 yr	0.00111	0.3025	0.1079	0.3701		
60-64 yr	0.00159	0.3938	0.1459	0.4510		
65-69 yr	0.00226	0.5127	0.1973	0.5495	52.4% (50-54 yr) 56.0% (55-59 yr) 59.6% (60-64 yr) 63.0% (65-69 yr)	68.4% (50-54 yr) 69.5% (55-59 yr) 70.6% (60-64 yr) 71.6% (65-69 yr)
Female						
50-54 yr	0.00053	0.3347	0.1094	0.3736		
55-59 yr	0.00076	0.4357	0.1479	0.4553		
60-64 yr	0.00108	0.5673	0.2001	0.5548		
65-69 yr	0.00154	0.7385	0.2706	0.6761		
Proximal						
Male						
50-54 yr	0.00027	0.3839	0.0852	0.3344		
55-59 yr	0.00038	0.4998	0.1153	0.4075		
60-64 yr	0.00055	0.6506	0.1559	0.4966		
65-69 yr	0.00078	0.8470	0.2108	0.6052	53.0% (50-54 yr) 56.6% (55-59 yr) 60.1% (60-64 yr) 63.6% (65-69 yr)	64.6% (50-54 yr) 65.8% (55-59 yr) 66.9% (60-64 yr) 68.0% (65-69 yr)
Female						
50-54 yr	0.00018	0.5529	0.1169	0.4115		
55-59 yr	0.00026	0.7199	0.1581	0.5014		
60-64 yr	0.00037	0.9372	0.2139	0.6110		
65-69 yr	0.00053	1.2200	0.2892	0.7446		

*sensitivity: logistic regression model with age, and location of CRC

doi:10.6342/NTU201901194

Table 6.4 Relative rates (RRs) of up-staging by location of CRC

Location of CRC	RR of up-staging ($^{\lambda_{21}}/_{\lambda_{31}}$)	95% CI of RR ($^{\lambda_{21}}/_{\lambda_{31}}$)
Distal	2.87	(2.80, 2.93)
Proximal	4.46	(4.27, 4.65)
Overall	3.19	(3.13, 3.26)

$$\text{Var}\left(^{\lambda_{21}}/_{\lambda_{31}}\right) = Var(\lambda_{21}) + Var(\lambda_{31}) - 2Cov(\lambda_{21}, \lambda_{31})$$

$$95\%\text{CI} - 2.87 \pm 1.96 \times \sqrt{\text{Var}\left(^{\lambda_{21}}/_{\lambda_{31}}\right)}$$

** The 95% CI for the total population obtained from the analytic form was 3.17-3.21.

doi:10.6342/NTU201901194

Table 6.5 Sex-age-specific relative rates of up-staging by location of CRC

Gender/Age groups	RR of up-staging ($\lambda_{21}/\lambda_{31}$)	95% CI of RR ($\lambda_{21}/\lambda_{31}$)
Distal		
Male		
50-54	2.91	(2.82, 3.02)
55-59	2.80	(2.73, 2.88)
60-64	2.70	(2.62, 2.79)
65-69	2.59	(2.50, 2.69)
Female		
50-54	3.05	(2.93, 3.18)
55-59	2.95	(2.84, 3.05)
60-64	2.83	(2.75, 2.95)
65-69	2.73	(2.63, 2.85)
Proximal		
Male		
50-54	4.52	(4.29, 4.75)
55-59	4.35	(4.14, 4.55)
60-64	4.17	(3.97, 4.37)
65-69	4.01	(3.81, 4.23)
Female		
50-54	4.72	(4.50, 5.00)
55-59	4.55	(4.33, 4.77)
60-64	4.38	(4.17, 4.59)
65-69	4.21	(4.03, 4.46)

doi:10.6342/NTU201901194

Table 6.6 The odds ratio for up-staging of CRC in proximal compared to distal assessed by Coxian Phase-type Markov process using analytic form

Inter-screening interval	Advanced Distal CRC	Advanced Proximal CRC	OR
Annual	31.41%	45.24%	1.80
Biennial	49.51%	65.43%	1.93
Triennial	59.95%	74.44%	1.95

Table 6.7 The odds ratio for up-staging of CRC in proximal compared to distal by age, and sex assessed by Coxian Phase-type Markov process using analytic form

Inter-screening interval	Advanced Distal CRC	Advanced Proximal CRC	OR
Female aged 50			
Annual	27.03%	40.30%	1.82
Biennial	44.36%	60.92%	1.96
Triennial	55.48%	71.48%	2.01
Female aged 60			
Annual	39.61%	55.67%	1.91
Biennial	57.99%	73.27%	1.99
Triennial	66.53%	78.84%	1.87
Male aged 50			
Annual	19.96%	30.64%	1.77
Biennial	34.57%	49.81%	1.88
Triennial	45.27%	61.80%	1.96
Male aged 60			
Annual	30.43%	44.66%	1.84
Biennial	48.17%	64.59%	1.96
Triennial	58.51%	73.49%	1.97

doi:10.6342/NTU201901194

Table 6.8 The odds ratio for up-staging of CRC in proximal compared to distal assessed by Coxian Phase-type Markov process using microsimulation approach

Inter-screening interval	Distal			Proximal			**OR**
	Total CRC	Advanced CRC	%	Total CRC	Advanced CRC	%	
Annual	71058	22320	31.41	70840	31910	45.05	**1.79**
Biennial	71311	35394	49.63	70464	46116	65.45	**1.92**
Triennial	70805	42510	60.04	70909	52915	74.62	**1.96**

Distal CRC: λ_1=0.00128, λ_{21}=0.4085, λ_{31}=0.1424, λ_{32}=0.5080

Proximal CRC: λ_1=0.00128, λ_{21}=0.6591, λ_{31}=0.1476, λ_{32}=0.5505

doi:10.6342/NTU201901194

Table 6.9 The odds ratio for up-staging of CRC with 1-year and 2-year inter-screening interval compared to triennial screening assessed by Coxian Phase-type Markov process using analytic form

Inter-screening interval	Advanced CRC	OR
$\lambda_{21} > \lambda_{31}$		
λ_{21}=0.45, λ_{31}=0.15		
Triennial	62.60%	Reference
Biennial	52.41%	0.66
Annual	33.84%	0.31
λ_{21}=0.4, λ_{31}=0.2		
Triennial	55.65%	Reference
Biennial	46.59%	0.70
Annual	30.08%	0.34
λ_{21}=0.45, λ_{31}=0.3		
Triennial	53.68%	Reference
Biennial	46.61%	0.75
Annual	31.66%	0.40
$\lambda_{21} < \lambda_{31}$		
λ_{21}=0.15, λ_{31}=0.45		
Triennial	20.87%	Reference
Biennial	17.47%	0.80
Annual	11.28%	0.48
λ_{21}=0.2, λ_{31}=0.4		
Triennial	27.82%	Reference
Biennial	23.29%	0.79
Annual	15.04%	0.46
λ_{21}=0.3, λ_{31}=0.45		
Triennial	35.78%	Reference
Biennial	31.07%	0.81
Annual	21.11%	0.48
$\lambda_{21} = \lambda_{31}$		
λ_{21}=0.3, λ_{31}=0.3		
Triennial	41.74%	Reference
Biennial	34.94%	0.75
Annual	22.56%	0.41

