

國立臺灣大學公共衛生學院
流行病學與預防醫學研究所

碩士論文

Institute of Epidemiology and Preventive Medicine

College of Public Health

National Taiwan University

Master's Thesis

元宇宙數位雙胞胎模型評估

免疫糞便潛血精準大腸癌篩檢

Metaverse Digital Twin Model for Evaluating Precision
Fecal Immunochemical Test (FIT) Service Screening for
Colorectal Cancer

魏仕翔

Wei, Shih-Hsiang

指導教授：陳秀熙 教授

Supervisor: Chen, Hsiu-Hsi, Professor

中華民國 113 年 7 月

July 2024



致謝

睽違校園多年後，還能重拾書本，我十分感謝生命給予機會，讓我能完成人生夢想。



在這段旅途中得到多位師長的悉心教導與支持，感謝陳秀熙教授悉心教導，以及嚴明芳教授、陳立昇教授、許辰陽醫師、邱月暇教授、范靜媛教授、林庭瑀博士關懷與指導。各位老師在這段期間不僅在學術上給予指導，還在生活上給予了我許多關懷。

最大的收穫莫過於踏入機器學習領域，要特別感謝陳教授與各位老師帶領我們探索與思考，並在遇到困難時不斷鼓勵我們前行。這段寶貴的經歷，使我在學術和專業技能上都有了長足的進步。

同時，也要感謝邱瀚模教授，有他的推薦、鼓勵與支持，我才能完成這段學業；感謝 BOC 團隊蘇秋文博士對我在入學前的準備給予幫助，以及 BOC 團隊同仁盈芊、力渝、楚翔、瑋萱、季芸等，給予了我支持與空間，讓我能夠順利完成學業。也特別感謝過往的同仁們，他們的協助與鼓勵也是我前行重要動力。

更要感謝我的家人志鴻，在這段期間內的包容與體諒，以及在日常生活中的照顧，我愛你。

中文摘要



背景：

糞便潛血免疫法 (Fecal immunochemical test, FIT)篩檢效益評估已經從族群層級的實證醫學證據轉向個人層級為中心的精準醫學。除了過去對影響評估效度的自我選擇偏差與干擾因子的校正之外，評估 N 對 1 試驗及考慮個人因子的大腸腺腫-大腸癌自然病史予以不同篩檢策略建議包括個人化篩檢間隔與個人風險分層篩檢建議評估及長程追蹤效益等挑戰亦興起新穎方法學考量。傳統上使用電腦模擬的評估模型無法適用此類複雜情境，其原因包括需要多重虛擬假設，且未能考慮真實世界數據 (Real World Data, RWD) 的屬性和特徵。

本論文旨在發展元宇宙數位雙胞胎模型，以評估以族群基礎之 FIT 篩檢效益，評估不同篩檢間隔等均一性篩檢與個人化風險為基礎篩檢建議之沉浸式篩檢政策。

方法：

元宇宙基礎架構乃透過數位雙胞胎設計創建 RWD 虛擬分身，經過 RWD 和虛擬世界數據 (Virtual World Data, VWD) 橋接合成，進行隨機化沉浸式介入評估。VWD 透過生成式人工智慧引擎利用兩種機器學習演算法生成。首先，利用馬可夫過程從 2004 年至 2022 年臺灣全國 FIT 篩檢數據中學習主宰不同部位大腸直腸癌 (近端和遠側部位) 自然病史參數，並評估對不同部位大腸直腸癌其不同篩檢間隔之篩檢效益之差別。其次，我們利用隨機森林機器學習模型辨別社區民眾個人特徵於大腸直腸腺瘤和癌症之個人化風險分位數，並嵌合於第一階段馬可夫過程模型，以評估個人化介入與篩檢政策之效益。



結果：

三階段馬可夫過程結果顯示近端大腸直腸癌從臨床前可偵測期(Preclinical Detectable Phase, PCDP)到臨床期(Clinical Phase, CP)進展速度較遠側大腸直腸癌快。若將癌症分為早期及晚期之五階段馬可夫過程結果亦也顯示相同結果，早期大腸癌在近端部位的 FIT 敏感度低於遠側部位，從腺腫到癌症近端部位轉移較快且敏感度低於遠側部份。若欲達到降低大腸直腸癌晚期發生率達 25%的效益，則對有較高近端大腸直腸癌風險個案需要每年篩檢一次，而對遠側大腸直腸癌潛在風險較高的個案而言，篩檢間隔可設定於每兩年篩檢一次。

利用隨機森林機器學習演算法找到可鑑別近端和遠側大腸直腸癌之個人特徵，並利用貝氏網路演算產生多階段個人化十分位風險評估模式，予以建立個人化篩檢策略，如高風險者(風險分數大於 75%)：每年篩檢一次，中高風險者(風險分數 50-75%)：每兩年篩檢一次，低風險者(風險分數 25%-50%)：每三年篩檢一次，最低風險者(風險分數 0%-25%)：每六年篩檢一次，此個人化篩檢策略於降低晚期大腸癌發生率可達 32%。

結論：

本論文顯示應用元宇宙數位雙胞胎模型可做為以病人為中心精準大腸直腸癌 FIT 篩檢策略設計，並用於評估篩檢效益。

關鍵字：大腸直腸癌、FIT、大規模癌症篩檢、元宇宙、數位雙胞胎



Background:

Evaluation of population-based Fecal immunochemical test (FIT) service screening has turned from universal evidence-based medicine to patient-centered personalized medicine. Therefore, a series of concerns and recommendations are raised and suggested including those threats to validity (self-selection and confounding factors) associated with the effectiveness of screening, N to 1 trial impasse, the individually-tailored disease natural course of colorectal neoplasia, various screening policies such inter-screening interval and risk-based screening strategy, and the logistics of long-term follow-up. The conventional simulation models cannot be accommodated to satisfy these complex situations because the use of this traditional approach requires numerous virtual assumptions and fails to take into account the real world data (RWD) property and inherent characteristics.

This thesis is therefore to develop a metaverse-based digital twin model for evaluating the effectiveness of population-based FIT service screening with a series of immersive screening policies with emphasis on inter-screening interval and risk-based screening strategies.

Methods:

The metaverse-based infrastructure was framed by adopting the digital twin design to spin out the avatar of the RWD, after synthetic bridging between the RWD and the virtual world data (VWD), which was further randomized to immersive interventions for evaluation. The VWD was generated by generative artificial intelligence engine by leveraging two machine learning algorithms. The first was to learn the parameters governing the site-(proximal and distal location)-based disease natural course of colorectal cancer with Markov process from the RWD of Taiwan nationwide FIT screening data between 2004 and 2022. The effectiveness of inter-screening interval by site was then evaluated. To develop the personalized disease natural course of colorectal adenoma and carcinoma, we leveraged the random forest of machine learning model to identify important variables based on community-based integrated screening data. The decile of personalized risk assessment model of colorectal adenoma and carcinoma was

therefore constructed. The effectiveness of individually-tailored intervention and risk-based screening policy was evaluated.



Result:

The results of transition rates and sensitivity of FIT test using three-state Markov process show the proximal CRCs progressed faster from pre-clinical detectable phase (PCDP) to clinical phase than the distal CRCs. The findings of five-state Markov process also show the proximal CRCs had faster transition from early to late AJCC stage during PCDP and further progressed to the late stage of clinical phase. The test sensitivity of early AJCC stage was lower in the proximal site than the distal site. The similar findings on faster transition rates from adenoma to carcinoma and lower sensitivity were also noted. If effectiveness of CRC mortality reduction is set at least 25% annual screening is required for those who are potential of developing proximal CRCs and biennial screening is sufficient for those who are potential of developing distal CRCs.

Individual features were extracted to distinguish the proximal CRCs from the distinguished CRCs by leveraging the random forest of machine learning algorithm. Risk-based-screening strategy (annual for 75% or higher risk, biennial for 50-75% moderate high risk four-yearly regime for 25%-50% lower risk and six-yearly regime for the lowest risk of 0-25%) in the light of the incorporation of these distinguished features into decile-based personalized risk model was demonstrated to reduce of incidence of advanced-stage CRC 32%.

Conclusion:

This thesis demonstrates the clinical usefulness of leveraging a metaverse digital twin model for evaluating and designing precision and patient-centered FIT screening strategies for CRC.

Keywords: Colorectal cancer, FIT, Mass cancer screening, Metaverse, Digital Twin

目次

致謝.....	i
中文摘要.....	ii
英文摘要.....	iv
第一章 、前言	1
第二章 、文獻探討	2
2.1 大腸癌流行病學.....	2
2.2 大腸癌風險因子.....	3
2.3 近端大腸癌 (Proximal Colon)及遠側大腸癌 (Distal Colon)篩檢效益	5
2.4 大腸直腸癌疾病自然史評估方法.....	6
2.5 臺灣大腸癌篩檢計畫沿革及效益評估.....	9
第三章 、材料與方法	14
3.1 元宇宙數位雙胞胎設計.....	14
3.2 真實世界資料設計與收集.....	16
3.2.1 臺灣大腸直腸癌篩檢.....	16
3.2.2 社區整合式篩檢.....	18
3.3 統計分析.....	19
3.3.1 多階段大腸直腸癌疾病進展.....	19
3.3.2 多階段大腸直腸腺腫--大腸直腸癌疾病進展	21
3.3.3 隨機森林鑑別不同部位大腸直腸癌發生	22
3.3.4 貝氏網絡模式鑑別不同階段大腸直腸腺腫--大腸直腸癌風險	23
3.3.5 元宇宙大腸癌篩檢數位分身效益評估	23
第四章 、結果	25
4.1 臺灣大腸癌篩檢.....	25
4.1.1 描述性分析結果.....	25
4.1.2 大腸直腸癌不同部位疾病進展病史.....	26
4.1.3 隨機森林 (Random Forest) 分析結果	30
4.2 社區整合式篩檢.....	33
4.2.1 描述性分析結果.....	33



4.2.2	隨機森林分析結果.....	37
4.2.3	大腸腺腫至侵襲癌個人化危險分層分析結果.....	42
4.2.4	大腸腺腫及侵襲癌個人化風險分數數位雙胞胎介入評估.....	43
第五章	、討論	45
5.1	近端及遠側大腸直腸癌自然病史.....	45
5.2	元宇宙數位雙胞胎模型用於精準 FIT 篩檢服務評估.....	47
5.3	研究限制.....	47
第六章	、結論	48
參考文獻.....		49



圖次

圖 3.1、數位雙胞胎設計用於元宇宙評估模型	15
圖 3.2、前瞻性定期篩檢設計與使用五階段馬可夫模型分析數據	18
圖 3.3、五階段馬可夫大腸癌進展模式	21
圖 3.4、八階段馬可夫大腸腺腫--大腸癌進展模式	22
圖 3.5、危險分層嵌入八階段馬可夫大腸腺腫-大腸癌進展模式	24
圖 4.1 (a)、近端大腸與正常個案之重要變項排序	30
圖 4.1 (b)、遠側大腸與正常個案之重要變項排序	30
圖 4.1 (c)、近側大腸與遠側大腸個案之重要變項排序	30
圖 4.1 (d)、近端大腸與正常個案 ROC 分析結果	31
圖 4.1 (e)、遠側大腸與正常個案 ROC 分析結果	32
圖 4.1 (f)、遠側大腸癌個案與近端大腸個案 ROC 分析結果	32
圖 4.2 (a)、近端大腸與正常個案之重要變項排序	37
圖 4.2 (b)、遠側大腸與正常個案之重要變項排序	38
圖 4.2 (c)、近側大腸與遠側大腸個案之重要變項排序	39
圖 4.2 (d)、近端大腸與正常個案 ROC 分析結果	40
圖 4.2 (e)、遠側大腸與正常個案 ROC 分析結果	40
圖 4.2 (f)、遠側大腸癌個案與近端大腸個案 ROC 分析結果	41

表次



表 4.1、臺灣大腸癌篩檢近端及遠側大腸癌描述性分析結果	25
表 4.2、大腸直腸侵襲癌近端及遠側疾病進展馬可夫三階段機器學習估計結果	28
表 4.3、大腸直腸侵襲癌近端及遠側疾病進展馬可夫五階段機器學習估計結果	28
表 4.4、大腸直腸腺腫-侵襲癌近端及遠側疾病進展馬可夫八階段機器學習估計 結果.....	29
表 4.5、社區整合式篩檢描述性分析結果	33
表 4.6、FIT 大腸癌檢篩檢間隔數位雙胞胎效益評估	44
表 4.7、FIT 大腸癌檢篩檢間隔數位雙胞胎精準篩檢政策	44

第一章、前言



大腸直腸癌糞便潛血大規模篩檢在經由臨床隨機分派證實可有效降低大腸直腸癌死亡率後，歐美許多國家已廣泛利用糞便潛血提供以族群為基礎之大規模篩檢服務，臺灣亦從 2004 年開始提供民眾兩年一次糞便潛血篩檢，然而對服務性篩檢效益評估面臨許多困難，包括因非隨機對照試驗設計於接受篩檢與未接受篩檢個案之間的比較會面臨方法學的難度，需要長期追蹤以得到結果，以及無法因應新興篩檢工具，不同政策等的及時評估。

過去研究探討上述相關問題多仰賴電腦模擬模型建立，然而透過電腦模擬產生虛擬族群的缺點主要在於缺乏與真實世界數據的介接，真實世界資料提供的參數包括篩檢率、轉介率和追蹤順從性等，再者觀察性服務篩檢評估會面臨選擇偏差和干擾因子之威脅，及所建構大腸直腸腺腫和癌症自然病史模型過於簡化等，舉例而言，大腸直腸腺腫及癌症在不同部位(近端和遠側部位)疾病自然進展病程幾乎未曾被探討。

近年來提出之個人化篩檢強調根據個人特徵所反應不同風險分層因而建議不同篩檢間隔、篩檢起始年齡等個人化篩檢規劃，由於其衍生出一系列篩檢參數與情境，因此對此類個人化篩檢效益評估也朝向如何創造如同傳統隨機臨床分派試驗情境進行實證等級最高之評估。數位雙胞胎設計結合元宇宙架構提供可能解決方案，數位雙胞胎設計橋接真實世界數據 (Real World Data, RWD)和虛擬世界數據 (Virtual World Data, VWD)，為真實世界數據創造虛擬分身，元宇宙架構則評估隨機分派予虛擬分身一系列沉浸式介入與評估。

本論文目的旨在發展以元宇宙架構為基礎之數位雙胞胎模型，創造臨床隨機分派情境，評估以族群為基礎大規模 FIT 篩檢不同政策，包括篩檢間隔、篩檢起始年齡與基於個人風險之個人化篩檢。

第二章、文獻探討



2.1 大腸癌流行病學

大腸直腸癌是全球公共衛生的問題，國際癌症組織 (International Agency for Research on Cancer , IARC) 2018 年發表 GLOBOCAN 資料顯示，估計全球有 180 萬新病例與 88 萬例左右死亡個案；常見癌症中，大腸直腸癌發生率男性中排名第三名(10.9%)，女性排名第二(9.5%)，大腸直腸癌死亡率在男性中排名第四(9.0%)、女性排名第三(9.5%)。(Chiu Han-Mo, 2021)

大腸直腸癌疾病負擔在全球各地區都有所不同，大腸直腸癌發生率最高地區為歐洲、北美及大洋洲。大腸直腸癌發生率與經濟發展型態習習相關，若以人類發展指數 (The Human Development Index , HDI)進行評估，該指數以預期壽命、教育程度與人均收入進行整體性呈現之指標；發現大腸直腸癌發生率與 HDI 指標進行相關，HDI 較高的地區其大腸直腸癌發生率也較高。

亞太地區大腸癌新發個案與死亡個案為全球最高，目前澳大利亞、日本、韓國、新加坡、紐西蘭與臺灣等國家皆有全國性篩檢計畫，這些計畫都由政府提供全部或部分經費；多數篩檢計畫起始年齡為 50 歲起提供篩檢服務，僅日本為 40 歲開始。只有澳大利亞與臺灣有限制年齡上限至 74 歲；由於大腸鏡檢查能力與人力資源分配，所有國家皆使用糞便潛血免疫法試驗(FIT)作為初篩測試，並且對篩檢結果為陽性之個案提供診斷性大腸鏡檢查。(Rabeneck et al., 2020)

