Study Report

P4-C2-017

DARWIN EU® - Time to onset of thromboembolic events in adults with selected types of cancer

07/05/2026

Version 4.0

Authors: Melissa Leung, Cesar Barboza, Ioanna Nika, Anton Barchuk, Talita Duarte-Salles

Public

CONTENTS

LIST OF ABBREVIATIONS **5**
1. TITLE **7**
2. DESCRIPTION OF THE STUDY TEAM **7**
3. ABSTRACT **8**
4. AMENDMENTS AND UPDATES **12**
5. MILESTONES **12**
6. RATIONALE AND BACKGROUND **12**
7. RESEARCH QUESTION AND OBJECTIVES **12**
8. RESEARCH METHODS **13**
8.1. Study design 13
8.2. Follow-up 13
8.3. Study population with inclusion and exclusion criteria 13
Figure 1. Graphical depiction of the study design. 14
8.4. Study setting and data sources 14
Table 1. Data sources 14
8.5. Study period 15
8.6. Variables 15
8.6.1. Covariates, including confounders, effect modifiers, intercurrent events, and other variables 16
8.7. Study size 16
8.8. Data transformation 16
8.9. Statistical methods 17
8.9.1. Missing values 17
9. RESULTS **17**
9.1. Participants 17
Table 2. Attrition of study participants 19
9.2. Main results 23
9.2.1. Cumulative incidence of thromboembolic events 5 years after the first cancer diagnosis 23
Figure 2a. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of bone cancer accounting for a competing risk of death. 25
Figure 2b. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of brain cancer accounting for a competing risk of death. 26
Figure 2c. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of breast cancer accounting for a competing risk of death. 27
Figure 2d. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of colorectal cancer accounting for a competing risk of death 28
Figure 2e. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of corpus uteri cancer accounting for a competing risk of death. 29
Figure 2f. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of oesophageal cancer accounting for a competing risk of death. 30
Figure 2g. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of kidney cancer accounting for a competing risk of death. 31
Figure 2h. Cumulative incidence of splanchnic vein thrombosis after the first diagnosis of liver cancer accounting for a competing risk of death. 32

Figure 2i. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of lung cancer accounting for a competing risk of death..... 33
Figure 2j. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of lymphoma and leukaemia accounting for a competing risk of death. 34
Figure 2k. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of melanoma accounting for a competing risk of death. 35
Figure 2l. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of ovarian cancer accounting for a competing risk of death. 36
Figure 2m. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of pancreatic cancer accounting for a competing risk of death..... 37
Figure 2n. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of prostate cancer accounting for a competing risk of death. 38
Figure 2o. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (TVE) after the first diagnosis of stomach cancer accounting for a competing risk of death..... 39
Table 3. Cumulative incidence of thromboembolic events (splanchnic vein thrombosis (SVT) in liver cancer and pulmonary embolism and deep vein thrombosis combined (VTE) in all other cancer types) at 6-month intervals up to 5 years after the first cancer diagnosis, % [95% confidence interval]..... 40
9.2.2. Cumulative incidence of thromboembolic events five years after the cancer diagnosis by sex, age, and study sub-period 54
9.2.3. Median time to thromboembolic event after the first cancer diagnosis..... 55
Table 4. Median time in days [IQR] to thromboembolic event after first cancer diagnosis. 56
9.2.4. Median time to thromboembolic event after the first cancer diagnosis by sex, age, and study sub-period 69

10. DISCUSSION 69
10.1. Key results 69
10.2. Strengths and limitations of the research methods..... 71
10.3. Interpretation 71
10.4. Generalisability..... 72

11. CONCLUSION..... 73

12. REFERENCES 74

13. ANNEXES..... 75
ANNEX I. Description of data sources..... 75
ANNEX II. Fitness for use assessment..... 87
ANNEX III. Operational and reporting considerations..... 90
ANNEX IV. List of stand-alone documents..... 92
Table S1. List of concepts used to define deep vein thrombosis (DVT). 92
Table S2. List of concepts used to define pulmonary embolism (PE)..... 96
Table S3. List of concepts used to define venous thromboembolism (VTE). 97
Table S4. List of concepts used to define pelvic vein thrombosis (PVT) (concept sets included all descendants of listed concepts). 102
Table S5. List of concepts used to splanchnic vein thrombosis (SVT). 103
Table S6. List of concepts used to define retinal vein thrombosis (RVT). 103
Table S7. List of concepts used to define disseminated intravascular coagulation (DIC). 104
ANNEX V. Glossary..... 105

Study title[1]	DARWIN EU®- Time to onset of thromboembolic events in adults with selected types of cancer
Study report version	V3.0
Date	07/05/2026
EUPAS number	EUPAS1000000814
Active substance	Not applicable
Medicinal product	Not applicable
Objectives	The aim of this study was to estimate time to onset of venous thromboembolic events in adults with each type of selected cancer. The specific objectives of the study were: 1. To estimate the cumulative incidence of thromboembolism within 5 years after the first cancer diagnosis in adults with each type of selected cancer, overall and stratified by age group, sex, and study subperiod. 2. To estimate the median time from the first cancer diagnosis to onset of thromboembolic events in individuals with thromboembolic events with each type of selected cancer, overall and stratified by age group, sex, and study subperiod.
Countries of study	Belgium, Denmark, Estonia, Finland, Germany, Spain, The Netherlands, The United Kingdom
Authors	Melissa Leung (m.leung@darwin-eu.org) Cesar Barboza (c.barboza@darwin-eu.org) Ioanna Nika (i.nika@darwin-eu.org) Anton Barchuk (a.barchuk@darwin-eu.org) Talita Duarte-Salles (t.duarte@darwin-eu.org)

[1] This is a routinely repeated study from P3-C3-005 with EUPAS1000000440.

LIST OF ABBREVIATIONS

Acronyms/terms	Description
AJCC/UICC	American Joint Committee on Cancer and the International Union Against Cancer
ATC	Anatomical Therapeutic Chemical
CDM	Common Data Model
CI	Confidence interval
CPRD	Clinical Practice Research Datalink
DA	Disease Analyser
DARWIN EU®	Data Analysis and Real World Interrogation Network
DK-DHR	Danish Data Health Registries
DTZ	Data Transfer Zone
DOI	Declaration Of Interests
DQD	Data Quality Dashboard
DRE	Digital Research Environment
DVT	Deep Venous Thrombosis
DIC	Disseminated Intravascular Coagulation
EHR	Electronic Health Record
EMA	European Medicines Agency
EBB	Estonian Biobank
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EU	European Union
EUPAS	EU Post-Authorisation Studies Register
FinOMOP-THL	Finnish National Health Registers
GDPR	General Data Protection Regulation
GP	General Practitioner
k	Thousands
ICD-O-3	International Classification of Diseases for Oncology, 3rd Edition
ICD-10	International Classification of Diseases, 10th revision
ICPC-1	International Classification of Primary Care
IP	Inpatient
IPCI	Integrated Primary Care Information Project
IRB	Institutional Review Board
LPD	Longitudinal Patient Database
M	Millions
OHDSI	Observational Health Data Sciences and Informatics
OMOP	Observational Medical Outcomes Partnership
OP	Outpatient

Acronyms/terms	Description
PE	Pulmonary Embolism
PVT	Pelvic Venous Thrombosis
RVT	Retinal vein thrombosis
SIDIAP	The Information System for Research in Primary Care
SNOMED	Systematized Nomenclature of Medicine
SVT	Splanchnic Vein Thrombosis
UKBB	UK Biobank
VTE	Venous Thromboembolism

1. TITLE

DARWIN EU® - Time to onset of thromboembolic events in adults with selected types of cancer

2. DESCRIPTION OF THE STUDY TEAM

Study team role	Names	Organisation
Principal Investigators	Melissa Leung Anton Barchuk	Erasmus MC
Co-principal Investigator	Talita Duarte-Salles	Erasmus MC
Data Scientists	Cesar Barboza Ioanna Nika Maarten van Kessel Ger Inberg Adam Black Ross Williams	Erasmus MC
Clinical Domain Expert	Anton Barchuk	Erasmus MC
Study Manager	Natasha Yefimenko	Erasmus MC
Data source	**Names**	**Data Partner Organisation***
IQVIA Longitudinal Patient Database Belgium (IQVIA LPD Belgium) IQVIA Disease Analyzer Germany (IQVIA DA Germany)	Gargi Jadhav Isabella Kacmarczy Akram Mendez Hanne van Ballegooijen Dina Vojinovic	IQVIA
Danish Data Health Registries (DK-DHR)	Elvira Bräuner Susanne Bruun	Danish Medicines Agency
Estonian Biobank (EBB)	Marek Oja Raivo Kolde Ami Sild	University of Tartu, Estonia
Finnish National Health Registers (FinOMOP-THL)	Toni Lehtonen Laura Salonen	Data and Analytics, Finnish Institute for Health and Welfare (THL) , Finland
The Information System for Research in Primary Care (SIDIAP)	Anna Palomar-Cros Irene López-Sánchez Agustina Giuliodori	IDIAPJGol
Integrated Primary Care Information (IPCI)	Katia Verhamme	Integrated Primary Care Information, The Netherlands
Clinical Practice Research Datalink GOLD (CPRD GOLD) UK BioBank (UKBB)	Antonella Delmestri	University of Oxford

*Data partners do not have an investigator role. Data partners execute code at their data source, review, and approve their results.

3. ABSTRACT

Title

DARWIN EU® - Time to onset of thromboembolic events in adults with selected types of cancer

Thromboembolic events are a common complication for individuals with cancer, with risks varying according to the cancer type, suggesting that mechanisms which play a role in the occurrence of these events may be specific to the cancer type itself or its treatment mechanisms. Haematological malignancies and lung, pancreas, stomach, bowel, and brain cancers are generally associated with a high risk of thrombosis, whilst prostate and breast cancers are associated with a low risk of thrombosis.

When a safety signal of a thromboembolic event appears in cancer populations, it can be challenging to assess a potential association with the oncologic treatment without reliable information on the background risk. This study is intended to address this knowledge gap by generating evidence on the time to onset of different venous thromboembolic events among adults with selected cancer types.

Objectives

The aim of this study was to estimate time to onset of venous thromboembolic events in adults with each type of selected cancer.

The specific objectives of the study were:

1. To estimate the cumulative incidence of thromboembolism within 5 years after the first cancer diagnosis in adults with each type of selected cancer, overall and stratified by age group, sex, and study subperiod.
2. To estimate the median time from the first cancer diagnosis to onset of thromboembolic events in individuals with thromboembolic events with each type of selected cancer, overall and stratified by age group, sex, and study subperiod.

Methods

Population-based cohort study. The index date was the date of the first cancer diagnosis. Individuals were followed up until the earliest of occurrence of a thromboembolic event, loss to follow-up, end of data availability, end of the study period, or death.

Population

The study population included all individuals aged 18 years and above with a primary diagnosis of one of the selected cancers (bone, brain, breast, colorectal, corpus uteri, kidney, leukaemia and lymphoma, liver, lung, melanoma, oesophageal, ovary, pancreas, prostate, stomach) during the inclusion period (from 01/01/2016 to 31/12/2022). Only individuals with an incident cancer diagnosis (excluding non-melanoma skin cancer), defined as a first cancer diagnosis after ≥365 days cancer-free history, were included. Cancer cases and thromboembolic events were identified based on appropriate computable phenotyping algorithms. Conditions in the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) used the Systematised Nomenclature of Medicine (SNOMED) as the standard vocabulary for diagnosis codes. The International Classification of Diseases for Oncology, 3rd Edition (ICD-O-3) was also considered for cancer diagnoses. Other eligibility criteria included at least 365 days of database history prior to index date and at least 365 days between index date and end of data availability in the data source.

Variables

Outcome:

The specific thromboembolic outcomes of interest were: deep vein thrombosis (DVT), pulmonary embolism (PE), venous thromboembolism (VTE, composite of DVT and PE), pelvic venous thrombosis (PVT), splanchnic vein thrombosis (SVT, including hepatic and extra-hepatic vein thrombosis), retinal vein thrombosis (RVT, including retinal central vein thrombosis), and disseminated intravascular coagulation (DIC).

Relevant covariates:

The following covariates were assessed at index date: age (categorised into age groups: 18–34, 35–44, 45–54, 55–64, 65–74, 75–84, and ≥85 years), sex, and calendar year of the first cancer diagnosis (categorised into study subperiods: 2016–2019 and 2020–2022). These variables were used to stratify the results.

Data sources

1. Belgium: IQVIA Longitudinal Patient Database Belgium (IQVIA LPD Belgium)
2. Denmark: Danish Data Health Registries (DK-DHR)
3. Estonia: Estonian Biobank (EBB)
4. Finland: Finnish Care Register for Health Care (FinOMOP-THL)
5. Germany: IQVIA Disease Analyzer Germany (IQVIA DA Germany)
6. Spain: The Information System for Research in Primary Care (SIDIAP)
7. The Netherlands: Integrated Primary Care Information (IPCI)
8. The United Kingdom: Clinical Practice Research Datalink GOLD (CPRD GOLD)
9. The United Kingdom: UK BioBank (UKBB)

Study size

No sample size was calculated, as this was a descriptive study which did not test a specific hypothesis. Based on the results of the study EUPAS1000000440, the expected number of person counts was the lowest for DIC (during 1-year follow-up: 5 in FinOMOP-THL – 171 in SIDIAP, with 0 counts in CPRD GOLD, EBB, IPCI, IQVIA DA Germany, and IQVIA LPD Belgium) and highest for VTE (during 1-year follow-up: 27 in IQVIA LPD Belgium – 4,597 in FinOMOP-THL).

Statistical analysis

Analyses were conducted separately for each data source and carried out in a federated manner, allowing analyses to be run locally without sharing individual-level data. The analyses were conducted overall, as well as by strata of age, sex, and calendar year of the first cancer diagnosis.

Absence of diagnosis codes was interpreted as the absence of the conditions themselves. A minimum cell count of 5 was used when reporting results, with any smaller count reported as “<5” and zero counts as “0”.

Objective 1

The cumulative incidence of thromboembolism in adults with each type of selected cancer was calculated in a competing risk survival analysis and reported at 6-month intervals within 5 years after the first cancer diagnosis, accounting for death as a competing risk.

Objective 2

Among the individuals with cancer who experienced a thromboembolic event, the median time from the first cancer diagnosis to the onset of the thromboembolic event was calculated.

Results

Overall, we included 975,962 individuals with a primary cancer across nine data sources (sum across all cancer types), among whom 64,650 thromboembolic events were used in total in the calculation of the cumulative incidence. The 6-month interval with the steepest increase in the cumulative incidence of thromboembolic events was within 6 months after the cancer diagnosis for VTE, DVT, PE, SVT, PVT, and DIC, but it varied between cancer types for RVT.

The highest percentages of the 5-year cumulative incidence of VTE, DVT, and PE were observed in ovarian cancer (VTE: ranging from 4.79% [95% CI: 3.95%–5.8%] in IQVIA DA Germany to 10.53% [95% CI: 9%–12.32%] in CPRD GOLD) and pancreatic cancer (VTE: ranging from 4.46% [95% CI: 3.41%–5.82%] in IPCI to 11.63% [95% CI: 9.57%–14.12%] in UKBB), while the lowest was observed in melanoma (VTE: ranging from 1.27% [95% CI: 1.09%–1.48%] in IQVIA DA Germany to 2.42% [95% CI: 2.03%–2.89%] in SIDIAP).