Assuming λ_1=0.00128, and λ_{32}=0.5

doi:10.6342/NTU201901194

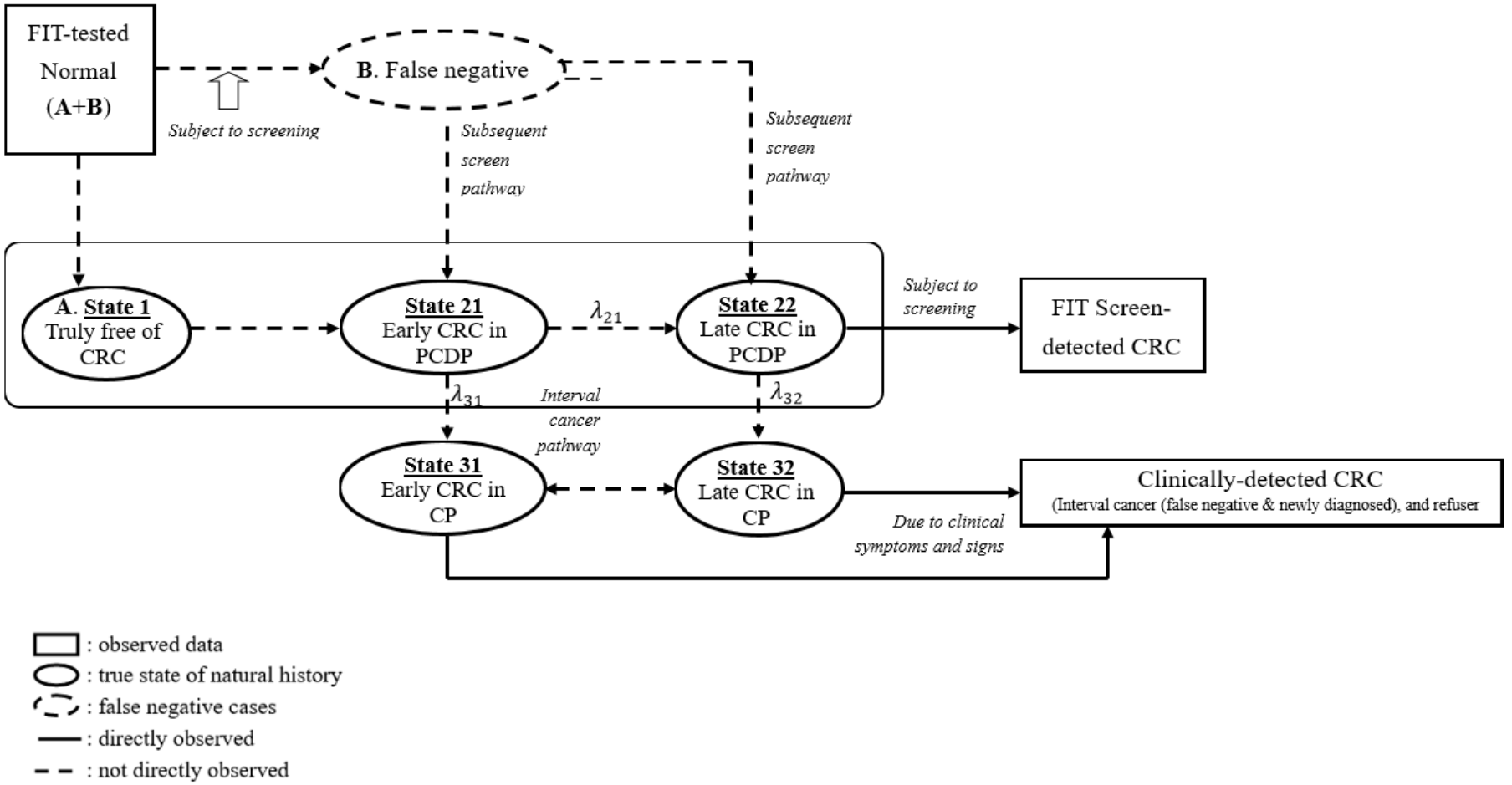
Table 6.10 The odds ratio for up-staging of CRC with 1-year and 2-year inter-screening interval compared to triennial screening assessed by Coxian Phase-type Markov process using microsimulation

Inter-screening interval	Total CRC	Advanced CRC	%	**OR**
$\lambda_{21} > \lambda_{31}$				
λ_{21}=0.45, λ_{31}=0.15				
Triennial	70820	44122	62.30%	Reference
Biennial	71011	37313	52.55%	0.67
Annual	70975	23988	33.80%	0.31
λ_{21}=0.4, λ_{31}=0.2				
Triennial	70207	39116	55.72%	Reference
Biennial	70216	32608	46.44%	0.69
Annual	70812	21293	30.07%	0.34
λ_{21}=0.45, λ_{31}=0.3				
Triennial	70988	38112	53.69%	Reference
Biennial	70564	32856	46.56%	0.75
Annual	70747	22725	32.12%	0.41
$\lambda_{21} < \lambda_{31}$				
λ_{21}=0.15, λ_{31}=0.45				
Triennial	70475	14854	21.08%	Reference
Biennial	71196	12510	17.57%	0.80
Annual	70916	8000	11.28%	0.48
λ_{21}=0.2, λ_{31}=0.4				
Triennial	70552	19598	27.78%	Reference
Biennial	70808	16428	23.20%	0.79
Annual	71162	10809	15.19%	0.47
λ_{21}=0.3, λ_{31}=0.45				
Triennial	70773	25210	35.62%	Reference
Biennial	70816	22101	31.21%	0.82
Annual	70617	14921	21.13%	0.48
$\lambda_{21} = \lambda_{31}$				
λ_{21}=0.3, λ_{31}=0.3				
Triennial	70717	29751	42.07%	Reference
Biennial	70441	24564	34.87%	0.74
Annual	70414	15707	22.31%	0.40

Assuming λ_1=0.00128, and λ_{32}=0.5

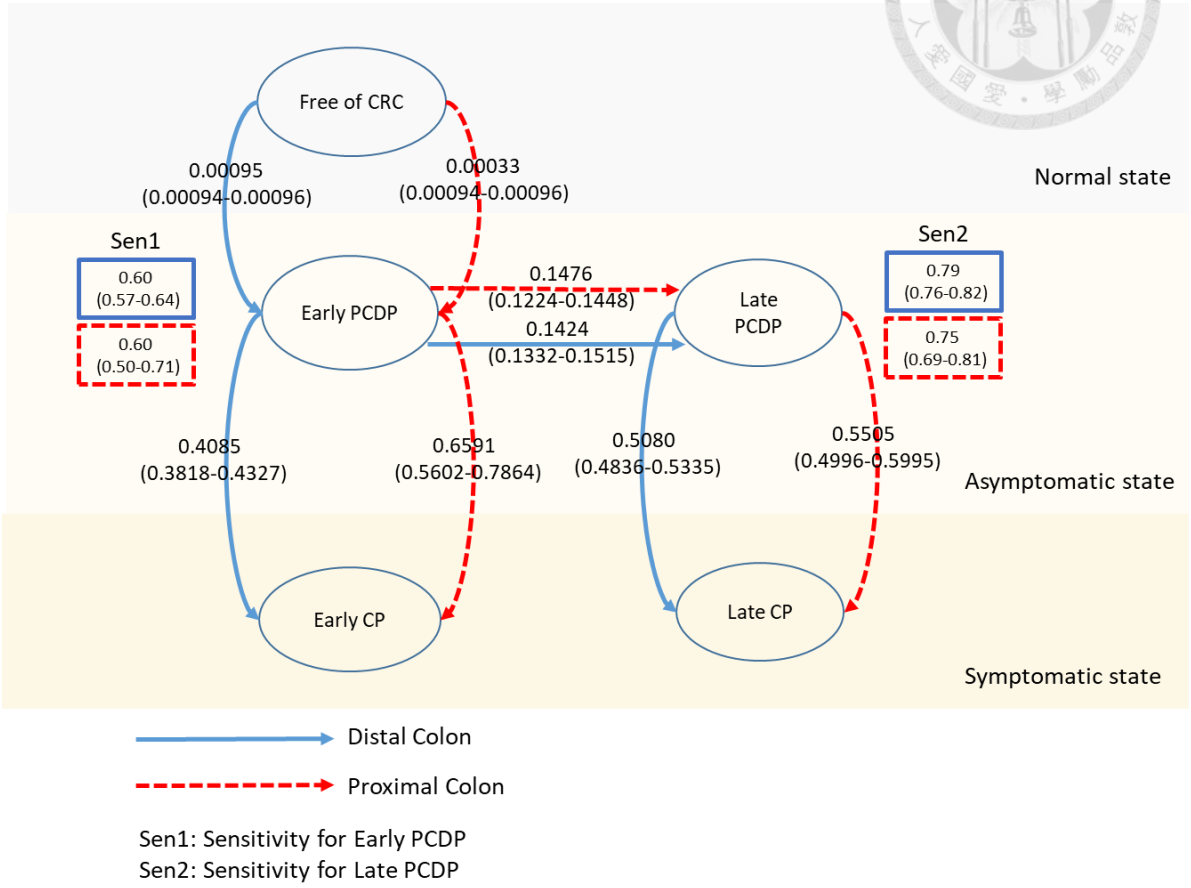
doi:10.6342/NTU201901194

Figure 6.1 Five-state Coxian phase-type Markov model with the corresponding detection model observed in screening



doi:10.6342/NTU201901194

Figure 6.2. Estimates of progression rates with five-state Markov model for distal and proximal cancer



doi:10.6342/NTU201901194

Figure 6.3 The correlation matrix of estimated parameters

(A) Estimation based on empirical data

	Estimate	Correlation matrix			
		λ_1	λ_{21}	λ_{31}	λ_{32}
λ_1	0.00131	1.0000	0.0601	0.0515	0.0714
λ_{21}	0.8917	0.0601	1.0000	0.7437	-0.1871
λ_{31}	0.2748	0.0515	0.7437	1.0000	-0.2137
λ_{32}	0.7485	0.0714	-0.1871	-0.2137	1.0000

