



TriNetX研究經驗分享

從真實世界資料到AI輔助分析的臨床應用



醫學研究部 基礎醫學科 生統小組 / 心臟血管研究中心
副研究員 陳韻仔 博士

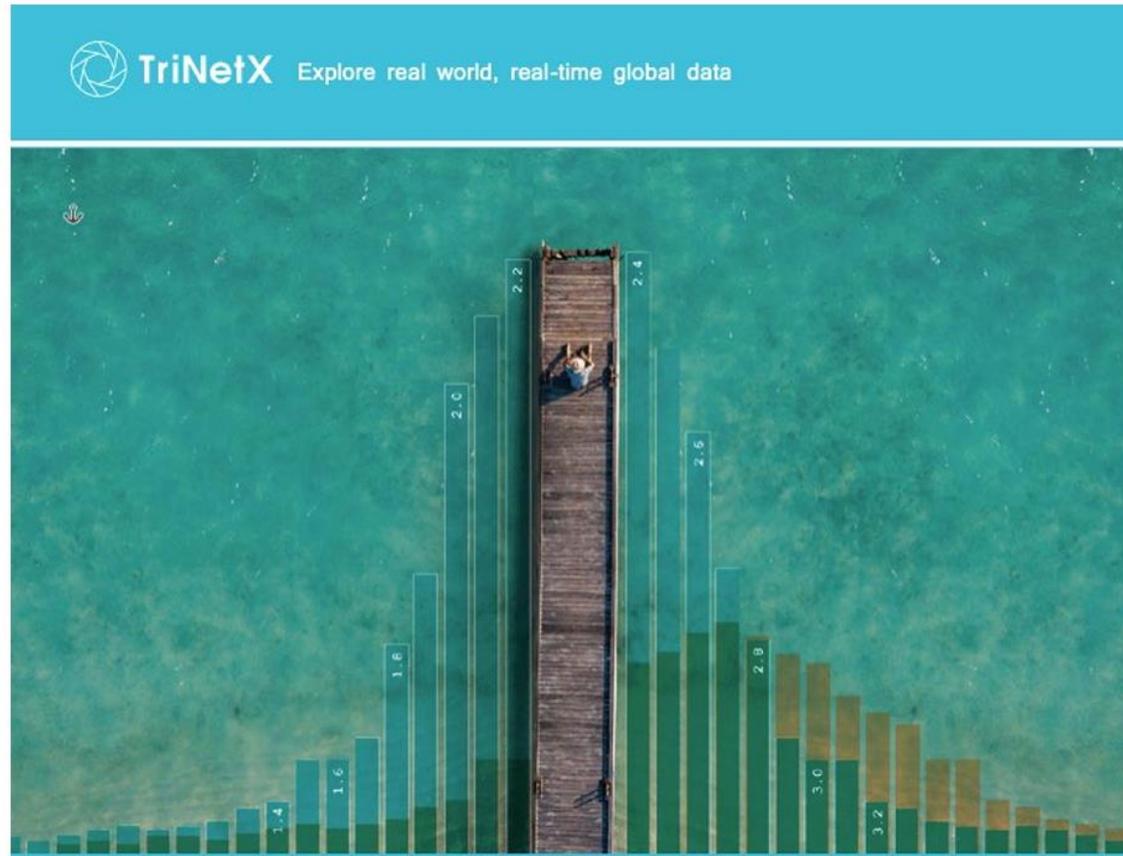
日期：114年11月17日

Outlines

- ▶ Part 1: Understanding TriNetX: Strengths & Limitations
- ▶ Part 2: My research
- ▶ Part 3: 善用AI : Chat GPT修改圖片
- ▶ Part 4: 投稿過程中的特殊問題
- ▶ Part 5: TriNetX其他注意事項

Part 1

Understanding TriNetX: Strengths & Limitations



Why Real-World Evidence Matters

- ◆ **RCT** 具有高度內部效度，但樣本侷限於「低風險、少共病、依從性高」的選擇性族群，難以反映真實臨床的多樣性。
- ◆ **RWE** 來自跨國、多中心、多族群的真實世界資料，能呈現不同醫療體系與族群差異，具更高外部效度。
- ◆ 可捕捉 **RCT** 無法呈現的長期自然史、實際用藥行為與治療持續性，包含停藥、換藥、依從性、共病負荷等真實臨床現象。
- ◆ **AF 搭配 T2DM** 屬高度複雜且高風險族群，現有 RCT 證據多為次族群 (subgroup) 分析，樣本有限。
- ◆ **RWE** 能補足臨床試驗的不足，提供長期追蹤、異質族群分析及真實臨床決策訊息，是 AF + T2DM 研究不可或缺的證據來源。

TriNetX Overview

- ◆ TriNetX 為全球最大之 聯邦式 Real-World Data 平台，整合來自 **19+** 國家、**120M+** 患者 的臨床 EMR/Claims 資料
- ◆ 全資料皆經 去識別化 (**de-identification**) 處理，並符合 **HIPAA**、**WIRB**、**GDPR** 等隱私與倫理規範
- ◆ 涵蓋 診斷 (**ICD**)、處置 (**CPT/OPS**)、用藥 (**ATC/NDC**)、實驗室 (**LOINC**) 與多維度人口學資訊
- ◆ 內建標準化分析模組：
 - ✓ **Propensity Score Matching (PSM)**
 - ✓ **Kaplan–Meier survival analysis**
 - ✓ **Cox proportional hazards models**
 - ✓ **Treatment Patterns / Time-on-Treatment / Treatment Pathways**
- ◆ 適用於 大型族群比較研究、治療效果評估、跨國差異分析 與 短期至中期預後的臨床研究

Strengths and Limitations of TriNetX

Strengths

- ▶ 資料量龐大、族群多元：涵蓋 >120M 患者、跨 19+ 國家，多中心異質性高。
- ▶ Cohort 建立快速且可重複：標準化 ICD / ATC / LOINC 模組，減少人工錯誤。
- ▶ 內建統計與分析工具：提供 1:1 PSM、Kaplan-Meier、Cox 模型等標準流程。
- ▶ 適合大型族群與跨國比較研究：包括藥物使用差異、族群差異、用藥趨勢。
- ▶ 提供治療行為分析模組：含 Time-on-Treatment、Treatment Patterns、Treatment Pathways，可用於觀察用藥持續性與治療序列。

Limitations

- ▶ 僅能進行 1:1 傾向分數配對 (PSM)，無 1:n 或多變項調整後的 Cox。
- ▶ 不支援 time-dependent Cox 或 competing risk 模型 (PSM 後無法再行事件模型調整)。
- ▶ 無法取得個別患者原始 EMR，無法檢視原始病歷或自訂演算法。
- ▶ 部分 ICD / 程序碼屬敏感代碼 (sensitive codes)，無法納入分析。
- ▶ 缺失值不補值 (no imputation)，死亡資訊依各醫療機構回報品質不一。

When Real-World Evidence Is Particularly Useful (1)

◆ 多重共病、臨床異質性高的族群

- ✓ 心房顫動合併第二型糖尿病患者，常同時具有腎臟疾病、心衰竭、高血壓、肥胖等多重共病。
- ✓ 這類患者在真實世界中比例極高，但在多數臨床試驗中被排除，因此試驗結果無法完全反映實際臨床的複雜度。
- ✓ 真實世界資料能觀察不同共病組合下的疾病自然史與治療反應，更貼近臨床照護情境。

◆ 需要長期追蹤的重要臨床事件

- ✓ 中風、失智、心衰竭惡化與死亡，多在多年後才會發生。
- ✓ 臨床試驗追蹤時間通常有限，較難完整呈現長期風險。
- ✓ 真實世界資料具有多年縱向紀錄，可更全面評估疾病發展與治療效果。

When Real-World Evidence Is Particularly Useful (2)

◆ 治療持續性、停藥與換藥行為 (真實世界特有)

- ✓ 心房顫動與糖尿病患者的藥物依從性、高風險併發症與副作用，使得治療中斷或換藥相當常見。
- ✓ 真實世界資料能分析：
 - 用藥持續時間
 - 停藥原因與頻率
 - 不同藥物之間的轉換情形
- ✓ 這些資訊反映「實際臨床決策」，是臨床試驗無法充分呈現的。

◆ 補足臨床試驗無法涵蓋的病人族群

- ✓ 多數臨床試驗排除高齡者、腎功能降低者、多重用藥者及病情不穩定者。
- ✓ 然而實際臨床上，這類患者才是心房顫動與糖尿病族群的主體。
- ✓ 真實世界資料提供更高外部效度，使結果能代表真正的臨床族群。

真實世界證據能補足臨床試驗在高異質族群、長期追蹤與實際臨床用藥行為上的不足，特別適用於**心房顫動合併第二型糖尿病**等複雜族群的研究。

Part 2

Year	Impact Factor (IF)	Total Articles	Total Cites
2024 (2025 update)	5.1	-	78854

> [J Clin Endocrinol Metab.](#) 2025 Aug 7;110(9):2605-2614. doi: 10.1210/clinem/dgae861.

SGLT2i and Cardiovascular Events in Patients With Concomitant Atrial Fibrillation and Diabetes: A TriNetX Cohort Study

Yu-Yu Hsiao ¹, Yun-Yu Chen ^{1 2 3 4 5}, Ming-Jen Kuo ^{1 2 5 6}, Yu-Shan Chien ^{1 2 5 6},
Guan-Yi Li ^{1 2 5}, Shang-Ju Wu ^{1 2 5 6}, Wei-Lun Lin ^{1 2 3 5}, Shu-Fen Chiu ¹,
Cheng-Hung Li ^{1 2 5 6}, Jiunn-Cherng Lin ^{1 2 5 6}, Ching-Heng Lin ³, Jin-Long Huang ^{1 2 5 7},
Yu-Cheng Hsieh ^{1 2 3 5 6}, Shih-Ann Chen ^{1 2 4 5}

Affiliations + expand

PMID: 39998428 DOI: [10.1210/clinem/dgae861](#)

[J Clin Endocrinol Metab.](#) 2025 Aug 7;110(9):2605-2614. doi: 10.1210/clinem/dgae861.

Abstract

Aims: Sodium-glucose co-transporter 2 inhibitors (SGLT2i) enhance cardiovascular outcomes in individuals with type 2 diabetes mellitus (T2DM). Whether such effects also occur in T2DM patients with atrial fibrillation (AF) remains unknown. We aimed to investigate SGLT2i use on cardiovascular outcomes in patients with concomitant AF and T2DM.

Methods: Patients with both AF and T2DM were identified from TriNetX, an international electronic medical record. Participants were divided into 2 groups according to their use of SGLT2i, at a 1:1 distribution through propensity score matching (PSM). The hazard ratio (HR) for clinical outcomes was determined using multivariate Cox hazards regression model.