The highest percentages of the 5-year cumulative incidence of SVT were observed in liver cancer (ranging from 1.16% [95% CI: 0.87%–1.54%] in FinOMOP-THL to 9.44% [95% CI: 8.58%–10.38%] in SIDIAP) and pancreatic cancer (ranging from 0.46% [95% CI: 0.25%–0.85%] in CPRD GOLD to 4.41% [95% CI: 2.22%–8.73%] in EBB), while the lowest was observed in breast cancer (ranging from 0.02% [95% CI: 0.01%–0.04%] in IQVIA DA Germany to 0.11% [95% CI: 0.07%–0.16%] in SIDIAP).

The 5-year cumulative incidence of PVT, RVT, and DIC remained below 1% across cancer types and data sources. The cumulative incidence of thromboembolic events was generally low in melanoma, breast cancer, and prostate cancer.

The shortest median time from cancer diagnosis to VTE, DVT, PE, and SVT was observed in pancreatic cancer and liver cancer, while the longest was observed in melanoma. Out of these four outcomes, the median [IQR] time was shortest for SVT in liver cancer (ranging from 6 days [0–238] in SIDIAP to 66 days [4–301] in DK-DHR), followed by PE in pancreatic cancer (ranging from 23 days [1–131] in EBB to 160 days [46–331] in IPCI), DVT in pancreatic cancer (ranging from 40 days [14–104] in UKBB to 218 days [123–482] in EBB), and VTE in pancreatic cancer (ranging from 42 days [0–116] in UKBB to 134 days [24–260] in IPCI).

The point estimates of the 5-year cumulative incidence were lower among females than males for SVT in liver cancer, which was reflected in a possible pattern towards longer event times among females than males from the diagnosis of liver cancer until the occurrence of SVT.

The point estimates of the 5-year cumulative incidence of VTE in pancreatic cancer were lower in the higher age groups.

The cumulative incidence results and time to onset of thromboembolic event were generally similar between the two study sub-periods, although the observed event times were generally longer among those diagnosed in 2016–2019 than among those diagnosed in 2020–2022. This was likely due to the shorter follow-up in the latter study sub-period.

Discussion

In this large multinational cohort study, which incorporated data from nine data sources across eight countries in Europe, we found the steepest increase in the cumulative incidence of thromboembolic events within the first six months after the first cancer diagnosis. This was observed for all thromboembolic events, except for RVT, for which the six-month interval with the steepest increase in the cumulative incidence varied by cancer type. The magnitude and timing of the increase in cumulative incidence varied by cancer type. Ovarian and pancreatic cancers had the highest percentages of 5-year cumulative incidence

of VTE, DVT, and PE, while melanoma had the lowest. Similarly, SVT had the highest 5-year cumulative incidence in liver and pancreatic cancers, while the lowest was observed in breast cancer.

These differences were reflected in time-to-event analyses, with pancreatic and liver cancers showing the shortest median times from cancer diagnosis to VTE, DVT, PE, and SVT, while the longest were observed in melanoma. The 5-year cumulative incidence of PVT, RVT, and DIC remained below 1% across cancer types and data sources. Overall, these findings highlight a critical early time period for the incidence of thromboembolic events after cancer diagnosis and demonstrate consistent patterns across multiple European populations, supporting their robustness and relevance for clinical and regulatory considerations.

4. AMENDMENTS AND UPDATES

None.

5. MILESTONES

Study deliverable	Timelines (planned)	Timelines (actual)
Final Study Protocol	10 December 2025	12 December 2025
Creation of Analytical code	December 2025	November 2025
Execution of Analytical Code on the data	December 2025	December 2025 – February 2026
Draft Study Report	January 2026	24 March 2026
Final Study Report	February 2026	April 2026

6. RATIONALE AND BACKGROUND

Thromboembolic events are a common complication for individuals with cancer, with risk varying according to the cancer type, suggesting that mechanisms that play a role in the occurrence of these events may be specific to the cancer type itself or its treatment.[1,2] Haematological malignancies and lung, pancreas, stomach, bowel, and brain cancers are generally associated with a high risk of thrombosis, whilst prostate and breast cancers are associated with low risk of thrombosis.[3]

When a safety signal of a thromboembolic event appears in cancer populations, it can be challenging to assess a potential association with the oncologic treatment without reliable information on the background risk. This study is intended to address this knowledge gap by generating evidence on the time to onset of different venous thromboembolic events among adults with selected cancer types.

This study is a routine repeated study of a previous DARWIN EU® study (EUPAS1000000440) that focused on estimating the incidence rates of venous thromboembolic events (VTE) in adults newly diagnosed with any of the following selected cancers (bone, brain, breast, colorectal, corpus uteri, kidney, leukaemia and lymphoma, liver, lung, melanoma, oesophageal, ovary, pancreas, prostate, stomach) during the period 2016–2022 and describing the individuals' characteristics at the time of cancer diagnosis. This study is now being repeated with the same study population and outcome, but with different objectives: to estimate the cumulative incidence of thromboembolism and the time to onset of thromboembolic events in adults with these selected types of cancer.

7. RESEARCH QUESTION AND OBJECTIVES

Objectives

The aim of this study was to estimate time to onset of venous thromboembolic events in adults with each type of selected cancer.

The specific objectives of the study were:

1. To estimate the cumulative incidence of thromboembolism within 5 years after the first cancer diagnosis in adults with each type of selected cancer, overall and stratified by age group, sex, and study subperiod.
2. To estimate the median time from the first cancer diagnosis to onset of thromboembolic events in individuals with thromboembolic events with each type of selected cancer, overall and stratified by age group, sex, and study subperiod.

8. RESEARCH METHODS

8.1. Study design

A cohort study was conducted.

8.2. Follow-up

For both objectives, follow-up started on the date of the first cancer diagnosis (index date) and ended on the date of occurrence of the outcome (thromboembolic event), loss to follow-up, end of data availability, 5 years of follow-up, or death, whichever came first.

8.3. Study population with inclusion and exclusion criteria

Objective 1

Inclusion criteria

- First diagnosis of a selected cancer (index date) between 01/01/2016 and 31/12/2022 (inclusion period)
- Age ≥18 years at cancer diagnosis
- Minimum of 365 days of available history before the cancer diagnosis date
- Cancer diagnosis date ≥365 days prior to end of data availability of the data source

Exclusion criteria

- History of any cancer diagnosis ever before the selected cancer diagnosis date

Objective 2

Inclusion criteria

- The subset of the cohort for objective 1 who experienced the outcome (i.e. thromboembolic event) during follow-up.

The study design to address objective 1, including assessment windows, is visualised in **Figure 1**. For objective 2, we included the subset of the cohort of adults with cancer who experienced a thromboembolic event during follow-up.

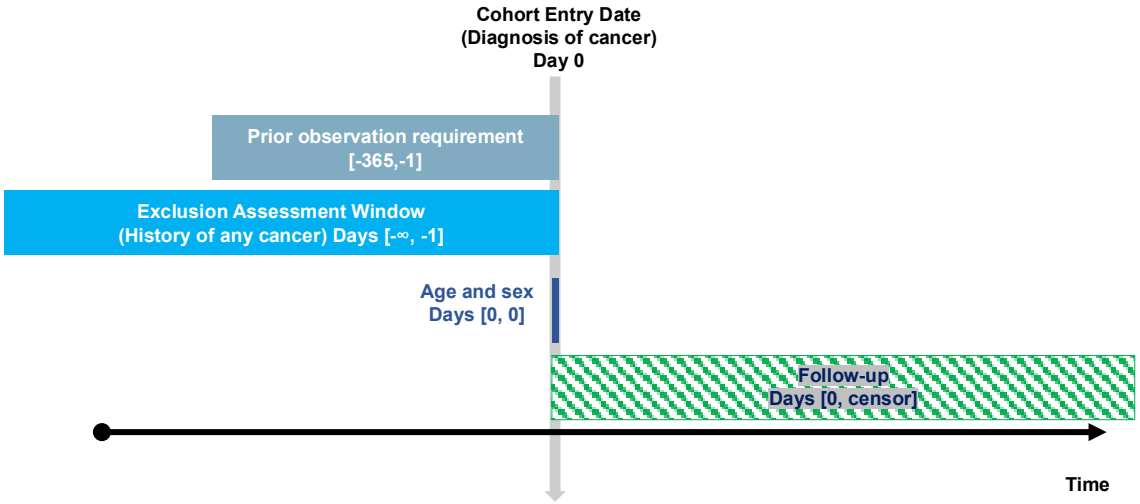

Figure 1. Graphical depiction of the study design.

The censor date was the earliest of occurrence of the outcome, loss to follow-up, end of data availability, or death.

8.4. Study setting and data sources

This study was conducted using data collected in routine care from different health care settings from 9 data sources part of the DARWIN EU® network and distributed in 8 different countries across Europe (**Table 1**). All data were a priori mapped to the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) .

Table 1. Data sources.

Country	Name of Data source	Health Care setting	Type of Data	Number of active individuals	Calendar period covered by each data source	Contributing to
Belgium	IQVIA LPD Belgium	Primary care	EHRs	189k	2015–2025	All objectives
Denmark	DK-DHR	Primary care (using prescription retrieval as a proxy), hospital care (IP and OP)	EHRs, registries	5.98M	1995–2025	All objectives
Estonia	EBB	Primary care, hospital care (IP and OP)	EHRs, claims, registries, biobank	212k	2004–2025	All objectives
Finland	FinOMOP-THL	Primary Care, Hospital care (IP and OP)	EHRs, registries	5.7M	2011–2025	All objectives

Country	Name of Data source	Health Care setting	Type of Data	Number of active individuals	Calendar period covered by each data source	Contributing to
Germany	IQVIA DA Germany	Primary care	EHRs	4.48M	1992–2025	All objectives
Spain	SIDIAP	Primary care (linked to hospital discharge data)	EHRs	5.95M	2006–2025	All objectives
The Netherlands	IPCI	Primary care	EHRs	1.33M	2006–2025	All objectives
The United Kingdom of Great Britain and Northern Ireland	CPRD GOLD	Primary care, hospital care (OP)	EHRs	2.83M	1987–2025	All objectives
The United Kingdom of Great Britain and Northern Ireland	UKBB	Primary care (up to 2017), hospital care (IP and OP, up to November 2022)	EHRs, registries, biobank	500k	1940–2025	All objectives

Data sources: IQVIA LPD=IQVIA Longitudinal Patient Database Belgium (IQVIA LPD Belgium); DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; IPCI=Integrated Primary Care Information; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank
Types of data: EHR=electronic health record, IP=inpatient, OP=outpatient
Number of active subjects: k=thousands, M=millions

Data sources selection

These data sources fulfil the criteria required in terms of data quality, completeness, timeliness, and representativeness for the cohort study while covering different regions of Europe (**Annex II**).

8.5. Study period

The study period was from 01/01/2016 up to the most recent data available for each contributing data source.

8.6. Variables

All objectives

The specific thromboembolic events of interest were:

- Deep vein thrombosis (DVT)
- Pulmonary embolism (PE)
- Venous thromboembolism (VTE, composite of DVT and PE)
- Pelvic venous thrombosis (PVT)

- Splanchnic vein thrombosis (SVT), including hepatic and extra-hepatic vein thrombosis
- Retinal vein thrombosis (RVT), including retinal central vein thrombosis
- Disseminated intravascular coagulation (DIC)

Each of the specific thromboembolic events was a binary outcome. Each outcome was assessed at any diagnosis position in the electronic health record and from any of the care settings in each data source.

The list of concepts was based on Systematised Nomenclature of Medicine codes and aligned with previous studies that used OMOP CDM and VTE as an outcome (Burn et al., 2022) and is provided in **Annex IV**.

8.6.1. Covariates, including confounders, effect modifiers, intercurrent events, and other variables

All Objectives

- Sex
    - Female/male
- Age (years) on the date of the first cancer diagnosis, categorised into age groups:
    - 18–34
    - 35–44
    - 45–54
    - 55–64
    - 65–74
    - 75–84
    - ≥85

  Age at index date was calculated using January 1st of the year of birth as proxy for the actual birthday. Date/month is either not present or cannot be made available for governance reasons. If available, date is often set to first of the month for personal privacy.
- Calendar year of the first cancer diagnosis, categorised into subperiods of the study period:
    - 2016–2019
    - 2020–2022

The status of each covariate was assessed on the date of the first cancer diagnosis. Each covariate was used to stratify the results.

8.7. Study size

No sample size was calculated, as this was a descriptive study which did not test a specific hypothesis. In addition, we used data from the study population that was included in EUPAS1000000440. Based on the results of the study EUPAS1000000440, the expected number of person counts was the lowest for DIC (during 1-year follow-up: 5 in FinOMOP-THL – 171 in SIDIAP, with 0 counts in CPRD GOLD, EBB, IPCI, IQVIA DA Germany, and IQVIA LPD Belgium) and highest for VTE (during 1-year follow-up: 27 in IQVIA LPD Belgium – 4,597 in FinOMOP-THL).

8.8. Data transformation

Analyses were conducted separately for each data source. Before study initiation, test runs of the analyses were performed on a subset of the data sources and quality control checks were performed. Once all the

tests passed (see **Annex III**), the final study codes package was released in the version-controlled Study Repository for execution against all the participating data sources.

The data partners locally executed the analytics against the OMOP CDM in R Studio and reviewed and approved the, by default, aggregated results.

The study results of all data sources were checked, after which they were made available to the team, and the dissemination phase started. All results were locked and timestamped for reproducibility and transparency

8.9. Statistical methods

Objective 1

The cumulative incidence of thromboembolic events at 6-month intervals within 5 years (6, 12, 18, 24, 30, 36, 42, 48, 54, and 60 months) in adults with each type of selected cancer was calculated based on OMOP CDM mapped data using the R package *CohortSurvival*,[4] developed by DARWIN EU®. The R package *CohortSurvival* was designed to work with data in the OMOP CDM format to extract and summarise survival data. A competing risk analysis was conducted to account for the risk of death. The 95% confidence intervals (CIs) for probabilities based on the Aalen-Johansen estimator were calculated using standard errors derived from the infinitesimal jackknife method.[5] Individuals with a thromboembolic event or death recorded on the same date as the first cancer diagnosis were excluded from the analysis. Individuals were censored on the date of loss to follow-up, the end of data availability, or the end of the study period. The *estimateCompetingRiskSurvival*() function was used in the analysis. The cumulative incidence was reported with the 95% CI. In addition, the overall cumulative incidence curve was plotted.

Objective 2

The median time to onset of venous thromboembolic events was calculated in days and based on OMOP CDM mapped data using the R package *CohortCharacteristics*,[6] developed by DARWIN EU®. The median time to onset was reported with the interquartile range (IQR). Thromboembolic events that were recorded on the same date as the first cancer diagnosis were included in the calculation of the median event times. However, the reported event counts corresponded to those with event times of at least one day, i.e. the event counts in objective 1, to maintain confidentiality by the suppression of results based on <5 events recorded at least one day after the first cancer diagnosis.

All objectives

The analyses were conducted overall, as well as by strata of age group, sex, and study subperiod. A minimum cell count of 5 was used when reporting results, with any smaller count reported as “<5” and zero counts as “0”.

8.9.1. Missing values

The absence of diagnosis codes was interpreted as the absence of the conditions themselves.

9. RESULTS

The main results based on data from the overall study population are presented below. Results stratified by sex, age, and study sub-period are described in text only. The full set of tables and figures for this study is available through an interactive web-application ShinyApp at EUPAS1000000814.