行政院主計處 2020 年評估，臺灣 HDI 位居全球第 21 名，發展程度屬極高等級，且多數民眾生活型態與其他發展程度極高的國家相似，西化的生活方式，如不良的生活習慣、久坐、肥胖、攝食紅肉等也被證實與大腸直腸癌的發生率及死亡率上升有關。

依據國民健康署 2023 年癌症登記年報資料，臺灣大腸直腸癌全國年齡標準化發生率為每十萬人口約 46.77 人(含原位癌)*，男性年齡標準化發生率為每十萬人口約 58.14 人(含原位癌)，為男性癌症第一名，而女性年齡標準化發生率為每十萬人口約 36.58 人，為女性癌症第三名。大腸直腸癌全國年齡標準化死亡率為每十萬人口約 14.46 人，男性年齡標準化死亡率為每十萬人口約 15.25 人，為男性癌症第三名，女性年齡標準化死亡率為每十萬人口約 11.21 人，為女性癌症第三名。(衛生福利部國民健康署, 2023)

整體而言，世界各國不論經濟發展程度，男性大腸直腸癌之發生率與死亡率皆高於女性。

2.2 大腸癌風險因子

導致大腸直腸癌之風險因子包含環境因素與遺傳因素，除家族史外，吸菸、肥胖、代謝症候群都可能導致大腸直腸癌；約有 6 成至 6 成 5 大腸直腸癌為偶發性，2 成為可歸因為家族史，約有 1 成為遺傳性癌症症候群(如 HNPCC、FAP)。

具有大腸直腸癌家族史被認為是重要的風險因子，通常一等親內患有大腸直腸癌者為具家族史高風險族群。與無家族史相比，具家族病史者大腸癌發病風險增加 1.8 倍，且發病年齡較偶發性大腸癌早，具家族史之個案，可能也具有遺傳有關腸道疾病，包括家族性腺瘤息肉症 (Familial adenomatous polyposis , FAP)、遺傳性非息肉大腸癌 (Hereditary nonpolyposis colorectal cancer, HNPCC or Lynch syndrome)、MUTYH 基因變異相關息肉症 (MUTYH-associated polyposis, MAP) 等。(Oh et al., 2018) 根據目前篩檢指引建議，對於一般族群篩檢年齡建議 50 歲開始，一等親內具家族史之個案建議 40 歲起接受篩檢，以降低晚期大腸癌之風險。(Henrikson et al., 2015)

* 結腸、直腸、乙狀結腸連結部與肛門，含原位癌與侵襲癌

吸菸已被證實與大腸癌相關，2013 年統合分析研究發現，與非吸菸者相比，重度吸菸者罹患大腸直腸癌之相對風險分別為 5 包年的 1.06 倍 (95% CI : 1.03-1.08) 與 30 包年之 1.26 倍 (95% CI : 1.17-1.36)。(Johnson et al., 2013)

除此之外，歐洲癌症營養調查進行長達 15 年追蹤研究，發現吸菸者近端大腸比遠側大腸有較高的罹癌風險；分子流行病學研究也顯示，微衛星不穩定性(Microsatellite instability, MSI)高，CpG 島高度甲基化(CIMP)，與 BRAF 基因突變容易罹患近端大腸癌；而吸菸致癌物會影響 DNA 修復系統，使微衛星不穩定性增加，相較於 MSI 低者，MSI 高者罹癌風險增加 1.94 倍(95% CI : 1.09-3.46)。(Murphy et al., 2019)

除了吸菸外，肥胖也被證實為大腸直腸癌風險因子，約有 11% 可歸因於過重與肥胖；腰圍與身體質量指數(BMI)都是測量肥胖之指標，2013 年研究發現，BMI 每增加一單位，大腸直腸癌發病風險增加 3%，且男性比女性也更高。(Bardou et al., 2013)，另外一項長期追蹤護理師健康的研究，調查 8 萬多名 25 歲至 42 歲之間的女性，結果發現與 BMI 介於 18.5 至 22.9 之間之年輕族群相比，過重(BMI 25.0-30)與肥胖(BMI ≥ 30)的風險分別增加 1.37 倍 (95% CI : 0.81-2.30) 與 1.93 倍(95% CI : 1.15-3.25)。(Liu et al., 2019)

針對肥胖族群的精準組織性篩檢，可作為降低該族群大腸直腸癌死亡率之方法。

前述提到肥胖為大腸癌風險因子之一，不過進行減重或減脂的方法可能對大腸直腸癌風險有保護作用。2016 年研究評估運動與大腸癌風險的關聯，透過代謝當量 (Metabolic equivalent of task, MET) 衡量運動強度，結果發現 MET 越高，罹癌風險越低。(Keum et al., 2016)

另一項研究收集 BMI 與家族史資訊以評估運動與大腸直腸癌風險之統合分析，研究證實運動可降低大腸直腸癌罹癌風險，且相較於 BMI 低者，運

動也可降低 BMI 高者之大腸直腸癌風險。(Shaw et al., 2018)

代謝症候群主要包括肥胖、高血糖、高血脂、高密度膽固醇偏低與血壓升高，一般與心血管或腦血管疾病風險相關，不過肥胖與糖尿病等代謝因子也是大腸直腸癌的風險因子。

臺灣 1999 年於基隆開始推行社區型整合式篩檢計畫，該計畫除了大腸癌篩檢外也同時篩檢其他慢性疾病，如糖尿病或高血壓，結果證實代謝症候群患者罹患大腸腺瘤的風險顯著提高 43% ($RR=1.43$, 95% CI : 1.01-2.02)。(T. H. Chen et al., 2004)

2007 年邱等人透過完整接受大腸鏡檢查個案 4,277 名，使用國家膽固醇教育計畫成人治療小組與修改後的亞洲標準定義代謝症候群，評估代謝症候群與大腸腫瘤 (Colorectal neoplasm) 之關聯；與無代謝症候群者比較，具有代謝症候群者大腸腫瘤風險增加 1.35 倍 ($OR=1.35$, 95% CI : 1.05-1.73)，此外也發現在不同部位其風險也有差異，遠側大腸勝算比為 0.96 (95% CI : 0.67-1.38)，近端大腸勝算比為 1.62 (95% CI : 1.14-2.30)。(Chiu et al., 2007)

綜上所述，大腸直腸癌為各國常見癌症之一，且經濟發展程度較高之國家其大腸直腸癌發生率也較高；不同性別也有明顯差異，男性發生率通常較女性高。研究證實吸菸、肥胖、代謝症候群等風險因子會顯著增加大腸直腸癌發病。有鑑於此，規劃公共衛生預防策略，以降低大腸直腸癌死亡率十分重要。

2.3 近端大腸癌 (Proximal Colon) 及遠側大腸癌 (Distal Colon) 篩檢效益

邱等人針對不同解剖部位之大腸癌期別進行分析，將大腸脾彎曲處以上為近端大腸 (Proximal Colon)，其餘部分至直腸部分為遠側大腸 (Distal Colon)。(H.-M. Chiu et al., 2021)

未篩組中近端大腸共有 7,175 例癌症，癌症期別 0 期至 IV 期分別為 5.7% (410 例)、8.9% (638 例)、28.9% (2,076 例)、28.0% (2,006 例)、28.5% (2,045 例)；篩檢組中近端大腸共 5,652 例癌症，癌症期別 0 期至 IV 期分別為 20.5% (1,160 例)、21.4% (1,211 例)、21.3% (1,205 例)、22.1% (1,241 例)、14.6% (827 例)。

未篩組中遠側大腸共有 2,2916 例癌症，癌症期別 0 期至 IV 期分別為 7% (1,953 例)、17.1% (3,919 例)、22.6% (5,170 例)、29.4% (6747 例)、23.9% (5,487 例)；篩檢組中近端大腸共 15,933 例癌症，癌症期別 0 期至 IV 期分別為 23% (3,665 例)、28.4% (4,524 例)、14.7% (2,346 例)、23.2% (3,699 例)、10.7% (1,699 例)。

該研究證明遠側大腸的篩檢效益比近端大腸效果更大，晚期大腸癌發生率降低遠側大腸與近端大腸分別為 39% ($aRR=0.61$, 95% CI 0.58-0.64)、16% ($aRR=0.84$, 95% CI 0.77-0.92)，死亡率降低 44% ($aRR=0.56$, 95% CI 0.53-0.69)、28% ($aRR=0.72$, 95% CI 0.66-0.80)。這個結果可能與大腸鏡檢查品質以及 FIT 對於近端大腸病灶得偵測敏感度較低有關。大腸鏡檢查不完整或是大腸鏡檢查品質不佳可能使發生率與死亡率降低效益減少，且會增加近端大腸間隔癌的可能；由於 FIT 主要以偵測糞便中血紅蛋白濃度，而近端部位的病灶其血紅蛋白可能因糞便在腸道移動的過程中降解，進而影響 FIT 對於近端部位病灶的偵測能力。另外，也有研究指出，FIT 對於 Advance adenoma 敏感度較低，而近端大腸主要存在容易被 FIT 或大腸鏡檢查遺漏之 Sessile serrated adenoma/polyps 之病灶。

2.4 大腸直腸癌疾病自然史評估方法

篩檢計畫中，間隔癌的發生將影響篩檢效益，間隔癌除了可能因偽陰性篩檢結果發生，也可能是因腫瘤生長速度較快，導致在篩檢間隔內發生癌症。



大腸直腸癌疾病進程中，會經歷由無疾病階段、腺瘤、癌症等階段，最終觀察到死亡，其過程中有許多階段是無法直接觀察，例如從無疾病階段轉移至腺瘤出現階段，為了評估篩檢計畫對於大腸直腸癌的影響，透過馬可夫模型 (Markov model)將疾病進展與疾病自然史結合，為一種可行之方式。

任等人透過馬可夫模型評估 COVID-19 對於大腸直腸癌篩檢之延遲所造成之影響；任等人開發之馬可夫模型，可模擬 FIT 篩檢過程，包括邀請率、參與率及篩檢率，以及整體篩檢族群陽性率、陽性個案轉介率、等待轉介時間與轉介完成率對於大腸直腸癌疾病自然史影響；疾病自然史包括一系列的大腸直腸腫瘤狀態，大腸直腸癌自然病史可由七階段之馬可夫模型建立，包括正常狀態、小腺瘤(小於 1 公分之腺瘤)、大腺瘤 (大於 1 公分之腺瘤)、早期 PCDP (第一期與第二期)、晚期 PCDP (第三期與第四期)、早期 CP、晚期 CP。根據大腸直腸癌階段，可能會發生大腸癌死亡之最終結果，或是在過程中發生其他原因之死亡結果。

篩檢服務架構、過程與結果與大腸直腸癌自然史模型結構，可轉化為馬可夫決策樹，以評估 COVID-19 疫情爆發對篩檢服務之影響；前述所列之大腸直腸癌七階段為決策樹骨幹，符合篩檢服務之個案，進入篩檢服務時都屬於疾病自然史中七種狀態其中之一；個案參與兩年一次之 FIT 篩檢，患有大、小腺瘤與早期、晚期 PCDP 之個案會依 FIT 對不同類型之腫瘤敏感度而被偵測為陽性，陽性個案將被安排至醫療院所接受確診大腸鏡檢查，在檢查中進行息肉切除，後續依據目前臨床追蹤指引進行監測。

結果證實，COVID-19 疫情導致篩檢延遲與定期每兩年進行一次 FIT 篩檢相比，延遲 0.5 年、1 年、1.5 年、2 年所造成晚期癌症風險分別增加 25%(RR = 1.25, 95% CI:=1.18-1.33)、29% (RR=1.29, 95% CI:1.21-1.37)、34% (RR=1.34,

95% CI:1.26-1.42)、39% (RR=1.39, 95% CI:1.31-1.48)；死亡風險分別增加 26% (RR=1.26, 95% CI:1.21-1.32)、28% (RR=1.28, 95% CI:1.22-1.33)、29% (RR=1.29, 95% CI:1.24-1.34)、30% (RR=1.30, 95% CI:1.25-1.36)。 (Jen et al., 2021)



2.5 臺灣大腸癌篩檢計畫沿革及效益評估



大腸癌篩檢可分為伺機性篩檢與組織性篩檢，主要差別在於是是否有條件限制參與對象及如何邀請符合條件對象進行篩檢，篩檢結果陽性個案將建議進行後續轉介確認診斷，以及癌前病變與早期偵測到之癌症是否有後續追蹤定期監測。

組織性篩檢計畫中，第一步是邀請目標族群參與篩檢，第二步是將篩檢結果為陽性之個案轉介進行確認檢查或診斷，並於追蹤期間對篩檢偵測到之腫瘤進行監測、或對早期偵測到之癌症提供治療；這些過程需透過不同公共衛生部門與醫療單位完成，也需要資金支持提供篩檢所需要之基礎設施。

與組織性篩檢不同，伺機性篩檢主要依靠臨床醫師對於次段預防的認知，或是患者在就醫時對其進行篩檢，刺激性篩檢缺乏邀請目標族群參加的階段，且通常沒有固定的篩檢工具、篩檢間隔時間規定，後續也缺乏組織化的轉介診斷體系與監測系統，很難評估其效益。

1992 年至 1997 年間，行政院衛生署(今衛生福利部)推出多中心癌症篩檢計畫(TMACHS)，以醫院作為執行單位，邀請 17 家醫學中心或區域醫院共同參與，針對高風險族群早期肝癌、大腸癌、乳癌為篩檢服務對象。就大腸癌計畫而言，以家族病史及個人病史等不同的風險之族群，邀請無大腸癌症狀的民眾參加，以糞便潛血檢查 (Fecal occult blood test, FOBT) 或鋇劑灌腸攝影檢查(Double-Contrast Barium Enema)結合大腸鏡檢查 (Colonoscopy)之單次篩檢 (one-shot)，評估結果顯示，與未篩檢對照組比較，每年一次、兩年一次、三年一次篩檢，分別可以降低 26% (95% CI : 0-50%)、23% (95% CI : 0-48%)、21% (95% CI : 0-47%)，大腸直腸癌死亡率。(Chen et al., 1999)

1999 年起，基隆推動「社區整合式篩檢」(Keelung community-based integrated screening, KCIS)(T. H. Chen et al., 2004)，針對 30 歲至 79 歲設籍於

基隆之居民提供篩檢服務，服務內容區分為腫瘤性疾病 (Neoplastic disease) 與非腫瘤性慢性疾病 (Nonneoplastic chronic diseases)，腫瘤性疾病包含子宮頸抹片、乳癌、大腸癌、口腔癌、肝癌，非腫瘤性慢性疾病包含糖尿病、高血壓、高血脂等代謝症候群之篩檢服務；其中大腸癌對 50 歲至 79 歲民眾提供糞便潛血檢查 (Fecal occult blood test, FOBT)，檢查結果為陽性之個案再進行內視鏡檢查，若有發現息肉，則依據美國 USPSTF (U.S. Preventive service task force) 指引進行追蹤。

結果發現大腸癌篩檢陽性率約為 6%，陽性個案轉介完成率約 67.1%，共發現 33 位大腸癌個案；更進一步發現，與無症狀者相比，罹患大腸癌之個案至少患有一種慢性疾病；具高血壓病史者與大腸癌腫瘤相關。

該研究成功建立一個社區多項疾病篩檢模型，透過此模型可有效提升篩檢參與率；而透過多項疾病整合式篩檢計畫，能建立多個風險因子與多種疾病間之關聯性；並且藉由已建立疾病篩檢計畫，可評估實證基礎較少的疾病篩檢計畫之效益。不過，雖然整合式篩檢計畫效率高，但仍需長時間追蹤發生率與死亡率是否降低，以評估其效益。