(B) Estimation based on independent sojourn times with confined total time

	Estimate	Correlation matrix			
		λ_1	λ_{21}	λ_{31}	λ_{32}
λ_1	0.00129	1.0000	0.0736	0.0356	0.0714
λ_{21}	0.9531	0.0736	1.0000	0.4693	-0.0805
λ_{31}	0.2815	0.0356	0.4693	1.0000	-0.1280
λ_{32}	0.7936	0.0714	-0.0805	-0.1280	1.0000

(C) Estimation based on data from Copula algorithm

	Estimate	Correlation matrix			
		λ_1	λ_{21}	λ_{31}	λ_{32}
λ_1	0.00129	1.0000	0.0195	0.0687	0.0887
λ_{21}	1.1594	0.0195	1.0000	0.5924	-0.0830
λ_{31}	0.4854	0.0687	0.5924	1.0000	-0.1183
λ_{32}	0.6438	0.0887	-0.0830	-0.1183	1.0000

doi:10.6342/NTU201901194

Chapter 7 Consolidated Coxian phase-type Markov (CPHM) Process

Summary

Background Whether population-based organized service screening can bring the benefit of reducing advanced cancer and death from specific-cause, i.e. colorectal cancer (CRC) is determined by a constellation of factors.

Methods We constructed a consolidated Coxian phase-type Markov (CPHM) model to assess the effectiveness of population-based FIT screening in CRC mortality reduction. The consolidated CPHM model include three parts: (1) structure-process—from the arrival until the referral system to undergo colonoscopy, (2) disease natural history of multistate CRC for three major phenotypes, free of CRC, CRC in the PCDP, and CRC in the CP with hidden phases, and (3) prognostic of CRC survival. Designed variables and non-designed variables in each part of CPHM process were considered. A series of simulation was adopted with the perturbation of coverage rate for FIT screening, positive rate of FIT, referral rate of confirmatory colonoscopy among cases with positive FIT.

Results The effectiveness of CRC mortality reduction is affected by coverage rate and sensitivity. A FIT program with 10% coverage rate resulted in mortality reduction by 5.6% in contrast to 58.8% for 100% coverage rate. The corresponding numbers needed to screen were 3433 and 339, respectively, indicating more efficiency with high coverage rate. Higher sensitivity also yielded higher benefit. The consolidated CPHM investigate the influence on effectiveness of the population-based program, and found that coverage rate and referral rate for confirmatory colonoscopy had a large impact on the effectiveness of FIT program. Higher positive rate of FIT may correlated with a larger benefit, but with a less magnitude compared with coverage and referral rate.

doi:10.6342/NTU201901194

Conclusion We developed a consolidated CPHM to incorporate structure-process factors, disease natural history of multistate CRC, and prognostic of CRC survival for the systematical evaluation of population-based FIT screening in CRC mortality reduction.

doi:10.6342/NTU201901194

7.1 Introduction

7.1.1 Rationale for the development of consolidated CPHM process

Whether population-based organized service screening can bring the benefit of reducing advanced cancer and death from specific-cause, i.e. colorectal cancer is determined by a constellation of factors including structure and process of screening program, multistate disease natural history, and screening policy such as inter-screening interval, age to begin with screen, and age to terminate screening. It should be noted that the determinants in structure of screening pertains to the capacity of professional facility, cost, and health behavior of participants. Factors responsible for the process of screening include the detectability of the screening method, referral rate, and waiting time for confirmatory diagnosis. After considering the structure and process, the elucidation of disease natural history model as seen in the previous chapters by the use of CPHM plays an important role in the assessment of how the adoption of screening interrupts the disease natural history in the absence of screening and how its effect is affected by the test sensitivity. The disease natural history of CRC by location can also decipher what is the force of upstaging transition or poor sensitivity accounts for the poor performance of detecting proximal CRCs, including higher advanced cancer rate, interval cancer rate, and smaller reduction of CRC mortality.

7.1.2 Property of CPHM for population-based screening program

Based on the context and methods of previous chapters, each phase of population-based screening program like Taiwanese colorectal cancer screening with FIT test can be very well integrated as a unified framework. This is partially due to the adequacy of phase-type process that accommodate serial transitions of screening process as seen in Chapters 3-4 and serial disease progression of disease natural history as seen in Chapter

doi:10.6342/NTU201901194

5-6 and partially due to the competing form of transient states in parallel with absorbing states that are well adapted to biological plausibility, health behavior, and clinical interpretation. The additional property of CPHM is featured with a series of clear temporal relationship. However, to consolidate different parts of CPHM require a Bayesian consolidated analysis to connect each part as a whole.

7.2 Study framework of consolidated CPHM for modelling population-based screening

Figure 7.1 displays how three Coxian phase-type Markov (CPHM) processes can be consolidated as a unified framework for modelling structure, process, disease natural history of multistate CRC, and prognostic of CRC survival based on each CPHM described in the previous Chapters 4-5.

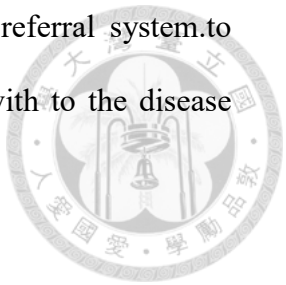
The first part of structure-process-based CPHM model is the modification of QH-CPH model that capture a complex screening program from the arrival until the referral system to undergo colonoscopy as proposed in Chapter 4. As described in Chapter 4, it is very interesting to note that Queue process for modelling the arrival rate of screenees and the rate of FIT positive can be integrated as one transient sate (FIT positive) and one absorbing state (no FIT test). The similar logic is also applied to non-compliance with colonoscopy (one absorbing) and compliance with colonoscopy with hidden phase of waiting time.

The second part is pertaining to CPHM process for multistate disease natural history of CRC that is capture by three main phenotypes, free of CRC, CRC in the PCDP, and CRC in the CP from the two latter of which hidden phases of PCDP are further derived by using tumour staging.

The third part is the prognosis of CRC. We first develop QH-CPH model to

doi:10.6342/NTU201901194

capture a complex screening program from the arrival until the referral system.to undergo colonoscopy Here, we then by linking QH-CPH model with to the disease natural history of CRC with Coxian-phase type model.



7.3 Designed and inherited non-designed variables

The variables in each part of CPHM process can be divided into two type, designed variables and non-designed variables. In the first part of the uptake of screening and positive results of FIT, the attendance rate is a non-designed variable that inherited from individual health behaviour. The positive FIT is the mixture of a design-based variable and a non-designed variable. The former is related to the cutoff of f-Hb that is dependent on the pre-determined cutoff and the latter is pertaining to individual susceptibility to CRC. Similarly, the compliance with colonoscopy and waiting time is also the mixture of a designed variable that is subject to the capacity of professional manpower and a non-designed variable that is determined by participant's health behavior. Most of variables pertaining to the disease natural history of CRC are non-designed variables inherited from individual disease properties.