9.1. Participants

Attrition of study participants per data source and cancer type is presented in **Table 2.** After excluding individuals with a prior history of cancer, the total number of participants per cancer type across all data sources was: bone cancer (n=3,328), brain cancer (n=16,019), breast cancer (n=191,151), colorectal cancer

(n=128,431), corpus uteri cancer (n=25,296), kidney cancer (n=32,881), liver cancer (n=18,767), lung cancer (n=105,527), lymphoma and leukaemia (n=126,499), melanoma (n=73,649), oesophageal cancer (n=15,547), ovarian cancer (n=16,794), pancreatic cancer (n=32,647), prostate cancer (n=169,739), and stomach cancer (n=19,687). Overall, we included 975,962 individuals with cancer across nine data sources.

The total number of thromboembolic events across all data sources was: VTE (n=28,802), PE (n=19,822), DVT (n=11,644), SVT (n= 1,828), RVT (n= 1,720), PVT (n= 650), and DIC (n=184). Overall, 64,650 thromboembolic events occurred during the study follow-up.

Table 2. Attrition of study participants.

		IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
Bone cancer	Qualifying initial records	21	586	50	960	1,291	1,796	0	470	122
	No prior history of cancer; included in the study	18	322	31	594	965	1,071	0	290	37
	Total number of thromboembolic events during follow-up	0	16	0	79	27	85	NC	15	0
Brain cancer	Qualifying initial records	140	4,650	175	4,073	3,920	4,985	491	1,155	664
	No prior history of cancer; included in the study	128	3,510	132	3,318	3,240	3,955	297	932	507
	Total number of thromboembolic events during follow-up	0	345	14	521	118	531	36	121	73
Breast cancer	Qualifying initial records	1,905	35,560	1,384	40,996	60,761	34,324	7,131	18,963	5,808
	No prior history of cancer; included in the study	1,853	31,873	1,243	38,047	56,337	31,931	6,553	18,134	5,180
	Total number of thromboembolic events during follow-up	14	1,495	54	2,207	1,413	1,571	272	1,180	214
Colorectal cancer	Qualifying initial records	409	35,264	950	27,788	29,207	35,059	6,527	13,393	4,871
	No prior history of cancer; included in the study	382	29,279	762	23,146	24,061	29,755	5,400	11,798	3,848
	Total number of thromboembolic events during follow-up	0	2,838	87	2,754	1,058	2,621	333	1,440	486

		IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
Corpus uteri cancer	Qualifying initial records	68	6,020	324	7,323	6,392	5,969	777	2,713	1,143
	No prior history of cancer; included in the study	60	4,910	260	6,121	5,113	4,858	622	2,453	899
	Total number of thromboembolic events during follow-up	0	420	15	627	189	506	38	225	102
Oesophageal cancer	Qualifying initial records	57	4,879	49	3,313	4,056	2,650	1,123	3,715	1,080
	No prior history of cancer; included in the study	53	3,168	35	2,487	3,121	1,860	837	3,223	763
	Total number of thromboembolic events during follow-up	<5	264	5	264	149	192	61	456	120
Kidney cancer	Qualifying initial records	40	8,594	381	8,293	11,714	9,849	1,412	1,283	1,380
	No prior history of cancer; included in the study	33	6,498	289	6,394	9,149	7,404	1,068	1,096	950
	Total number of thromboembolic events during follow-up	0	588	22	604	303	754	58	117	86
Liver cancer	Qualifying initial records	37	4,775	121	5,563	4,282	7,622	832	1,538	753
	No prior history of cancer; included in the study	35	3,338	75	4,177	3,214	5,618	557	1,297	456
	Total number of thromboembolic events during follow-up	0	297	11	328	145	747	24	78	78
Lung cancer	Qualifying initial records	540	37,091	505	21,507	25,951	27,635	7,232	14,037	3,765

		IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
	No prior history of cancer; included in the study	486	27,757	332	16,828	20,343	20,068	5,288	11,844	2,581
	Total number of thromboembolic events during follow-up	16	2,955	40	1,930	877	2,833	335	1,163	396
Lymphoma and Leukaemia	Qualifying initial records	865	30,833	1,083	24,930	39,020	33,585	3,739	11,523	3,959
	No prior history of cancer; included in the study	815	25,172	884	20,690	33,640	28,831	3,009	10,429	3,029
	Total number of thromboembolic events during follow-up	<5	1,684	61	1,374	949	1,714	182	658	213
Melanoma	Qualifying initial records	138	20,197	632	10,571	32,386	9,181	3,658	5,855	2,401
	No prior history of cancer; included in the study	130	17,145	526	8,820	28,972	7,928	3,096	5,235	1,797
	Total number of thromboembolic events during follow-up	0	580	11	350	397	319	70	193	47
Ovarian cancer	Qualifying initial records	61	4,171	279	4,134	5,784	4,161	628	1,997	880
	No prior history of cancer; included in the study	54	2,842	217	3,085	4,603	3,244	473	1,704	572
	Total number of thromboembolic events during follow-up	0	456	26	430	241	404	44	317	116
Pancreatic cancer	Qualifying initial records	136	9,014	266	10,210	8,220	8,830	1,697	2,718	1,343

		IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
	No prior history of cancer; included in the study	123	6,587	193	8,139	6,387	6,705	1,252	2,354	907
	Total number of thromboembolic events during follow-up	0	965	40	1,150	334	1,066	107	318	220
Prostate cancer	Qualifying initial records	952	33,709	1,256	40,495	57,911	26,690	5,125	16,562	8,241
	No prior history of cancer; included in the study	920	29,476	1,067	36,620	51,871	22,704	4,527	15,485	7,069
	Total number of thromboembolic events during follow-up	16	1,678	84	1,910	1,237	1,085	225	773	274
Stomach cancer	Qualifying initial records	46	6,434	293	4,137	6,240	6,882	922	1,279	851
	No prior history of cancer; included in the study	42	4,137	191	3,300	4,737	5,059	678	1,055	488
	Total number of thromboembolic events during follow-up	0	438	32	384	167	621	38	125	91

No prior history of cancer: counts shown after excluding individuals with prior history of cancer; Total number of thromboembolic events during follow-up: sum of the thromboembolic event outcomes that occurred at least one day after the first cancer diagnosis (only >5 counts of each outcome were included and individuals with >1 outcome contributed to the total as many times as the number of outcomes that they experienced).
NC=not captured. IQVIA LPD=IQVIA Longitudinal Patient Database Belgium (IQVIA LPD Belgium); DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; IPCI=Integrated Primary Care Information; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank.

9.2. Main results

9.2.1. Cumulative incidence of thromboembolic events 5 years after the first cancer diagnosis

The cumulative incidence of SVT in individuals with liver cancer and the cumulative incidence of VTE in individuals with each of the other cancer types up to five years after the first cancer diagnosis are shown in **Figures 2a–o** and **Table 3**. These outcomes were selected as in the previous study (Incidence rates of venous thromboembolic events in patients with selected cancers) SVT was the outcome with the highest combined incidence rate across data sources one year after the first diagnosis of liver, and VTE was the outcome with the highest combined incidence rate across data sources across the other cancer types.

The steepest increase in the cumulative incidence of VTE, DVT and PE was observed within six months after the cancer diagnosis across all cancer types (although the point estimate remained <1% across data sources six months after the diagnosis of melanoma and prostate cancer), after which the cumulative incidence was relatively stable in bone cancer, oesophageal cancer, liver cancer, brain cancer, pancreatic cancer, and stomach cancer.

The highest 5-year cumulative incidence of VTE was observed in ovarian cancer (ranging from 4.79% [95% CI: 3.95%–5.80%] in IQVIA DA Germany to 10.53% [95% CI: 9.00%–12.32%] in CPRD GOLD) and pancreatic cancer (ranging from 4.46% [95% CI: 3.41%–5.82%] in IPCI to 11.63% [95% CI: 9.57%–14.12%] in UKBB), while the lowest was observed in melanoma (ranging from 1.27% [95% CI: 1.09%–1.48%] in IQVIA DA Germany to 2.42% [95% CI: 2.03%–2.89%] in SIDIAP).

Similar to VTE, the highest 5-year cumulative incidence of DVT was observed in pancreatic cancer (ranging from 1.85% [95% CI: 1.54%–2.22%] in DK-DHR to 4.61% [95% CI: 2.33%–9.13%] in EBB) and ovarian cancer (ranging from 1.79% [95% CI: 0.90%–3.55%] in IPCI to 3.98% [95% CI: 3.08%–5.15%] in CPRD GOLD), while the lowest was observed in melanoma (ranging from 0.48% [95% CI: 0.24%–0.96%] in UKBB to 1.46% [95% CI: 1.17%–1.84%] in SIDIAP).

The highest 5-year cumulative incidence of PE was also observed in ovarian cancer (ranging from 2.98% [95% CI: 2.34%–3.79%] in IQVIA DA Germany to 7.98% [95% CI: 5.95%–10.70%] in UKBB) and pancreatic cancer (ranging from 2.54% [95% CI: 1.78%–3.62%] in IPCI to 9.05% [95% CI: 7.24%–11.32%] in UKBB), while the lowest was observed in melanoma (ranging from 0.72% [95% CI: 0.45%–1.16%] in IPCI to 1.47% [95% CI: 1.22%–1.77%] in FinOMOP-THL).

The highest 5-year cumulative incidence of SVT was observed in liver cancer (ranging from 1.16% [95% CI: 0.87%–1.54%] in FinOMOP-THL to 9.44% [95% CI: 8.58%–10.38%] in SIDIAP) and pancreatic cancer (ranging from 0.46% [95% CI: 0.25%–0.85%] in CPRD GOLD to 4.41% [95% CI: 2.22%–8.73%] in EBB), while the lowest was observed in breast cancer (ranging from 0.02% [95% CI: 0.01%–0.04%] in IQVIA DA Germany to 0.11% [95% CI: 0.07%–0.16%] in SIDIAP). The steepest increase in the cumulative incidence of SVT was within six months after the first cancer diagnosis in colorectal, kidney, liver, ovarian, pancreatic, and stomach cancers; and six to twelve months in oesophageal cancer.

The cumulative incidence of SVT remained below 0.3% five years after cancer diagnosis across data sources in breast, corpus uteri, lung, lymphoma and leukaemia, melanoma, and prostate cancers. Event counts of SVT were 0 or <5 after five years of follow-up among individuals with bone and brain cancers across all data sources.

The highest 5-year cumulative incidence of PVT was observed in ovarian cancer (0.46% [95% CI: 0.27%–0.77%] in FinOMOP-THL) and prostate cancer (ranging from 0.02% [95% CI: 0.01%–0.04%] in DK-DHR to 1.21% [95% CI: 0.66%–2.22%] in EBB), while the lowest was observed in lung cancer (0.07% [95% CI: 0.04%–0.13%] in SIDIAP and 0.15% [95% CI: 0.10%–0.22%] in FinOMOP-THL).

The number of PVT events within five years of follow-up among individuals with bone, oesophageal, and liver cancers across all data sources were 0 or <5; similarly, 0 or <5 PVT events were recorded in most data

sources across all the other cancer types. The steepest increase in the cumulative incidence of PVT was generally 0–6 months after the first cancer diagnosis, and the cumulative incidence remained below 1% across most cancer types and data sources within five years after the first cancer diagnosis. The only exception was the 5-year cumulative incidence of PVT in prostate cancer (point estimate 1.21%), but the point estimate of the cumulative incidence of PVT in prostate cancer remained <1% in the three other data sources with at least five events.

The highest 5-year cumulative incidence of RVT was observed in kidney cancer (ranging from 0.17% [95% CI: 0.08%–0.34%] in IQVIA DA Germany to 0.69% [95% CI: 0.31%–1.56%] in CPRD GOLD) and prostate cancer (ranging from 0.12% [95% CI: 0.08%–0.17%] in IQVIA DA Germany to 0.67% [95% CI: 0.59%–0.77%] in FinOMOP-THL), while the lowest was observed in pancreatic cancer (0.08% [95% CI: 0.03%–0.17%] in FinOMOP-THL, 0.09% [95% CI: 0.04%–0.21%] in SIDIAP and 0.09% [95% CI: 0.04%–0.22%] in DK-DHR). Event counts of RVT were 0 or <5 after five years of follow-up among individuals with bone cancer and brain cancer across all data sources. The 6-month interval with the steepest increase in the cumulative incidence of RVT varied between cancer types and data sources, and the cumulative incidence remained below 1% across cancer types and data sources.

The highest 5-year cumulative incidence of DIC was observed in oesophageal cancer (0.31% [95% CI: 0.13%–0.75%] in SIDIAP) and liver cancer (0.28% [95% CI: 0.16%–0.50%] in SIDIAP), while the lowest was observed in breast cancer (0.03% [95% CI: 0.02%–0.07%] in SIDIAP). Event counts of DIC were 0 or <5 after five years of follow-up among individuals with bone, brain, corpus uteri, melanoma, and ovarian cancers across all data sources. Three data sources had five or more records of DIC in lymphoma and leukaemia, and two data sources had five or more records of DIC in lung cancer. Across the eight other types of cancer, five or more records of DIC were present only in SIDIAP. The steepest increase in the cumulative incidence of DIC was generally within six months after the cancer diagnosis, and the cumulative incidence remained below 1% across cancer types and data sources five years after cancer diagnosis.

The cumulative incidence of death, which was considered a competing event, was higher than the cumulative incidence of each of the outcomes. This was the case for all data sources, except for IQVIA DA Germany, in which the death records were incomplete. For example, the 5-year cumulative incidence of death of ovarian cancer ranged from 21.66% [95% CI: 16.12%–29.09%] in EBB to 51.90% [95% CI: [49.93%–53.95%] in DK-DHR, while the 5-year cumulative incidence of VTE ranged from 4.79% [95% CI: 3.95%–5.80%] in IQVIA DA Germany to 10.53% [95% CI: 9.00%–12.32%] in CPRD GOLD.

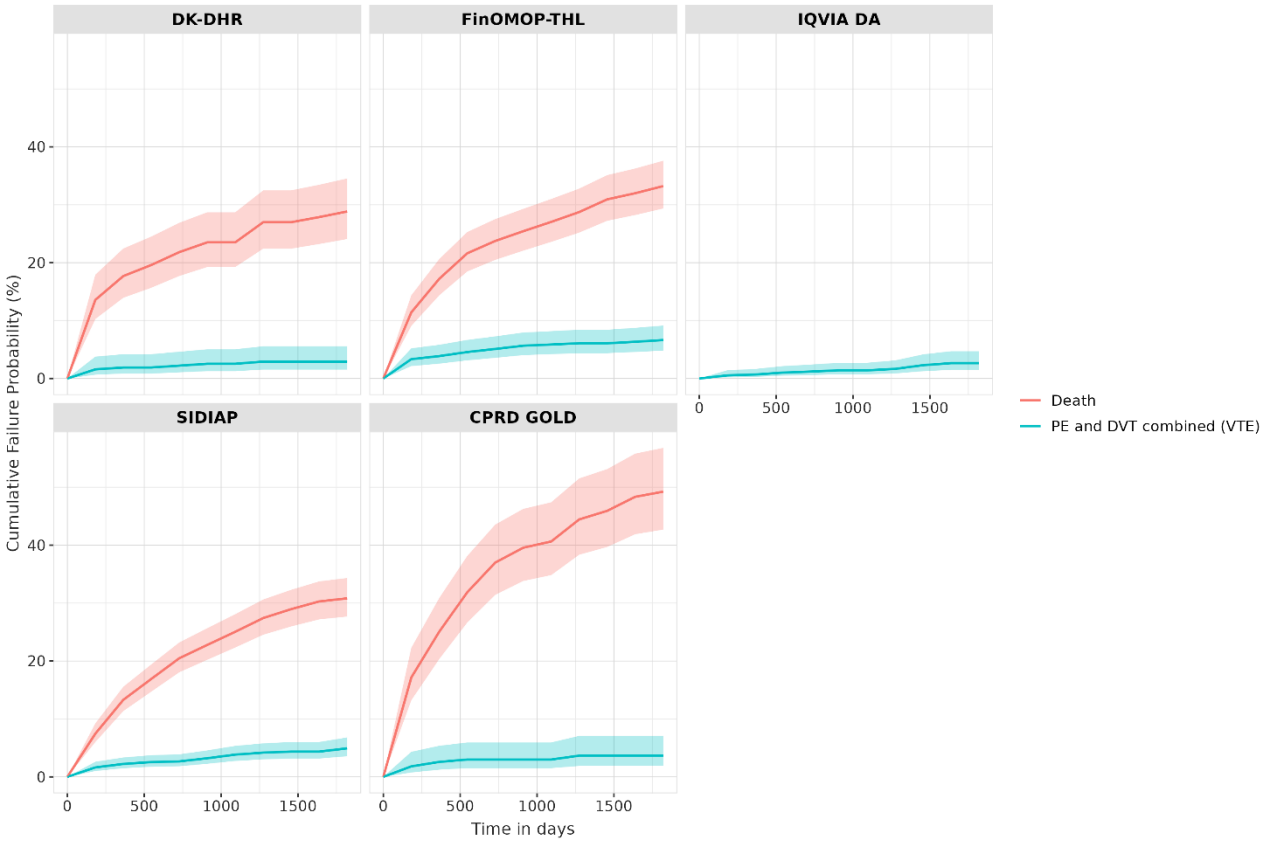

Figure 2a. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of bone cancer accounting for a competing risk of death.