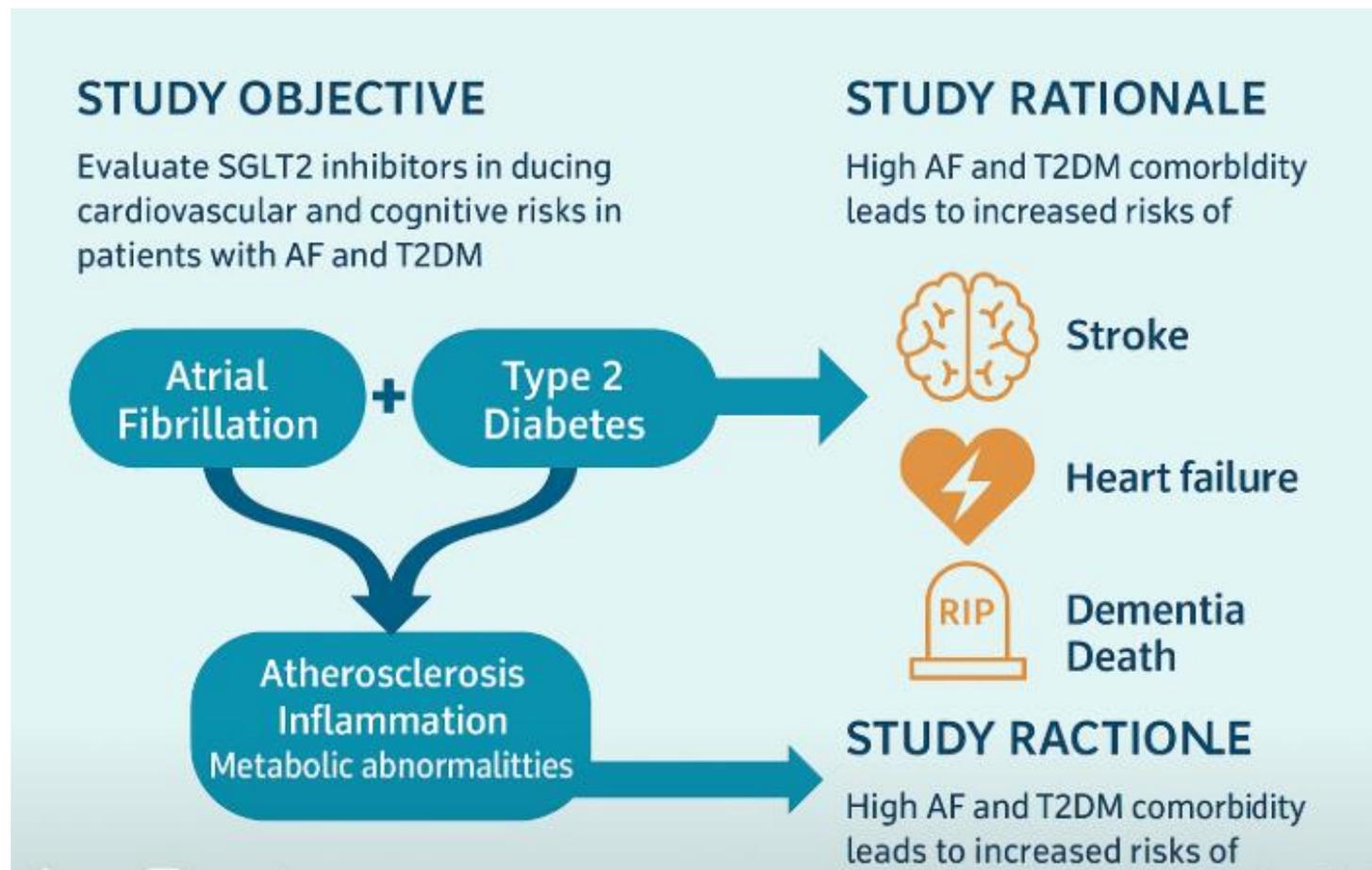
Results: We studied 339 792 patients with AF and T2DM, with 32 945 (9.70%) SGLT2i users. Following PSM, 17 011 patients aged 68.4 ± 7.9 years were included in each group. After a 3-year follow-up, patients treated with SGLT2i showed significantly reduced risks of stroke (adjusted HR: 0.830, $P < .001$), dementia (adjusted HR: 0.662, $P < .001$), long-standing persistent AF (adjusted HR: 0.917, $P < .001$), heart failure (adjusted HR: 0.833, $P < .001$), and all-cause mortality (adjusted HR: 0.532, $P < .001$).

Conclusion: The use of SGLT2i was associated with reduced risks of stroke, dementia, long-standing persistent AF, heart failure, and mortality in patients with both AF and T2DM. SGLT2i may be considered as a potential first-line therapy for this population.

Keywords: SGLT2 inhibitor; atrial fibrillation; dementia; heart failure; stroke; type 2 diabetes.

Study Objective

- ◆ 本研究旨在評估 鈉-葡萄糖協同轉運蛋白第二型抑制劑 (SGLT2 抑制劑) 是否能降低 心房顫動合併第二型糖尿病患者之 重大臨床不良事件風險，包含：
 - 缺血性中風
 - 失智症
 - 長期持續性心房纖維顫動 (心律不整惡化)
 - 心衰竭發生或惡化
 - 冠狀動脈疾病事件
 - 全因死亡



Study Rationale

- ◆ **心房顫動與第二型糖尿病高度共病**，兩者皆具促發血管硬化、發炎反應、代謝異常等特性，導致中風、失智、心衰竭等重大事件風險顯著增加。
- ◆ **目前臨床試驗極少以「心房顫動 + 糖尿病」作為主要研究族群**，多為次族群分析 (subgroup analysis)，因此缺乏直接證據評估此複雜族群的治療效益。
- ◆ **SGLT2 抑制劑在心腎代謝方面具有跨器官保護作用**，包含減少心衰竭住院、改善腎臟功能、降低死亡率等，但其在「同時患有心房顫動與糖尿病」的高風險族群中之效益仍有限研究。
- ◆ **大型真實世界資料庫 (如 TriNetX) 能補足臨床試驗的證據缺口：**
 - 可納入多重共病與高齡族群
 - 具足夠事件數量 (stroke、dementia 等長期結局)
 - 能反映真實臨床用藥行為→ 因此非常適合探討 SGLT2 抑制劑在此雙重疾病族群中的實際臨床效果。

Analysis Process

- ◆ Defining the cohorts through query criteria.
- ◆ Setting up and running the analysis.
 - **Setting up the Analysis**
 - Index event, outcomes criteria, and the time frame.
 - **Compare outcomes**
 - Measures of Association, Survival, Number of Instances and Lab result distribution



Cohorts definition

This section lists all terms used in the definitions of the two cohorts.

Query Criteria for Cohort 1 (query name: A SGLT2i)

This query was run on the network Research with 81 HCO(s) queried and 81 HCO(s) responded. A total of 67 provider(s) responded with patients. The final cohort included 33,100 patients who matched the query criteria listed in the table below. For the text representation of the query criteria please see Appendix A.

Ungrouped terms					
must have		demographics	Age		Age (between 55 and 85 years (most recent occurrence))
Group 1					
		AF			
must have	any of	diagnosis	UMLS:ICD10CM:I48.0		Paroxysmal atrial fibrillation
		diagnosis	UMLS:ICD10CM:I48.1		Persistent atrial fibrillation
		diagnosis	UMLS:ICD10CM:I48.2		Chronic atrial fibrillation
		diagnosis	UMLS:ICD10CM:I48.19		Other persistent atrial fibrillation
		diagnosis	UMLS:ICD10CM:I48.21		Permanent atrial fibrillation
		diagnosis	UMLS:ICD10CM:I48		Atrial fibrillation and flutter
		diagnosis	UMLS:ICD10CM:I48.20		Chronic atrial fibrillation, unspecified
		diagnosis	UMLS:ICD10CM:I48.11		Longstanding persistent atrial fibrillation
		diagnosis	UMLS:ICD10CM:I48.9		Unspecified atrial fibrillation and atrial flutter
		diagnosis	UMLS:ICD10CM:I48.91		Unspecified atrial fibrillation

date constraint

The terms in this group occurred on or after Jan 1, 2016

Cohorts definition

This section lists all terms used in the definitions of the two cohorts.

Query Criteria for Cohort 1 (query name: A SGLT2i)

This query was run on the network Research with 81 HCO(s) queried and 81 HCO(s) responded. A total of 67 provider(s) responded with patients. The final cohort included 33,100 patients who matched the query criteria listed in the table below. For the text representation of the query criteria please see Appendix A.

Group 2

DM+visit

must have	diagnosis	UMLS:ICD10CM:E08-E13	Diabetes mellitus
	and	visit	TNX:Visit
number of occurrences	Greater than or equal to 3 instances		
date constraint	The terms in this group occurred on or after Jan 1, 2016		

Group 3

Group 3A DM

must have	diagnosis	UMLS:ICD10CM:E08-E13	Diabetes mellitus
date constraint	The terms in this group occurred on or after Jan 1, 2016		
event relationship	Any instance of SGLT2i occurred within 3 years on or after the first instance of DM		

Group 3B SGLT2i

must have	any of	medication	NLM:RXNORM:1488564	dapagliflozin
		medication	NLM:RXNORM:1545653	empagliflozin
		medication	NLM:RXNORM:1373458	canagliflozin

Query Criteria for Cohort 2 (query name: B Non-SGLT2i)

This query was run on the network Research with 81 HCO(s) queried and 81 HCO(s) responded. A total of 73 provider(s) responded with patients. The final cohort included 308,774 patients who matched the query criteria listed in the table below.

Ungrouped terms					
must have		demographics	Age		Age (between 55 and 85 years (most recent occurrence))
Group 1					
	AF				
must have	any of	diagnosis	UMLS:ICD10CM:I48.0		Paroxysmal atrial fibrillation
		diagnosis	UMLS:ICD10CM:I48.1		Persistent atrial fibrillation
		diagnosis	UMLS:ICD10CM:I48.2		Chronic atrial fibrillation
		diagnosis	UMLS:ICD10CM:I48.19		Other persistent atrial fibrillation
		diagnosis	UMLS:ICD10CM:I48.21		Permanent atrial fibrillation
		diagnosis	UMLS:ICD10CM:I48		Atrial fibrillation and flutter
		diagnosis	UMLS:ICD10CM:I48.20		Chronic atrial fibrillation, unspecified
		diagnosis	UMLS:ICD10CM:I48.11		Longstanding persistent atrial fibrillation
		diagnosis	UMLS:ICD10CM:I48.9		Unspecified atrial fibrillation and atrial flutter
		diagnosis	UMLS:ICD10CM:I48.91		Unspecified atrial fibrillation
date constraint		The terms in this group occurred on or after Jan 1, 2016			

Query Criteria for Cohort 2 (query name: B Non-SGLT2i)

This query was run on the network Research with 81 HCO(s) queried and 81 HCO(s) responded. A total of 73 provider(s) responded with patients. The final cohort included 308,774 patients who matched the query criteria listed in the table below.