7.4 Simulation

We simulated 300,000 cohort with 100 replications by the 5-state CRC natural history model to generate the number of CRC (including early stage cancer and advanced cancer) and the number of CRC death in the invited group and pseudo control group (uninvited group) which did not exposed in FIT screening during the periods of 12 years. The base case parameters of CRC natural history were derived from the estimated results of five-state model of CRC natural history and parameters from CRC to CRC death were modelled by AFT model considering the type of detected CRC

doi:10.6342/NTU201901194

(PCDP: screened-detected CRC or CP: clinical detected CRC) and staging of CRC (Early CRC or late CRC). The base case parameters are shown in the Figure 7.2.

7.4.1 The projection of effectiveness with varying screening parameters

We set the basecase of screening program was 57% screening rate, 2-year inter-screening interval with sensitivity of early CRC and advanced CRC as 60.68% and 77.98%, respectively. The three scenarios were like the follows.

- Scenario I: changing screening rate from 10% to 100%
- Scenario II: changing inter-screening interval with 1-, 2-, and 3-yearly
- Scenario III: with improved sensitivity between 60% and 80% for early CRC and between 80% and 100% for advanced CRC, respectively.

Given the Taiwanese scenario (basecase), the estimated CRC mortality reduction was 33.52% (95% CI=28.05%-38.99%).

Table 7.1 shows the projection of effectiveness in terms of mortality reduction from CRC. Given 10% coverage rate, CRC mortality would be reduced by 5.76%. With increasing coverage rate, the benefit of CRC mortality reduction gets larger. A CRC mortality reduction by 58.80% is anticipated with a 100% coverage rate (Figure 7.3). Higher coverage rate also results in lower number needed to screen (NNS) in order to reduce one CRC death, hence higher efficiency. The NNS decreases from 3433 (10% coverage rate) to 339 (100% coverage rate). When sensitivity of early stage cancer being improved from 60% to 80% and the sensitivity of advanced cancer from 80% to 100%, the CRC mortality reduction is increased by 5% (from 35% to 40%), and the number needed to screen decrease from 568 to 497.

doi:10.6342/NTU201901194

7.4.2 The projection of effectiveness with consolidated CPHM underpinning

In light of the consolidated CPHM framework, we investigate the effectiveness of screening with different arrival rate (attendance to screen), positive rate of FIT, and process to undertaking colonoscopy (Figure 7.1). The influence of these factors were examined under three designed inter-screening intervals: 1-, 2-, and 3-yearly.

Figure 7.4 shows the results of effectiveness of FIT screening using consolidated CPHM. The effectiveness is larger affected by coverage rate and referral rate, but slightly affected by positive rate. Given 7% positive rate and 60% referral rate, screening program with 40%, 60%, and 80% yielded 14%, 21%, and 28% mortality reduction, respectively. Given 7% positive rate and 60% screening rate, 21%, 28%, and 36% mortality can be anticipated with 40%, 60%, and 80% compliance for confirmatory colonoscopy, respectively. Compared with their counterparts in a program given 4% positive rate, screening program with higher positivity rate generally can result in higher effectiveness.

7.5 Discussion

We developed a consolidated CPHM to incorporate structure-process factors, disease natural history of multistate CRC, and prognostic of CRC survival for the systematical evaluation of population-based FIT screening in CRC mortality reduction. This approach can be used as a guide to construct the decision analysis for evaluation of population-based screening program, as shown in Figure 7.5. Through this decision analysis, systematic evaluation of effectiveness in reducing advanced cancer and cause-specific mortality can be done from participant behavior factors such as attendance rate, the selection of cutoff of the screening method, compliance rate with confirmatory diagnosis, quality assurance like program sensitivity to screening policy such a kind of

doi:10.6342/NTU201901194

CPHM not only aids in health policy-maker for designing evidence based screening policy and physician for sharing decision-making with patient but provide the capacity of population-based screening program given limited resources.

The framework of consolidated CPHM model is the backbone of economic evaluation of screening policy. More importantly, decision analysis play an important role in big data analysis for the development of an artificial intelligent system for population-based screening program.