基於社區整合式篩檢經驗，衛生福利部國民健康署自 2004 年起開始推動社區到點篩檢服務 (Out-reach screening service)，全面補助 50 至 69 歲民眾每兩年一次免疫化學法糞便潛血篩檢檢查 (Fecal immunochemical test, FIT)，篩檢採用國際認證兩種試劑，包括 OC-Sensor (Eiken Chemical Co. Ltd, Tokyo, Japan) 與 HM-JACK (Kyowa Medex Co. Ltd, Tokyo, Japan)，並於 2015 年起 Kyowa 公司導入二代檢驗試劑 HM-JACKarc 以取代原本的 HM-JACK 試劑。隨著國人平均餘命延長，2013 年修改年齡上限至 74 歲。目前，臺灣大腸癌篩檢計畫邀請 50 至 74 歲民眾參與篩檢，亦追蹤其後續陽性轉介與確診等結果，以達到『早期發現，早期治療』的效果。

篩檢計畫執行初期，為有效管理各地衛生局的篩檢資料，因此成立篩檢資料監測中心，以紙本申報方式收集篩檢及陽性個案，及陽性個案其後續轉介狀況。為提升資料處理能量與時效性，自 2002 年推行媒體申報作業、建置網路版申報、回饋系統，更成為現今篩檢資料庫系統之起源；2004 年開始完備網路版申報及回饋系統，逐步建構篩檢資料庫系統，以提供決策者包括篩檢成效評估、資料及品質監測、組織因素、共病因子、流行病學及社會與經濟學之考量。

所有篩檢相關數據，包含人口統計學資料、FIT 檢查結果、內視鏡檢查結果與病理報告結果都儲存於中央資料庫，透過與臺灣癌症登記資料和死亡統計資料可得到參加者後續癌症診斷結果，期別資訊以及死因代碼，以此評估大腸癌之發生率與死亡率。

邱等人於 2015 年評估大腸癌篩檢效益相關結果 (Chiu et al., 2015)，研究比較 2004 年至 2009 年參與篩檢與未參與篩檢之大腸癌死亡率。此期間內共有 1,160,895 名 50 歲至 69 歲個案參與篩檢，平均追蹤 3.09 年，最長追蹤 6 年，涵蓋率約 21.4%，重複篩檢率為 28.3%；整體陽性率為 4%，後續篩檢陽性率為 3.8%；首次參與篩檢且結果為陽性之個案，約有 80% 完成確診檢查，其中 85.5% 以大腸鏡檢查，利用乙狀大腸鏡與鋇劑攝影完成檢查或其他方式完成確診者分別為 10.2% 與 4.3%。透過確診檢查找到 19,398 例腺瘤 (Adenoma)，5,500 位進行性腺瘤 (Advance adenoma) 以及 2,805 例大腸癌 (Colorectal cancer) 個案，陽性預測值分別為 39.5%、11.7%、6.1%，首次篩檢偵測率腺瘤為 16.0%、進行性腺瘤為 11.7%、大腸癌 6.1%。大腸癌 2,805 例中有 77% 具有完整期別資訊，其中 13.3%(286 例) 為原位癌，I 期至 IV 期分別為 35.5%(755 例)、20.9%(450 例)、23.7%(510 例)、7.2%(154 例)。根據癌症登記資料，未參與大腸癌篩檢者有 18,639 例大腸癌個案，其中約 5%(928 例) 為原位癌，I 期至 IV 期分別為 16.1%(3,009 例)、17.8%(5,188 例)、31.8%(5,935

例)、19.2%(3,579 例)。以期別資訊而言，兩組間晚期癌症的明顯差異值得注意。

篩檢組與未篩檢組間死亡率累積死亡風險比為 0.38 (95% CI: 0.35-0.42)，顯示大腸癌死亡風險降低 62%；而在調整自我選擇偏差(self-selection bias)後，大腸癌死亡率風險比為 0.90 (95% CI : 0.84-0.95)，表示在 21.4%涵蓋率下大腸癌死亡率降低 10%。

研究證實了 FIT 相較於 gFOBT (Guaiac-based Faecal Occult Blood Test) 之優勢，與 gFOBT 相比，FIT 具有較高得敏感度與特異度，且無需限制飲食。此外，在篩檢涵蓋率為兩成左右即可下降 10%大腸癌死亡率，若涵蓋率可提升至 60%，則預估可下降 36%大腸癌死亡率。

邱等人於 2021 年發表了長期追蹤成果 (H.-M. Chiu et al., 2021)，針對 2004 年至 2009 年參與大腸癌篩檢之個案，追蹤至 2014 年，評估暴露篩檢與未暴露篩檢之大腸癌死亡風險。

本篇研究共納入 5,417,699 名個案，其中 3,067,853 名至少參與一次大腸癌篩檢，約有 2,349,846 名個案未接受任何大腸癌篩檢，篩檢涵蓋率為 56.6%。參與篩檢個案中約 52.3%參加了兩次或以上的大腸癌篩檢；首次參與篩檢之陽性率為 7.0%，重複篩檢之陽性率為 6.4%；首次篩檢組轉介完成率為 70%，其中 89.9%接受大腸鏡檢查；重複篩檢組中轉介完成率為 62.5%，其中 94.1%接受大腸鏡檢查；首次篩檢組確診 6,756 例癌症，偵測率為 2.20‰，重複篩檢組 3,118 例，偵測率為 1.94‰。

未篩檢組共有 30,295 例癌症，癌症期別 0 期為 6.7%(2,019 例)，I 期至 IV 期分別為 15.1%(4,587 例)、24.2%(7,317 例)、29.1%(8,812 例)、25%(7,560 例)；篩檢組共有 15,933 例癌症，癌症期別 0 期為 22.3%(4,832 例)，I 期至 IV 期分別為 26.6%(5,753 例)、16.5%(3,585 例)、22.9%(4,962 例)、11.7%(2,533 例)。

研究結果證實經調整篩檢參與率、轉介確診率及大腸鏡檢查品質後，晚期大腸癌（II 期以上）發生率顯著降低 34% ($aRR=0.66$, 95% CI 0.63-0.70)，死亡率降低 40% ($aRR=0.60$, 95% CI 0.57-0.64)，這項結果與 2015 年評估結果相近。



第三章、材料與方法



本研究以元宇宙數位雙胞胎基礎之架構下評估篩檢效益，進一步結合社區篩檢之大腸直腸癌篩檢計畫進行增加個人特徵收集後不同部位大腸直腸癌風險鑑別分析，再後，以此為基礎進行危險分層篩檢建議 (Risk-based Screening)效益評估。所用資料為為前瞻性世代研究，針對臺灣大腸直腸癌篩檢計畫參與世代進行分析，先估計不同部位大腸直腸癌疾病進展及 FIT 篩檢工具表現，並利用機器學習方法進行不同部位大腸癌發生的個人特徵萃取，以建立數位雙胞胎所需個人化資訊。

3.1 元宇宙數位雙胞胎設計

圖 3.1 為本研究採用之元宇宙數位雙胞胎設計整體架構，進行步驟有六：

1. 真實世界資料 (Real World Data, RWD)

本研究始於自以族群為基礎之大腸直腸癌 FIT 篩檢計畫所獲得之真實世界資料 (RWD)。圖 3.1 左上角顯示 4 位受試者代表不同的篩檢偵測模式，包括盛行篩檢個案 (受試者 1)、後續篩檢個案 (受試者 2)、間隔癌個案 (受試者 3) 和拒絕篩檢個案 (受試者 4)。

2. 人工智慧馬可夫生成引擎 (AI Markov Machine Learning Engine)

將上述 RWD 利用 Markov 機器學習算法得到與大腸直腸腺瘤和癌症疾病進展自然病程的相關參數。

3. 生成虛擬世界資料 (Virtual world data, VWD)

產生虛擬世代並利用前項步驟所得之參數產生每個個案從出生後所開始經歷大腸直腸腺瘤和癌症之自然病程，並形成虛擬世界數據 (VWD)。

4. 橋接合成真實世界數位分身 (Digital Twin)

將 RWD 和 VWD 進行橋接，創建具備大腸直腸腺瘤和癌症自然病程完整信息的 RWD 數位雙胞胎。

5. 介入評估研究設計

針對 RWD 數位分身進行“假設性”隨機臨床分派研究，針對不同的介入措施，包括篩檢間隔、篩檢起始年齡和風險分層的篩檢政策進行沈浸式隨機臨床分派研究 (Immersive Randomized Controlled Trial)，如圖 3.1 所示，以不同篩檢間隔為例，我們可將 RWD 數位分身分成 1 年、2 年、3 年與不篩檢，以進行不同篩檢間隔計畫下相對於不篩檢所能得到的效益。

6. 效益評估

針對不同介入措施組合，對降低晚期大腸直腸癌發生與大腸直腸癌死亡之成效。

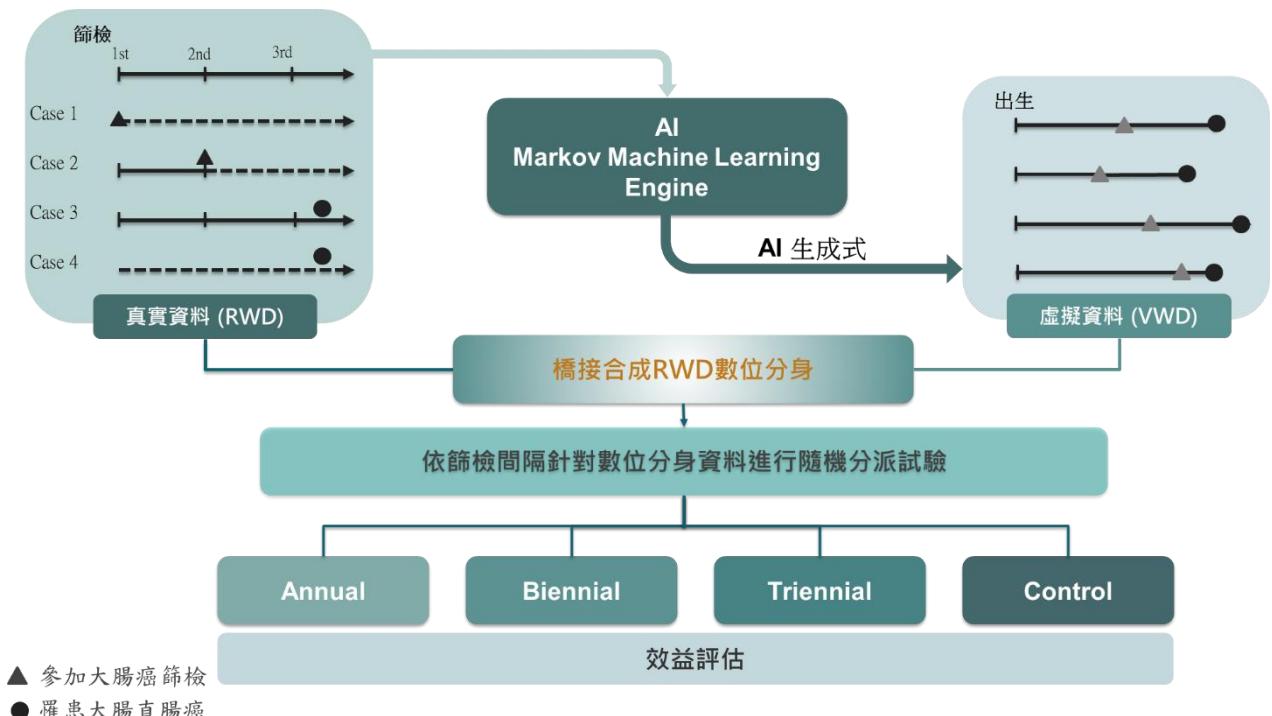


圖 3.1、數位雙胞胎設計用於元宇宙評估模型



3.2 真實世界資料設計與收集

3.2.1 臺灣大腸直腸癌篩檢

臺灣大腸直腸癌篩檢計劃始自 2004 年起，針對 50 至 69 歲民眾(自 2012 年起，篩檢年齡延長至 74 歲) 提供每兩年一次的糞便免疫化學法試驗(Fecal Immunochemical Test, FIT)，並於 2012 年起，將篩檢年齡上限展延至 74 歲)。參加篩檢個案若 FIT 篩檢結果為陽性，則進一步將個案轉介至合格醫療院所進行進一步確診，主要建議確診方法為大腸鏡檢查(Chiu et al., 2015; H. M. Chiu et al., 2021)

歷次篩檢紀錄(包括篩檢日期、FIT 試劑廠牌、FIT 定量數值)、陽性者之轉介確診紀錄(包括轉介日期、轉診醫療院所、檢查方式與確診結果)均由國民健康署癌症資料監測中心進行系統性資料維護。為確定參加篩檢個案在篩檢計畫之外的確診情形，國民健康署定期追蹤參與篩檢個案所有大腸癌發生情形與大腸癌死亡狀況，本研究所有大腸癌確診及死亡分別以國家癌症登記與死亡登記認定為之。

為了比較不同解剖部位與間隔癌 (Interval cancer, IC)的分期差異，位於脾曲以上的大腸部位被視為近端大腸 (Proximal Colon)，其餘部分和直腸被視為遠側大腸(Distal Colon)，分期部分使用 AJCC 分期系統將 II 期或更嚴重的情況分類為晚期(進階)大腸直腸癌。

所有篩檢紀錄可依篩檢結果區分為

- A. 無病個案 (Disease-free)：FIT 陰性或陽性且接受轉介者未有大腸癌確診紀錄個案
- B. 篩檢偵測大腸腺腫個案 (Screen-detected Adenoma)：FIT 陽性且接受轉介並經診斷為大腸腺腫個案
- C. 篩檢偵測大腸癌個案 (Screen-detected CRC)：FIT 陽性且接受轉



- D. FIT 間隔癌個案 (Interval Cancer)：FIT 陰性但在下次接受 FIT 篩檢之前被確診為大腸癌
- E. 大腸鏡檢後大腸癌個案 (Post-colonoscopy CRC, PCCRC)：FIT 陽性且接受轉介診斷結果非大腸癌但卻於後續被確診為大腸癌個案
- F. 陽性未轉介大腸癌個案：FIT 陽性未接受轉介並於後續被確診為大腸癌個案
- G. 拒絕篩檢大腸癌個案 (Refuser CRC)：未參加 FIT 篩檢但後續被確診為大腸癌個案

上述 B 與 C 為篩檢偵測個案 (Screen-detected)，由於篩檢計畫服務對象限定為第一次篩檢之前未有臨床大腸直腸癌診斷者為對象，會產生存活分析所謂左側截切 (Left Truncation) 現象，因此本研究進一步針對篩檢偵測個案分為：若個案在第一次參與 FIT 篩檢被偵測的盛行個案 (Prevalence screening detection, PSD)，及後續參與篩檢後被偵測的後續個案 (Subsequence screening detection, SSD)。

上述 D, E, F 為篩檢計畫內之臨床偵測個案 (Clinically detected)，在篩檢時因不同原因(D：FIT 陰性；E：大腸鏡未偵測出；F：未接受轉介)未在當次篩檢被偵測診斷，而在後續因出現臨床症狀就醫確診之大腸癌個案。

圖 3.2 顯示本研究不同偵測模式大腸癌個案示意圖，ICs(1)代表那些在第一次篩檢時結果正常 (□)，但在第二次篩檢前出現大腸癌症狀 (▲)；ICs(2)代表那些在第一次篩檢 (□) 及後續篩檢 (○) 的結果均為正常，但在 2014 年下一次篩檢前出現症狀 (▲)。在符合篩檢年齡組中，患有大腸直腸癌但未參加篩檢的人被視為在研究期間拒絕篩檢 (Refuser) 的個案。

本研究使用的資料為 2004-2022 年參加篩檢紀錄，癌症個案與死因確

認追蹤至 2020 年年底。



定期篩檢設計						用於分析五階段馬可夫模型之數據		
						轉移時間	偵測模式	AJCC 期別
						首次篩檢年齡	PCDP	早期
PSD	■							
SSD 1	□	●				2年	PCDP	早期
SSD 2	□	○	○		●	2年	PCDP	早期
IC 1	□	▲				2004年起TSLN	CP	晚期
IC 2	□	○			○ ▲	2012年起TSLN	CP	晚期
Refuser	◇			◆		罹癌時間	CP	晚期