Group 2				
DM+visit				
must have		visit	TNX:Visit	Visit
	and	diagnosis	UMLS:ICD10CM:E08-E13	Diabetes mellitus
number of occurrences		Greater than or equal to 3 instances		
date constraint		The terms in this group occurred on or after Jan 1, 2016		
Group 3				
Group 3A DM				
must have		diagnosis	UMLS:ICD10CM:E08-E13	Diabetes mellitus
date constraint		The terms in this group occurred on or after Jan 1, 2016		
event relationship		Any instance of SGLT2i occurred on or after the first instance of DM		
Group 3B SGLT2i				
cannot have		medication	NLM:RXNORM:1488564	dapagliflozin
	or	medication	NLM:RXNORM:1545653	empagliflozin
	or	medication	NLM:RXNORM:1373458	canagliflozin

Analysis Setup : 分析設定重點

- ◆ Defining the cohorts through query criteria.
- ◆ Setting up and running the analysis.
 - **Setting up the Analysis**
 - Index event, outcomes criteria, and the time frame.
 - **Compare outcomes**
 - Measures of Association, Survival, Number of Instances and Lab result distribution

Analysis Process

TriNetX Studies Connect Browse Network What's New 5

My Studies / Untitled Study / My Analyses / New

Analytics

My Analyses

Analyses that are currently available to me.

Analyze Outcomes
How do patients in a cohort experience outcomes?
Explore Characteristics | Review Outcomes
Single Cohort

Compare Outcomes
How do outcomes compare between cohorts?
Compare Characteristics | Review Outcomes
Two Cohorts

Compare Cohorts
How do patient characteristics compare between cohorts?
Compare Characteristics
Two Cohorts

Competing Risks
What is the probability for a patient to have competing risks?
Review Characteristics | Review Competing Risks
Single Cohort

Treatment Pathways
In what order do patients receive treatments following a diagnosis?
Review Characteristics | Review Treatment Pathways
Single Cohort

Incidence and Prevalence
What are the incidence and prevalence of events of interest in a cohort?
Review Incidence and Prevalence
Single Cohort

Advanced Explore Cohort
What are the characteristics of my cohort in different time periods?
Explore Characteristics | Explore Outcomes
Single Cohort

Cox Proportional Hazards Model
What impact do a collection of predictors have on an outcome?
Select Covariates | Review Outcomes
Two Cohorts

Analysis Setup：分析設定重點

◆ 指標事件 (Index Event)

- ✓ 決定分析起始點指標事件是每位患者「正式進入分析的起點」。
- ✓ 患者第一次符合 cohort 定義的那一天。
- ✓ 第一次同時符合「心房顫動」+「第二型糖尿病」的日期。
- ✓ 每位患者的 index date 都會自動由 TriNetX 根據 cohort 條件計算。

◆ 追蹤期間 (Time Window)

- ✓ 本研究設定追蹤期間為：指標事件後第 1 天 → 第 1,095 天 (3 年) 任何事件 (診斷、用藥、處置、檢驗) 只要發生在這個期間內，才會被算進本研究。
- ✓ 追蹤期開始在第 1 天是為了避免「同一天的診斷」造成暴露混淆。

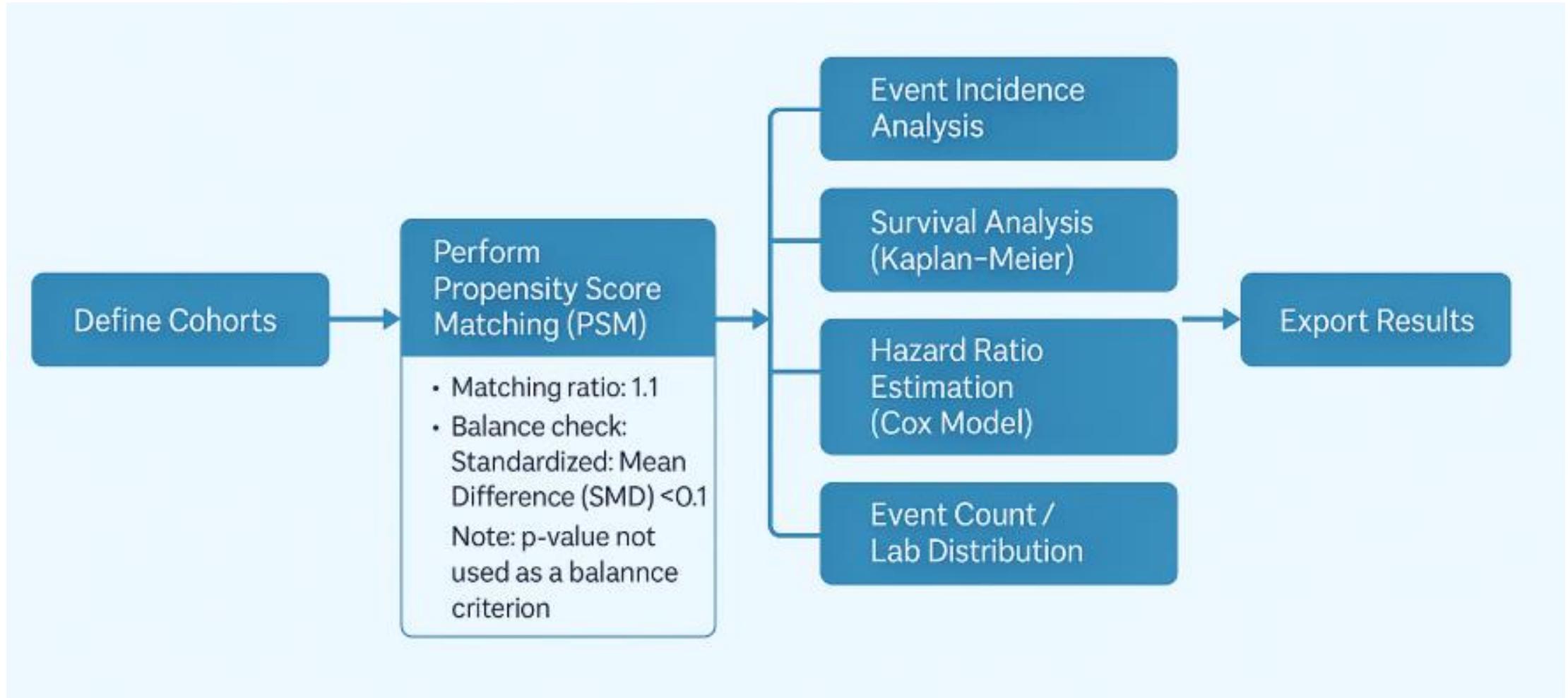
◆ TriNetX 只能回溯 最多 20 年內的指標事件。若病人首次符合條件是在 20 年前 → 會被排除。

Outcome Definitions

Table below outlines the definitions for each outcome and the analysis specifications. For outcome definitions consisting of more than one term, at least one term must match. Please see Appendix C for the text representation of the outcome definitions.

All-cause mortality		
Outcome definition		
Demographics	Deceased	Deceased
Settings for the performed analyses		
Risk analysis		excluding patients with outcome prior to the time window
Kaplan - Meier survival analysis		excluding patients with outcome prior to the time window
Stroke		
Outcome definition		
Diagnosis	UMLS:ICD10CM:I60	Nontraumatic subarachnoid hemorrhage
Diagnosis	UMLS:ICD10CM:I61	Nontraumatic intracerebral hemorrhage
Diagnosis	UMLS:ICD10CM:I62	Other and unspecified nontraumatic intracranial hemorrhage
Diagnosis	UMLS:ICD10CM:I63	Cerebral infarction
Diagnosis	UMLS:ICD10CM:I65	Occlusion and stenosis of precerebral arteries, not resulting in cerebral infarction
Diagnosis	UMLS:ICD10CM:I66	Occlusion and stenosis of cerebral arteries, not resulting in cerebral infarction
Diagnosis	UMLS:ICD10CM:I67	Other cerebrovascular diseases
Diagnosis	UMLS:ICD10CM:I68	Cerebrovascular disorders in diseases classified elsewhere
Diagnosis	UMLS:ICD10CM:G46	Vascular syndromes of brain in cerebrovascular diseases
Settings for the performed analyses		
Risk analysis		excluding patients with outcome prior to the time window
Kaplan - Meier survival analysis		excluding patients with outcome prior to the time window