DK-DHR=Danish Data Health Registries; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; CPRD GOLD=Clinical Practice Research Datalink GOLD.
DVT=deep vein thrombosis; PE=pulmonary embolism; VTE=composite outcome of pulmonary embolism and deep vein thrombosis.
Death data were incomplete in IQVIA DA Germany.

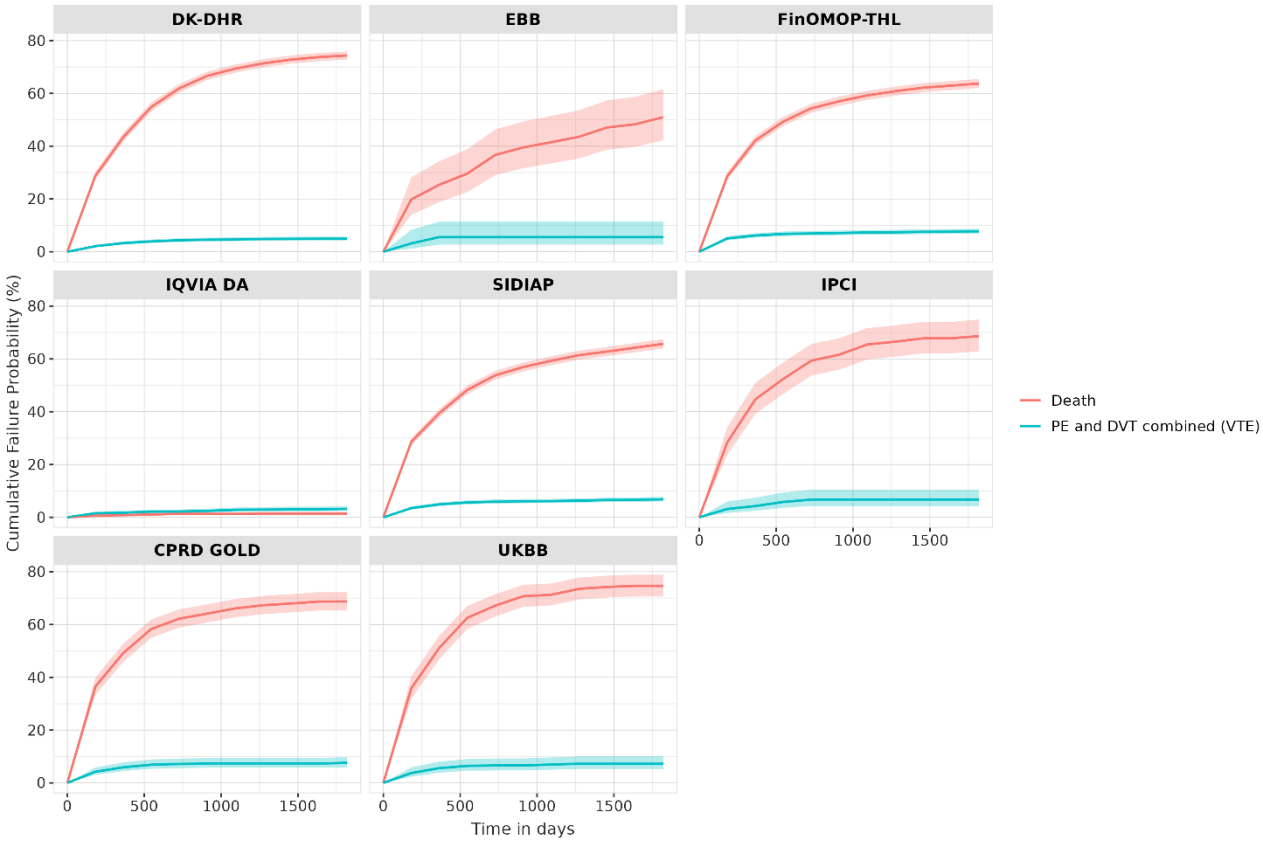

Figure 2b. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of brain cancer accounting for a competing risk of death.

DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; IPCI=Integrated Primary Care Information; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank.
DVT=deep vein thrombosis; PE=pulmonary embolism; VTE=composite outcome of pulmonary embolism and deep vein thrombosis.
Death data were incomplete in IQVIA DA Germany.

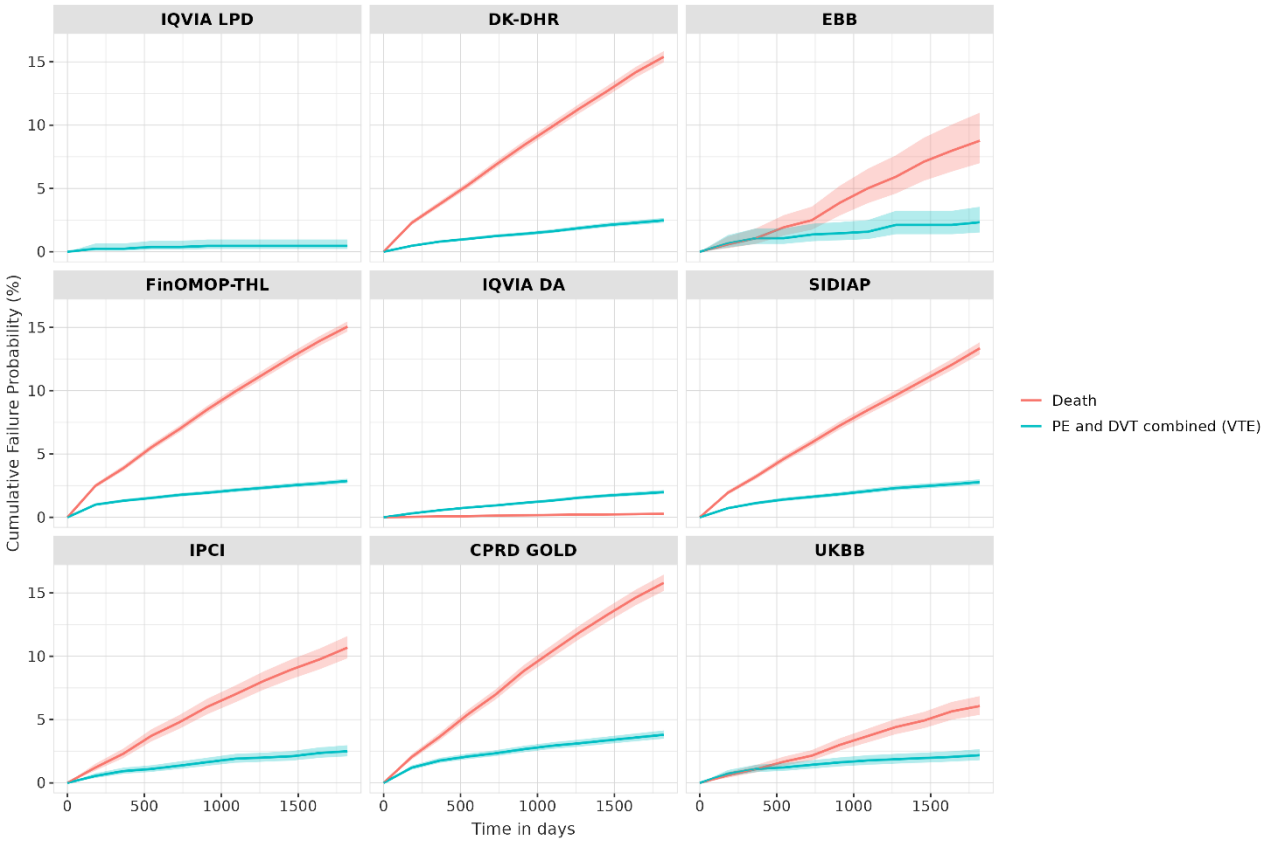

Figure 2c. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of breast cancer accounting for a competing risk of death.

IQVIA LPD=IQVIA Longitudinal Patient Database Belgium (IQVIA LPD Belgium); DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; IPCI=Integrated Primary Care Information; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank.
DVT=deep vein thrombosis; PE=pulmonary embolism; VTE=composite outcome of pulmonary embolism and deep vein thrombosis.
Death data were incomplete in IQVIA LPD Belgium and IQVIA DA Germany.

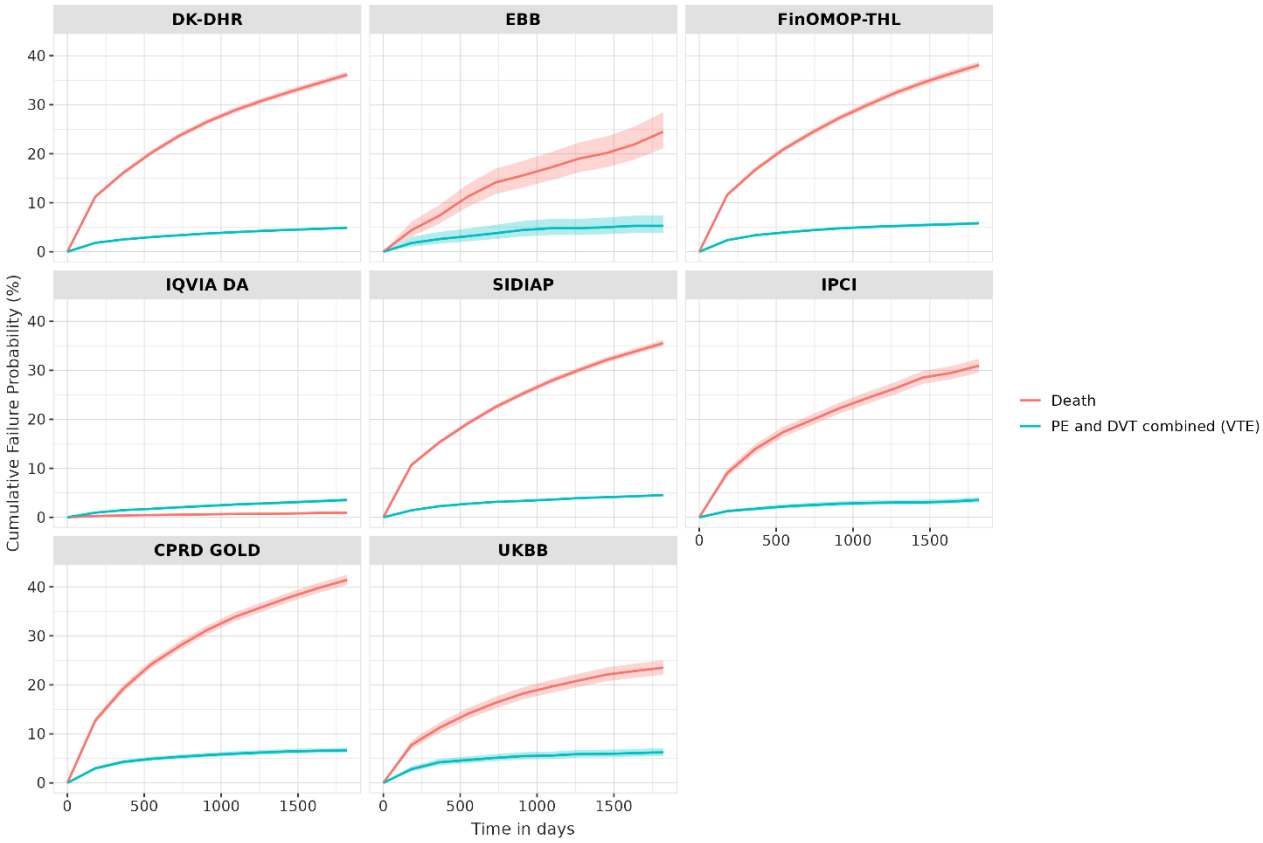

Figure 2d. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of colorectal cancer accounting for a competing risk of death.

DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; IPCI=Integrated Primary Care Information; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank.
DVT=deep vein thrombosis; PE=pulmonary embolism; VTE=composite outcome of pulmonary embolism and deep vein thrombosis.
Death data were incomplete in IQVIA LPD Belgium and IQVIA DA Germany.

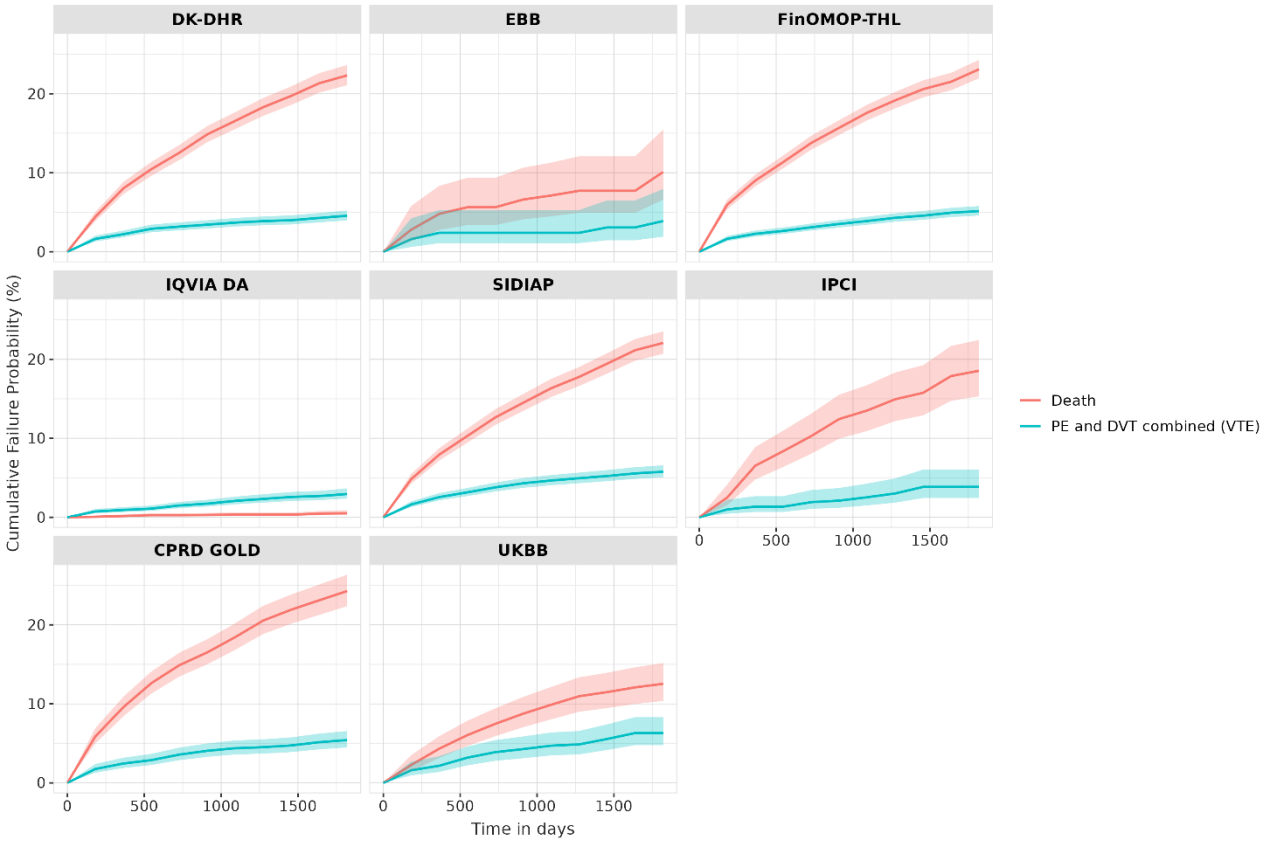

Figure 2e. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of corpus uteri cancer accounting for a competing risk of death.

DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; IPCI=Integrated Primary Care Information; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank.
DVT=deep vein thrombosis; PE=pulmonary embolism; VTE=composite outcome of pulmonary embolism and deep vein thrombosis.
Death data were incomplete in IQVIA DA Germany.

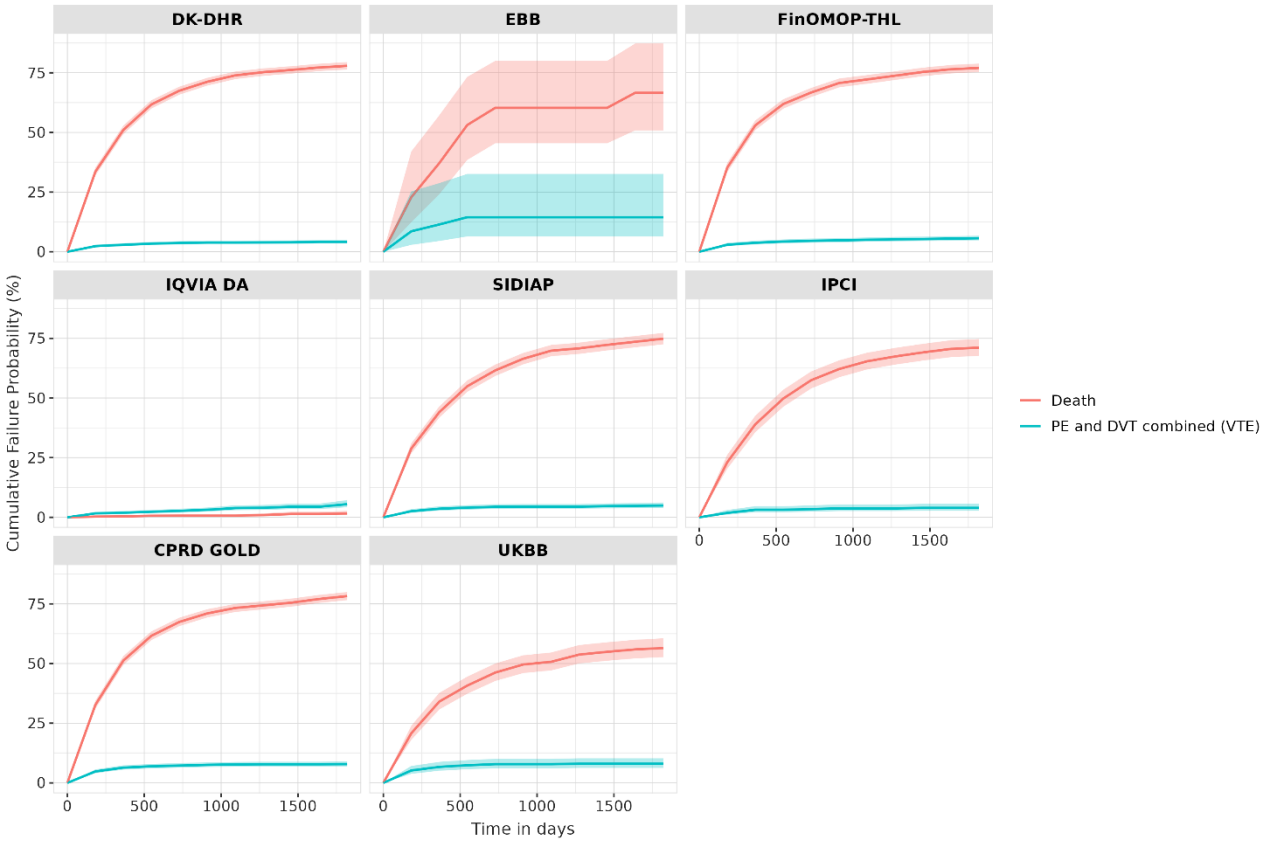

Figure 2f. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of oesophageal cancer accounting for a competing risk of death.

DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; IPCI=Integrated Primary Care Information; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank.
DVT=deep vein thrombosis; PE=pulmonary embolism; VTE=composite outcome of pulmonary embolism and deep vein thrombosis.
Death data were incomplete in IQVIA DA Germany.

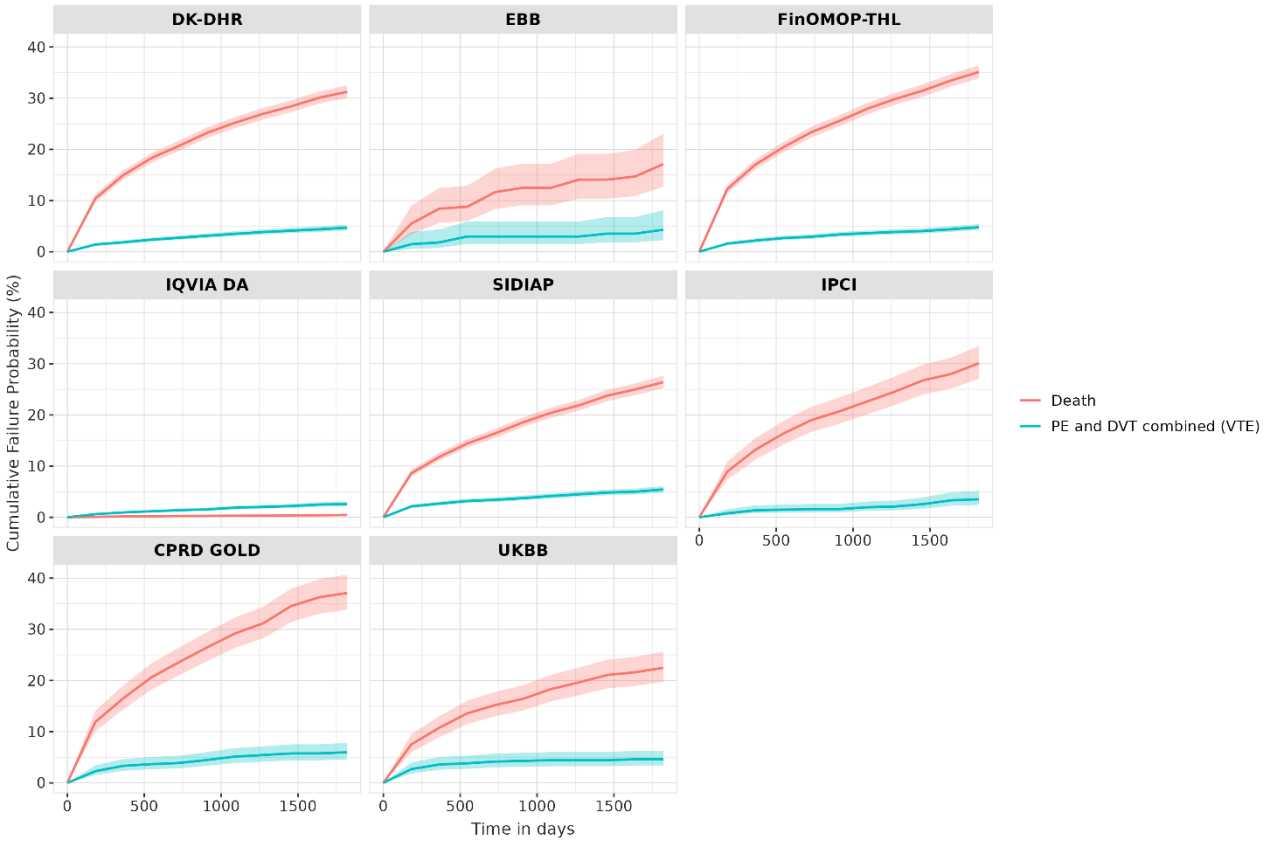

Figure 2g. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of kidney cancer accounting for a competing risk of death.

DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; IPCI=Integrated Primary Care Information; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank.
DVT=deep vein thrombosis; PE=pulmonary embolism; VTE=composite outcome of pulmonary embolism and deep vein thrombosis.
Death data were incomplete in IQVIA DA Germany.

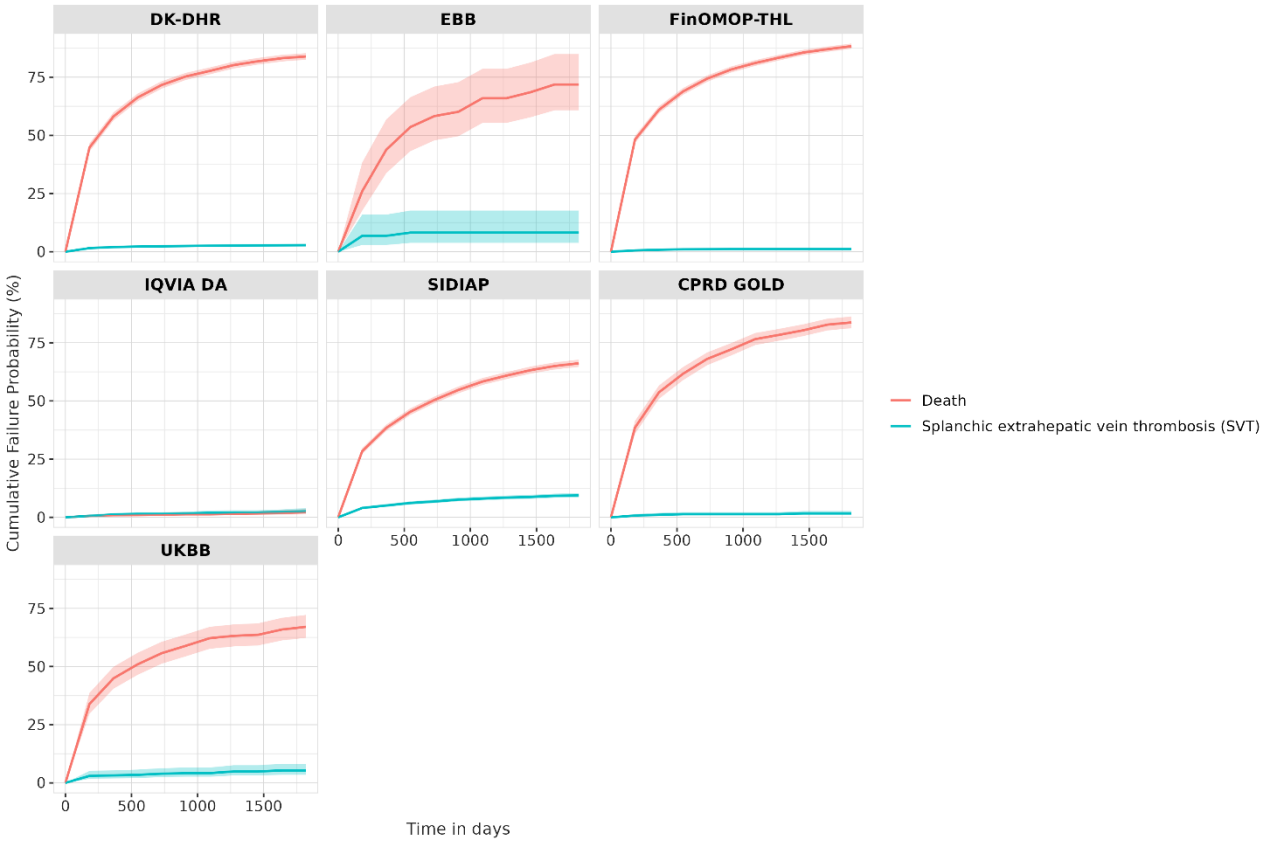

Figure 2h. Cumulative incidence of splanchnic vein thrombosis after the first diagnosis of liver cancer accounting for a competing risk of death.

DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank.
Death data were incomplete in IQVIA DA Germany.

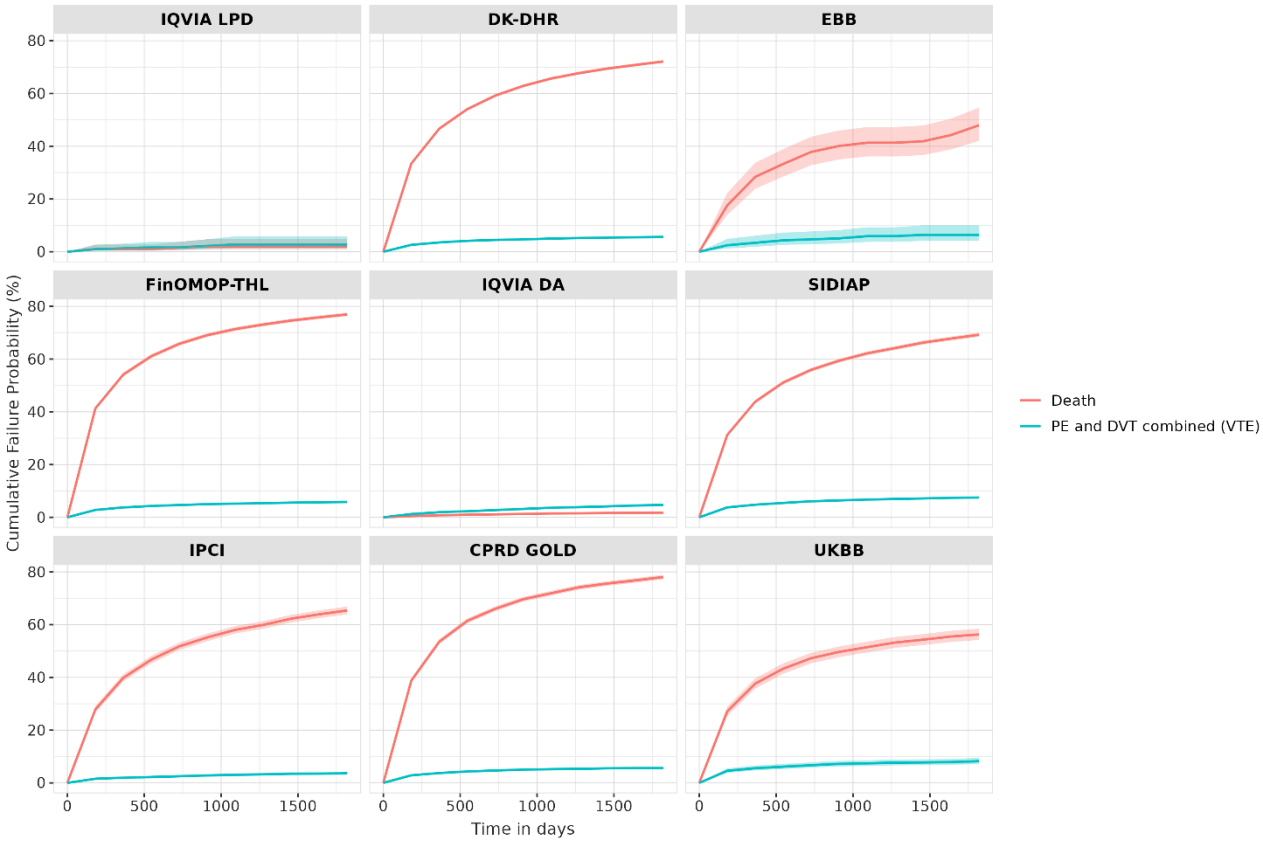

Figure 2i. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of lung cancer accounting for a competing risk of death.