PSD：盛行偵測個案 (Prevalence Screening-detected)

灰色代表罹患大腸直腸癌

SSD：發生偵測個案 (Subsequence Screening-detected)

IC：間隔癌 (Interval cancer)

Refuser：拒絕篩檢個案

TSLN：距上次篩檢陰性時間 (Time since last negative screen)

圖 3.2、前瞻性定期篩檢設計與使用五階段馬可夫模型分析數據

3.2.2 社區整合式篩檢

除了臺灣大腸直腸癌篩檢資料之外，本研究亦納入臺灣某社區自 2005 年持續進行之社區整合式篩檢資料，以檢視增加納入個人特徵後對鑑別不同部位大腸癌之影響。該整合式篩檢資料之架構可參考(T. H. H. Chen et al., 2004)，其篩檢服務中結合了五大癌症（大腸直腸癌、乳癌、子宮頸癌、口腔癌與肝癌)與三項重要慢性病（血壓、血糖、血脂），為了增加民眾的可近性與便利性，每年定期辦理外展式的到點篩檢服務，使參加者在同一時間同一場次可同時接受衛生局所提供的健康服務，並在服務後的 2 週後於相同地點發放檢查報告，並提供醫療專業諮詢服務。

參加篩檢民眾接受身體測量（身高、體重、腰圍及臀圍）、血壓測量、血液生化檢查及尿液檢查，並以結構式問卷詢問個人生活習慣（吸

菸、飲酒、嚼食檳榔、運動習慣)、教育程度、飲食習慣(蔬菜、肉類、咖啡、水果、牛奶、蛋、豆類等攝食頻率)、是否會接受其他健康檢查、對常見慢性病健康識能、個人病史與家族病史等。



3.3 統計分析

3.3.1 多階段大腸直腸癌疾病進展

為了估計從無疾病狀態、經過臨床症前期 (Preclinical Detectable Phase, PCDP)，最終到臨床階段 (Clinical Phase, CP)的疾病自然史轉移率，之前在大腸癌和乳癌研究中使用過的三階段馬可夫模型估計進入PCDP 的發病率和從 PCDP 到 CP 的轉移率(Chen et al., 1999; T. H. H. Chen et al., 2004)。

值得注意的是，後者是無法直接觀察，因醫療措施介入，當大腸直腸癌在 PCDP 中被檢測到時，則無法觀察後續轉移機率。三階段的模式型顯示如下：

$$\text{Disease free} \xrightarrow{\lambda_1} \text{PCDP} \xrightarrow{\lambda_2} \text{CP}$$

該模式狀態空間為 $\Omega = \{1: \text{Disease - free}, 2: \text{PCDP}, 3: \text{CP}\}$ ，在馬可夫過程 (Markov Process)的前提之下，上述模式由兩個瞬間轉移速率參數 λ_1 與 λ_2 所主宰，分別代表從無疾病狀態至 PCDP 之發生率，與從 PCDP 至 CP 之轉移速率，此 2 參數可以以轉移速率矩陣表示如下，

$$\begin{matrix} & & 1 & 2 & 3 \\ 1 & \left[\begin{matrix} -\lambda_1 & \lambda_1 & 0 \\ 0 & -\lambda_2 & \lambda_2 \\ 0 & 0 & 0 \end{matrix} \right] \\ 2 & & & & \\ 3 & & & & \end{matrix}$$

在某段時間間隔之間的轉移機率可以以下列轉移機率矩陣表示，

$$P = \begin{matrix} 1 & 1 & 2 & 3 \\ 2 & \left[\begin{matrix} P_{11}(t) & P_{12}(t) & P_{13}(t) \\ 0 & P_{22}(t) & P_{23}(t) \\ 0 & 0 & 1 \end{matrix} \right] \\ 3 & \end{matrix}$$

其中



$$P_{11}(t) = e^{-\lambda_1 t}$$

$$P_{12}(t) = \frac{\lambda_1(e^{-\lambda_1 t} - e^{-\lambda_2 t})}{(\lambda_2 - \lambda_1)}$$

$$P_{13}(t) = 1 - \frac{\lambda_2 e^{-\lambda_1 t}}{\lambda_2 - \lambda_1} + \frac{\lambda_1 e^{-\lambda_2 t}}{\lambda_2 - \lambda_1}$$

$$P_{11}(t) = e^{-\lambda_2 t}$$

$$P_{12}(t) = 1 - e^{-\lambda_2 t}$$

以上由轉移機率矩陣乃利用 Kolmogorov equation (Cox, 1965)，條件於 $P(0) = I$ 之下所推導出

$$\frac{d}{dt} P(t) = Q_t P(t) \quad t \geq 0$$

以此為基礎，依據不同篩檢偵測模式推導其對應之概似函數，並利用蒙地卡羅馬可夫模擬進行得到各參數之事後分佈。

本研究進一步將三階段馬可夫模型擴展到五階段，將上述 PCDP 與 CP 依 AJCC 分期分為第 1 期與第 2 期以上，並根據近端部位和遠側部位大腸識別從早期 PCDP 到晚期 PCDP 再到 CP 的疾病進展，以深入了解不同解剖部位的疾病自然史。其模式型表示如下

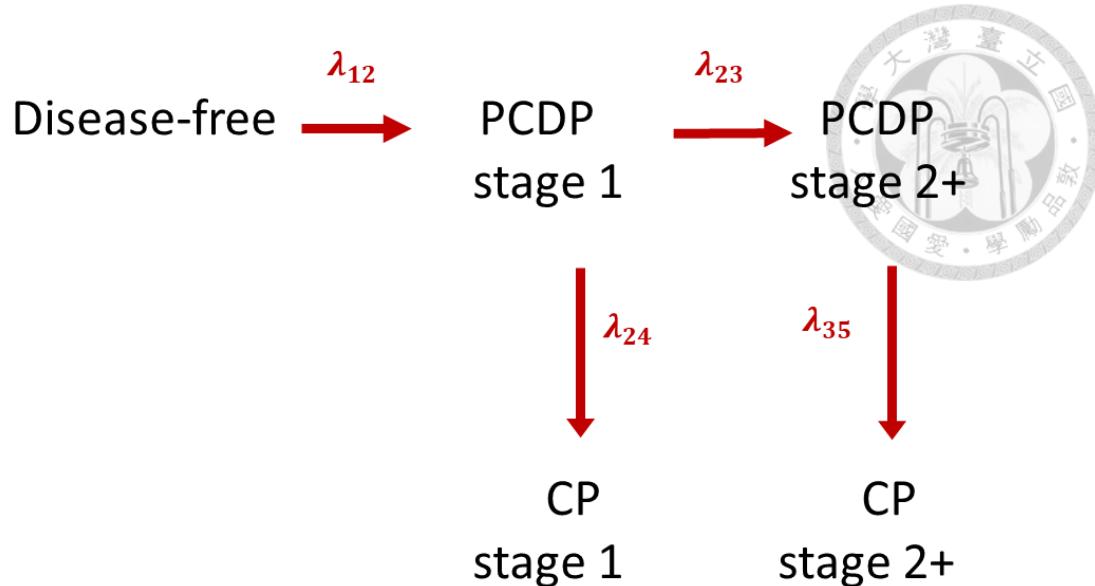


圖 3.3、五階段馬可夫大腸癌進展模式

五階段馬可夫模型之前在乳癌 (考慮腫瘤大小和節狀態等腫瘤屬性)和大腸直腸癌 (使用 Duke's 分期)的研究中使用過。早期和晚期大腸癌的 FIT 的特異度與敏感度也被納入考慮。為了估計發病率、進展率和敏感度的參數，需要知道相應的轉移時間 (首次篩檢的年齡、篩檢間隔或自上次觀察以來的時間)、偵測模式 (PCDP、CP)和 AJCC 分期 (早期、晚期)納入模型。上節圖 3.2 右側為不同偵測模式對應之五階段馬可夫模型所中使用的資料。

3.3.2 多階段大腸直腸腺腫--大腸直腸癌疾病進展

進一步考慮大腸癌癌前病變--大腸腺腫在癌症形成之角色，本研究利用八階段馬可夫過程進行估計，其模式型如下，

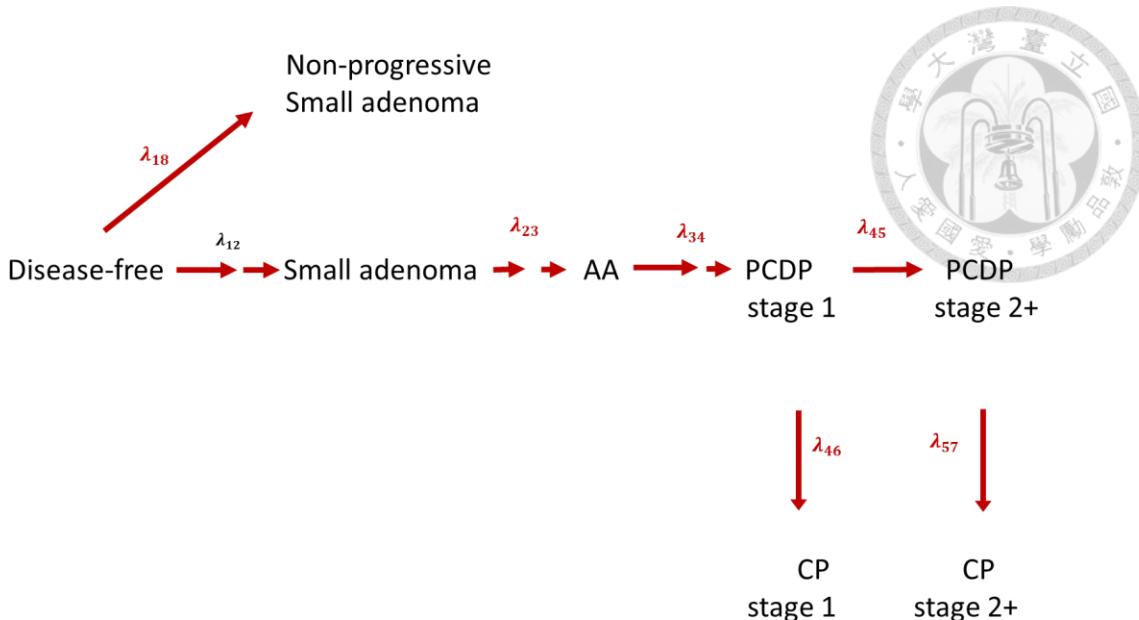


圖 3.4、八階段馬可夫大腸腺腫--大腸癌進展模式

3.3.3 隨機森林鑑別不同部位大腸直腸癌發生

本研究利用隨機森林(Random Forest)機器學習方法分類無病個案與不同部位大腸癌(遠側與近端)，我們將分析資料集隨機分成 70%訓練資料集 (Train Dataset)與 30%測試資料集 (Test Dataset)構建決策樹 (Decision Tree)，每個決策樹針對目標變項，從選取特徵中進行資料分類，以最大資訊獲得 (Information Gain)作為根部變項 (Root Variable)，並依被分類之分枝再進行下一個重要變項之選取及分類，直至決策樹葉部資料少於 5 為止，本研究亦設定決策樹最大層數為 10。此步驟重覆 1000 次，得到 1000 顆決策樹形成隨機森林，最後，綜合所有決策樹的預測結果，通過多數投票來確定最終分類。

由於每棵樹都是從訓練資料集的隨機子集中訓練而成，有助於提升模型的穩定性，同時減少過度擬合 (Overfitting)的風險。決策樹分析結果將以變數重要性 (Variable Importance Plot, VIMP)評估每個變項對模型預測的貢獻，各變項在決策樹出現愈接近根部愈重要，VIMP 以在

隨機森林中接近根部的平均表現比較各變項之重要性。並報導其接受者作業操作特徵曲線 (Receiver Operating Characteristic Curve, ROC Curve) 及曲線下面積 (Area Under the Curve, AUC)。

3.3.4 貝氏網絡模式鑑別不同階段大腸直腸腺腫--大腸直腸癌風險

本研究利用貝氏網路 (Bayesian Network) 模型進行不同階段大腸腺腫與大腸癌之分類，貝氏網絡模型透過有向無循環圖模型 (Directed Acyclic Graphic Model, DAG) 建構解釋特徵與目標變項之最佳關聯架構，以節點代表個人特徵，邊 (Edge) 表示變項之間的條件依賴 (Conditional Dependency) 關係。模型的構建過程包括確定網絡結構和估計條件概率分佈。

本研究以 Tree Augmented Naïve Bayesian Network (TAN) 方法進行，在 TAN 中，除了 Naïve Bayesian Network 之下使各特徵變項在目標變項存在之下形成條件獨立之外，另外允許各特徵變項之間具條件相依性，以增加模型彈性。

本研究利用 70% 訓練資料集進行模型訓練，估計每個變項的條件概率分佈，訓練好的貝氏網絡模型對 30% 測試資料集進行分類，並計算給定觀察資料之事後概率分佈 (Posterior Distribution)，據此得到個案發生不同部位大腸癌之風險。

3.3.5 元宇宙大腸癌篩檢數位分身效益評估

本研究將利用以上機器學習語法得到不同階段大腸腺腫、大腸癌之預測機率，利用累積風險函數轉換為相對分數並作族群內風險分層，以上述八階段馬夫鏈模式為基礎，族群風險中位數即為估計所得轉移速率，相對風險為轉移速率之相對倍率 (見圖 3.5)，據此生成虛擬世界

資料，並橋接合成真實世界數位分身，隨機分派為對照組、不同間隔篩檢組與風險分層大腸癌篩檢策略。

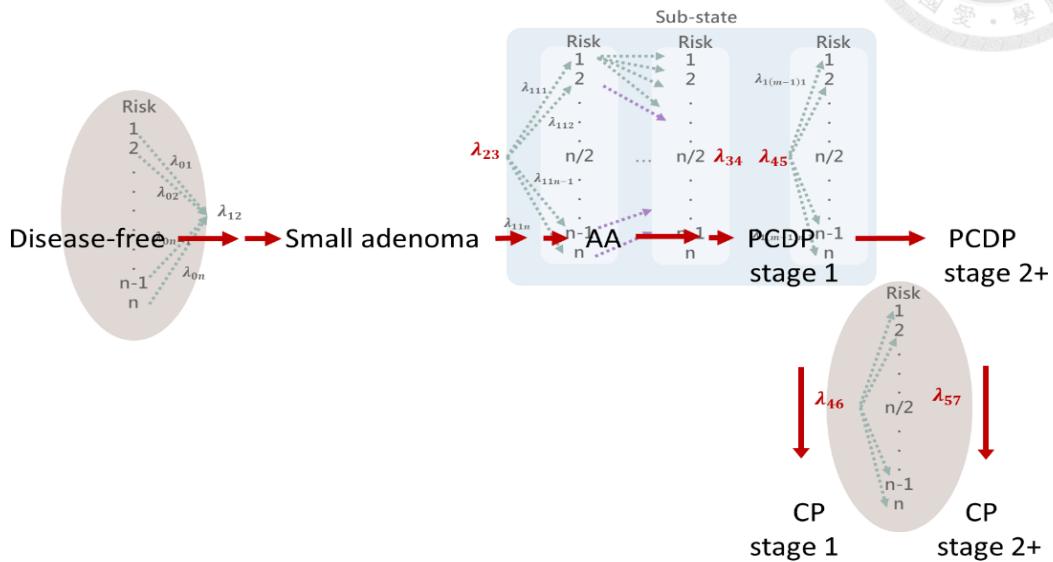


圖 3.5、危險分層嵌入八階段馬可夫大腸腺腫-大腸癌進展模式

風險分層在前 25%個案授於 1 年篩檢間隔、25-50%個案授予 2 年篩檢間隔、50-75%賦予 3 年篩檢間隔，最低風險的 25%個案授予 6 年篩檢間隔，比較此篩檢策略之下晚期大腸直腸癌個案，並與對照組進行比較。