Analysis Process

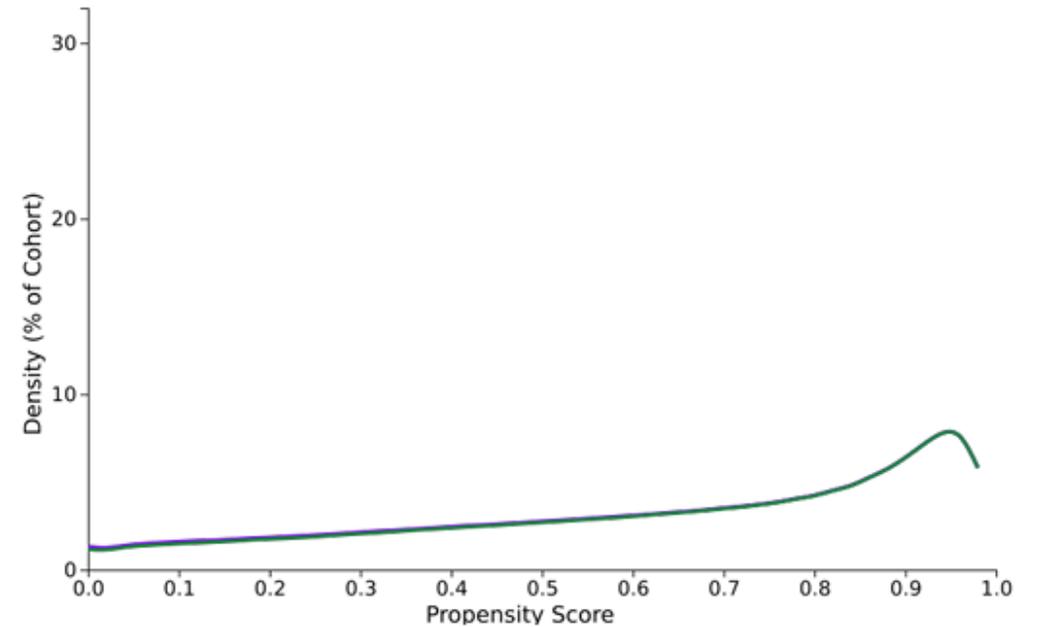
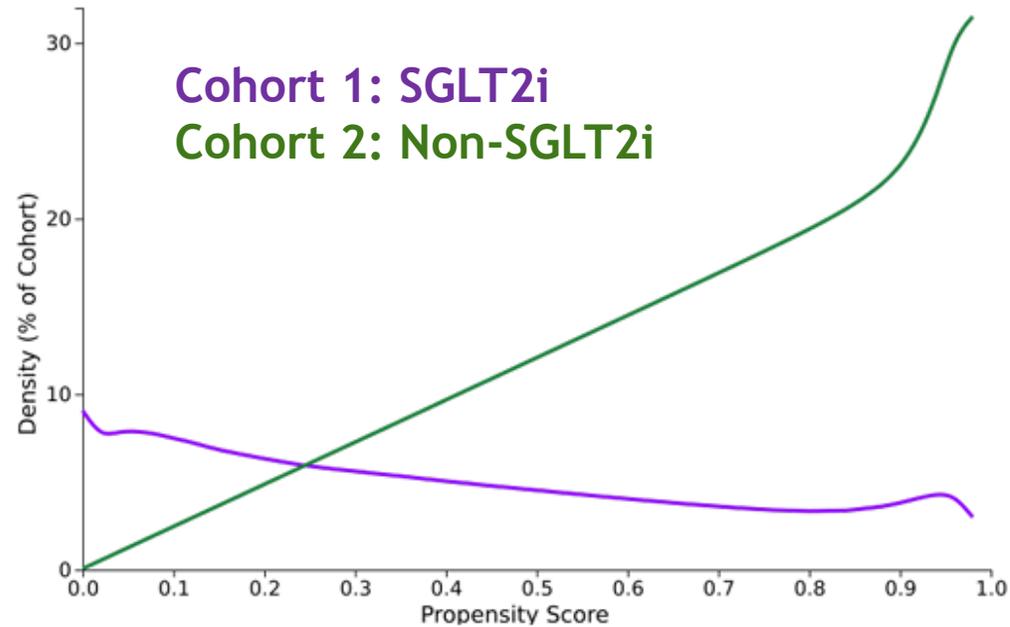


Propensity Score 1:1 Matching

Cohort 1 and cohort 2 patient count before and after propensity score matching

Cohort	Patient count before matching	Patient count after matching
1 - A SGLT2i	32,945	17,011
2 - B Non-SGLT2i	306,847	17,011

Propensity score density function - Before and after matching (cohort 1 - purple, cohort 2 - green)



Propensity Score 1:1 Matching

**Balance
SMD < 0.1**

Cohort 1 (N = 17,011) and cohort 2 (N = 17,011) characteristics after propensity score matching

Demographics

Cohort			Mean ± SD	Patients	% of Cohort	P-Value	Std diff.
1	AI	Age at Index	68.4 +/- 7.9	17,011	100%	0.611	0.006
2			68.4 +/- 7.8	17,011	100%		
1	F	Female		5,226	30.7%	0.144	0.016
2			5,102	30.0%			
1	2054-5	Black or African American		1,843	10.8%	0.563	0.006
2			1,810	10.6%			

Diagnosis

Cohort			Mean ± SD	Patients	% of Cohort	P-Value	Std diff.
1	I10	Essential (primary) hypertension		15,073	88.6%	0.560	0.006
2			15,107	88.8%			
1	I15	Secondary hypertension		702	4.1%	0.935	0.001
2			705	4.1%			
1	E78	Disorders of lipoprotein metabolism and other lipidemias		14,659	86.2%	0.741	0.004
2			14,680	86.3%			
1	N18	Chronic kidney disease (CKD)		6,140	36.1%	0.182	0.014
2			6,022	35.4%			

After Propensity Score 1:1 Matching

The Journal of Clinical Endocrinology & Metabolism, 2024, Vol. 00, No. 0

Table 1. Baseline characteristics of the study cohort before and after propensity score matching

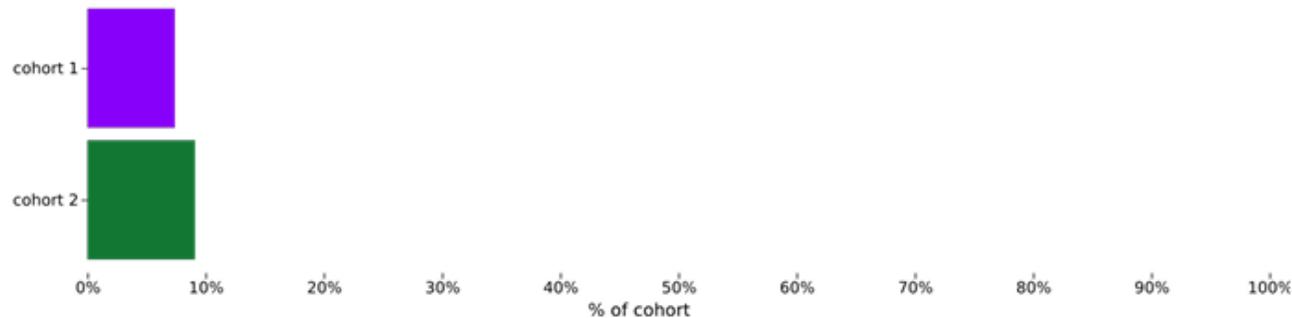
Variables	Initial group				Propensity score-matched group			
	SGLT2i (n = 32 945)	Non-SGLT2i (n = 306 847)	P-value	SDM	SGLT2i (n = 17 011)	Non-SGLT2i (n = 17 011)	P-value	SDM
Demographics								
Age at baseline (years)	68.1 ± 7.83	69.6 ± 7.71	<.001	0.19	68.4 ± 7.88	68.4 ± 7.81	.61	0.006
Male, (%)	21 666 (65.8)	175 370 (57.2)	<.001	0.18	11 007 (64.7)	11 160 (65.6)	.08	0.02
Race (n, %)								
Asian	1044 (3.17)	9188 (2.99)	.08	0.01	570 (3.35)	560 (3.29)	.76	0.00
White	22 932 (69.61)	209 219 (68.18)	<.001	0.03	11 746 (69.05)	11 880 (69.84)	.11	0.02
Black or African American	3547 (10.77)	37 918 (12.36)	<.001	0.05	1843 (10.83)	1810 (10.64)	.56	0.01
Native Hawaiian or Other Pacific Islander	269 (0.82)	2528 (0.82)	.89	0.00	161 (0.95)	155 (0.91)	.73	0.00
American Indian or Alaska Native	88 (0.27)	863 (0.28)	.64	0.00	40 (0.24)	32 (0.19)	.35	0.01
Unknown race	4393 (13.33)	41 326 (13.47)	.50	0.00	2310 (13.58)	2245 (13.2)	.3	0.01
Other race	672 (2.04)	5805 (1.89)	.06	0.01	341 (2.01)	329 (1.93)	.64	0.01
Underlying diseases (n, %)								
Essential (primary) hypertension	28 855 (87.59)	260 692 (84.96)	<.001	0.08	15 073 (88.61)	15 107 (88.81)	.56	0.01
Secondary Hypertension	1389 (4.22)	11 687 (3.81)	<.001	0.02	702 (4.13)	705 (4.14)	.93	0.00
Hyperlipidemia	28 251 (85.75)	238 857 (77.84)	<.001	0.21	14 659 (86.17)	14 680 (86.3)	.74	0.00
Chronic kidney disease	11 027 (33.47)	115 271 (37.57)	<.001	0.09	6140 (36.09)	6022 (35.4)	.18	0.01
Medication uses (n, %)								
Oral hypoglycemic agents	28 498 (86.5)	145 923 (47.56)	<.001	0.91	12 564 (73.86)	12 904 (75.86)	<.001	0.05
Metformin	20 599 (62.53)	115 726 (37.72)	<.001	0.51	9975 (58.64)	10 289 (60.48)	<.001	0.04
Glipizide	5776 (17.53)	34 654 (11.29)	<.001	0.18	2788 (16.39)	2899 (17.04)	.11	0.02
Sitagliptin	6146 (18.66)	25 535 (8.32)	<.001	0.31	2773 (16.3)	2858 (16.8)	.22	0.01
Empagliflozin	11 828 (35.9)	0 (0.00)	NA	NA	1447 (8.51)	0 (0.00)	NA	NA
Dapagliflozin	6161 (18.7)	0 (0.00)	NA	NA	790 (4.64)	0 (0.00)	NA	NA
Canagliflozin	4110 (12.48)	0 (0.00)	NA	NA	1046 (6.15)	0 (0.00)	NA	NA

Risk Analysis After Propensity Score 1:1 Matching

2 Stroke

Risk analysis excluding patients with outcome prior to the time window

Cohort	Patients in cohort	Patients with outcome	Risk	
1 A SGLT2i	12,803	946	0.074	
2 B Non-SGLT2i	12,495	1,137	0.091	
		95% CI	z	p
Risk Difference	-0.017	(-0.024, -0.010)	-4.949	0.000
Risk Ratio	0.812	(0.748, 0.882)	N/A	N/A
Odds Ratio	0.797	(0.728, 0.872)	N/A	N/A



4,208 patients in Cohort 1 and 4,516 patients in Cohort 2 were excluded from results because they had the outcome prior to the time window.

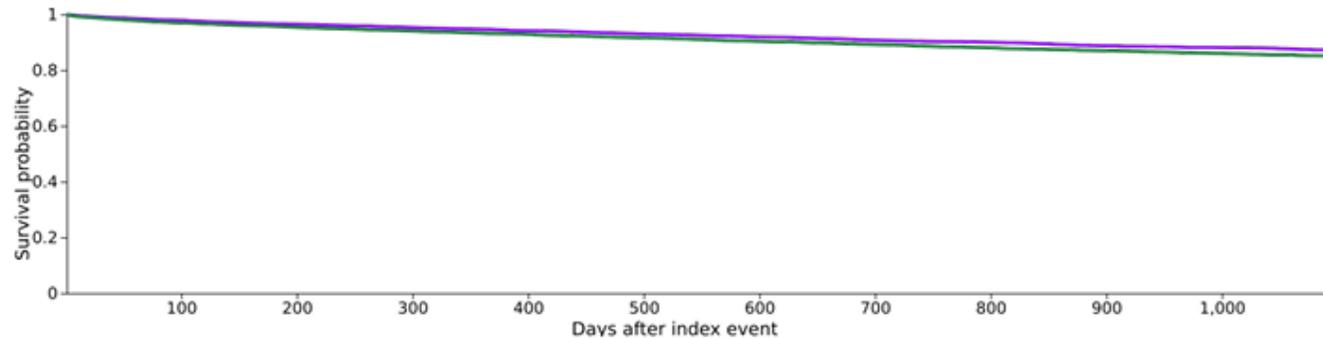
Risk Analysis After Propensity Score 1:1 Matching

Kaplan - Meier survival analysis excluding patients with outcome prior to the time window

Cohort	Patients in cohort	Patients with outcome	Median survival (days)	Survival probability at end of time window
1 A SGLT2i	12,803	946	--	87.14%
2 B Non-SGLT2i	12,495	1,137	--	85.16%

	χ^2	df	p
Log-Rank Test	17.940	1	0.000

	Hazard Ratio	95% CI	χ^2	df	p
Hazard Ratio and Proportionality	0.830	(0.761, 0.905)	3.456	1	0.000



4,208 patients in Cohort 1 and 4,516 patients in Cohort 2 were excluded from results because they had the outcome prior to the time window.