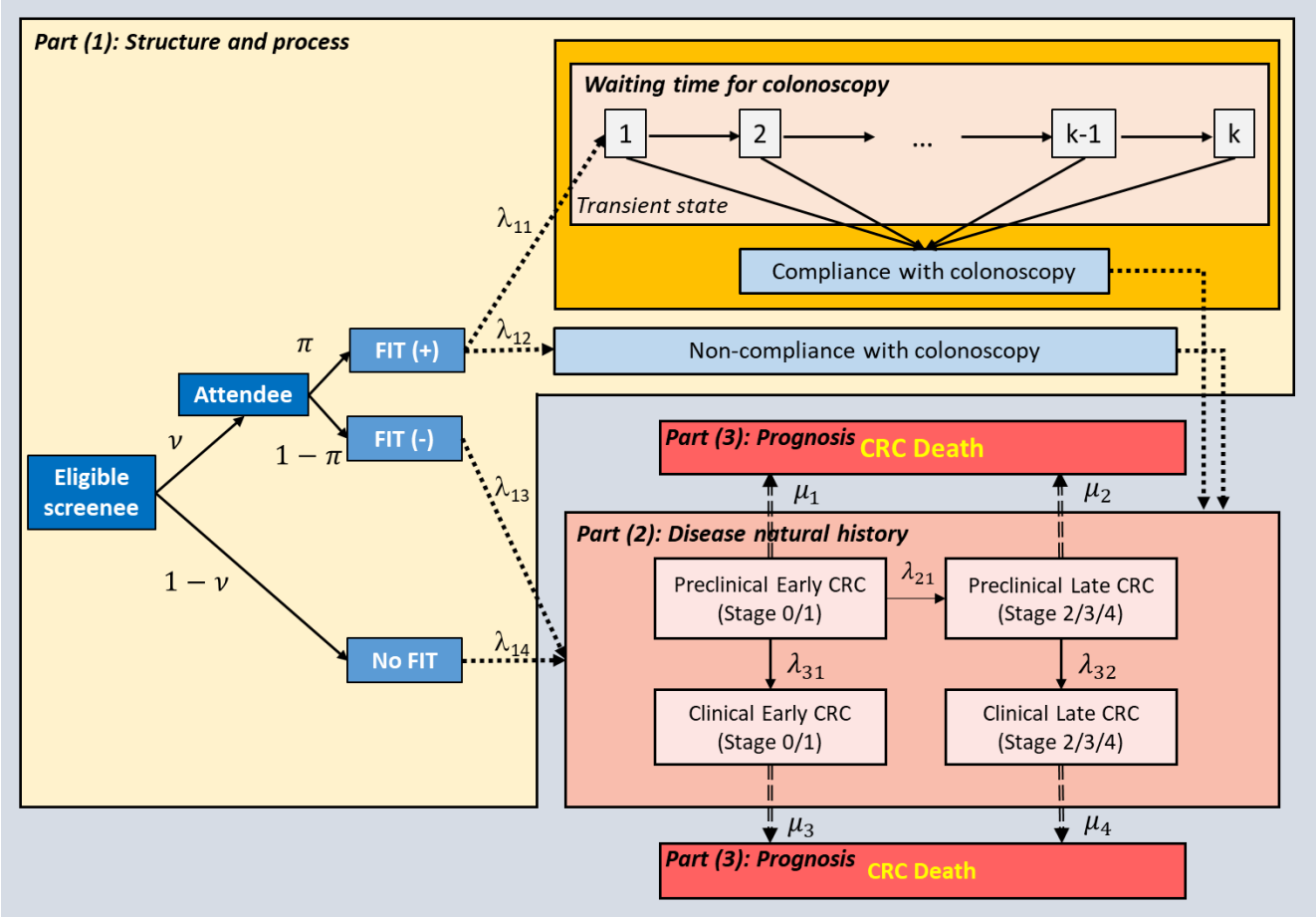
doi:10.6342/NTU201901194

Table 7.1 The projection of effectiveness in terms of CRC mortality reduction

Parameters	Number of CRC Death		CRC Mortality reduction (95% CI; %)	Number needed to be screened
	Control group (Uninvited)	Invited group		
Scenario I: Change coverage rates				
10	1500	1413	5.76 (-1.08-12.60)	3433
20	1505	1327	11.76 (5.11-18.42)	1686
30	1503	1238	17.53 (11.48-23.58)	1135
40	1507	1151	23.60 (17.96-29.25)	841
50	1502	1060	29.40 (23.80-35.00)	678
60	1503	973	35.21 (30.15-40.27)	566
70	1501	884	41.02 (36.11-45.93)	487
80	1502	797	46.93 (42.48-51.38)	425
90	1499	709	52.72 (48.77-56.66)	379
100	1504	619	58.80 (54.94-62.66)	339
Scenario II: Change inter-screening interval				
1-yearly	1497	880	41.18 (36.26-46.09)	486
2-yearly	1501	973	35.09 (29.50-40.68)	569
3-yearly	1505	1022	32.01 (26.76-37.27)	622
Scenario III: sensitivity of early and advanced CRC				
60% 80%	1500	971	35.19 (30.08-40.30)	568
65% 80%	1502	960	36.01 (30.77-41.25)	554
70% 85%	1505	948	36.97 (31.86-42.09)	538
75% 85%	1500	939	37.36 (32.21-42.50)	535
75% 90%	1499	926	38.22 (33.23-43.22)	523
80% 95%	1502	910	39.38 (34.40-44.37)	507
80% 100%	1502	898	40.15 (35.47-44.83)	497

doi:10.6342/NTU201901194

Figure 7.1. The consolidated QH-CPH model with consolidated CPHM underpinning



doi:10.6342/NTU201901194

Figure 7.2 Base case parameters for the disease natural history of CRC and the hazard to CRC death

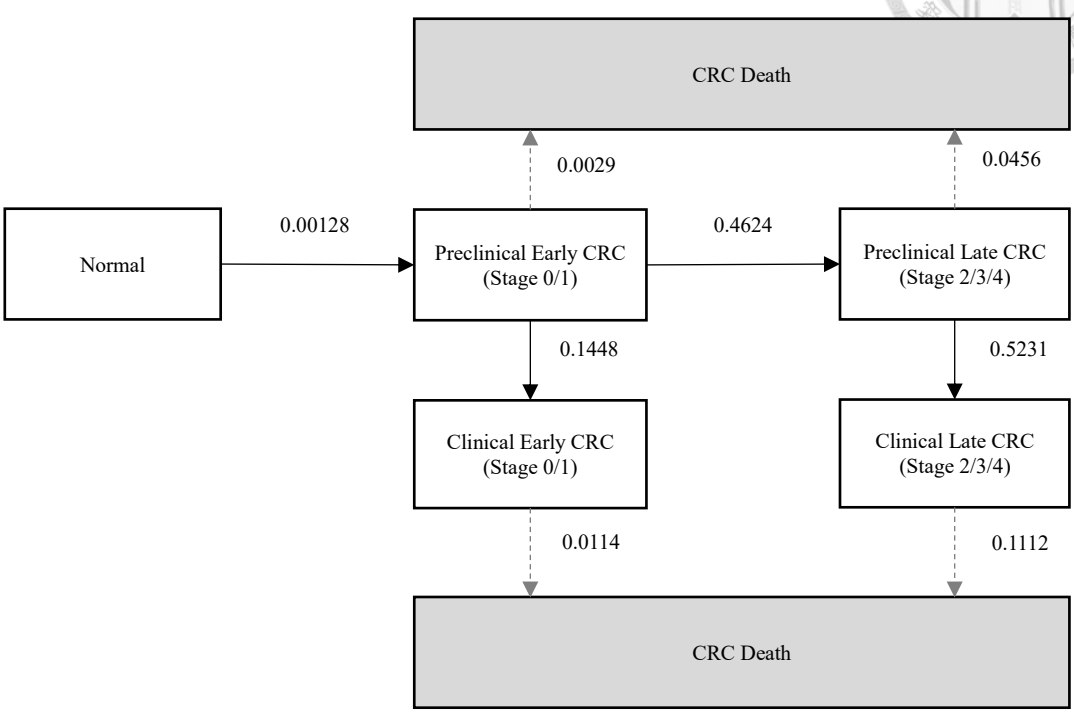
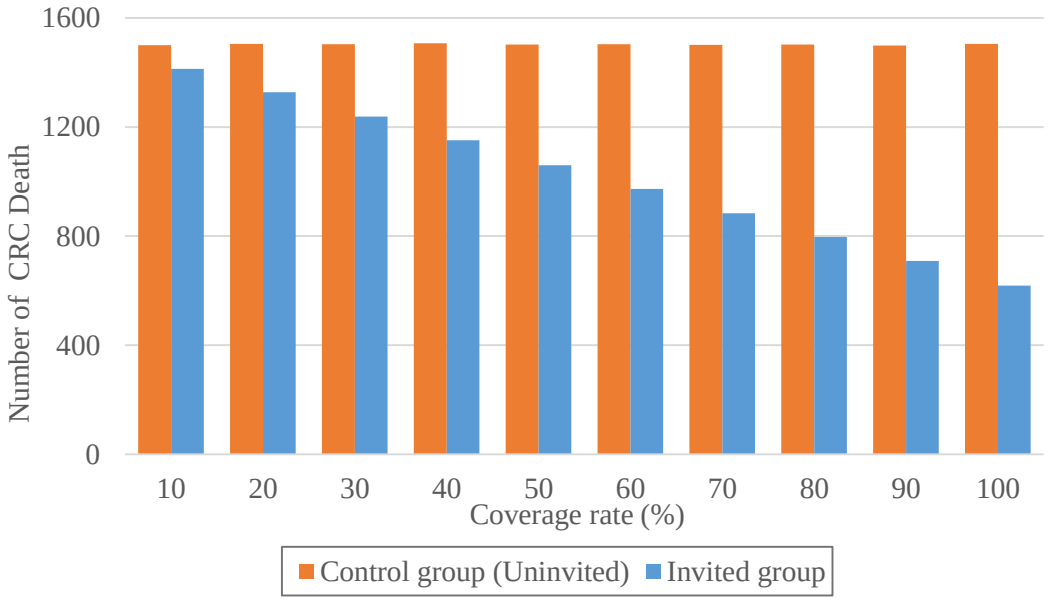


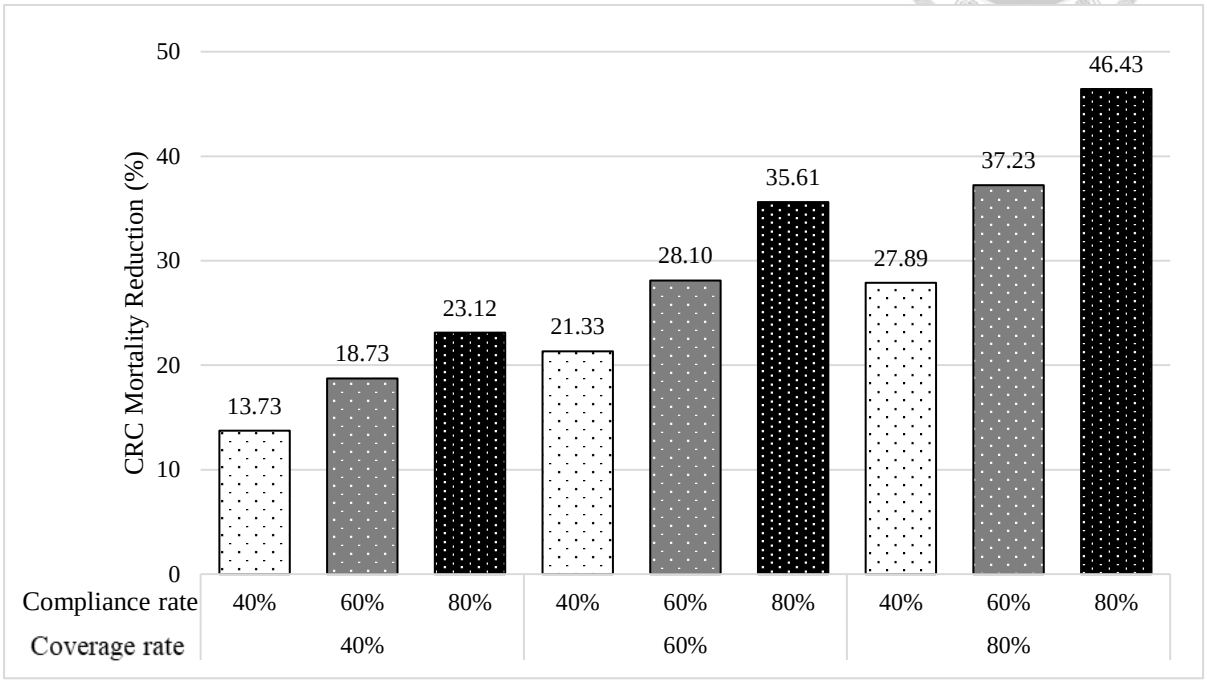
Figure 7.3. The projected number of CRC death with different coverage rate