所有統計分析都使用 SAS 9.4 版本軟體進行。

第四章、結果



4.1 臺灣大腸癌篩檢

4.1.1 描述性分析結果

表 4.1 顯示了曾參與全國篩檢計畫中具大腸癌診斷結果者依近端和遠側部位之描述性分析結果。近端大腸直腸癌個案更可能是女性、診斷年齡較高、間隔癌個案或後續臨床症狀癌症與 AJCC 分期晚期之個案。

表 4.1、臺灣大腸癌篩檢近端及遠側大腸癌描述性分析結果

	所有參與篩檢人次	大腸癌症人數		大腸癌症人數	
		近端部位		遠側部位大腸	
		(Proximal Colon)	(Distal Colon)		
	12,318,580		30,264		64,413
		N	%	n	%
性別					
	男性	5,189,497	42.13	14,557	48.10
	女性	7,129,083	57.87	15,707	51.90
篩檢年齡					
	50-54	3,263,975	26.50	4,820	15.93
	55-59	3,360,142	27.28	7,285	24.07
	60-64	3,140,421	25.49	9,328	30.82
	65-69	2,554,042	20.73	8,831	29.18
教育程度					
	無	660,616	9.60	1,017	10.47
	小學	1,145,106	16.64	2,064	21.25
	國中	1,288,377	18.73	1,908	19.64
	高中	1,819,214	26.44	2,322	23.91
	大學	1,180,673	17.16	1,511	15.56
	研究所	206,245	3.00	222	2.29
	拒答	580,227	8.43	669	6.89
篩檢試劑廠牌					
	Eiken	8,099,582	65.75	19,592	64.74
	Kyowa	3,359,552	27.27	9,714	32.10
	Kyowa2	858,209	6.97	955	3.16

	所有參與篩檢人次		大腸癌症人數 近端部位 (Proximal Colon)		大腸癌症人數 遠側部位 (Distal Colon)	
	12,318,580		30,264		64,413	
	N	%	n	%	n	%
Others	1,169	0.01	3	0.01	10	0.02
大腸癌家族病史						
無	10,571,439	87.17	25,227	83.36	54,575	84.73
有	764,316	6.30	2,734	9.03	4,827	7.49
不確定	791,639	6.53	1,483	4.90	3,269	5.08
其他癌症家族病史						
無	6,257,193	69.43	10,957	68.59	24,321	69.41
有	1,692,431	18.78	2,770	17.34	5,862	16.73
不確定	1,062,029	11.79	2,247	14.07	4,856	13.86
FIT 濃度 (ng/ml)						
1-19	11,103,836	90.14	17,061	56.37	31,605	49.07
20-49	667,647	5.42	3,137	10.37	6,397	9.93
50-99	203,827	1.65	2,060	6.81	4,699	7.30
100-149	102,326	0.83	1,220	4.03	2,941	4.57
150-249	110,539	0.90	2,804	9.27	7,370	11.44
250-449	49,738	0.40	1,296	4.28	3,653	5.67
>450	80,599	0.65	2,686	8.88	7,748	12.03
偵測模式						
篩檢偵測個案(PSD)			3,798	23.94	12,172	32.21
後續篩檢偵測案(SSD)			1,946	12.27	4,970	13.15
兩年內間隔癌			2,625	16.55	6,030	15.96
兩年後間隔癌			7,496	47.25	14,612	38.67
AJCC 分期						
Stage I			2,656	22.25	7,727	27.74
Stage II			2,431	20.37	4,539	16.29
Stage III			3,893	32.62	9,552	34.29
Stage IV			2,421	20.28	4,484	16.10
Not known			534	4.47	1,557	5.59

4.1.2 大腸直腸癌不同部位疾病進展病史

表 4.2 與表 4.3 為考慮到敏感度後侵襲性大腸直腸癌之三階段和

五階段疾病自然史。前者的結果顯示，遠側大腸直腸癌發生率為近端大腸直腸癌三倍，而近端大腸直腸癌從 PCDP 到 CP 的進展速度比遠側大腸直腸癌更快。五階段疾病自然史模型也觀察到了類似的結果。與遠側大腸直腸癌相比，近端大腸直腸癌在 PCDP 期間從早期到晚期以及從晚期 PCDP 到 CP 的進展速率更快。

進一步將大腸直腸腺腫階段納入疾病自然史模型，進一步估計腺瘤 (Adenoma) 進展至進行性腺瘤 (Advanced Adenoma, AA) 之速率。(表 4.4)

整體而言，遠側大腸腺腫發生率高於近端大腸，不過不論遠側部位或近端部位，從腺瘤轉移至進行性腺瘤轉移速率並無差異，不過近端大腸直腸癌在 PCDP 期間從早期到晚期以及從晚期 PCDP 到 CP 的進展速率比遠端大腸癌更快，與五階段疾病自然史模型估計結果相同。

表 4.2、大腸直腸侵襲癌近端及遠側疾病進展馬可夫三階段機器學習估計結果

	近端部位大腸 (Proximal Colon)		遠側部位大腸 (Distal Colon)	
	估計值	95% 信賴區間	估計值	95% 信賴區間
發生率 (per 100,000)	29.8	(29.5, 30.2)	74.2	(73.7, 74.7)
年轉移率				
PCDP → CP	0.47	(0.45, 0.49)	0.34	(0.33, 0.35)
Sensitivity PCDP (%)	0.87	(0.93, 0.87)	0.83	(0.81, 0.85)

表 4.3、大腸直腸侵襲癌近端及遠側疾病進展馬可夫五階段機器學習估計結果

	近端部位大腸 (Proximal Colon)		遠側部位大腸 (Distal Colon)	
	估計值	95% 信賴區間	估計值	95% 信賴區間
發生率 (per 100,000)	29.9	(29.5, 30.2)	74.4	(73.9, 74.9)
年轉移率				
Early PCDP → Late PCDP	1.10	(0.96, 1.27)	0.77	(0.71, 0.83)
Early PCDP → Early CP	0.14	(0.12, 0.16)	0.14	(0.12, 0.15)
Late PCDP → Late CP	0.53	(0.50, 0.56)	0.41	(0.39, 0.42)
Sensitivity of Early PCDP (%)	59.49	(51.4, 69.11)	65.1	(60.1, 70.09)
Sensitivity of Late PCDP (%)	82.27	(78.73, 85.55)	79.1	(77.17, 81.02)

表 4.4、大腸直腸腺腫-侵襲癌近端及遠側疾病進展馬可夫八階段機器學習估計結果

	近端部位大腸 (Proximal Colon)		遠側部位大腸 (Distal Colon)	
	估計值	95% 信賴區間	估計值	95% 信賴區間
發生率 (per 100,000)				
大腸腺腫	48.33	(48.13, 48.54)	145.00	(144.40, 145.61)
非進展性大腸腺腫	6.59	(6.50, 6.67)	19.76	(19.51, 20.00)
年轉移率				
Adenoma → AA	0.33	(0.33, 0.34)	0.33	(0.33, 0.34)
AA → Early PCDP	0.23	(0.23, 0.24)	0.19	(0.19, 0.20)
Sensitivity of AA	0.23	(0.18, 0.28)	0.32	(0.27, 0.37)
Early PCDP → Late PCDP	1.10	(0.96, 1.27)	0.77	(0.71, 0.83)
Early PCDP → Early CP	0.14	(0.12, 0.16)	0.14	(0.12, 0.15)
Late PCDP → Late CP	0.53	(0.50, 0.56)	0.41	(0.39, 0.42)
Sensitivity of Early PCDP (%)	59.49	(51.4, 69.11)	65.1	(60.19, 70.09)
Sensitivity of Late PCDP (%)	82.27	(78.73, 85.55)	79.1	(77.17, 81.02)

4.1.3 隨機森林 (Random Forest) 分析結果



圖 4.(a)利用隨機森林鑑別近端大腸直腸癌個案與正常個案比較所得到結果，依重要變項排序為 FIT 濃度、篩檢年齡、FIT 廠牌、教育程度、其他癌症家族史、大腸癌家族史與性別。圖 4.2 (b)-(c) 為遠側大腸直腸個案與正常個案以及遠側與近端癌症病人比較也觀察到類似的結果。