Risk Analysis After Propensity Score 1:1 Matching

Table 2. Risks of primary and secondary outcomes

Events	Total number for analysis	3-year cumulative rate (%)	Hazard ratio, 95% CI	P-value
Primary outcome				
Stroke ^a				
Non-SGLT2i	12 495	1137 (9.10)	1.0	Ref.
SGLT2i	12 803	946 (7.39)	0.830 (0.761, 0.905)	<.001
Incident dementia ^a				
Non-SGLT2i	16 219	444 (2.74)	1.0	Ref.
SGLT2i	16 453	287 (1.74)	0.662 (0.571, 0.768)	<.001
Long-standing persistent AF ^a				
Non-SGLT2i	17 011	12 244 (72.0)	1.0	Ref.
SGLT2i	17 011	12 013 (70.62)	0.917 (0.894, 0.940)	<.001
Secondary outcomes				
Heart failure ^a				
Non-SGLT2i	8896	1450 (16.30)	1.0	Ref.
SGLT2i	7156	1026 (14.34)	0.833 (0.769, 0.902)	<.001
Coronary artery disease ^a				
Non-SGLT2i	6771	946 (13.97)	1.0	Ref.
SGLT2i	5885	51 (6.30)	0.927 (0.851, 1.010)	.085
All-cause death ^a				
Non-SGLT2i	16 815	3238 (19.26)	1.0	Ref.
SGLT2i	16 883	1685 (0.10)	0.532 (0.502, 0.564)	<.001

Abbreviations: AF, atrial fibrillation; CI, confidence interval; Ref, reference; SGLT2i, sodium-glucose cotransporter-2 inhibitor.

^aConditional Cox proportional hazard model was adjusted for all relevant covariates, including baseline characteristics, comorbidities, and medications used.

Risk Analysis After Propensity Score 1:1 Matching

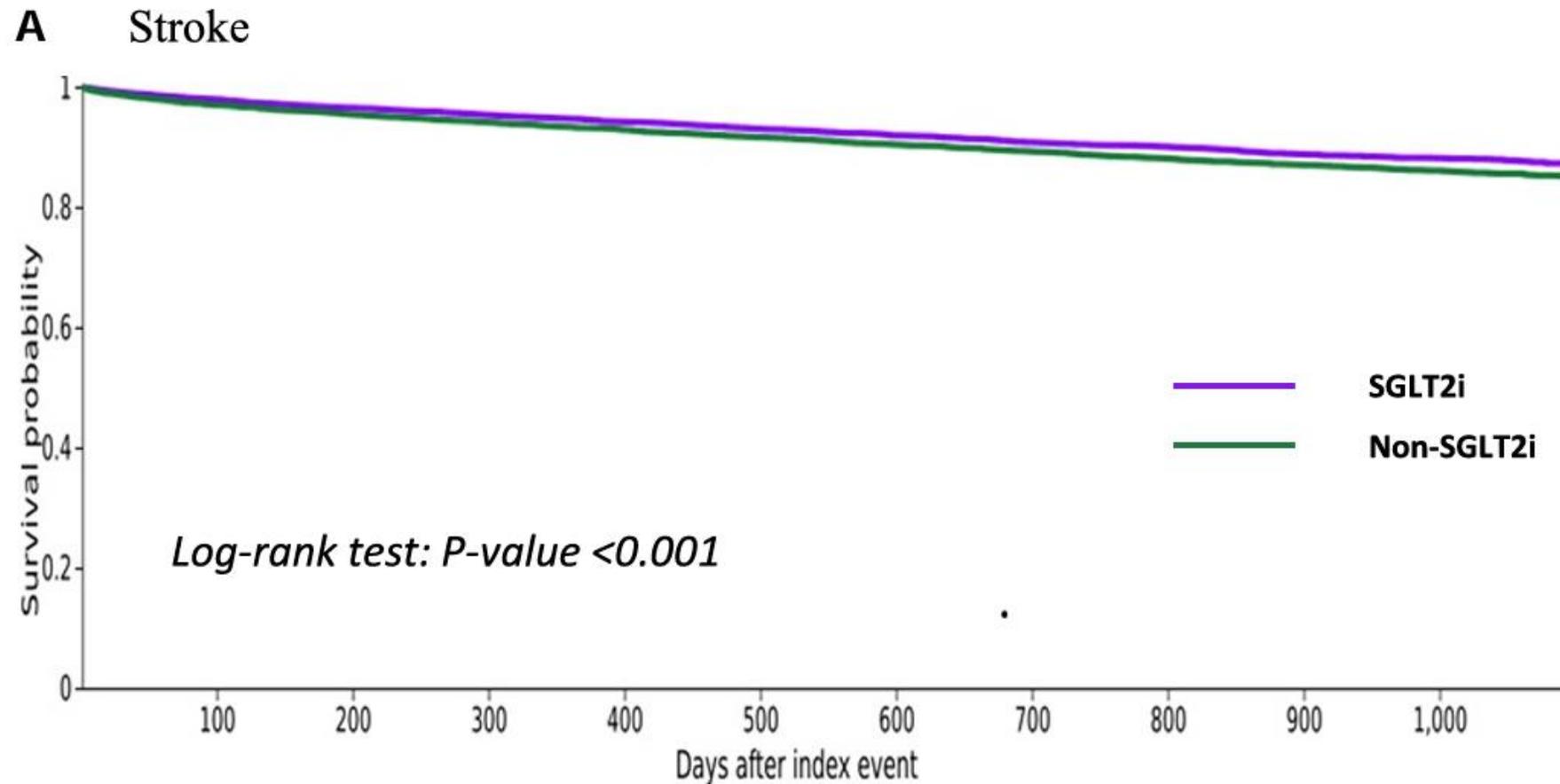
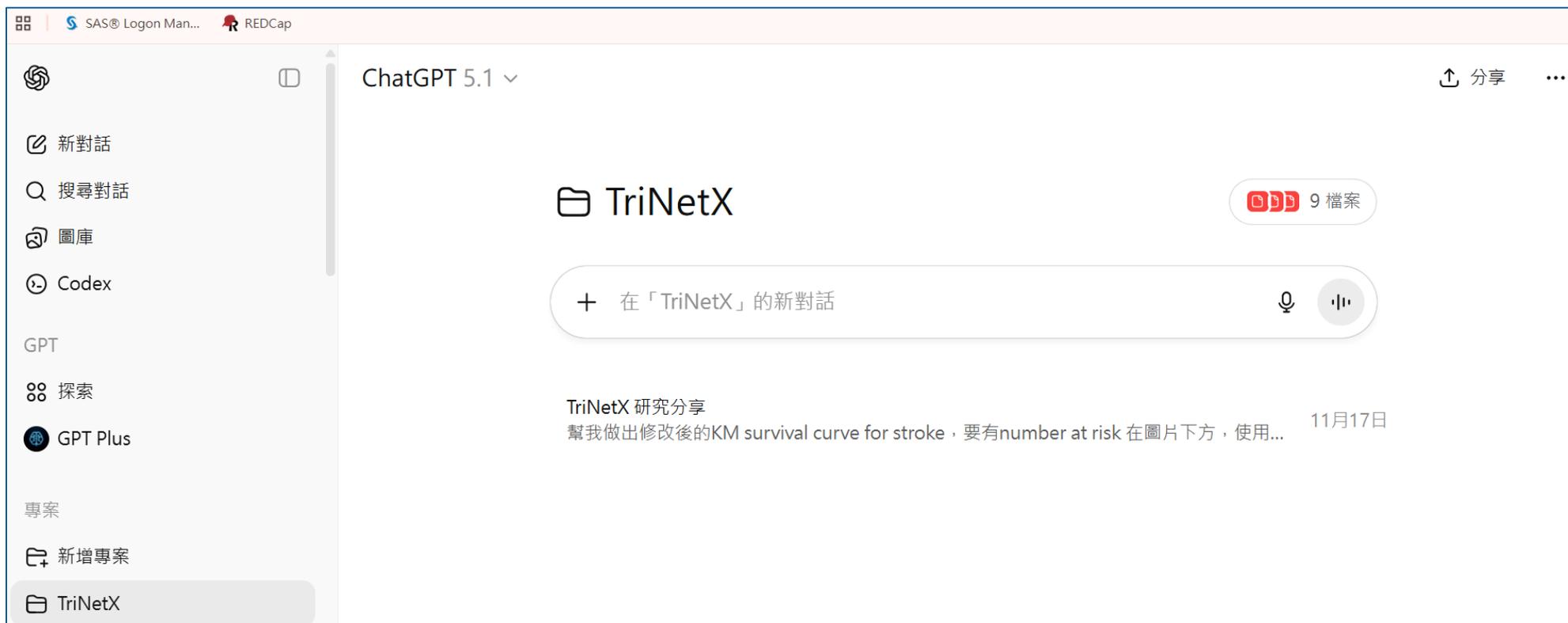


Figure 1. Kaplan–Meier survival curves for cardiovascular events of (A) stroke, (B) incident dementia, and (C) long-standing persistent atrial fibrillation.