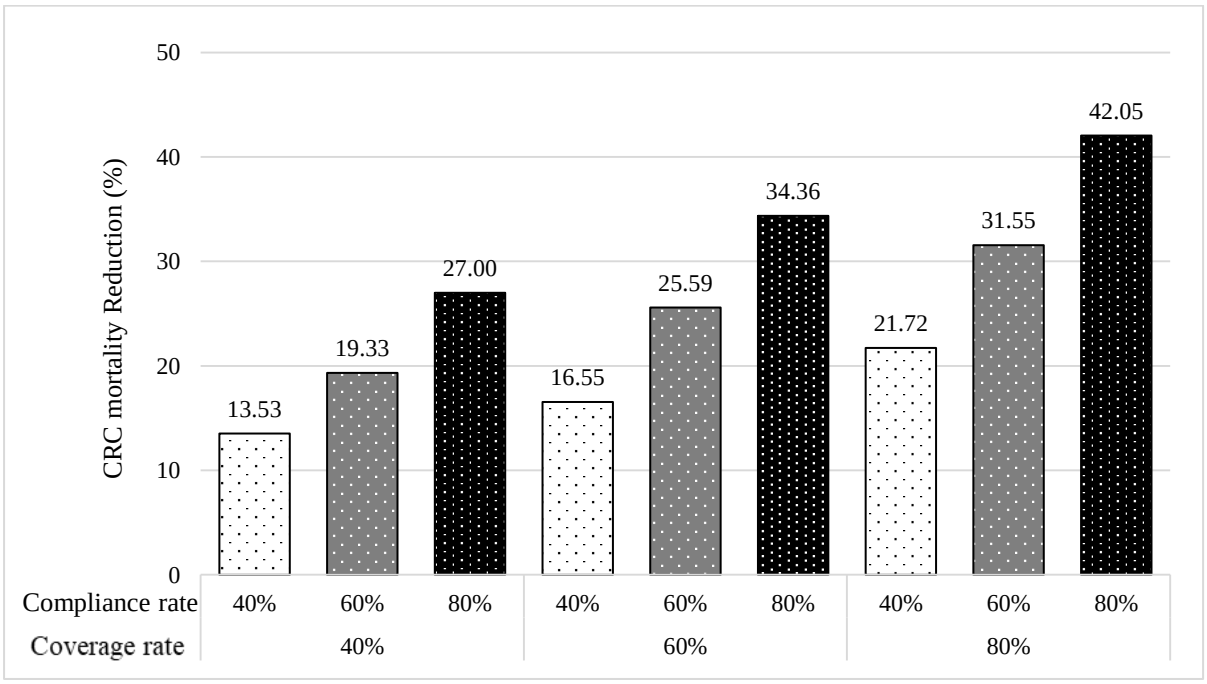
doi:10.6342/NTU201901194

Figure 7.4 Predicted effectiveness on mortality reduction by different scenarios in the biennial screening regime