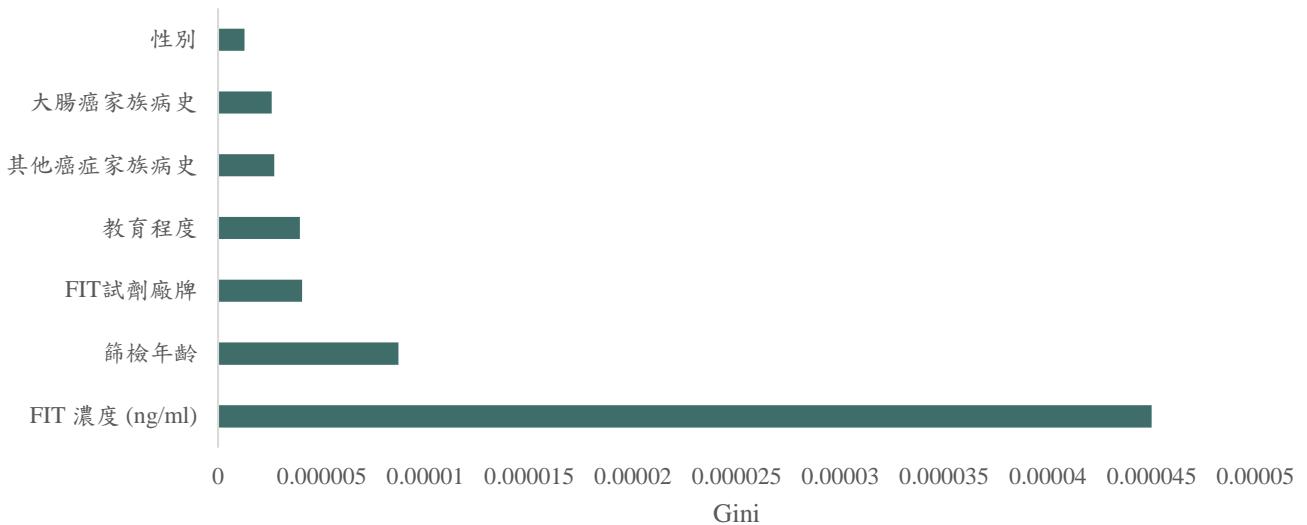


圖 4.1 (a)、近端大腸與正常個案之重要變項排序

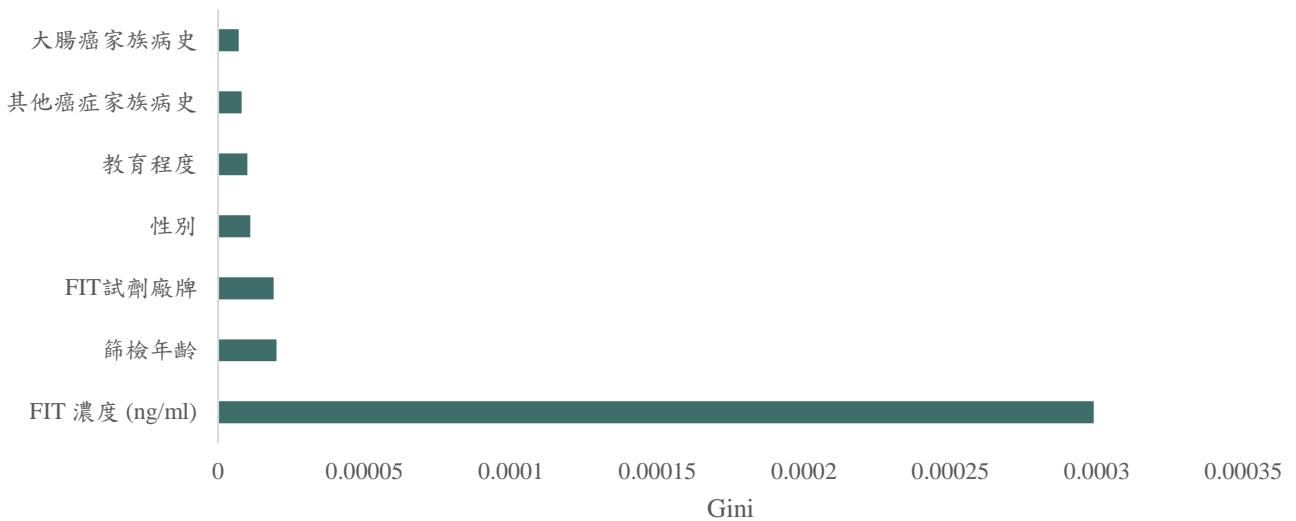


圖 4.1 (b)、遠側大腸與正常個案之重要變項排序

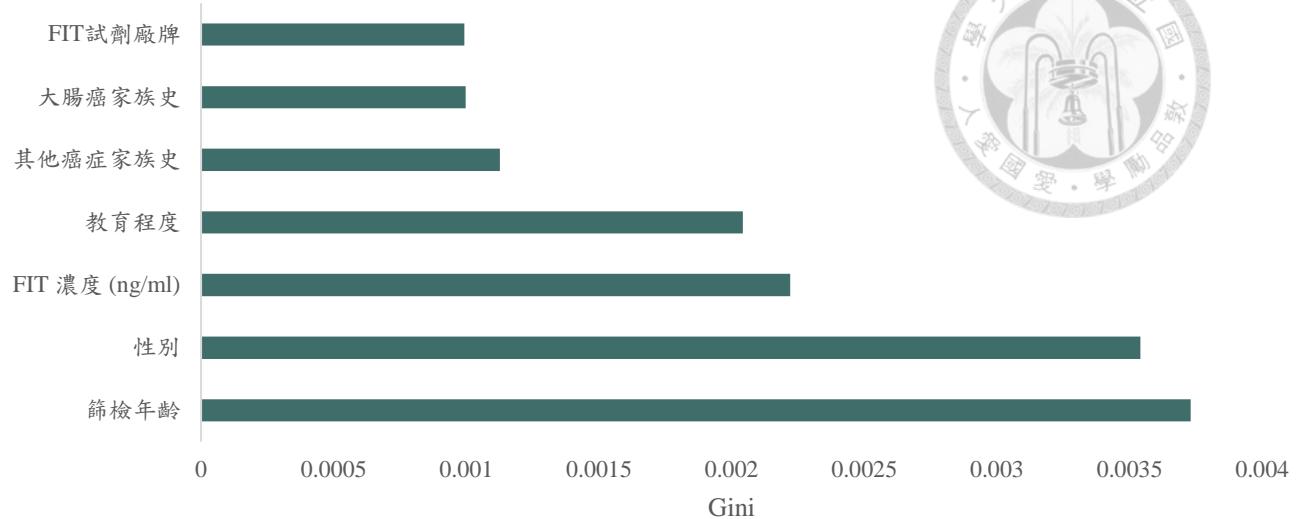


圖 4.1 (c)、近側大腸與遠側大腸個案之重要變項排序

圖 4.1(d)-(f) 則為所對應 ROC 分析, AUC 結果與正常個案比較為 0.68, 遠側部位大腸與近端部位大腸比較為 0.61。

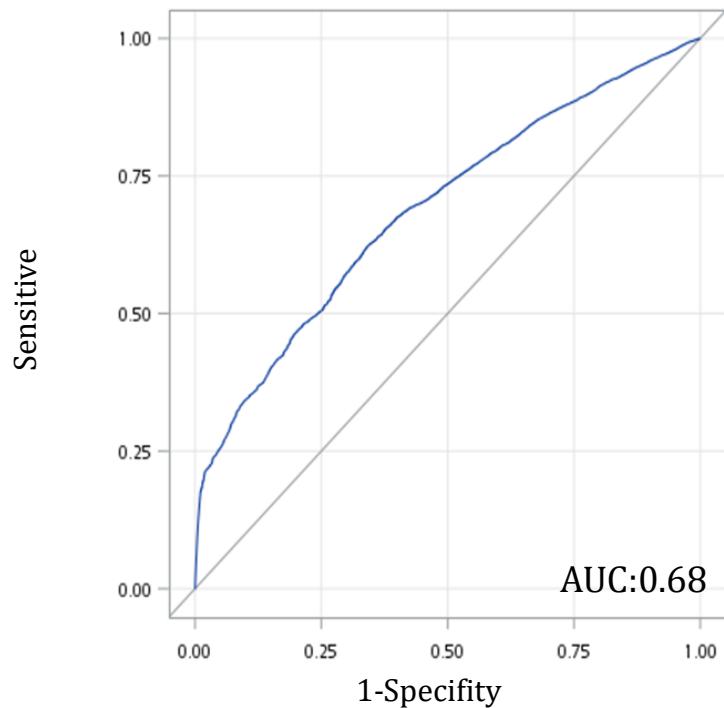


圖 4.1 (d)、近端大腸與正常個案 ROC 分析結果

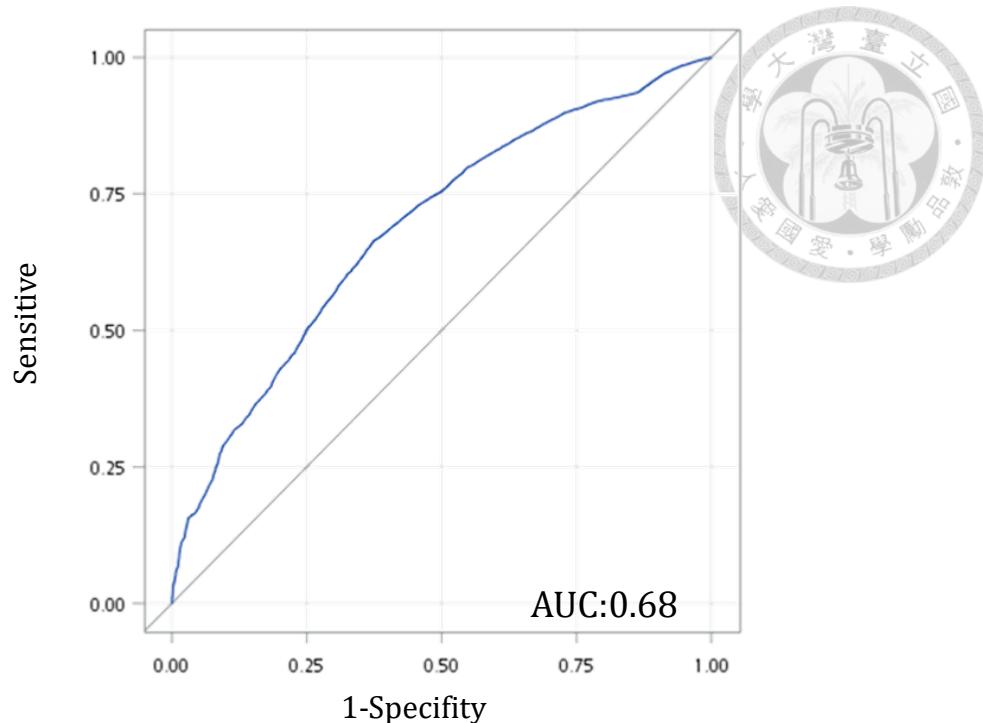


圖 4.1 (e)、遠側大腸與正常個案 ROC 分析結果

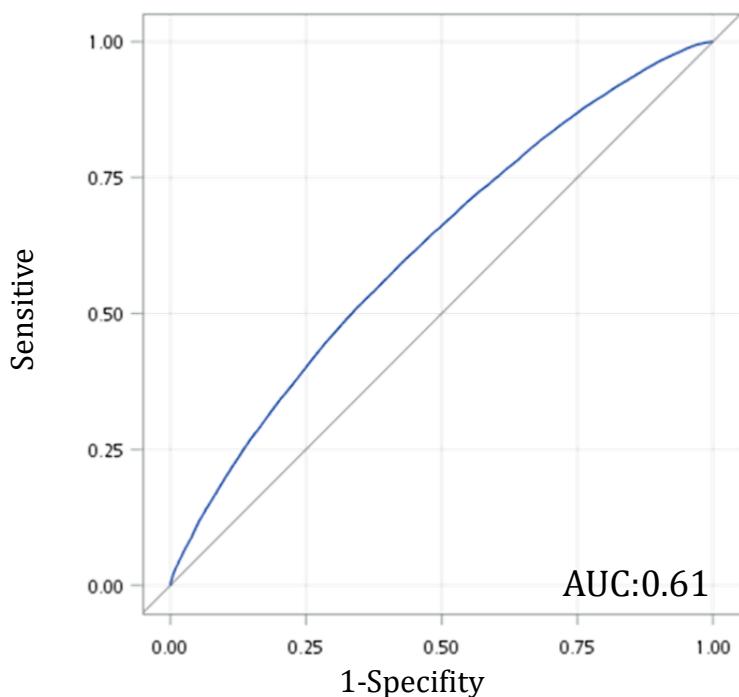


圖 4.1 (f)、遠側大腸癌個案與近端大腸個案 ROC 分析結果



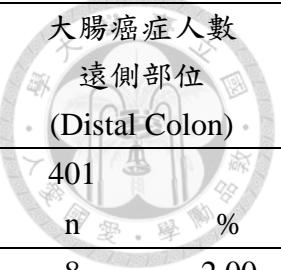
4.2 社區整合式篩檢

4.2.1 描述性分析結果

表 4.5 為臺灣社區整合式篩檢描述性分析結果，其中人口學變項結果與全國一致，在其他變項方面近端大腸癌血糖與高血壓似乎較遠側大腸癌多，而吸菸及飲酒和檳榔似乎遠側較近端多。

表 4.5、社區整合式篩檢描述性分析結果

	篩檢人數		大腸癌症人數		大腸癌症人數	
			近端部位 (Proximal Colon)		遠側部位 (Distal Colon)	
	N	%	n	%	n	%
性別						
男性	10,111	25.92	49	26.78	163	40.65
女性	28,904	74.08	134	73.22	238	59.35
篩檢年齡						
50-54	18,328	46.98	46	25.14	130	32.42
55-59	8,193	21.00	28	15.30	84	20.95
60-64	6,092	15.61	43	23.50	73	18.20
65-69	5,954	15.26	63	34.43	109	27.18
71-75	448	1.15	3	1.63	5	1.25
教育程度						
不識字	5,870	15.05	38	20.77	80	19.95
小學	15,753	40.38	85	46.45	184	45.89
國中	6,289	16.12	26	14.21	54	13.47
高中	7,413	19.00	20	10.93	53	13.22
大學以上	3,521	9.02	13	7.10	30	7.48
大腸癌家族病史						
無	37,718	96.68	167	91.26	385	96.01
有	1,297	3.32	16	8.74	16	3.99
FIT 濃度 (ng/ml)						
Undetected	20,757	53.20	70	38.25	186	46.38
1-19	15,105	38.72	69	37.70	115	28.68
20-49	1,592	4.08	13	7.10	23	5.74
50-99	647	1.66	9	4.92	19	4.74

	篩檢人數	大腸癌症人數		大腸癌症人數			
		近端部位 (Proximal Colon)		遠側部位 (Distal Colon)			
		N	%	n	%		
	39,015			183		401	
100-149	257	0.66		2	1.09	8	
150-249	206	0.53		1	0.55	8	
250-449	156	0.40		4	2.19	9	
>450	295	0.76		15	8.20	33	
蔬菜							
不吃	143	0.37		1	0.55	2	
少量(1/2-2 碗)	33,678	86.46		153	83.61	350	
大量(超過 3 碗)	5,132	13.17		29	15.85	49	
肉類							
不吃	14,598	37.51		79	43.41	156	
少量(1/2-2 碗)	18,286	46.99		79	43.41	185	
大量(超過 3 碗)	6,031	15.50		24	13.19	60	
蛋、豆類							
不吃	7,782	19.98		45	24.59	94	
少量(1/2-2 碗)	24,872	63.87		111	60.66	245	
大量(超過 3 碗)	6,288	16.15		27	14.75	62	
水果							
不吃	2,510	6.45		17	9.29	25	
少量(1/2-2 碗)	13,575	34.86		64	34.97	164	
大量(超過 3 碗)	22,854	58.69		102	55.74	212	
咖啡							
很少或不吃	28,902	74.22		150	81.97	312	
每周 1-4 次	6,744	17.32		24	13.11	62	
每周超過 5 次	3,297	8.47		9	4.92	27	
牛奶							
很少或不吃	10,287	26.41		42	22.95	112	
每周 1-4 次	13,729	35.25		59	32.24	138	
每周超過 5 次	14,935	38.34		82	44.81	151	
吸菸							
無	33,158	85.08		159	86.89	307	
已戒	2,053	5.27		11	6.01	34	
持續	3,761	9.65		13	7.10	60	

	篩檢人數	大腸癌症人數		大腸癌症人數	
		近端部位 (Proximal Colon)		遠側部位 (Distal Colon)	
		39,015	183	401	%
		N	n	n	%
飲酒					
	無	30,124	77.30	149	81.42
	已戒	579	1.49	4	2.18
	持續	8,267	21.21	30	16.39
嚼檳					
	無	35,942	92.23	171	93.44
	已戒	1,808	4.64	8	4.37
	持續	1,219	3.13	4	2.19
運動					
	無	16,114	41.35	73	39.89
	有	22,852	58.65	110	60.11
一年內參與健康檢查					
	無	35,024	89.87	164	89.62
	有	3,946	10.13	19	10.38
腰圍					
	正常	23,103	59.26	101	55.19
	異常 ¹	15,886	40.74	82	44.81
三酸甘油酯					
	正常	29,469	76.15	129	71.67
	異常 ²	9,230	23.85	51	28.33
高密度膽固醇					
	正常	23,430	60.21	107	58.47
	異常 ³	15,485	39.79	76	41.53
血壓					
	正常	16,952	43.48	52	28.42
	異常 ⁴	22,036	56.52	131	71.58
血糖					
	正常	25,707	66.28	105	57.38
	異常 ⁵	13,080	33.72	76	41.53
尿酸					
	正常	34,402	88.76	160	88.4
	異常 ⁶	4,355	11.24	21	11.6

篩檢人數	大腸癌症人數		大腸癌症人數	
	近端部位 (Proximal Colon)		遠側部位 (Distal Colon)	
	N	%	n	%
39,015			183	
				401
				n %
肌酸酐				
正常	37,758	97.42	170	93.92
異常 ⁷	999	2.58	11	6.08
糖尿病健康意識*				
無	780	21.22	7	22.58
有	2,895	78.78	24	77.42
高血壓健康意識**				
無	5,262	45.13	33	40.24
有	6,398	54.87	49	59.76

¹腰圍異常：男性≥90 公分，女性≥80 公分

²三酸甘油酯異常：空腹三酸甘油酯≥150 mg/dL

³高密度膽固醇異常：男性<40 mg/dL，女性<50 mg/dL

⁴血壓異常：收縮壓≥130 mmHg 且收縮壓≥85 mmHg

⁵血糖異常：空腹血糖≥100mg/dL

⁶尿酸異常：男性>13 mg/dL，女性>10 mg/dL

⁷肌酸酐異常： $\geq 1.3 \text{ mg/dL}$

*糖尿病健康意識：血糖值高於 126 mg/dL 且自述罹患糖尿病者

**高血壓健康意識：收縮壓≥130 mmHg 且收縮壓≥85 mmHg 且自述罹患高血壓者



4.2.2 隨機森林分析結果

圖 4.(a)-(c)為隨機森林 VIMP 分析上述表 4.4 相關變項排列結果，
圖 4.3(d)-(f)為 ROC 分析結果，其 AUC 與正常個案比較皆高於 90%，
而近端部位大腸與遠側部位大腸比較則為 89%。