Part 3

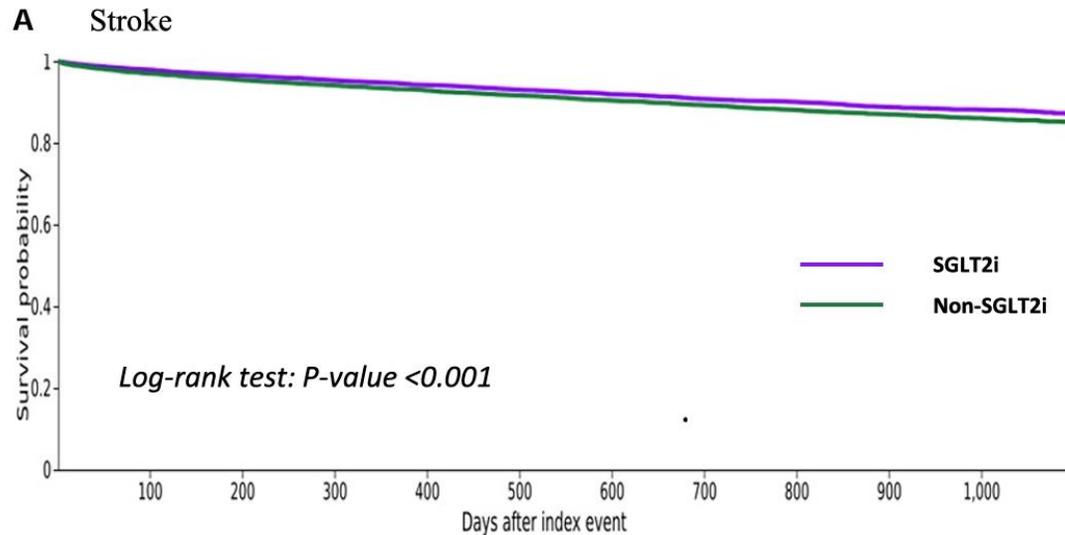
善用AI：Chat GPT修改圖片



善用AI：Chat GPT修改圖片

- ▶ 修改這張KM圖，X軸尺規改365天為一個單位；Y軸尺規改0.7-1，並且每0.1為一個單位；使這張圖更適合專業期刊發表，並根據兩組真實的存活比例，計算number at risk，每365天給一個人數數字，並給我number at risk的計算公式

Events	Total number for analysis	3-year cumulative rate (%)	Hazard ratio, 95% CI	P-value
Primary outcome				
Stroke ^a				
Non-SGLT2i	12 495	1137 (9.10)	1.0	Ref.
SGLT2i	12 803	946 (7.39)	0.830 (0.761, 0.905)	<.001



Number at risk

盡量不要用「存活率 × 初始人數」去估，審稿時會被看出來

要算出真實的 number at risk，必須有：
每位病人的 time-to-event / censor 資料，
或至少 TriNetX 匯出的 KM summary table
(含 time, n.risk, n.event, n.censor)

AI後製輔助應用 (Post-processing / Visualization)

SAS® Logon Man... REDCap

ChatGPT 5.1

1137 (9.10) st: P-value <0.001

946 (7.39)

Days after index event

請做出接近真實的number at risk: 修改這張KM圖，X軸尺規改365天為一個單位；Y軸尺規改0.7-1，並且每0.1為一個單位；使這張圖更適合專業期刊發表，並根據兩組真實的存活比例，計算number at risk，每365天給一個人數數字

已思考 13s >

已思考 45s >

好，我幫你做到三件事：

1. 給一組「接近真實」的 number at risk (每 365 天)
2. 說明我是用什麼假設計出來的 (讓你在心裡有底)
3. 整理出 number at risk 的正式計算公式，方便寫 Methods

AI後製輔助應用 (Post-processing / Visualization)

📁 > ChatGPT 5.1 ▾

1 依你提供的真實三年發生率，推估的 number at risk (每 365 天)

你給的結果是：

- **Non-SGLT2i 組**
 - Total number for analysis : 12,495
 - 3-year cumulative rate : 9.10% (1137 件 stroke)
- **SGLT2i 組**
 - Total number for analysis : 12,803
 - 3-year cumulative rate : 7.39% (946 件 stroke)

這代表 3 年時大約：

- Non-SGLT2i 存活比例 $\approx 1 - 0.0910 = 0.909$
- SGLT2i 存活比例 $\approx 1 - 0.0739 = 0.9261$

AI後製輔助應用 (Post-processing / Visualization)

📁 > ChatGPT 5.1 ▾

- Non-SGLT2i 存活比例 $\approx 1 - 0.0910 = 0.909$
- SGLT2i 存活比例 $\approx 1 - 0.0739 = 0.9261$

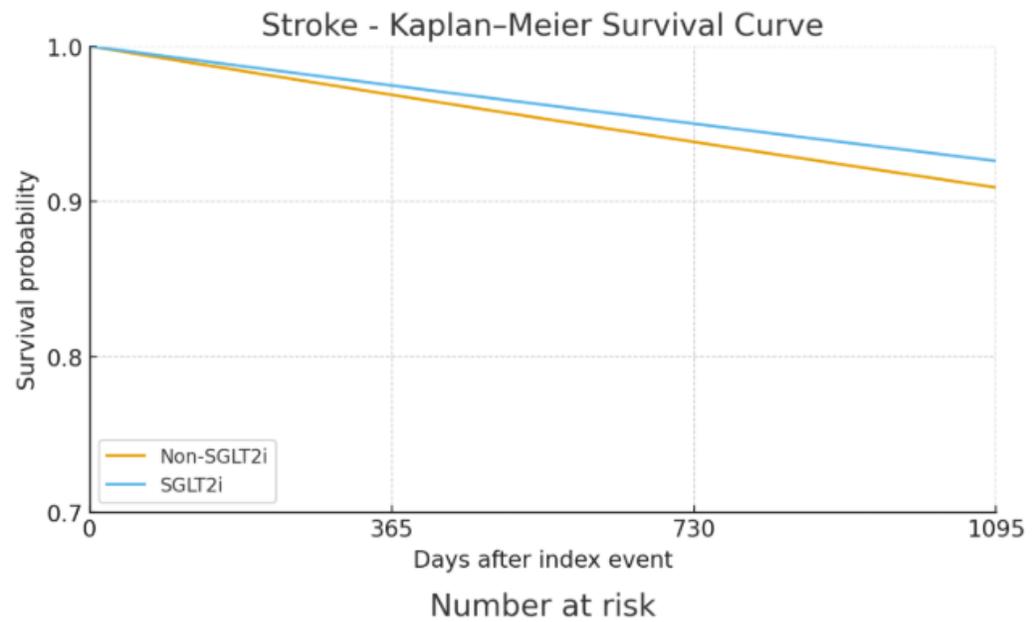
我用「常數 hazard (指數分佈) + 幾乎不考慮其他截尾」這個假設，讓 3 年時的存活率剛好對到 0.909、0.9261，然後往前推 365、730 天的存活率，再乘以 baseline 人數，得到**接近真實**的 number at risk (與三年事件數可以完全對上)：

(四捨五入到整數)

Time after index (days)	Non-SGLT2i — n at risk	SGLT2i — n at risk
0	12,495	12,803
365	12,104	12,480
730	11,725	12,164
1,095	11,358	11,857

AI後製輔助應用 (Post-processing / Visualization)

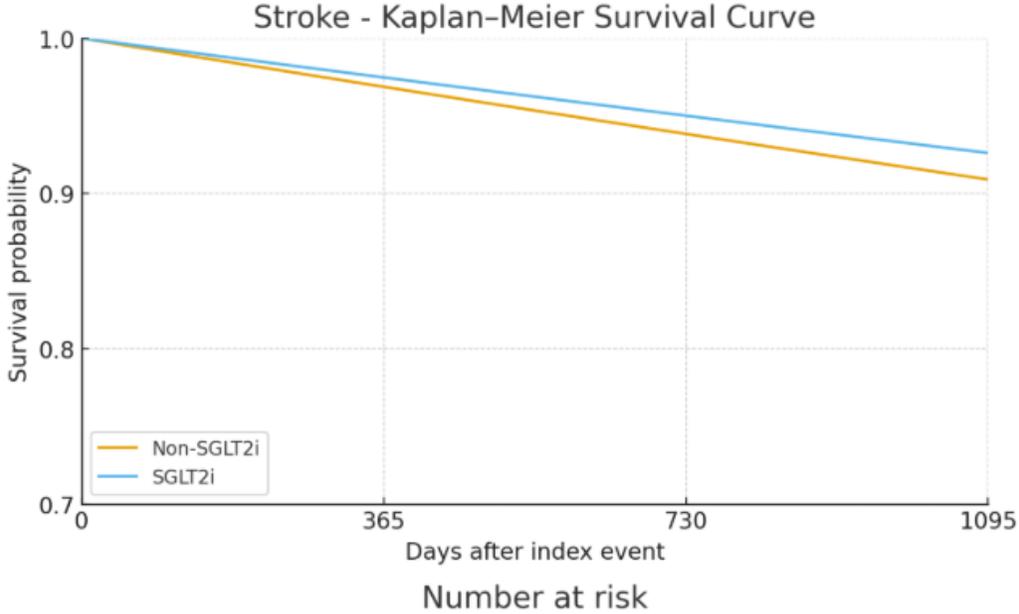
📁 > ChatGPT 5.1 ▾



	0	365	730	1095
Non-SGLT2i	12495	12104	11725	11358
SGLT2i	12803	12480	12164	11857

幫我做出修改後的KM survival curve for stroke，要有number at risk在圖片下方，使用英文，符合專業期刊發表的規範，做出接近真實的number at risk: 修改這張KM圖，X軸尺規改365天為一個單位；Y軸尺規改0.7-1，並且每0.1為一個單位；使這張圖更適合專業期刊發表，並根據兩組真實的存活比例，計算number at risk，每365天給一個人數數字

AI後製輔助應用 (Post-processing / Visualization)

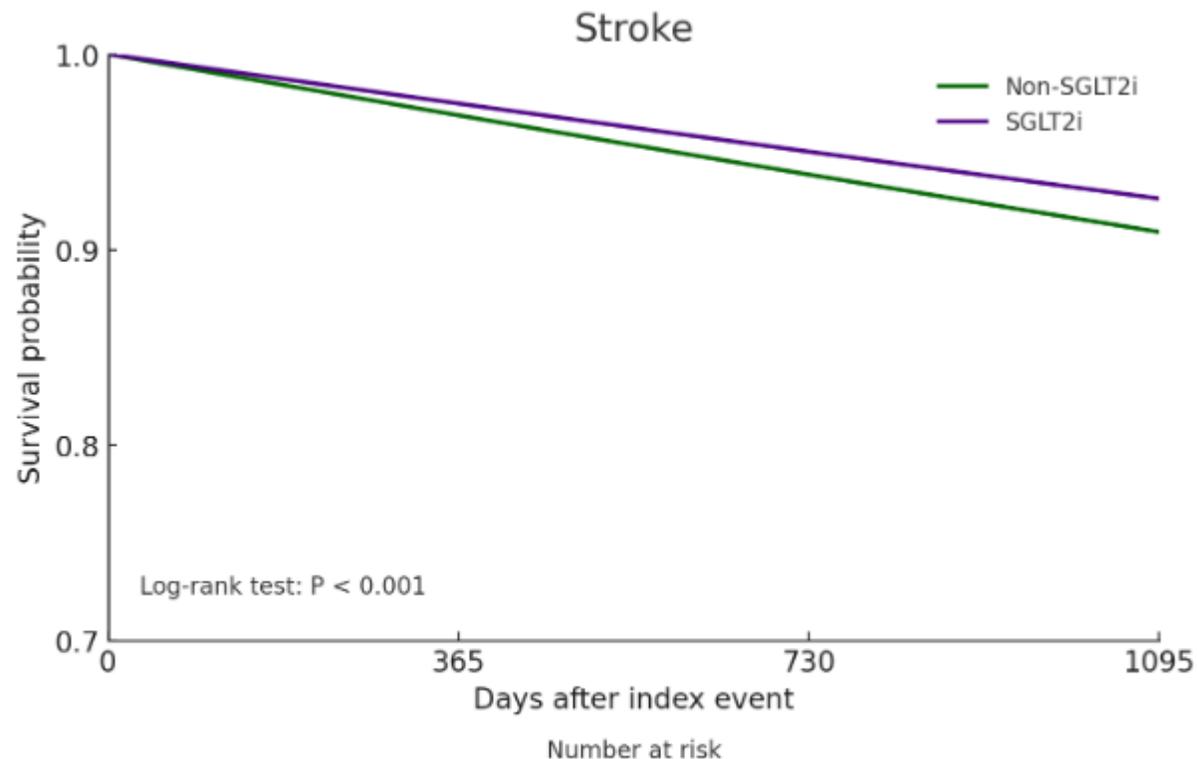


	0	365	730	1095
Non-SGLT2i	12495	12104	11725	11358
SGLT2i	12803	12480	12164	11857

AI後製輔助應用 (Post-processing / Visualization)