IQVIA LPD=IQVIA Longitudinal Patient Database Belgium (IQVIA LPD Belgium); DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; IPCI=Integrated Primary Care Information; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank.
DVT=deep vein thrombosis; PE=pulmonary embolism; VTE=composite outcome of pulmonary embolism and deep vein thrombosis.
Death data were incomplete in IQVIA LPD Belgium and IQVIA DA Germany.

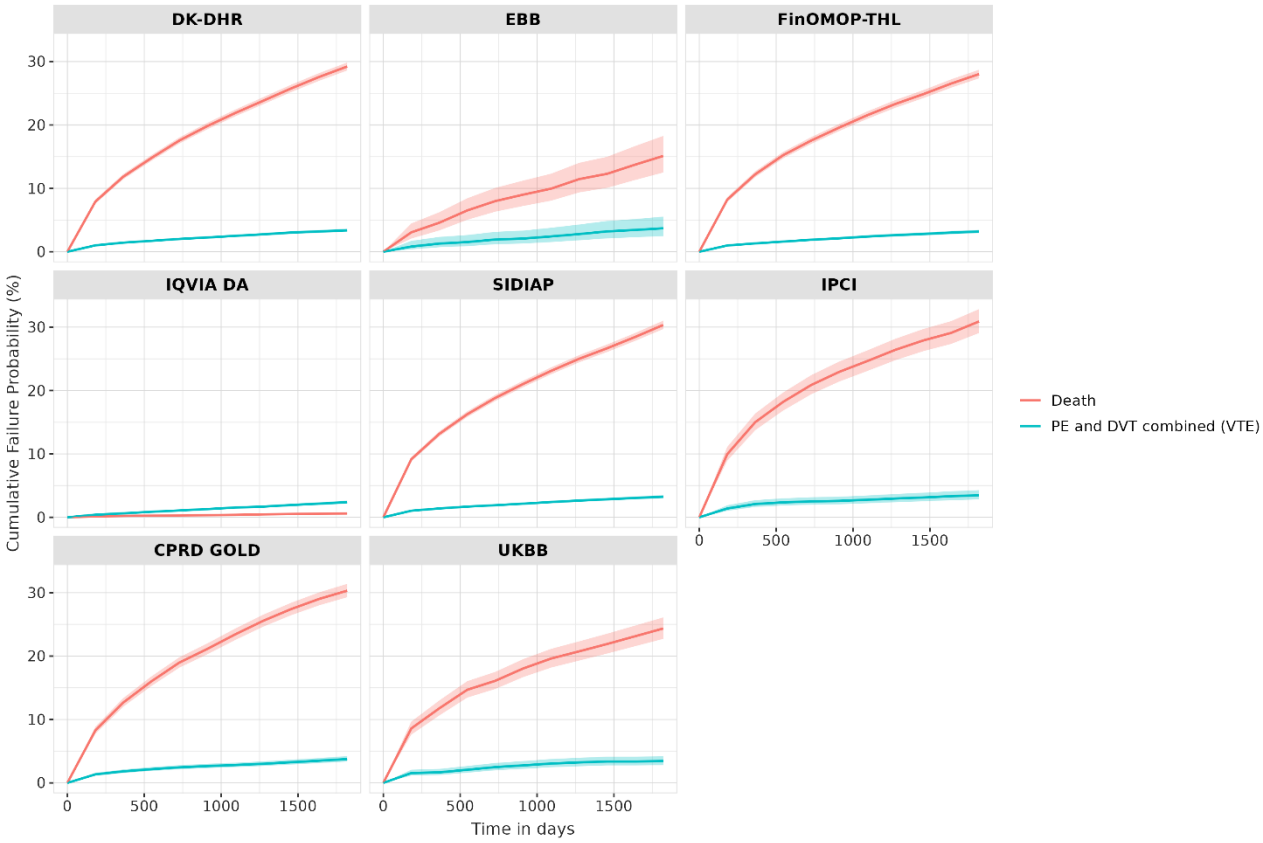

Figure 2j. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of lymphoma and leukaemia accounting for a competing risk of death.

DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; IPCI=Integrated Primary Care Information; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank.
DVT=deep vein thrombosis; PE=pulmonary embolism; VTE=composite outcome of pulmonary embolism and deep vein thrombosis.
Death data were incomplete in IQVIA DA Germany.

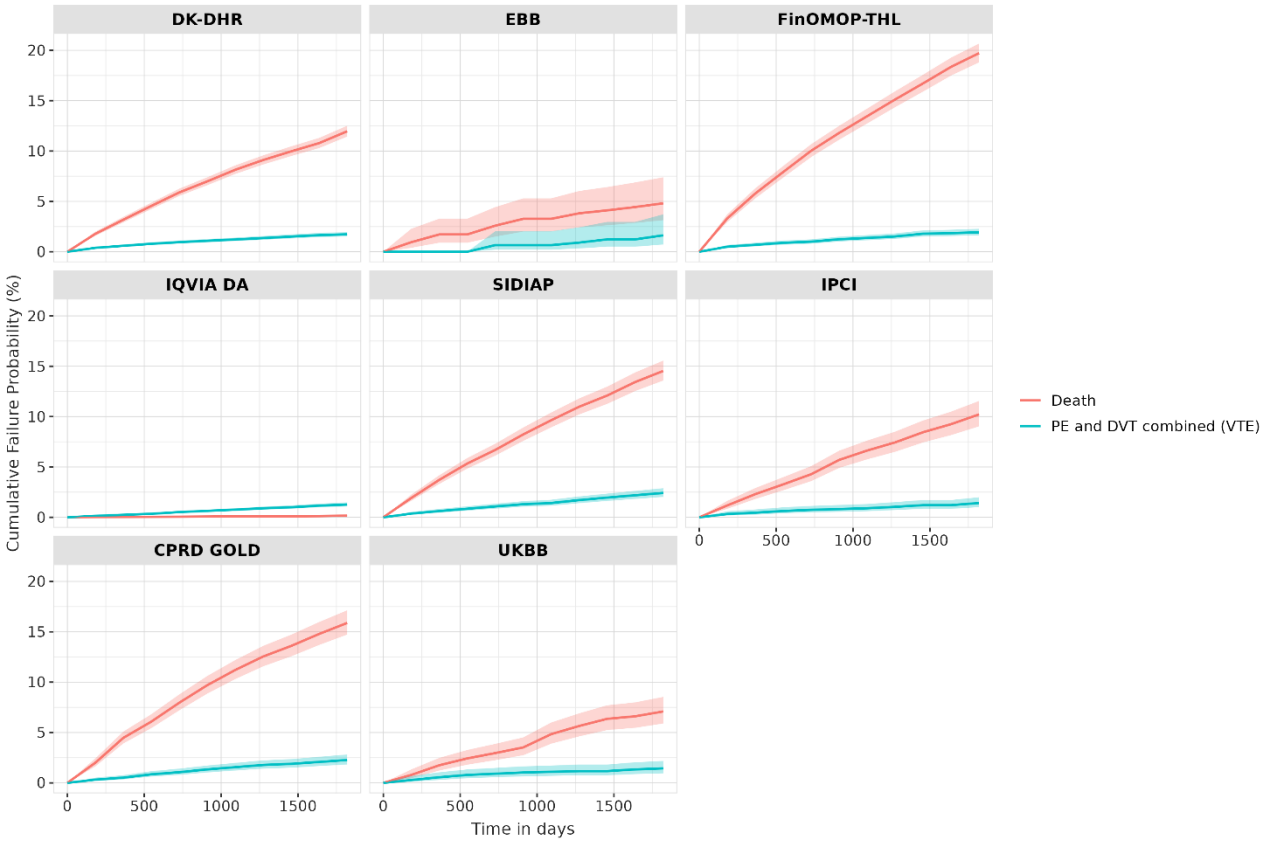

Figure 2k. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of melanoma accounting for a competing risk of death.

DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; IPCI=Integrated Primary Care Information; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank.
DVT=deep vein thrombosis; PE=pulmonary embolism; VTE=composite outcome of pulmonary embolism and deep vein thrombosis.
Death data were incomplete in IQVIA DA Germany.

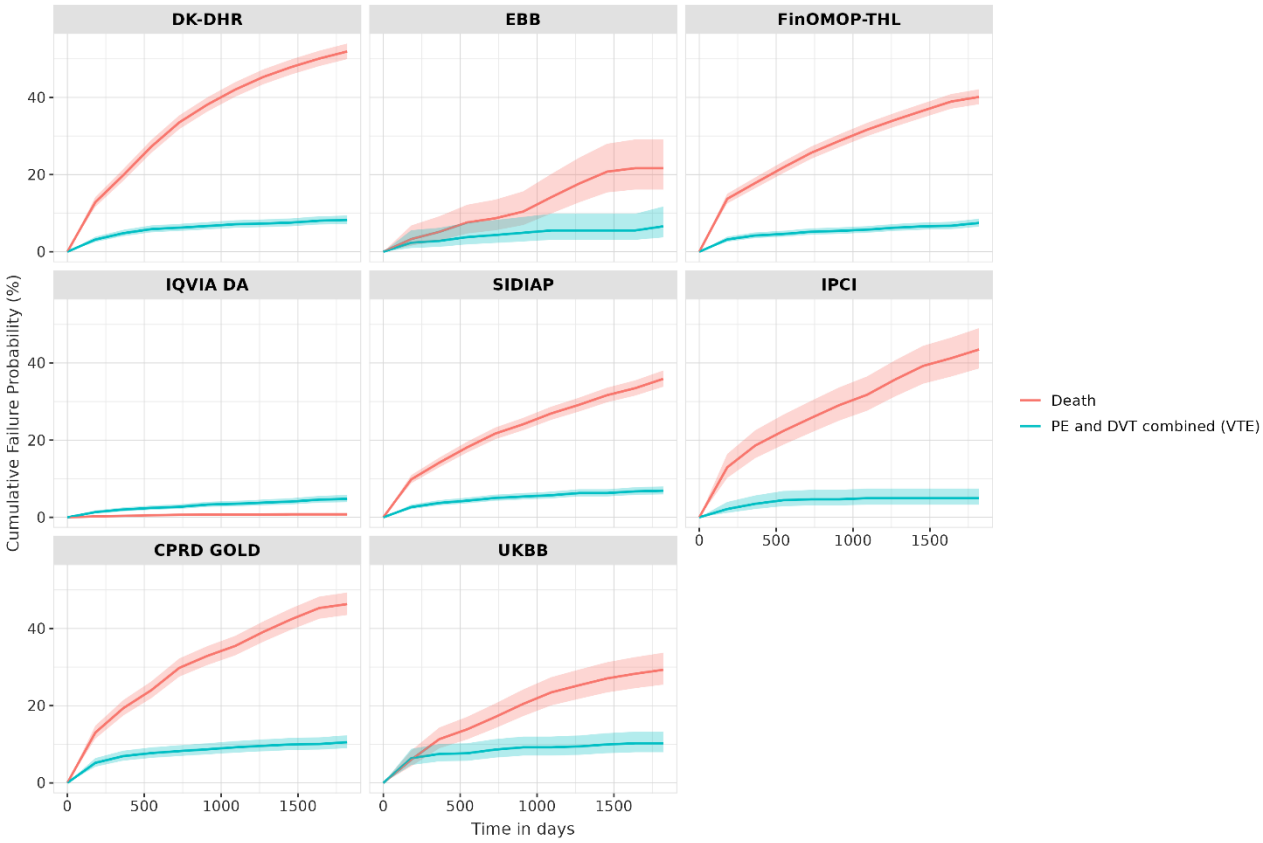

Figure 2l. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of ovarian cancer accounting for a competing risk of death.

DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; IPCI=Integrated Primary Care Information; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank.
DVT=deep vein thrombosis; PE=pulmonary embolism; VTE=composite outcome of pulmonary embolism and deep vein thrombosis.
Death data were incomplete in IQVIA DA Germany.

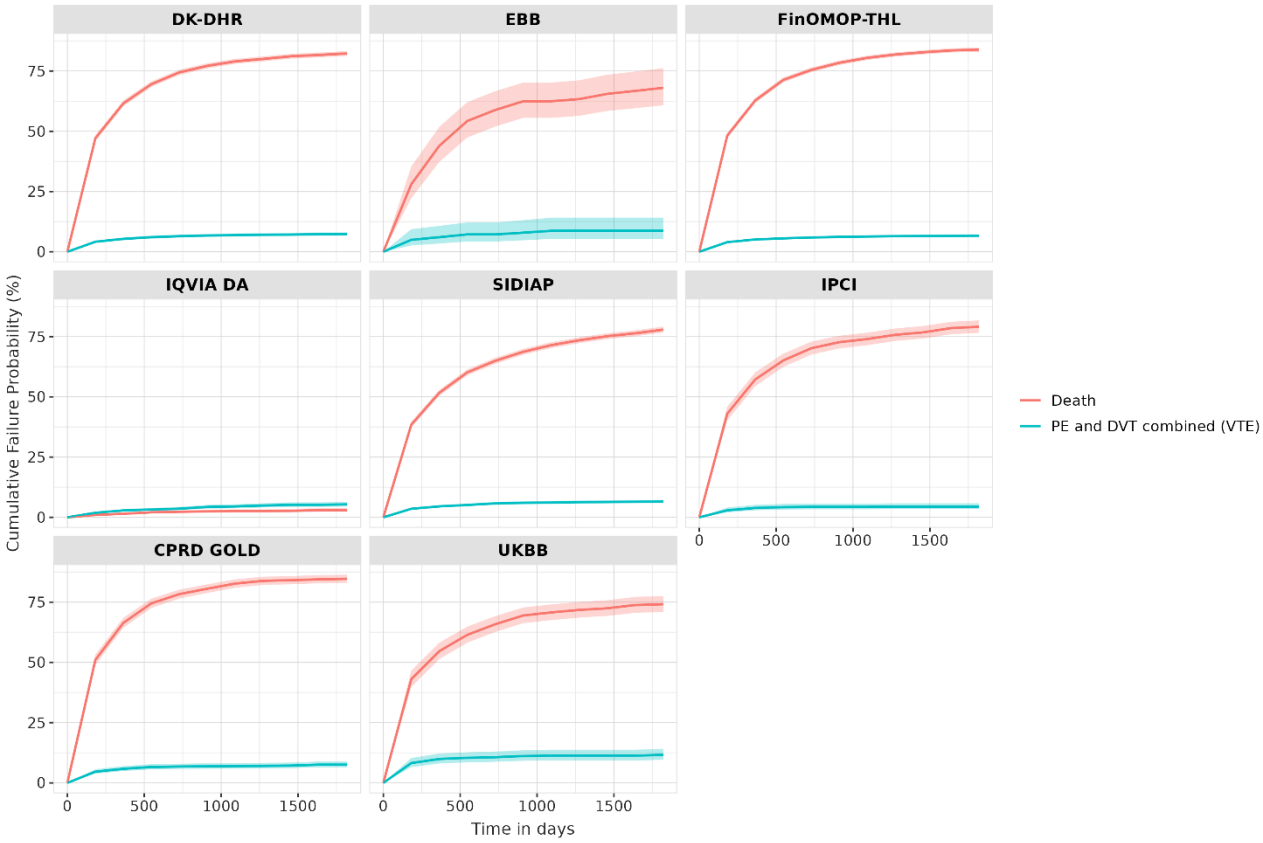

Figure 2m. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of pancreatic cancer accounting for a competing risk of death.

DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; IPCI=Integrated Primary Care Information; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank.
DVT=deep vein thrombosis; PE=pulmonary embolism; VTE=composite outcome of pulmonary embolism and deep vein thrombosis.
Death data were incomplete in IQVIA DA Germany.

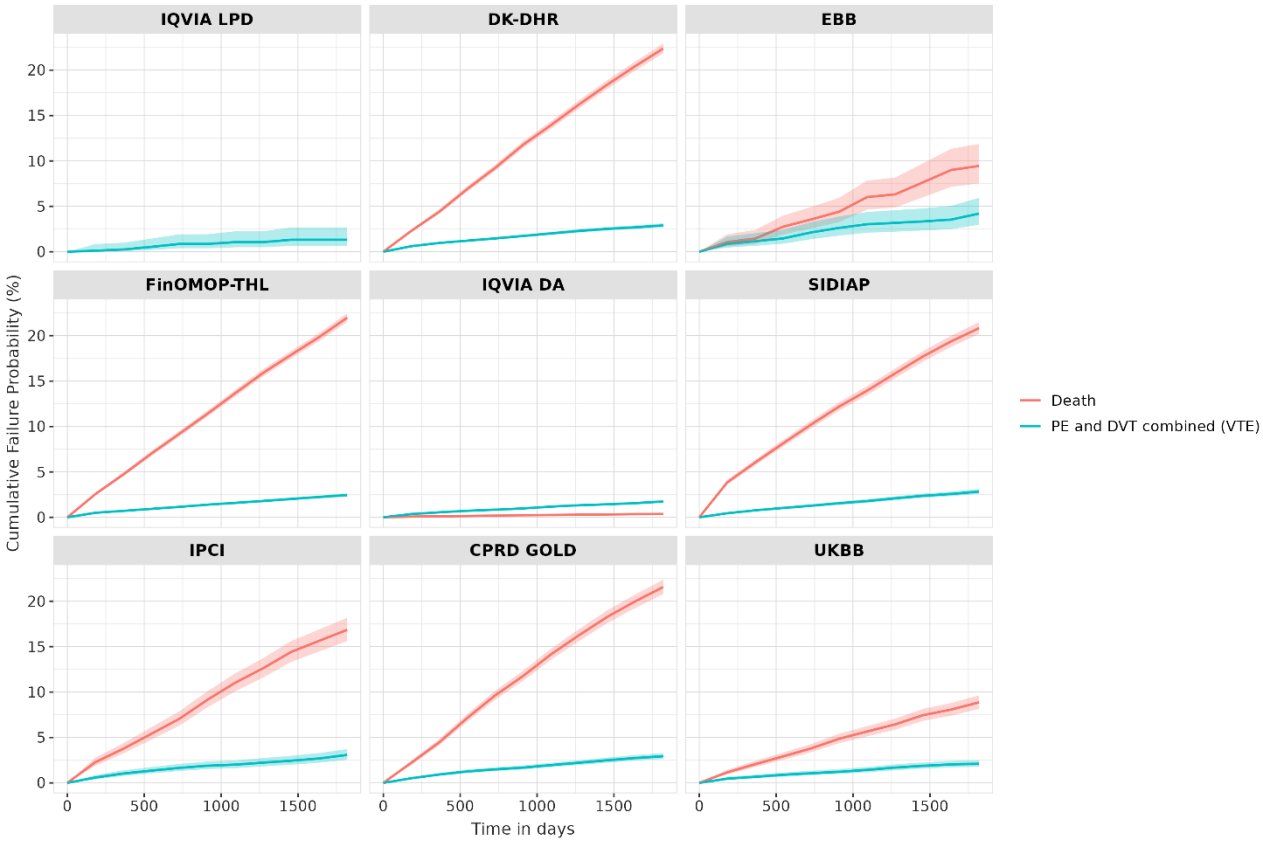

Figure 2n. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (VTE) after the first diagnosis of prostate cancer accounting for a competing risk of death.

IQVIA LPD=IQVIA Longitudinal Patient Database Belgium (IQVIA LPD Belgium); DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; IPCI=Integrated Primary Care Information; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank.
DVT=deep vein thrombosis; PE=pulmonary embolism; VTE=composite outcome of pulmonary embolism and deep vein thrombosis.
Death data were incomplete in IQVIA LPD Belgium and IQVIA DA Germany.

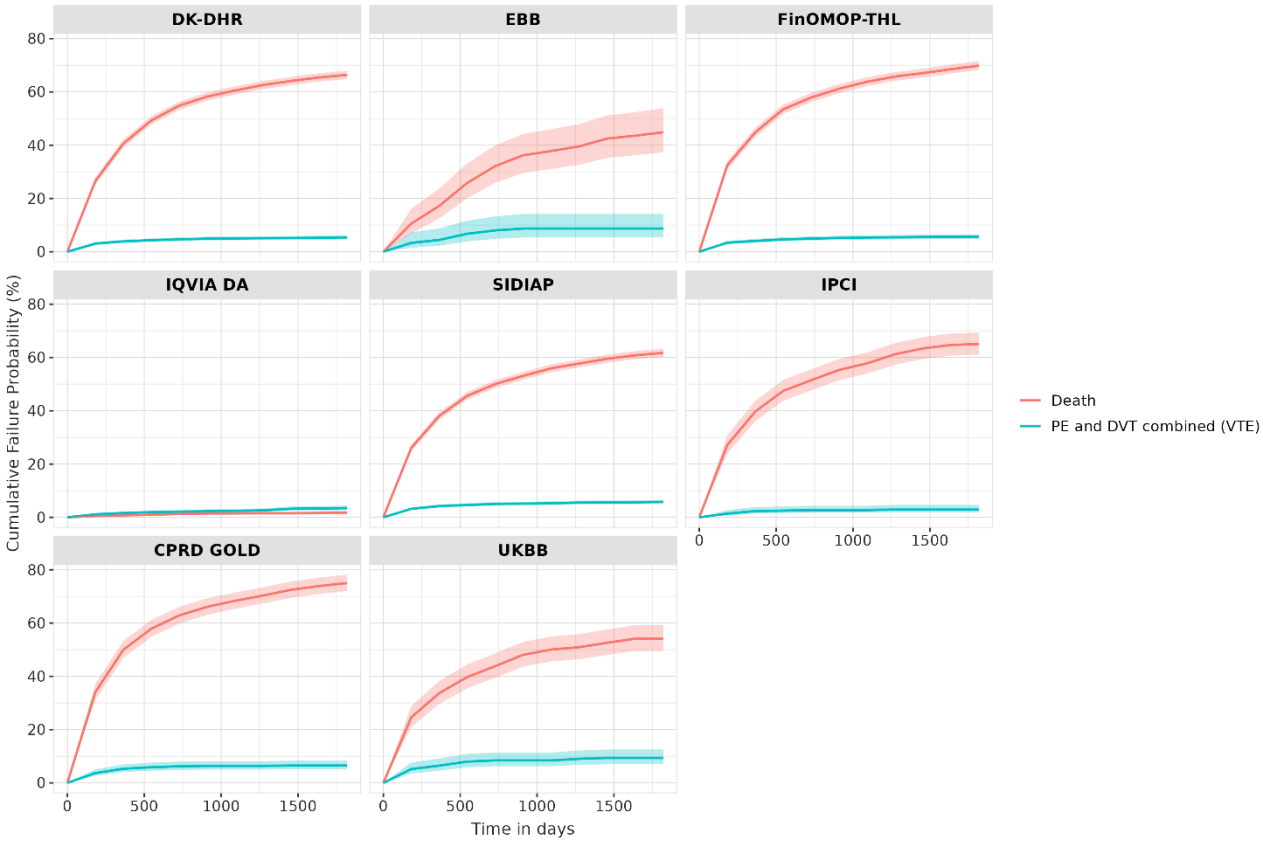

Figure 2o. Cumulative incidence of pulmonary embolism and deep vein thrombosis combined (TVE) after the first diagnosis of stomach cancer accounting for a competing risk of death.

DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; IPCI=Integrated Primary Care Information; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank.
DVT=deep vein thrombosis; PE=pulmonary embolism; VTE=composite outcome of pulmonary embolism and deep vein thrombosis.
Death data were incomplete in IQVIA DA Germany.

Table 3. Cumulative incidence of thromboembolic events (splanchnic vein thrombosis (SVT) in liver cancer and pulmonary embolism and deep vein thrombosis combined (VTE) in all other cancer types) at 6-month intervals up to 5 years after the first cancer diagnosis, % [95% confidence interval].

Time since first cancer diagnosis (months)	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
Liver cancer (N of SVT)	0	86	6	47	33	412	0	18	21
6	NA	1.57 [1.19–2.07]	6.85 [2.94–15.96]	0.59 [0.40–0.88]	0.63 [0.37–1.06]	4.01 [3.51–4.59]	NA	0.74 [0.38–1.41]	3.00 [1.75–5.12]
12	NA	2.01 [1.58–2.56]	6.85 [2.94–15.96]	0.84 [0.60–1.17]	1.20 [0.80–1.82]	5.07 [4.49–5.71]	NA	1.09 [0.63–1.87]	3.23 [1.93–5.40]
18	NA	2.23 [1.77–2.81]	8.22 [3.82–17.69]	1.06 [0.79–1.42]	1.45 [0.98–2.15]	6.16 [5.52–6.87]	NA	1.36 [0.83–2.21]	3.46 [2.10–5.68]
24	NA	2.36 [1.88–2.95]	8.22 [3.82–17.69]	1.11 [0.83–1.48]	1.45 [0.98–2.15]	6.81 [6.13–7.56]	NA	1.36 [0.83–2.21]	3.93 [2.47–6.27]
30	NA	2.46 [1.97–3.06]	8.22 [3.82–17.69]	1.16 [0.87–1.54]	1.66 [1.13–2.44]	7.60 [6.88–8.4]	NA	1.36 [0.83–2.21]	4.19 [2.66–6.58]
36	NA	2.57 [2.07–3.19]	8.22 [3.82–17.69]	1.16 [0.87–1.54]	1.93 [1.32–2.82]	8.05 [7.30–8.88]	NA	1.36 [0.83–2.21]	4.19 [2.66–6.58]
42	NA	2.65 [2.15–3.28]	8.22 [3.82–17.69]	1.16 [0.87–1.54]	2.07 [1.42–3.04]	8.48 [7.70–9.34]	NA	1.36 [0.83–2.21]	4.91 [3.19–7.58]
48	NA	2.70 [2.18–3.34]	8.22 [3.82–17.69]	1.16 [0.87–1.54]	2.07 [1.42–3.04]	8.77 [7.97–9.65]	NA	1.67 [1.04–2.68]	4.91 [3.19–7.58]
54	NA	2.75 [2.23–3.4]	8.22 [3.82–17.69]	1.16 [0.87–1.54]	2.29 [1.55–3.39]	9.24 [8.40–10.16]	NA	1.67 [1.04–2.68]	5.35 [3.49–8.20]
60	NA	2.81 [2.28–3.47]	8.22 [3.82–17.69]	1.16 [0.87–1.54]	2.60 [1.72–3.95]	9.44 [8.58–10.38]	NA	1.67 [1.04–2.68]	5.35 [3.49–8.20]
Bone cancer	0	9	0	36	13	40	NC	9	0

Time since first cancer diagnosis (months)	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
(N of VTE)									
6	NA	1.58 [0.66–3.77]	NA	3.34 [2.15–5.20]	0.54 [0.20–1.42]	1.63 [1.01–2.6]	NC	1.81 [0.76–4.32]	NA
12	NA	1.90 [0.86–4.19]	NA	3.87 [2.57–5.82]	0.69 [0.29–1.64]	2.23 [1.49–3.34]	NC	2.58 [1.24–5.36]	NA
18	NA	1.90 [0.86–4.19]	NA	4.57 [3.14–6.65]	1.02 [0.49–2.14]	2.55 [1.75–3.73]	NC	2.99 [1.51–5.93]	NA
24	NA	2.22 [1.07–4.62]	NA	5.1 [3.58–7.27]	1.20 [0.60–2.41]	2.67 [1.84–3.88]	NC	2.99 [1.51–5.93]	NA
30	NA	2.54 [1.28–5.04]	NA	5.66 [4.04–7.93]	1.41 [0.73–2.71]	3.21 [2.26–4.56]	NC	2.99 [1.51–5.93]	NA
36	NA	2.54 [1.28–5.04]	NA	5.87 [4.21–8.17]	1.41 [0.73–2.71]	3.84 [2.75–5.34]	NC	2.99 [1.51–5.93]	NA
42	NA	2.92 [1.53–5.56]	NA	6.08 [4.38–8.42]	1.66 [0.88–3.12]	4.18 [3.03–5.78]	NC	3.66 [1.89–7.07]	NA
48	NA	2.92 [1.53–5.56]	NA	6.08 [4.38–8.42]	2.30 [1.27–4.16]	4.36 [3.17–6.01]	NC	3.66 [1.89–7.07]	NA
54	NA	2.92 [1.53–5.56]	NA	6.34 [4.59–8.74]	2.65 [1.49–4.71]	4.36 [3.17–6.01]	NC	3.66 [1.89–7.07]	NA
60	NA	2.92 [1.53–5.56]	NA	6.64 [4.82–9.13]	2.65 [1.49–4.71]	4.92 [3.56–6.80]	NC	3.66 [1.89–7.07]	NA
Brain cancer (N of VTE)	0	166	7	245	56	236	18	59	34
6	NA	2.10 [1.67–2.64]	3.17 [1.21–8.33]	5.04 [4.34–5.85]	1.54 [1.11–2.14]	3.48 [2.94–4.11]	3.15 [1.65–5.99]	4.22 [3.05–5.84]	3.73 [2.37–5.86]