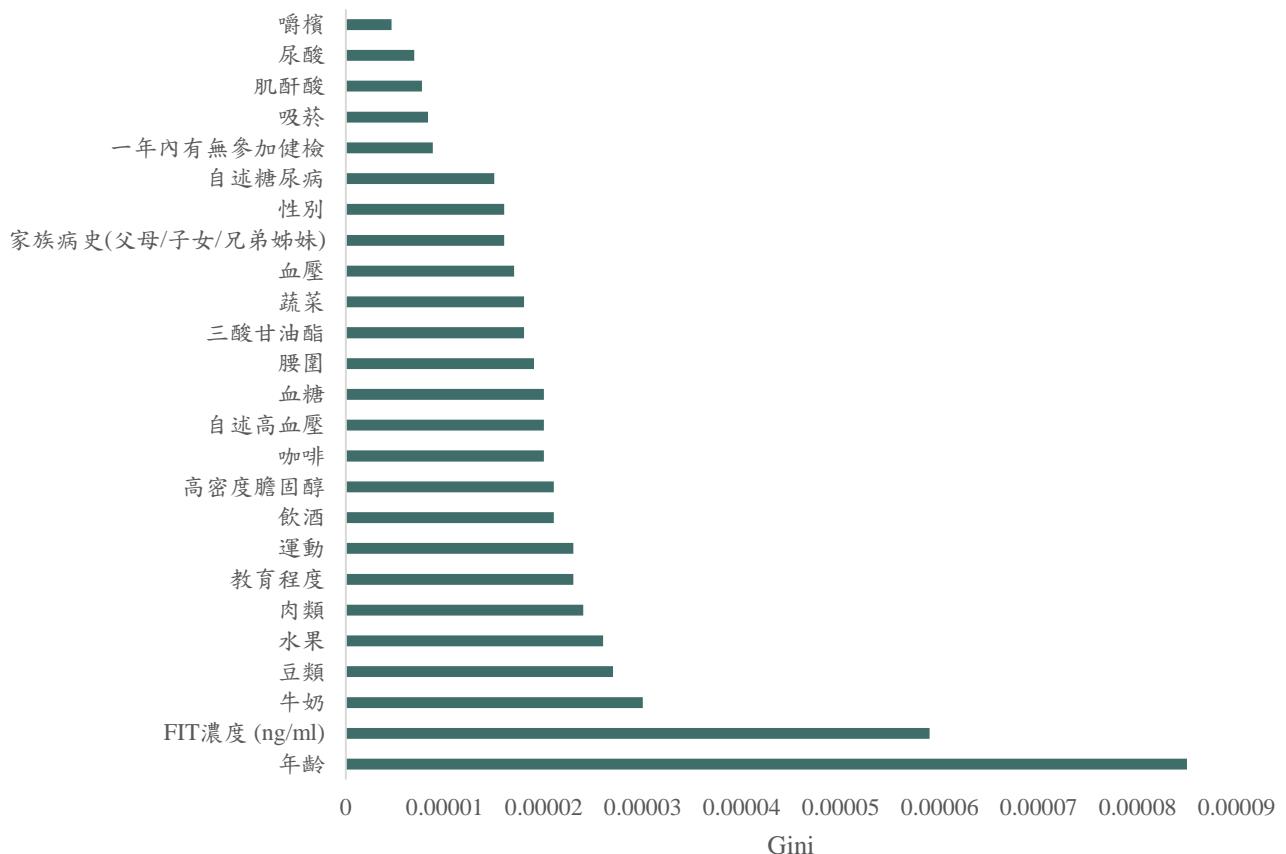


圖 4.2 (a)、近端大腸與正常個案之重要變項排序

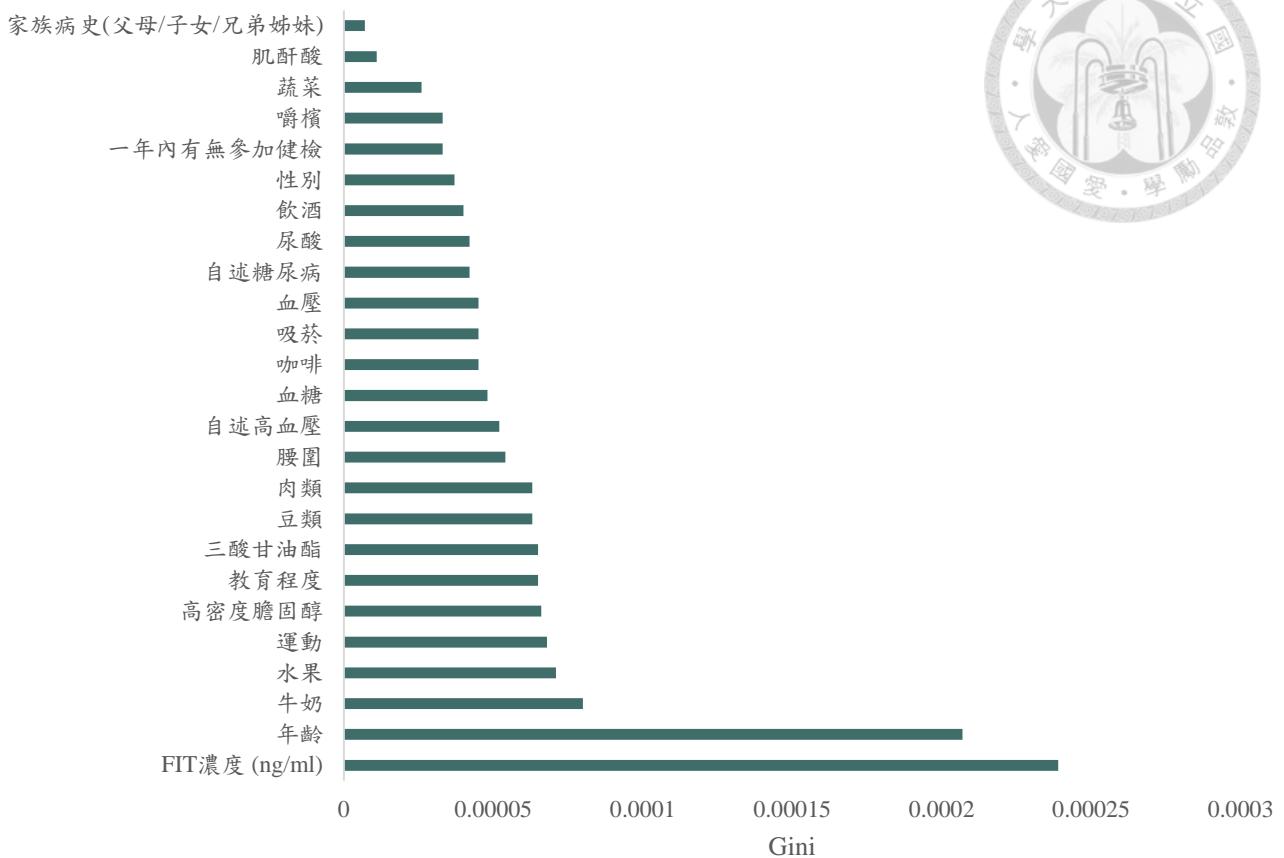


圖 4.2 (b)、遠側大腸與正常個案之重要變項排序

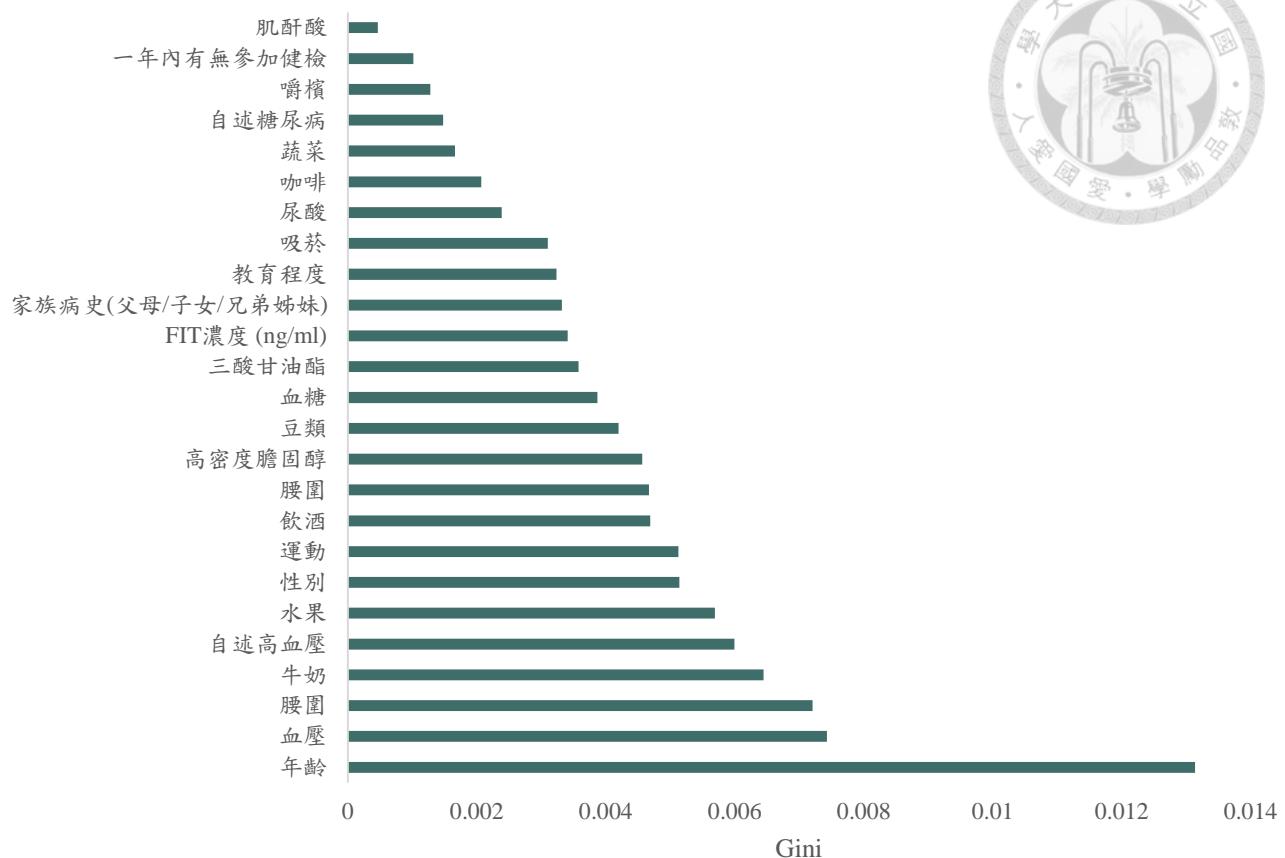


圖 4.2 (c)、近側大腸與遠側大腸個案之重要變項排序

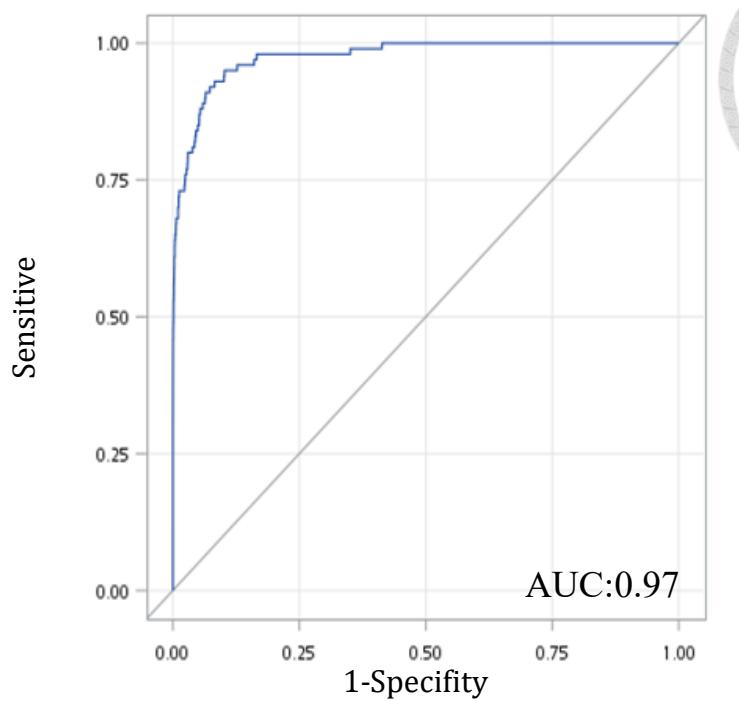


圖 4.2 (d)、近端大腸與正常個案 ROC 分析結果

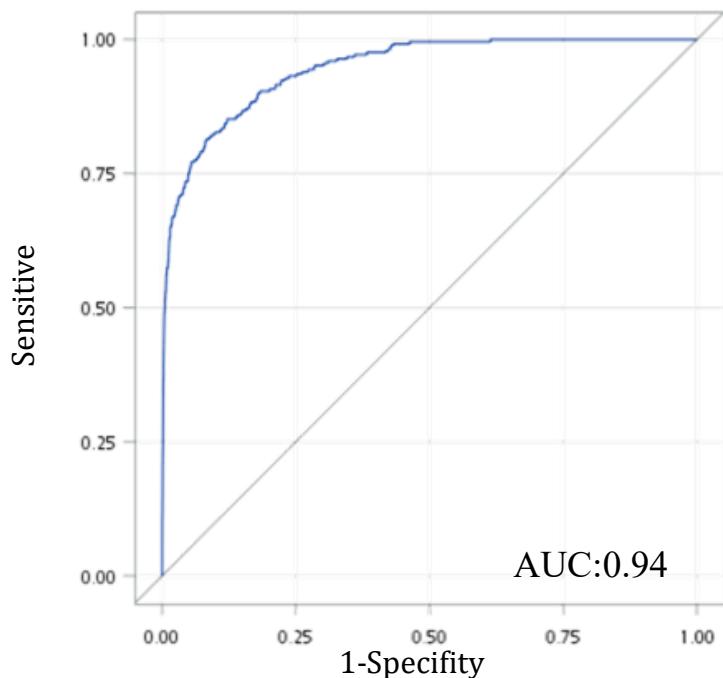


圖 4.2 (e)、遠側大腸與正常個案 ROC 分析結果

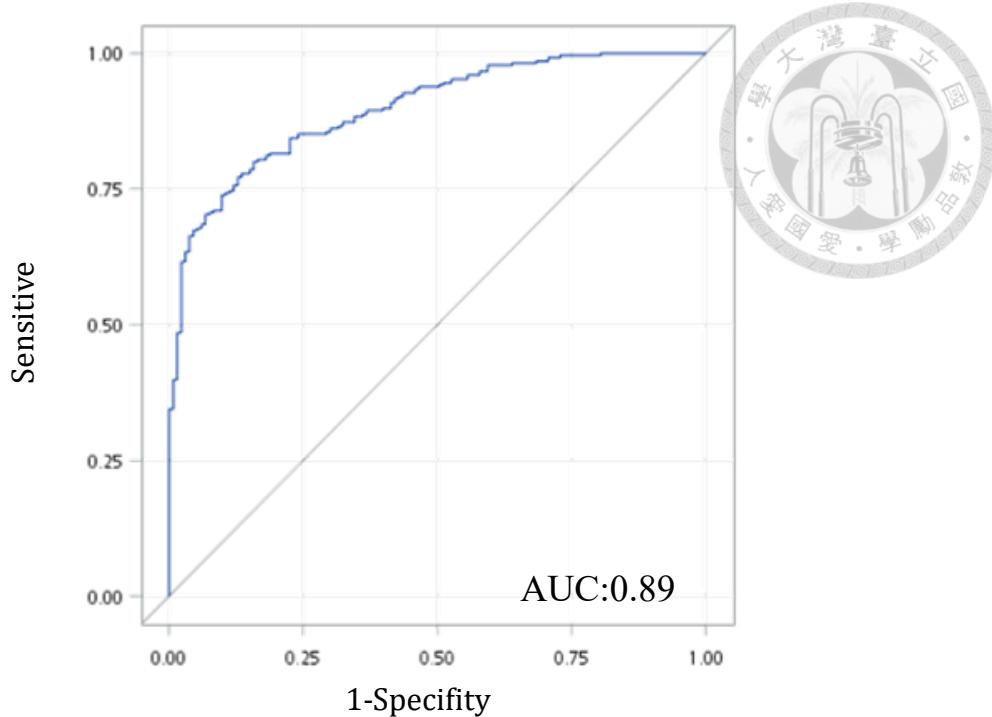


圖 4.2 (f)、遠側大腸癌個案與近端大腸個案 ROC 分析結果

4.2.3 大腸腺腫至侵襲癌個人化危險分層分析結果

利用上節隨機森林中各項個人特徵，可以利用貝氏網路進一步描繪這些特徵在大腸腺腫至侵襲癌進展過程中相關性，依不同疾病過程可分為腺腫發生特徵 (Initiator for Advanced Adenoma Features, IAF)、惡性轉變特徵 (Malignant Transformation Features, MTF) 及 臨床 進展 特徵 (Progressive Promotor Features, PPF)，並利用貝氏網路分析後所產生之個人化十分位危險分數。

腺腫發生特徵包含年齡、性別、教育程度、大腸癌家族史、與個人高血壓健康意識外，生活習慣中的嚼食檳榔、飲酒、抽菸等習慣，生化指標中尿酸濃度、FIT 濃度、血糖、血壓、腰圍與三酸甘油酯等都是影響進行性腺瘤發生之特徵；而惡性轉變特徵中，則是年齡、高血壓、FIT 濃度、飲食習慣，是否參加健康檢查等為影響進行性腺瘤轉變為無症狀之 CP 之特徵；影響無症狀之 CP 至有症狀之 CP 臨床進展特徵則包含年齡、肌酸酐、FIT 濃度、飲酒習慣與嚼食檳榔習慣等特徵。

藉由貝氏網路展現不同疾病進展階段之特徵，可產生多階段個人化十分位風險分數，並且進一步與疾病自然史結合。



4.2.4 大腸腺腫及侵襲癌個人化風險分數數位雙胞胎介入評估

利用上節貝氏網路分析後所產生之個人化十分位危險分數與大腸直腸癌疾病自然史結合，即可針對個案不同特徵計算風險分數，並進一步模擬當提供各項不同介入措施後，評估其介入效益。

舉例一位 64 歲男性，低教育程度，無家族病史，生活習慣中有飲酒習慣，無規律運動及無吸菸行為，並考量其飲食習慣與生化檢驗資料後，進一步規劃個人化介入措施，例如改變生活習慣、血糖控制與高血壓衛教等，透過個人化風險分數數位雙胞胎評估其介入後的發病風險。若無任何介入措施下，該個案 10 年晚期癌症發生風險約 14%，若利用個人化風險分數數位雙胞胎提供個人化措施後，10 年晚期癌症發生風險僅 4.7%，個人化介入措施對預防晚期大腸癌發生之效益達 66%。

表 4.6 列出不同篩檢間隔依遠側及近端數位雙胞胎評估效益，間隔越短篩檢效益越高，若以篩檢效益至少降低 25%晚期癌症發生率為閾值，對近端大腸直腸癌風險較高之個案建議需要每年篩檢一次，遠側大腸直腸癌潛在風險較高的個案而言，則可建議設定於每兩年篩檢一次。

表 4.7 顯示數位雙胞胎評估精準篩檢政策可達到降低 32% (RR:0.68, 95% CI:0.67-0.68)晚期大腸癌發生之效益。

表 4.6、FIT 大腸癌篩檢間隔數位雙胞胎效益評估

部位	未篩檢組	篩檢組	RR	95% 信賴區間
1 年一次				
遠側部位	72,766	45,014	0.62	(0.61, 0.63)
近端部位	27,612	18,716	0.68	(0.67, 0.69)
整體	100,378	63,730	0.63	(0.63, 0.64)
2 年一次				
遠側部位	72,766	53,225	0.73	(0.72, 0.74)
近端部位	27,612	21,834	0.79	(0.78, 0.80)
整體	100,378	75,059	0.75	(0.73, 0.75)
3 年一次				
遠側部位	72,766	57,396	0.79	(0.78, 0.80)
近端部位	27,612	23,145	0.84	(0.82, 0.85)
整體	100,378	80,541	0.80	(0.80, 0.81)

表 4.7、FIT 大腸癌篩檢間隔數位雙胞胎精準篩檢政策

篩檢策略	未篩檢組	篩檢組	RR	95% 信賴區間
Risk-based*	100,394	68,052	0.68	(0.67, 0.68)