ChatGPT 專業版指令→

幫我做出修改後的KM survival curve for stroke，要有number at risk 在圖片下使用英文，符合專業期刊發表的規範。做出接近真實的number at risk: 修改這張KM圖，X軸尺規改365天為一個單位，Y軸尺規改0.7-1，並且每0.1為一個單位。使這張圖更適合專業期刊發表，並根據兩組真實的存活比例，計算number at risk，每365天給一個人數數字；設定Non-SGLT2i組為深綠線，SGLT2i組為深紫線，並注意圖式標示的顏色，標示不要遮隱到圖片線條。



	0	365	730	1095
Non-SGLT2i	12495	12104	11725	11358
SGLT2i	12803	12480	12164	11857

Part 4

投稿過程中的特殊問題

Q3:

Time of exposure to the i-SGLT2 medication is not stated or controlled in the statistical analysis.

↵

Reply:

Thank you for raising the important issue regarding the time of exposure to i-SGLT2 medication. We acknowledge that the duration of exposure could impact outcomes and should ideally be considered in the analysis.

↵

In this study, we were unable to adjust directly for the duration of exposure within the statistical model due to the constraints of our dataset and methodology. However, we have analyzed the treatment pathways using TriNetX, which provides insight into the typical durations of medication use among our study cohorts.

↵

To further illustrate this point, we have included a new supplemental figure (Figure S 2), which details the distribution and median duration of i-SGLT2 medication use in our cohorts. This figure can be referenced for a clearer understanding of how exposure time has been considered indirectly through our treatment pathway analysis.

↵

We appreciate your attention to this detail and hope that the supplemental Figure S2 helps clarify the Methods (Lines 143-144) and Limitations (Lines 272-276).

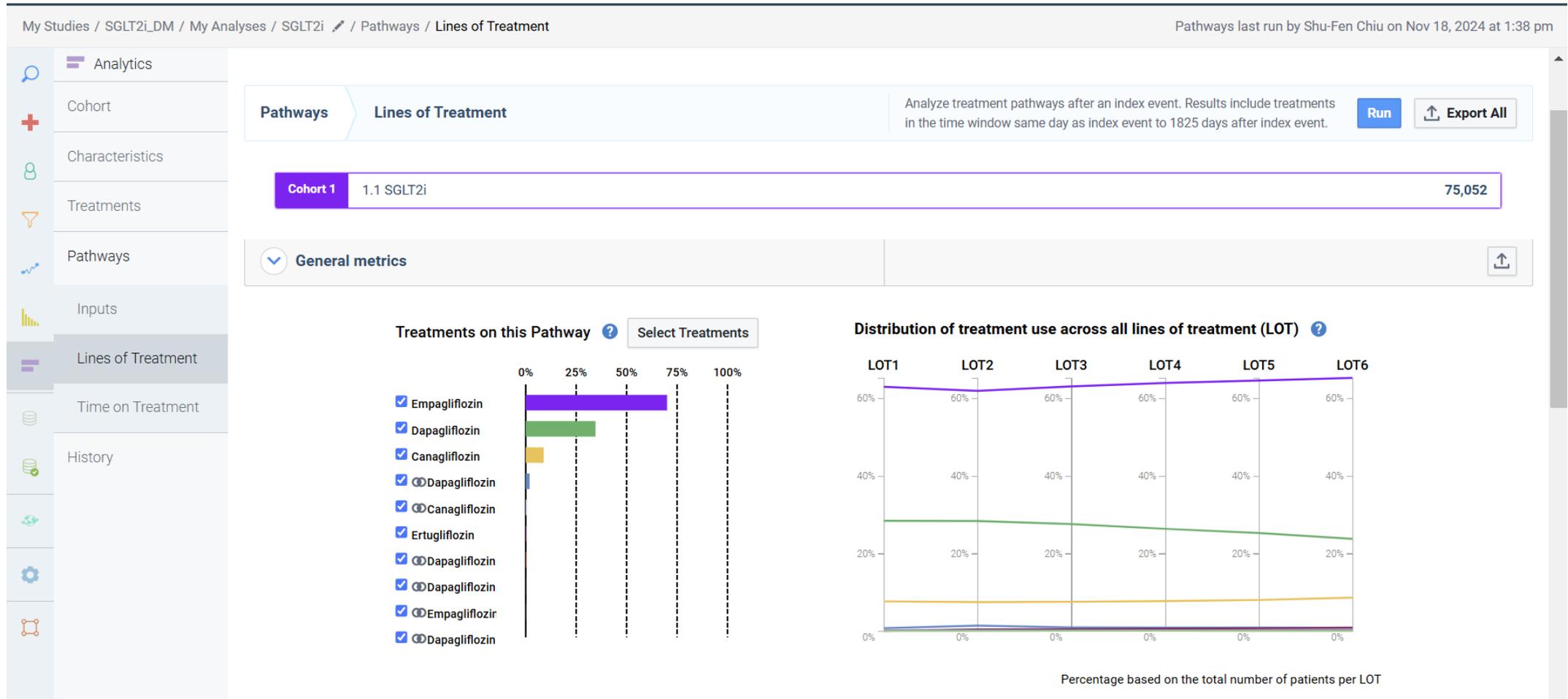
治療途徑分析 (Treatment Pathways)

治療途徑分析可以辨識出某個群體中最常用的治療方法、聯合治療方式及治療順序，並追蹤病人在發生指標事件後接受治療的順序與治療時間(Time on Treatment)。

The screenshot displays the TriNetX Analytics dashboard. The top navigation bar includes the TriNetX logo, a 'Studies' dropdown menu, and links for 'Connect', 'Browse Network', and 'What's New' (with a notification badge for 5). The breadcrumb trail shows 'My Studies / Untitled Study / My Analyses / New'. The main content area is titled 'My Analyses' and lists several analysis options, each with a brief description, a representative chart, and available actions. The 'Treatment Pathways' option is highlighted with a red border. It is described as 'In what order do patients receive treatments following a diagnosis?' and includes a 'Review Treatment Pathways' action. Other visible options include 'Analyze Outcomes', 'Compare Outcomes', 'Compare Cohorts', 'Competing Risks', 'Incidence and Prevalence', 'Advanced Explore Cohort', and 'Cox Proportional Hazards Model'.

一般指標(General metrics)：

包括按降序排列的每種藥物的患者比例，以及所有治療線中的治療使用分佈。



追蹤指標 (Follow-up Metrics) : 患者人數以及平均追蹤時間

My Studies / SGLT2i_DM / My Analyses / SGLT2i / Pathways / Time on Treatment Pathways last run by Shu-Fen Chiu on Nov 18, 2024 at 1:38 pm

- Analytics
- Cohort
- Characteristics
- Treatments
- Pathways
- Inputs
- Lines of Treatment
- Time on Treatment**
- History

Pathways
Time on Treatment

Analyze treatment time after an index event. Results include treatments same day as index event to 1825 days after index event.

Run

Cohort 1
1.1 SGLT2i
75,052

i Results are incomplete; one or more HCOs failed to return results.

v
Follow-up Metrics
↑

Cohort	Patients in Cohort	Follow-up Metrics				
		Mean Follow-up (Days)	Standard Deviation	Median Follow-up (Days)	Interquartile Range	
1	1.1 SGLT2i	75,052	808.537	581.168	704	919

v
Follow-up Time
↑

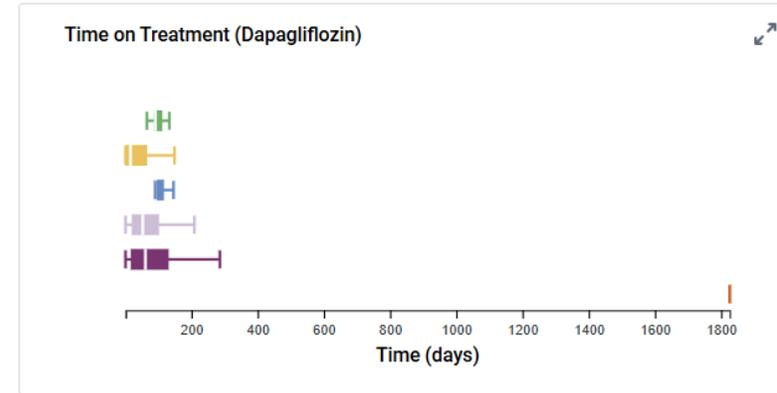
Dapagliflozin 的治療持續時間和保留率

Time on Treatment

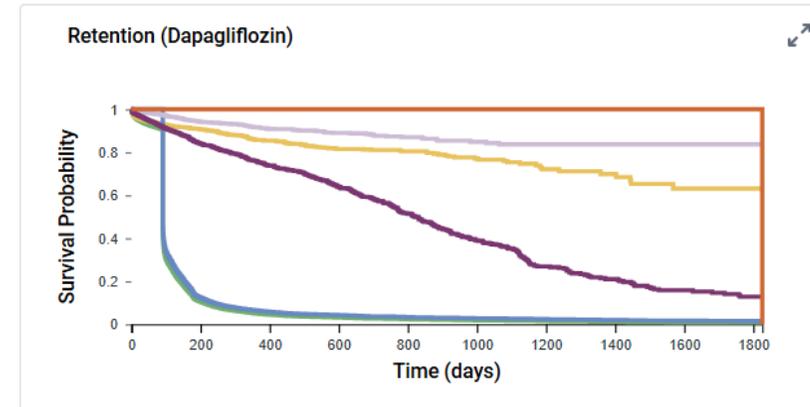
Time patients stayed on the treatment



Treatment	n	Mean ± SD (days)	Median	Min	Max
Dapagliflozin (all)	21,201	121.785 ± 143.741	91	1	1,826
All switches	1,647	63.795 ± 152.231	16	1	1,568
Treatment end	16,436	123.766 ± 98.53	91	91	1,735
Death	704	93.813 ± 137.887	53	1	1,064
End of record	2,393	157.071 ± 272.198	61	1	1,762
End of time window	21	1,826 ± 0	1,826	1,826	1,826