(a) Given 7% positive rate



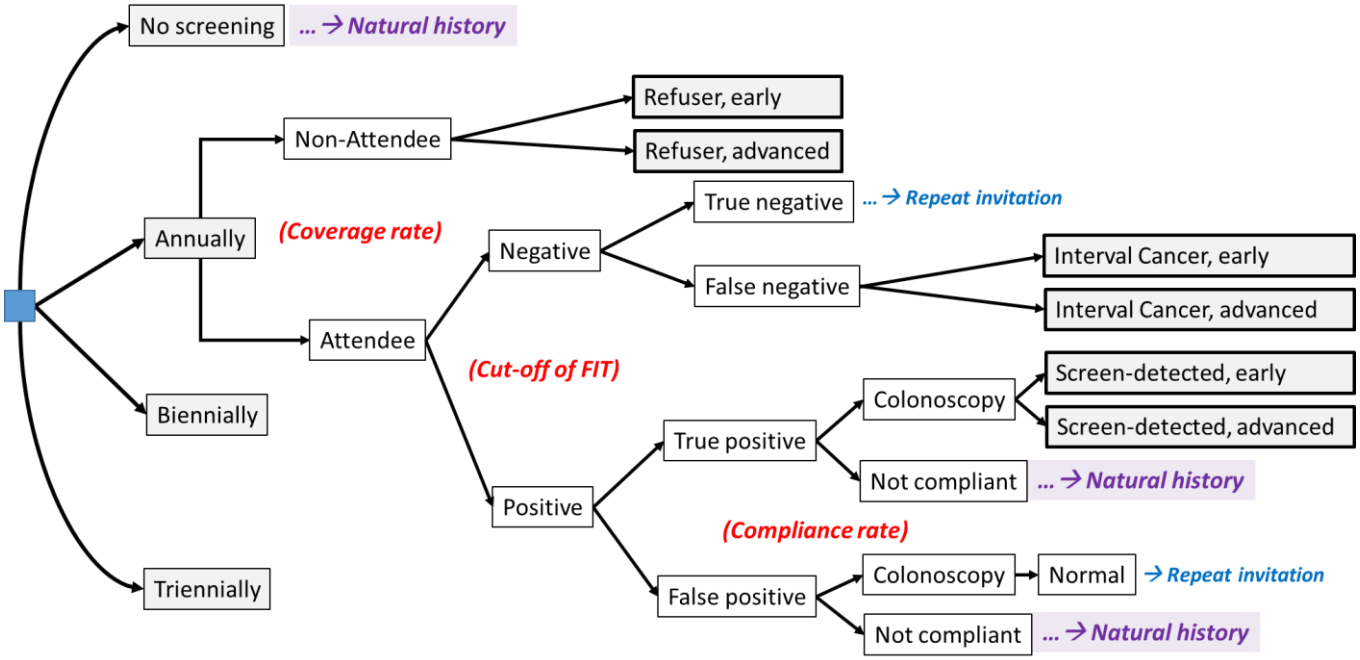
(b) Given 4% positive rate



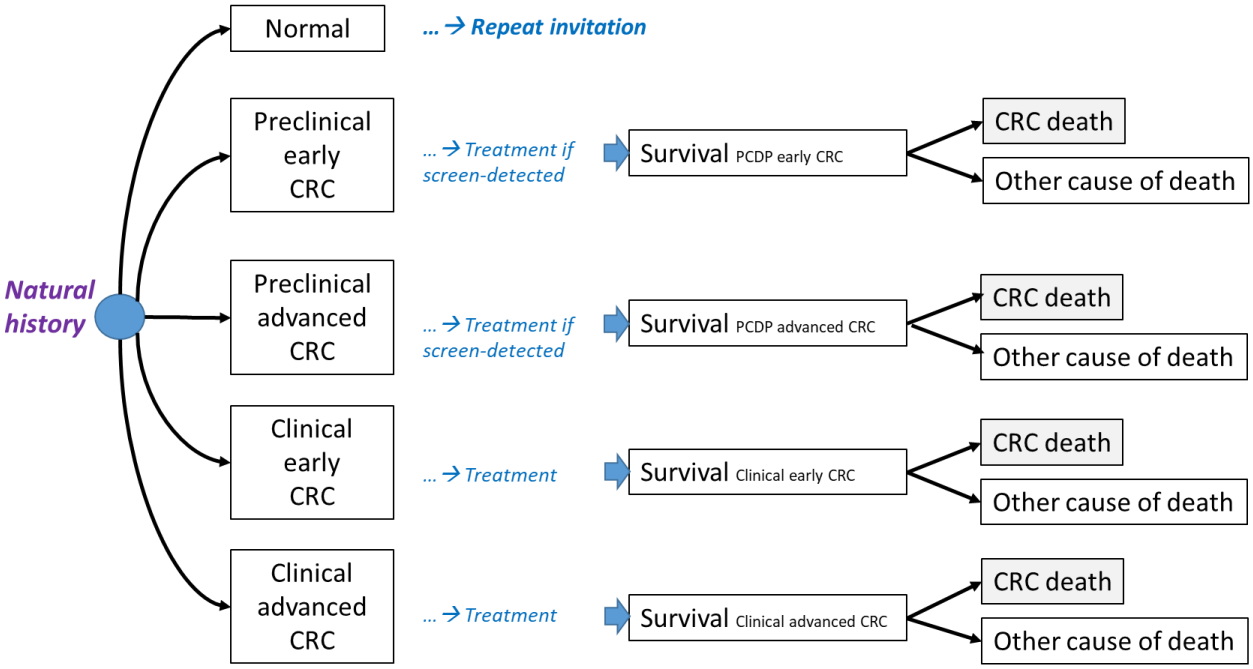
doi:10.6342/NTU201901194

Figure 7.5 Decision framework of population-based CRC screening with consolidated Coxian phase-type Markov (CPHM) process

(A) Decision tree for the consolidated CPHM



(B) Natural history model for multi-state CRC



doi:10.6342/NTU201901194

Reference

Blanks, R., Burón Pust, A., Alison, R., He, E., Barnes, I., Patnick, J., ... & Green, J. (2019). Screen-detected and interval colorectal cancers in England: Associations with lifestyle and other factors in women in a large UK prospective cohort. International journal of cancer.

Brown, E. R., & Ibrahim, J. G. (2003). Bayesian approaches to joint cure-rate and longitudinal models with applications to cancer vaccine trials. Biometrics, 59(3), 686-693.

Buskermolen, M., Gini, A., Naber, S. K., Toes-Zoutendijk, E., de Koning, H. J., & Lansdorp-Vogelaar, I. (2018). Modeling in Colorectal Cancer Screening: Assessing External and Predictive Validity of MISCAN-Colon Microsimulation Model Using NORCCAP Trial Results. Medical Decision Making, 38(8), 917-929.

Caldarella, A., Crocetti, E., Messerini, L., & Paci, E. (2013). Trends in colorectal incidence by anatomic subsite from 1985 to 2005: a population-based study. International journal of colorectal disease, 28(5), 637-641.

Carot, L., Castells, A., Hernández, C., Alvarez-Urturi, C., Balaguer, F., Lanas, A., ... & Carballo, F. (2018). Detection of serrated lesions in proximal colon by simulated sigmoidoscopy vs faecal immunochemical testing in a multicentre, pragmatic, randomised controlled trial. United European gastroenterology journal, 6(10), 1527-1537.

Chen, H. H., Prevost, T. C., & Duffy, S. W. (1999). Evaluation of screening for nasopharyngeal carcinoma: trial design using Markov chain models. British journal of cancer, 79(11), 1894.

Chen, L. S., Liao, C. S., Chang, S. H., Lai, H. C., & Chen, T. H. H. (2007). Cost-effectiveness analysis for determining optimal cut-off of immunochemical faecal

doi:10.6342/NTU201901194

occult blood test for population-based colorectal cancer screening (KCIS 16). Journal of medical screening, 14(4), 191-199.

Chen, L. S., Yen, A. M. F., Duffy, S. W., Tabar, L., Lin, W. C., & Chen, H. H. (2010). Computer-aided system of evaluation for population-based all-in-one service screening (CASE-PASS): From study design to outcome analysis with bias adjustment. Annals of epidemiology, 20(10), 786-796.

Chen, S. L. S., Hsu, C. Y., Yen, A. M. F., Young, G. P., Chiu, S. Y. H., Fann, J. C. Y., ... & Chen, H. H. (2018). Demand for colonoscopy in colorectal cancer screening using a quantitative fecal immunochemical test and age/sex-specific thresholds for test positivity. Cancer Epidemiology and Prevention Biomarkers, 27(6), 704-709.

Chen, T. H. H., Yen, M. F., Lai, M. S., Koong, S. L., Wang, C. Y., Wong, J. M., ... & Duffy, S. W. (1999). Evaluation of a selective screening for colorectal carcinoma: the Taiwan Multicenter Cancer Screening (TAMCAS) project. Cancer: Interdisciplinary International Journal of the American Cancer Society, 86(7), 1116-1128.

Chen, T. H., Kuo, H. S., Yen, M. F., Lai, M. S., Tabar, L., & Duffy, S. W. (2000). Estimation of sojourn time in chronic disease screening without data on interval cases. Biometrics, 56(1), 167-172.

Chen, V. C., Tan, H. K., Chen, C. Y., Chen, T. H., Liao, L. R., Lee, C. T., ... & Cheng, A. T. (2011). Mortality and suicide after self-harm: community cohort study in Taiwan. The British Journal of Psychiatry, 198(1), 31-36.

Cheung, Y. B. (2002). Zero-inflated models for regression analysis of count data: a study of growth and development. Statistics in medicine, 21(10), 1461-1469.

Chiu, H. M., Chen, S. L. S., Yen, A. M. F., Chiu, S. Y. H., Fann, J. C. Y., Lee, Y. C., ... & Koong, S. L. (2015). Effectiveness of fecal immunochemical testing in reducing

doi:10.6342/NTU201901194

colorectal cancer mortality from the O ne M illion T aiwanese S creening P rogram. Cancer, 121(18), 3221-3229.

Chiu, H. M., Lee, Y. C., Chou, C. K., & Wu, M. S. (2015). Mo1929 Comparison of One-Day and Two-Day Sampling on Detection of Significant Colorectal Neoplasm by Fecal Immunochemical Test-Preliminary Results From a Population-Based Randomized Controlled Trial. Gastroenterology, 148(4), S-742.

Chiu, H. M., Lee, Y. C., Tu, C. H., Chen, C. C., Tseng, P. H., Liang, J. T., ... & Wu, M. S. (2013). Association between early stage colon neoplasms and false-negative results from the fecal immunochemical test. Clinical Gastroenterology and Hepatology, 11(7), 832-838.

Chiu, S. Y. H., Malila, N., Yen, A. M. F., Anttila, A., Hakama, M., & Chen, H. H. (2011). Analytical decision model for sample size and effectiveness projections for use in planning a population-based randomized controlled trial of colorectal cancer screening. Journal of evaluation in clinical practice, 17(1), 123-129.

Chiu, S. Y. H., Malila, N., Yen, A. M. F., Chen, S. L. S., Fann, J. C. Y., & Hakama, M. (2017). Predicting the effectiveness of the Finnish population-based colorectal cancer screening programme. Journal of medical screening, 24(4), 182-188.

Consul, P., & Famoye, F. (1992). Generalized Poisson regression model. Communications in Statistics-Theory and Methods, 21(1), 89-109.

Conway, R. W., & Maxwell, W. L. (1962). A queuing model with state dependent service rates. Journal of Industrial Engineering, 12(2), 132-136.

Cox, D. R., & Miller, H. D. (1965). The theory of stochastic processes, Methuen & Co. Ltd, London, UK.

doi:10.6342/NTU201901194

CUZICK, J., EDWARDS, R., & SEGNAN, N. (1997). Adjusting for non-compliance and contamination in randomized clinical trials. Statistics in medicine, 16(9), 1017-1029.

Deb, P., & Norton, E. C. (2018). Modeling health care expenditures and use. Annual review of public health, 39, 489-505.

Denters, M. J., Deutekom, M., Bossuyt, P. M., Fockens, P., & Dekker, E. (2013). Patient burden of colonoscopy after positive fecal immunochemical testing for colorectal cancer screening. Endoscopy, 45(05), 342-349.

Donnelly, C., McFetridge, L. M., Marshall, A. H., & Mitchell, H. J. (2018). A two-stage approach to the joint analysis of longitudinal and survival data utilising the Coxian phase-type distribution. Statistical methods in medical research, 27(12), 3577-3594.