*風險分數分為四組，並給予不同之篩檢間隔，如高風險組 (風險分數大於 75%) 應每 1 年篩檢一次，中風險組 (風險分數介於 50-75%) 應每 2 年篩檢一次，低風險組 (風險分數介於 25-50%) 應每 3 年篩檢一次，最低風險組 (風險分數為 0%-25%) 應每 6 年篩檢一次。

第五章、討論

本論文首次提出利用元宇宙數位雙胞胎模型評估 FIT 篩檢在個人和族群層面之應用，以實現精準個人化醫學於篩檢政策應用實用性。



5.1 近端及遠側大腸直腸癌自然病史

本研究之臨床價值為利用馬可夫機器學習區分近端與遠側大腸直腸癌自然病史，剖析不同部位之疾病進展速度與 FIT 早期檢測大腸直腸癌敏感度問題，解釋為何近端部位大腸有較高間隔癌 (Interval cancer)風險，且比遠側部位大腸更容易出現晚期癌症。

邱等人透過 2005 年至 2010 年間於臺大醫院健康管理中心接受大腸癌篩檢之數據評估 FIT 敏感度，結果發現 FIT 偵測非進行性腺瘤、進行性腺瘤、癌症之敏感度分別為 10.6% (95% CI : 10.2% – 12.3%)、28.0% (95% CI : 24.6% – 31.7%) 和 78.6% (95% CI : 58.5% – 91.0%)，且對於近端部位之進行性腺瘤或非息肉性腺瘤其敏感度低於遠側部位；對於偵測病灶尺寸小於 15mm 之腺瘤或非息肉型腺瘤時偽陽性率較高。以上結果顯示 FIT 對於小型或非息肉性腺瘤敏感度較差，且具有部位差異性。(Chiu et al., 2013)

本論文證明遠側部位大腸癌發生率是近端部位大腸的三倍，但遠側部位大腸癌無症狀之 PCDP 到有症狀之 CP 的進展速度比近端部位大腸慢，此結果可能代表遠側部位大腸出現之病灶更容易被預防，因為遠側部位疾病進展較慢，就算 FIT 篩檢偽陰性個案也可能在後續篩檢過程中被重新偵測。本研究亦與先前研究所示近端大腸更容易進展為晚期癌症(Blanks et al., 2019) 相呼應。

過去研究提出近端部位與遠側部位大腸癌的流行病學和腫瘤發展途徑差異。形成間隔癌的途徑之一是透過鋸齒狀腺瘤 (Serrated neoplasia) 途

徑，這些腺瘤起源於鋸齒狀息肉 (Serrated polyps)，即使利用大腸鏡檢查 SSAs (Sessile Serrated Adenoma)也難以被檢測，從而導致間隔癌。臺灣前瞻性研究顯示，FIT 對 SSAs 的敏感度顯著低於傳統腺瘤。(Chang et al., 2017) 本論文結果顯示近端部位病灶較易出現晚期病變肇因於疾病進展較快而非敏感度較差之結果支持 SSAs 在近端部位大腸更容易被發現的結果，以及 SSAs 從異型增生到惡性的過渡期較迅速。英國 FOBT 大腸癌篩檢中也觀察到近端大腸相同趨勢，在近端大腸中，間隔癌更可能出現晚期病灶，且更可能是非腺癌形態，這與鋸齒狀途徑癌症的分子和形態特徵相似。(Blanks et al., 2019)

有關遠側部位大腸癌疾病自然史，本論文發現從早期 PCDP 到晚期 PCDP (0.77)和從晚期 PCDP 到晚期 CP (0.41)的年轉移率，與其他大腸直腸癌篩檢結果相似。邱等人針對 Nottingham 和 Funen 隨機分派試驗之統合分析研究也顯示，與臨床 Dukes' Stage A/B 相比，從 PCDP Dukes' Stage A/B 到 PCDP Dukes' Stage C/D 的期別上升機率較高，儘管上述研究使用 FOBT 而非 FIT。(Chiu et al., 2011)

另一篇美國利用三階段馬可夫模型估計不同部位的大腸直腸癌平均滯留期 (Mean Sojourn Time, MST)，結果發現遠側部位大腸平均滯留期比近端部位大腸來的更短，且隨著年齡增加，近端部位之平均滯留期未隨年齡而增加；近端部位具較長的平均滯留期，表示該部位之病程較不易進展，此結果與本論文結論相異。這一差異可能來自於該研究之限制，首先美國研究無法區分篩檢偵測個案與臨床個案，且除了 FOBT 外，也使用乙狀大腸鏡或大腸鏡劑攝影做為篩檢工具，如此可能將臨床個案錯誤歸類為篩檢偵測個案，進而導致較長的 MST。(Zheng & Rutter, 2012)

此外，特定部位 FIT 敏感度會受到篩檢計畫轉介率和大腸鏡品質影響，

與其他高轉介率和高大腸鏡敏感度研究相比，對 FIT 敏感度結果的解釋應十分謹慎。

5.2 元宇宙數位雙胞胎模型用於精準 FIT 篩檢服務評估



除了考慮大腸直腸腺瘤和癌症在近端和遠端部位的自然病史差異外，本論文還將個人特徵加入大腸直腸腫瘤多階段自然病史地圖中。第一步是利用機器學習演算法中隨機森林鑑別近端和遠側大腸直腸癌之特徵，這些特徵包括人口特徵、FIT濃度、生活方式因素和生物標記。然後使用貝氏網路建立這些個人特徵在疾病過程三個不同階段之關係，分別為腺腫發生特徵、惡性轉化特徵及臨床進展特徵。這些特徵為個人化風險分位數提供基礎，利用元宇宙基礎的數位雙胞胎結合個人化風險分類，使我們能對個人層面之個人化介入進行量化評估。這種元宇宙數位雙胞胎設計提供了一個良好之方法評估沉浸式介入措施（如篩檢間隔和個人化風險分層之篩檢策略）對減少晚期大腸直腸癌或大腸直腸癌死亡的效果。

5.3 研究限制

本論文有三個限制。首先，本研究之馬可夫機器學習 AI 生成式引擎目前只能應用於臺灣族群層面，當應用於其他族群時，需針對 AI 生成式引擎參數進行校正。其次，元宇宙數位雙胞胎設計應該是動態的，以便產生依時序性動態更新涵蓋初段到末段預防的沉浸式介入。第三，元宇宙數位雙胞胎設計還需進一步考慮成本效益對政策之影響。例如考量不同減少晚期大腸直腸癌效果閾值時，需要參考各種篩檢政策涉及的成本效益。

第六章、結論



本論文提出一個創新的元宇宙數位雙胞胎模型，應用於臺灣族群 FIT 篩檢服務和社區整合式篩檢服務，評估精準 FIT 篩檢服務之效益，如不同篩檢策略、篩檢間隔等個人化篩檢策略之評估。

參考文獻

- Bardou, M., Barkun, A. N., & Martel, M. (2013). Obesity and colorectal cancer. *Gut*, 62(6), 933-947. <https://doi.org/10.1136/gutjnl-2013-304701>
- Blanks, R., Burón Pust, A., Alison, R., He, E., Barnes, I., Patnick, J., Reeves, G. K., Floud, S., Beral, V., & 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*, 145(3), 728-734. <https://doi.org/10.1002/ijc.32168>
- Chang, L.-C., Shun, C.-T., Hsu, W.-F., Tu, C.-H., Tsai, P.-Y., Lin, B.-R., Liang, J.-T., Wu, M.-S., & Chiu, H.-M. (2017). Fecal Immunochemical Test Detects Sessile Serrated Adenomas and Polyps With a Low Level of Sensitivity. *Clinical Gastroenterology and Hepatology*, 15(6), 872-879.e871. <https://doi.org/10.1016/j.cgh.2016.07.029>
- Chen, T. H.-H., Yen, M.-F., Lai, M.-S., Koong, S.-L., Wang, C.-Y., Wong, J.-M., Prevost, T. C., & Duffy, S. W. (1999). Evaluation of a selective screening for colorectal carcinoma. *Cancer*, 86(7), 1116-1128. [https://doi.org/10.1002/\(sici\)1097-0142\(19991001\)86:7<1116::aid-cncr4>3.0.co;2-d](https://doi.org/10.1002/(sici)1097-0142(19991001)86:7<1116::aid-cncr4>3.0.co;2-d)
- Chen, T. H., Chiu, Y. H., Luh, D. L., Yen, M. F., Wu, H. M., Chen, L. S., Tung, T. H., Huang, C. C., Chan, C. C., Shiu, M. N., Yeh, Y. P., Liou, H. H., Liao, C. S., Lai, H. C., Chiang, C. P., Peng, H. L., Tseng, C. D., Yen, M. S., Hsu, W. C., . . . Taiwan Community-Based Integrated Screening, G. (2004). Community-based multiple screening model: design, implementation, and analysis of 42,387 participants. *Cancer*, 100(8), 1734-1743. <https://doi.org/10.1002/cncr.20171>
- Chen, T. H. H., Chiu, Y. H., Luh, D. L., Yen, M. F., Wu, H. M., Chen, L. S., Tung, T. H., Huang, C. C., Chan, C. C., & Shiu, M. N. (2004). Community-based multiple screening model: Design, implementation, and analysis of 42,387 participants Taiwan community-based integrated screening group. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 100(8), 1734-1743.
- Chiu, H.-M., Jen, G. H.-H., Wang, Y.-W., Fann, J. C.-Y., Hsu, C.-Y., Jeng, Y.-C., Yen, A. M.-F., Chiu, S. Y.-H., Chen, S. L.-S., & Hsu, W.-F. (2021). Long-term effectiveness of faecal immunochemical test screening for proximal and distal colorectal cancers. *Gut*, 70(12), 2321-2329.
- Chiu Han-Mo, C. H.-H. (2021). *CRC Screening Theory and Practice*. <https://doi.org/10.1007/978-981-15-7482-5>
- Chiu, H. M., Chen, S. L., Yen, A. M., Chiu, S. Y., Fann, J. C., Lee, Y. C., Pan, S. L., Wu, M. S., Liao, C. S., Chen, H. H., Koong, S. L., & Chiou, S. T. (2015). Effectiveness of fecal immunochemical testing in reducing colorectal cancer mortality from the One Million Taiwanese Screening Program. *Cancer*, 121(18), 3221-3229.



<https://doi.org/10.1002/cncr.29462>

Chiu, H. M., Jen, G. H., Wang, Y. W., Fann, J. C., Hsu, C. Y., Jeng, Y. C., Yen, A. M., Chiu, S. Y., Chen, S. L., Hsu, W. F., Lee, Y. C., Wu, M. S., Wu, C. Y., Jou, Y. Y., & Chen, T. H. (2021). Long-term effectiveness of faecal immunochemical test screening for proximal and distal colorectal cancers. *Gut*, 70(12), 2321-2329.

<https://doi.org/10.1136/gutjnl-2020-322545>

Chiu, H. M., Lee, Y. C., Tu, C. H., Chen, C. C., Tseng, P. H., Liang, J. T., Shun, C. T., Lin, 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.e832. <https://doi.org/10.1016/j.cgh.2013.01.013>

Chiu, H. M., Lin, J. T., Shun, C. T., Liang, J. T., Lee, Y. C., Huang, S. P., & Wu, M. S. (2007). Association of Metabolic Syndrome With Proximal and Synchronous Colorectal Neoplasm. *Clinical Gastroenterology and Hepatology*, 5(2), 221-229.

<https://doi.org/10.1016/j.cgh.2006.06.022>

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.

<https://doi.org/10.1111/j.1365-2753.2010.01378.x>

Cox, D. R., Miller, H. D. (1965). *The theory of stochastic processes*.

Henrikson, N. B., Webber, E. M., Goddard, K. A., Scrol, A., Piper, M., Williams, M. S., Zallen, D. T., Calonge, N., Ganiats, T. G., Janssens, A. C. J. W., Zauber, A., Lansdorp-Vogelaar, I., Van Ballegooijen, M., & Whitlock, E. P. (2015). Family history and the natural history of colorectal cancer: systematic review. *Genetics in Medicine*, 17(9), 702-712. <https://doi.org/10.1038/gim.2014.188>

Hsu, W. F., Hsu, C. Y., Yen, A. M., Chen, S. L., Chiu, S. Y., Fann, J. C., Lee, Y. C., Chiu, H. M., & Chen, H. H. (2021). Classifying interval cancers as false negatives or newly occurring in fecal immunochemical testing. *J Med Screen*, 28(3), 286-294.

<https://doi.org/10.1177/0969141320986830>

Jen, G. H.-H., Yen, A. M.-F., Hsu, C.-Y., Chiu, H.-M., Chen, S. L.-S., & Chen, T. H.-H. (2021). Modelling the impacts of COVID-19 pandemic on the quality of population-based colorectal cancer screening. *Preventive Medicine*, 151, 106597.

<https://doi.org/10.1016/j.ypmed.2021.106597>

Johnson, C. M., Wei, C., Ensor, J. E., Smolenski, D. J., Amos, C. I., Levin, B., & Berry, D. A. (2013). Meta-analyses of colorectal cancer risk factors. *Cancer Causes & Control*, 24(6), 1207-1222. <https://doi.org/10.1007/s10552-013-0201-5>

Keum, N., Bao, Y., Smith-Warner, S. A., Orav, J., Wu, K., Fuchs, C. S., & Giovannucci, E. L. (2016). Association of Physical Activity by Type and Intensity With Digestive



- System Cancer Risk. *JAMA Oncology*, 2(9), 1146.
<https://doi.org/10.1001/jamaoncol.2016.0740>
- Liu, P.-H., Wu, K., Ng, K., Zauber, A. G., Nguyen, L. H., Song, M., He, X., Fuchs, C. S., Ogino, S., Willett, W. C., Chan, A. T., Giovannucci, E. L., & Cao, Y. (2019). Association of Obesity With Risk of Early-Onset Colorectal Cancer Among Women. *JAMA Oncology*, 5(1), 37. <https://doi.org/10.1001/jamaoncol.2018.4280>
- Murphy, N., Ward, H. A., Jenab, M., Rothwell, J. A., Boutron-Ruault, M.-C., Carbonnel, F., Kvaskoff, M., Kaaks, R., Kühn, T., Boeing, H., Aleksandrova, K., Weiderpass, E., Skeie, G., Borch, K. B., Tjønneland, A., Kyrø, C., Overvad, K., Dahm, C. C., Jakobsen, P., . . . Gunter, M. J. (2019). Heterogeneity of Colorectal Cancer Risk Factors by Anatomical Subsite in 10 European Countries: A Multinational Cohort Study. *Clinical Gastroenterology and Hepatology*, 17(7), 1323-1331.e1326.
<https://doi.org/10.1016/j.cgh.2018.07.030>
- Oh, M., McBride, A., Yun, S., Bhattacharjee, S., Slack, M., Martin, J. R., Jeter, J., & Abraham, I. (2018). BRCA1 and BRCA2 Gene Mutations and Colorectal Cancer Risk: Systematic Review and Meta-analysis. *JNCI: Journal of the National Cancer Institute*, 110(11), 1178-1189. <https://doi.org/10.1093/jnci/djy148>
- Rabeneck, L., Chiu, H. M., & Senore, C. (2020). International Perspective on the Burden of Colorectal Cancer and Public Health Effects. *Gastroenterology*, 158(2), 447-452.
<https://doi.org/10.1053/j.gastro.2019.10.007>
- Shaw, E., Farris, M. S., Stone, C. R., Derksen, J. W. G., Johnson, R., Hilsden, R. J., Friedenreich, C. M., & Brenner, D. R. (2018). Effects of physical activity on colorectal cancer risk among family history and body mass index subgroups: a systematic review and meta-analysis. *BMC Cancer*, 18(1).
<https://doi.org/10.1186/s12885-017-3970-5>
- Zheng, W., & Rutter, C. M. (2012). Estimated Mean Sojourn Time Associated with Hemoccult SENSA for Detection of Proximal and Distal Colorectal Cancer. *Cancer Epidemiology Biomarkers & Prevention*, 21(10), 1722-1730.
<https://doi.org/10.1158/1055-9965.epi-12-0561>
- 衛生福利部國民健康署. (2023). 110 年癌症登記年報.