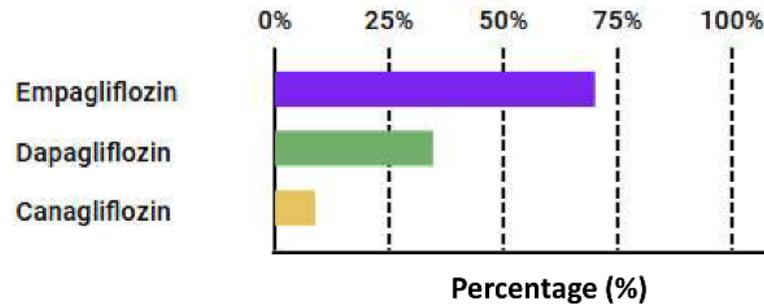


Retention	n	Median Survival
<input checked="" type="checkbox"/> Dapagliflozin (all)	21,201	91
<input checked="" type="checkbox"/> All switches	1,647	-
<input checked="" type="checkbox"/> Treatment end	16,436	91
<input checked="" type="checkbox"/> Death	704	-
<input checked="" type="checkbox"/> End of record	2,393	820
<input checked="" type="checkbox"/> End of time window	21	1,826

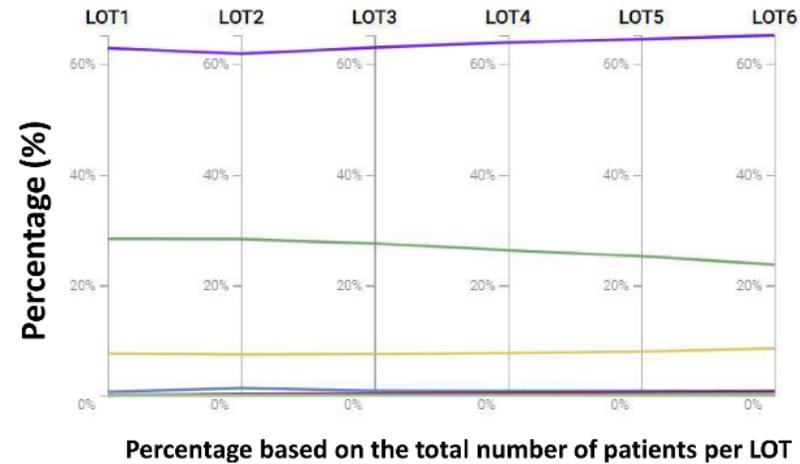


治療途徑分析 (Treatment Pathways)

(A) Treatments on pathway



(B) Distribution of treatment use across all lines of treatment (LOT)

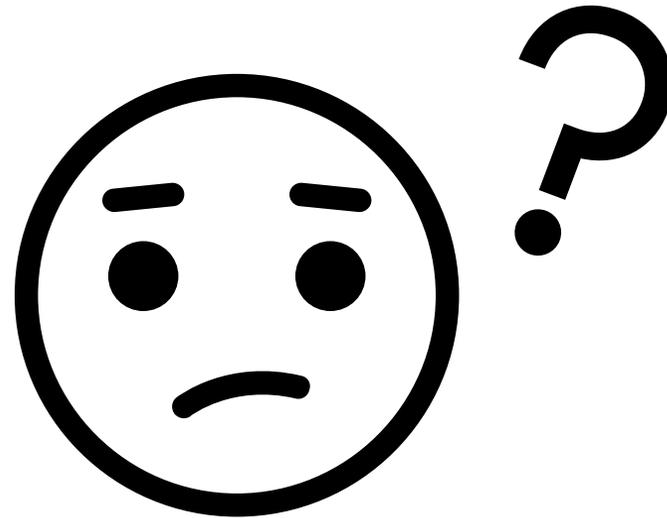


(C) Follow-up metrics

	Mean (days)	Standard deviation (days)	Median (days)
Whole period			
SGLT2i	808.5	581.2	704
LOT1			
Dapagliflozin	121.8	143.7	91
Empagliflozin	124.7	141.7	91
Canagliflozin	122.1	142.4	91

Part 5

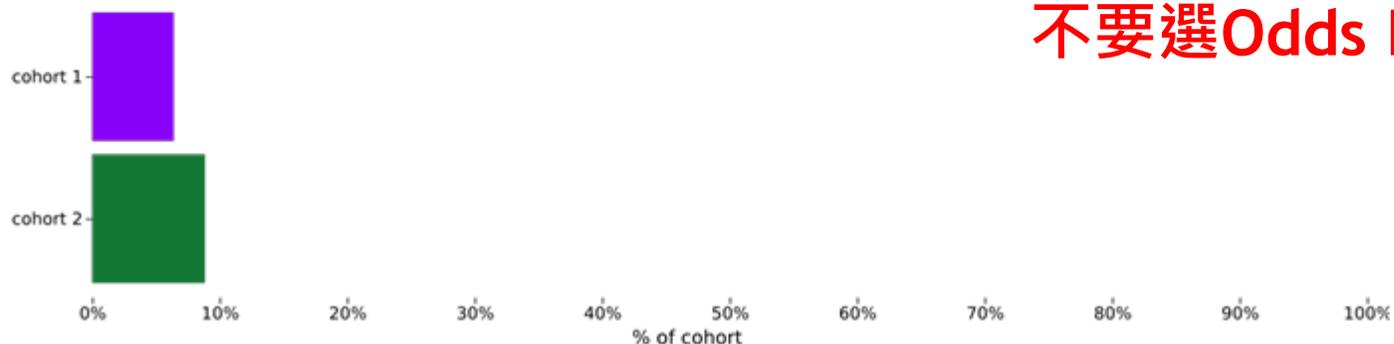
TriNetX其他注意事項



8 Pneumonia

Risk analysis excluding patients with outcome prior to the time window

Cohort	Patients in cohort	Patients with outcome		Risk	
1 A SGLT2i	14,104	901		0.064	
2 B Non-SGLT2i	13,710	1,211		0.088	
			95% CI	z	p
Risk Difference	-0.024	(-0.031, -0.018)	-7.695	0.000	
Risk Ratio	0.723	(0.666, 0.786)	N/A	N/A	
Odds Ratio	0.704	(0.644, 0.770)	N/A	N/A	



不要選Odds Ratio做存活分析的評估

2,907 patients in Cohort 1 and 3,301 patients in Cohort 2 were excluded from results because they had the outcome prior to the time window.

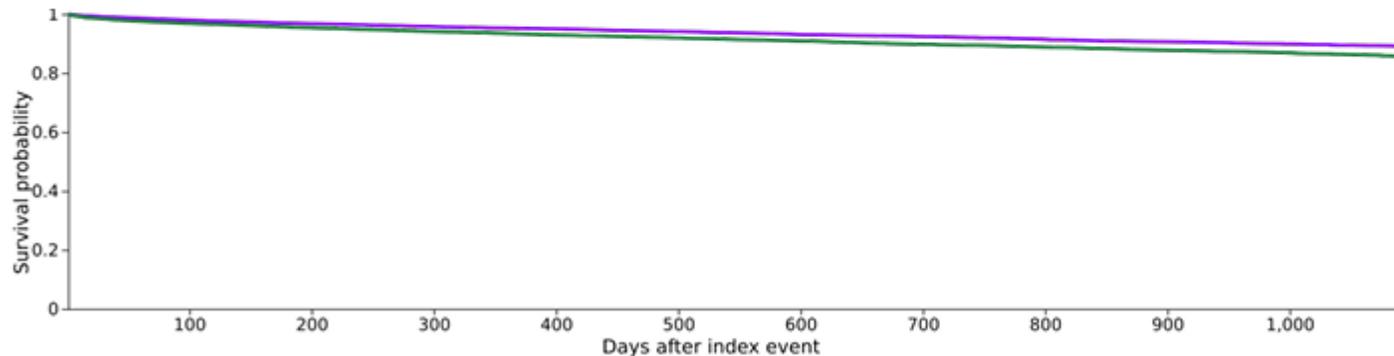
Kaplan - Meier survival analysis excluding patients with outcome prior to the time window

Cohort	Patients in cohort	Patients with outcome	Median survival (days)	Survival probability at end of time window
1 A SGLT2i	14,104	901	--	89.07%
2 B Non-SGLT2i	13,710	1,211	--	85.73%
Log-Rank Test	χ^2	df	p	
	49.485	1	0.000	

Hazard Ratio and Proportionality

Hazard Ratio	95% CI	χ^2	df	p
0.735	(0.674, 0.801)	2.014	1	0.156

選Harzrd Ratio，但是注意95% CI與P-value的對應問題



2,907 patients in Cohort 1 and 3,301 patients in Cohort 2 were excluded from results because they had the outcome prior to the time window.

Summary

- 本研究以 TriNetX 執行 AF + 第二型糖尿病族群之 SGLT2 抑制劑分析，結果顯示 SGLT2i 與較低的中風、失智、心衰竭與死亡風險相關。
- TriNetX 提供標準化、可重複的真實世界資料分析流程，能快速建立大規模多中心 cohort，特別適合多重共病族群的臨床研究。
- TriNetX 為動態資料庫，各醫療端持續更新病例，因此每次登入平台後，cohort 大小與事件數可能會有所變動，統計數值須重新確認。
- TriNetX 線上分析以匯總資料為主，不提供個案層級資料；若需進階統計模型（如 time-dependent Cox、competing risk）或學術發表，皆需取得 IRB 核可後，向 TriNetX 申請患者層級資料下載。
- AI 工具能協助強化圖表呈現、視覺化結果與審查回覆，但不取代 TriNetX 本身的分析邏輯與流程。
- TriNetX 帳號具有嚴格授權規範，不可與非本院人員、未授權人員、或特定地區/國家合作對象共享、借用或共同操作帳號，以避免違反資料隱私與跨境資料規範。

基礎醫學科生統小組：統計方法教育訓練



臺中榮民總醫院 心臟血管研究中心



Thank you for listening