Duffy, S. W., Cuzick, J., Tabar, L., Vitak, B., Chen, T. H. H., Yen, M. F., & Smith, R. A. (2002). Correcting for non-compliance bias in case–control studies to evaluate cancer screening programmes. Journal of the Royal Statistical Society: Series C (Applied Statistics), 51(2), 235-243.

Fackrell, M. (2009). Modelling healthcare systems with phase-type distributions. Health care management science, 12(1), 11.

Gordon, A. S., Marshall, A. H., & Cairns, K. J. (2016). A conditional approach for modelling patient readmissions to hospital using a mixture of Coxian phase-type distributions incorporating Bayes' theorem. Statistics in medicine, 35(21), 3810-3826.

Haenszel, W., & Correa, P. (1973). Cancer of the large intestine: epidemiologic findings. Diseases of the Colon & Rectum, 16(5), 371-377.

Jen, H. H., Hsu, C. Y., Chen, S. L. S., Yen, A. M. F., Chiu, S. Y. H., Fann, J. C. Y., ... & Chen, H. H. (2018). Rolling-out Screening Volume Affecting Compliance Rate and

doi:10.6342/NTU201901194

Waiting Time of FIT-based Colonoscopy. Journal of clinical gastroenterology, 52(9), 821-827.

Langsetmo, L., Kats, A. M., Cawthon, P. M., Cauley, J. A., Vo, T. N., Taylor, B. C., ... & Schousboe, J. T. (2017). The association between objectively measured physical activity and subsequent health care utilization in older men. The Journals of Gerontology: Series A, 74(6), 820-826.

Larsen, I. K., & Bray, F. (2010). Trends in colorectal cancer incidence in Norway 1962–2006: an interpretation of the temporal patterns by anatomic subsite. International journal of cancer, 126(3), 721-732.

Levin, T. R., Corley, D. A., Jensen, C. D., Schottinger, J. E., Quinn, V. P., Zauber, A. G., ... & Lee, A. T. (2018). Effects of organized colorectal cancer screening on cancer incidence and mortality in a large community-based population. Gastroenterology, 155(5), 1383-1391.

Lin, C. H., Schiøler, K. L., Ekstrøm, C. T., & Konradsen, F. (2018). Location, seasonal, and functional characteristics of water holding containers with juvenile and pupalAedes aegypti in Southern Taiwan: A cross-sectional study using hurdle model analyses. PLoS neglected tropical diseases, 12(10), e0006882.

Marshall, A. H., & McClean, S. I. (2003). Conditional phase-type distributions for modelling patient length of stay in hospital. International Transactions in Operational Research, 10(6), 565-576.

Marshall, A. H., & McCrink, L. (2009, August). Discrete Conditional Phase-type model (DC_Ph) for patient waiting time with a logistic regression component to predict patient admission to hospital. In 2009 22nd IEEE International Symposium on Computer-Based Medical Systems (pp. 1-6). IEEE.

doi:10.6342/NTU201901194

Marshall, A. H., Burns, L., & Shaw, B. (2007). Patient activity in hospital using discrete conditional phase-type (DC-PH) models. In Recent Advances in Stochastic Modeling and Data Analysis(pp. 154-161).

Marshall, A. H., Shaw, B., & McClean, S. I. (2007). Estimating the costs for a group of geriatric patients using the Coxian phase-type distribution. Statistics in medicine, 26(13), 2716-2729.

McClean, S., Garg, L., Barton, M., & Fullerton, K. (2010, October). Using mixed phase-type distributions to model patient pathways. In 2010 IEEE 23rd International Symposium on Computer-Based Medical Systems (CBMS) (pp. 172-177). IEEE.

McDonald, P. J., Strachan, J. A., Digby, J., Steele, R. J., & Fraser, C. G. (2012). Faecal haemoglobin concentrations by gender and age: implications for population-based screening for colorectal cancer. Clinical chemistry and laboratory medicine, 50(5), 935-940.

Mullahy, J. (1986). Specification and testing of some modified count data models. Journal of econometrics, 33(3), 341-365.

Neelon, B., Ghosh, P., & Loebs, P. F. (2013). A spatial Poisson hurdle model for exploring geographic variation in emergency department visits. Journal of the Royal Statistical Society: Series A (Statistics in Society), 176(2), 389-413.

Neuts, M. F. (1975). Probability distributions of phase type. Liber Amicorum Prof. Emeritus H. Florin.

Rossi, P. G., Federici, A., Bartolozzi, F., Farchi, S., Borgia, P., & Guasticchi, G. (2005). Understanding non-compliance to colorectal cancer screening: a case control study, nested in a randomised trial [ISRCTN83029072]. BMC Public Health, 5(1), 139.

doi:10.6342/NTU201901194

Rossi, P. G., Vicentini, M., Sacchettini, C., Di Felice, E., Caroli, S., Ferrari, F., ... & Sassatelli, R. (2015). Impact of screening program on incidence of colorectal cancer: a cohort study in Italy. The American journal of gastroenterology, 110(9), 1359.

Schwarz, G. (1978). Estimating the dimension of a model. The annals of statistics, 6(2), 461-464.

Siddiqui, A. A., Patel, A., & Huerta, S. (2006). Determinants of compliance with colonoscopy in patients with adenomatous colon polyps in a veteran population. Alimentary pharmacology & therapeutics, 24(11-12), 1623-1630.

Singh, S., Singh, P. P., Murad, M. H., Singh, H., & Samadder, N. J. (2014). Prevalence, risk factors, and outcomes of interval colorectal cancers: a systematic review and meta-analysis. The American journal of gastroenterology, 109(9), 1375.

Slattery, M. L., Friedman, G. D., Potter, J. D., Edwards, S., Caan, B. J., & Samowitz, W. (1996). A description of age, sex, and site distributions of colon carcinoma in three geographic areas. Cancer: Interdisciplinary International Journal of the American Cancer Society, 78(8), 1666-1670.

Tabar, L., Yen, M. F., Vitak, B., Chen, H. H. T., Smith, R. A., & Duffy, S. W. (2003). Mammography service screening and mortality in breast cancer patients: 20-year follow-up before and after introduction of screening. The Lancet, 361(9367), 1405-1410.

Tang, X., Luo, Z., & Gardiner, J. C. (2012). Modeling hospital length of stay by Coxian phase-type regression with heterogeneity. Statistics in Medicine, 31(14), 1502-1516.

Titman, A. C., & Sharples, L. D. (2010). Semi-Markov models with phase-type sojourn distributions. Biometrics, 66(3), 742-752.

Winkelmann, R. (2004). Health care reform and the number of doctor visits—an econometric analysis. Journal of Applied Econometrics, 19(4), 455-472.

doi:10.6342/NTU201901194

Wu, G. H. M., Wang, Y. M., Yen, A. M. F., Wong, J. M., Lai, H. C., Warwick, J., & Chen, T. H. H. (2006). Cost-effectiveness analysis of colorectal cancer screening with stool DNA testing in intermediate-incidence countries. BMC cancer, 6(1), 136.

Wu, J. C. Y., Anttila, A., Yen, A. M. F., Hakama, M., Saarenmaa, I., Sarkeala, T., ... & Chen, T. H. H. (2010). Evaluation of breast cancer service screening programme with a Bayesian approach: mortality analysis in a Finnish region. Breast cancer research and treatment, 121(3), 671-678.

Yang, K. C., Liao, C. S., Chiu, Y. H., Yen, A. M. F., & Chen, T. H. H. (2006). Colorectal cancer screening with faecal occult blood test within a multiple disease screening programme: an experience from Keelung, Taiwan. Journal of medical screening, 13(4).

Yu, D., Hopman, W. M., & Paterson, W. G. (2008). Wait time for endoscopic evaluation at a Canadian tertiary care centre: Comparison with Canadian Association of Gastroenterology targets. Canadian Journal of Gastroenterology and Hepatology, 22(7), 621-626.

Zhen, Z., Shao, L., & Zhang, L. (2018). Spatial hurdle models for predicting the number of children with lead poisoning. International journal of environmental research and public health, 15(9), 1792.

doi:10.6342/NTU201901194