Time since first cancer diagnosis (months)	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
12	NA	3.27 [2.73–3.93]	5.56 [2.70–11.41]	6.15 [5.38–7.04]	1.75 [1.28–2.4]	4.91 [4.26–5.65]	4.27 [2.45–7.43]	5.90 [4.48–7.76]	5.59 [3.87–8.06]
18	NA	3.92 [3.32–4.63]	5.56 [2.70–11.41]	6.74 [5.93–7.66]	2.14 [1.59–2.86]	5.62 [4.93–6.41]	5.82 [3.62–9.38]	6.88 [5.33–8.88]	6.42 [4.57–9.02]
24	NA	4.36 [3.73–5.10]	5.56 [2.70–11.41]	6.95 [6.13–7.89]	2.21 [1.65–2.95]	5.92 [5.20–6.74]	6.71 [4.29–10.50]	7.17 [5.58–9.22]	6.64 [4.75–9.28]
30	NA	4.57 [3.92–5.33]	5.56 [2.70–11.41]	7.12 [6.28–8.06]	2.45 [1.85–3.26]	6.07 [5.34–6.90]	6.71 [4.29–10.50]	7.34 [5.72–9.41]	6.64 [4.75–9.28]
36	NA	4.67 [4.01–5.43]	5.56 [2.70–11.41]	7.29 [6.45–8.25]	2.81 [2.13–3.71]	6.15 [5.42–6.99]	6.71 [4.29–10.50]	7.34 [5.72–9.41]	6.93 [4.98–9.64]
42	NA	4.85 [4.18–5.63]	5.56 [2.70–11.41]	7.37 [6.52–8.33]	2.93 [2.22–3.86]	6.29 [5.54–7.15]	6.71 [4.29–10.50]	7.34 [5.72–9.41]	7.26 [5.24–10.06]
48	NA	4.89 [4.21–5.67]	5.56 [2.70–11.41]	7.54 [6.67–8.51]	3.05 [2.31–4.02]	6.56 [5.78–7.44]	6.71 [4.29–10.50]	7.34 [5.72–9.41]	7.26 [5.24–10.06]
54	NA	4.93 [4.25–5.72]	5.56 [2.70–11.41]	7.63 [6.76–8.61]	3.05 [2.31–4.02]	6.62 [5.83–7.51]	6.71 [4.29–10.50]	7.34 [5.72–9.41]	7.26 [5.24–10.06]
60	NA	4.93 [4.25–5.72]	5.56 [2.70–11.41]	7.73 [6.85–8.72]	3.22 [2.43–4.27]	6.85 [6.03–7.78]	6.71 [4.29–10.50]	7.60 [5.93–9.75]	7.26 [5.24–10.06]
Breast cancer (N of VTE)	7	684	23	973	661	692	134	550	101
6	0.24 [0.09–0.65]	0.47 [0.40–0.55]	0.66 [0.33–1.31]	1.00 [0.91–1.11]	0.31 [0.26–0.37]	0.72 [0.64–0.82]	0.55 [0.39–0.76]	1.20 [1.04–1.37]	0.73 [0.53–1.01]
12	0.24 [0.09–0.65]	0.80 [0.70–0.90]	1.07 [0.62–1.84]	1.31 [1.20–1.43]	0.56 [0.50–0.64]	1.12 [1.01–1.24]	0.93 [0.72–1.20]	1.76 [1.58–1.97]	1.10 [0.85–1.43]
18	0.38 [0.17–0.86]	1.01 [0.91–1.13]	1.07 [0.62–1.84]	1.53 [1.41–1.66]	0.77 [0.69–0.86]	1.41 [1.28–1.55]	1.10 [0.87–1.39]	2.08 [1.88–2.31]	1.22 [0.95–1.56]

Time since first cancer diagnosis (months)	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
24	0.38 [0.17–0.86]	1.24 [1.13–1.37]	1.36 [0.83–2.21]	1.77 [1.64–1.91]	0.94 [0.85–1.04]	1.62 [1.48–1.77]	1.36 [1.10–1.69]	2.34 [2.12–2.58]	1.43 [1.13–1.79]
30	0.46 [0.22–0.97]	1.41 [1.28–1.55]	1.46 [0.91–2.34]	1.94 [1.81–2.09]	1.14 [1.04–1.25]	1.83 [1.68–1.99]	1.64 [1.35–1.99]	2.66 [2.42–2.92]	1.62 [1.31–2.01]
36	0.46 [0.22–0.97]	1.61 [1.48–1.76]	1.58 [1.00–2.50]	2.15 [2.00–2.30]	1.32 [1.21–1.44]	2.07 [1.91–2.25]	1.91 [1.59–2.29]	2.93 [2.68–3.20]	1.77 [1.43–2.18]
42	0.46 [0.22–0.97]	1.87 [1.72–2.03]	2.12 [1.39–3.23]	2.33 [2.18–2.49]	1.55 [1.43–1.68]	2.31 [2.14–2.5]	2.00 [1.67–2.39]	3.13 [2.87–3.42]	1.87 [1.52–2.29]
48	0.46 [0.22–0.97]	2.11 [1.95–2.29]	2.12 [1.39–3.23]	2.52 [2.36–2.69]	1.72 [1.59–1.86]	2.46 [2.27–2.66]	2.10 [1.76–2.51]	3.36 [3.08–3.66]	1.96 [1.60–2.39]
54	0.46 [0.22–0.97]	2.29 [2.12–2.47]	2.12 [1.39–3.23]	2.68 [2.51–2.85]	1.85 [1.71–2.01]	2.61 [2.42–2.82]	2.36 [1.98–2.80]	3.58 [3.29–3.90]	2.05 [1.68–2.50]
60	0.46 [0.22–0.97]	2.48 [2.30–2.67]	2.33 [1.52–3.55]	2.87 [2.69–3.05]	1.99 [1.84–2.15]	2.78 [2.57–3.00]	2.50 [2.10–2.97]	3.80 [3.49–4.14]	2.19 [1.80–2.66]
Colorectal cancer (N of VTE)	0	1,298	35	1,226	482	1,115	163	665	220
6	NA	1.82 [1.67–1.98]	1.77 [1.03–3.04]	2.36 [2.16–2.56]	0.95 [0.81–1.10]	1.42 [1.29–1.57]	1.26 [0.99–1.60]	2.97 [2.67–3.31]	2.79 [2.31–3.37]
12	NA	2.51 [2.34–2.7]	2.59 [1.66–4.04]	3.38 [3.15–3.62]	1.46 [1.29–1.64]	2.27 [2.11–2.45]	1.73 [1.41–2.13]	4.29 [3.92–4.69]	4.21 [3.61–4.90]
18	NA	3.00 [2.81–3.21]	3.16 [2.12–4.73]	3.92 [3.68–4.19]	1.73 [1.55–1.94]	2.75 [2.57–2.95]	2.20 [1.84–2.64]	4.90 [4.51–5.33]	4.66 [4.03–5.39]
24	NA	3.37 [3.17–3.59]	3.78 [2.61–5.48]	4.39 [4.13–4.67]	2.05 [1.84–2.27]	3.15 [2.95–3.36]	2.50 [2.11–2.97]	5.32 [4.91–5.77]	5.11 [4.45–5.87]

Time since first cancer diagnosis (months)	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
30	NA	3.73 [3.52–3.96]	4.45 [3.15–6.29]	4.77 [4.50–5.06]	2.33 [2.10–2.57]	3.35 [3.15–3.58]	2.77 [2.35–3.26]	5.67 [5.24–6.13]	5.47 [4.79–6.26]
36	NA	4.00 [3.77–4.23]	4.80 [3.43–6.71]	5.04 [4.76–5.34]	2.62 [2.37–2.88]	3.61 [3.39–3.84]	2.94 [2.50–3.45]	5.97 [5.53–6.45]	5.58 [4.88–6.37]
42	NA	4.27 [4.03–4.51]	4.80 [3.43–6.71]	5.25 [4.96–5.56]	2.83 [2.57–3.11]	3.92 [3.69–4.16]	3.02 [2.58–3.54]	6.21 [5.76–6.70]	5.89 [5.17–6.71]
48	NA	4.48 [4.24–4.74]	5.02 [3.61–6.99]	5.45 [5.16–5.76]	3.05 [2.78–3.35]	4.12 [3.88–4.37]	3.05 [2.60–3.57]	6.44 [5.97–6.94]	5.93 [5.20–6.76]
54	NA	4.68 [4.43–4.94]	5.30 [3.82–7.37]	5.61 [5.31–5.93]	3.30 [3.00–3.62]	4.29 [4.04–4.56]	3.22 [2.76–3.77]	6.55 [6.08–7.06]	6.08 [5.34–6.92]
60	NA	4.88 [4.62–5.14]	5.30 [3.82–7.37]	5.82 [5.51–6.15]	3.54 [3.22–3.89]	4.55 [4.29–4.83]	3.54 [3.03–4.13]	6.65 [6.17–7.17]	6.24 [5.48–7.10]
Corpus uteri cancer (N of VTE)	0	202	8	277	89	224	19	106	49
6	NA	1.63 [1.31–2.04]	1.60 [0.61–4.23]	1.64 [1.35–2.00]	0.74 [0.52–1.06]	1.63 [1.30–2.03]	1.00 [0.45–2.21]	1.76 [1.29–2.39]	1.60 [0.95–2.68]
12	NA	2.22 [1.84–2.68]	2.40 [1.09–5.29]	2.27 [1.92–2.68]	0.93 [0.68–1.28]	2.56 [2.15–3.06]	1.35 [0.68–2.68]	2.45 [1.89–3.18]	2.17 [1.39–3.38]
18	NA	2.91 [2.47–3.43]	2.40 [1.09–5.29]	2.64 [2.26–3.08]	1.10 [0.82–1.49]	3.16 [2.69–3.71]	1.35 [0.68–2.68]	2.88 [2.26–3.67]	3.19 [2.22–4.60]
24	NA	3.18 [2.72–3.72]	2.40 [1.09–5.29]	3.10 [2.69–3.57]	1.50 [1.16–1.95]	3.79 [3.27–4.40]	1.92 [1.07–3.45]	3.58 [2.88–4.46]	3.89 [2.80–5.41]
30	NA	3.43 [2.95–3.99]	2.40 [1.09–5.29]	3.52 [3.07–4.02]	1.73 [1.36–2.22]	4.31 [3.74–4.97]	2.11 [1.21–3.71]	4.06 [3.30–4.99]	4.28 [3.12–5.87]

Time since first cancer diagnosis (months)	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
36	NA	3.69 [3.19–4.27]	2.40 [1.09–5.29]	3.89 [3.42–4.43]	2.08 [1.65–2.62]	4.66 [4.06–5.35]	2.54 [1.51–4.27]	4.39 [3.60–5.37]	4.71 [3.48–6.38]
42	NA	3.88 [3.36–4.47]	2.40 [1.09–5.29]	4.30 [3.80–4.87]	2.33 [1.87–2.92]	4.95 [4.32–5.66]	3.02 [1.86–4.92]	4.52 [3.71–5.51]	4.87 [3.61–6.57]
48	NA	3.99 [3.46–4.60]	3.08 [1.46–6.50]	4.57 [4.05–5.16]	2.58 [2.07–3.21]	5.23 [4.58–5.98]	3.85 [2.45–6.04]	4.74 [3.90–5.76]	5.56 [4.17–7.41]
54	NA	4.28 [3.72–4.91]	3.08 [1.46–6.50]	4.94 [4.39–5.56]	2.69 [2.17–3.34]	5.56 [4.87–6.35]	3.85 [2.45–6.04]	5.15 [4.25–6.23]	6.31 [4.78–8.32]
60	NA	4.56 [3.97–5.22]	3.89 [1.90–7.94]	5.15 [4.58–5.79]	2.95 [2.38–3.66]	5.77 [5.05–6.58]	3.85 [2.45–6.04]	5.42 [4.48–6.56]	6.31 [4.78–8.32]
Oesophageal cancer (N of VTE)	<5	124	5	131	73	81	30	222	58
6	NA	2.35 [1.87–2.95]	8.57 [2.90–25.29]	2.94 [2.34–3.70]	1.65 [1.19–2.28]	2.56 [1.93–3.41]	1.86 [1.13–3.07]	4.79 [4.08–5.63]	5.20 [3.81–7.08]
12	NA	2.91 [2.37–3.57]	11.43 [4.54–28.74]	3.77 [3.08–4.61]	1.92 [1.41–2.60]	3.60 [2.83–4.59]	3.15 [2.14–4.63]	6.39 [5.56–7.34]	6.70 [5.11–8.78]
18	NA	3.40 [2.81–4.10]	14.48 [6.43–32.59]	4.27 [3.53–5.15]	2.31 [1.74–3.08]	4.05 [3.22–5.08]	3.15 [2.14–4.63]	6.95 [6.08–7.94]	7.39 [5.71–9.55]
24	NA	3.69 [3.08–4.42]	14.48 [6.43–32.59]	4.60 [3.83–5.52]	2.70 [2.05–3.55]	4.40 [3.53–5.49]	3.44 [2.37–4.98]	7.26 [6.37–8.28]	7.84 [6.11–10.06]
30	NA	3.87 [3.24–4.61]	14.48 [6.43–32.59]	4.77 [3.99–5.71]	3.15 [2.41–4.10]	4.48 [3.60–5.57]	3.74 [2.62–5.36]	7.55 [6.64–8.59]	7.84 [6.11–10.06]
36	NA	3.87 [3.24–4.61]	14.48 [6.43–32.59]	5.01 [4.21–5.97]	3.87 [3.00–5.00]	4.48 [3.60–5.57]	3.74 [2.62–5.36]	7.73 [6.81–8.78]	7.84 [6.11–10.06]

Time since first cancer diagnosis (months)	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
42	NA	3.91 [3.27–4.66]	14.48 [6.43–32.59]	5.22 [4.39–6.19]	3.99 [3.10–5.14]	4.48 [3.60–5.57]	3.74 [2.62–5.36]	7.79 [6.86–8.84]	8.05 [6.28–10.31]
48	NA	3.95 [3.31–4.70]	14.48 [6.43–32.59]	5.34 [4.50–6.33]	4.44 [3.44–5.72]	4.68 [3.77–5.82]	3.96 [2.78–5.65]	7.79 [6.86–8.84]	8.05 [6.28–10.31]
54	NA	4.14 [3.48–4.92]	14.48 [6.43–32.59]	5.53 [4.67–6.54]	4.44 [3.44–5.72]	4.80 [3.86–5.97]	3.96 [2.78–5.65]	7.79 [6.86–8.84]	8.05 [6.28–10.31]
60	NA	4.14 [3.48–4.92]	14.48 [6.43–32.59]	5.67 [4.79–6.70]	5.51 [4.23–7.18]	4.95 [3.98–6.15]	3.96 [2.78–5.65]	7.86 [6.92–8.93]	8.05 [6.28–10.31]
Kidney cancer (N of VTE)	0	266	10	266	142	316	28	53	40
6	NA	1.41 [1.15–1.74]	1.47 [0.55–3.88]	1.58 [1.30–1.92]	0.62 [0.46–0.83]	2.13 [1.82–2.49]	0.78 [0.39–1.55]	2.28 [1.52–3.42]	2.69 [1.82–4.00]
12	NA	1.84 [1.54–2.21]	1.83 [0.77–4.36]	2.18 [1.84–2.58]	0.95 [0.75–1.20]	2.68 [2.33–3.08]	1.38 [0.82–2.33]	3.33 [2.38–4.66]	3.59 [2.56–5.05]
18	NA	2.36 [2.01–2.76]	2.98 [1.51–5.90]	2.67 [2.30–3.11]	1.17 [0.94–1.45]	3.18 [2.79–3.62]	1.49 [0.90–2.46]	3.66 [2.66–5.05]	3.82 [2.74–5.31]
24	NA	2.73 [2.35–3.16]	2.98 [1.51–5.90]	2.93 [2.54–3.39]	1.39 [1.14–1.70]	3.42 [3.01–3.88]	1.60 [0.98–2.61]	3.89 [2.85–5.32]	4.17 [3.04–5.72]
30	NA	3.13 [2.72–3.59]	2.98 [1.51–5.90]	3.37 [2.94–3.85]	1.54 [1.27–1.87]	3.75 [3.32–4.23]	1.60 [0.98–2.61]	4.49 [3.35–6.02]	4.30 [3.15–5.87]
36	NA	3.5 [3.07–3.99]	2.98 [1.51–5.90]	3.64 [3.19–4.14]	1.89 [1.58–2.26]	4.16 [3.69–4.68]	1.97 [1.26–3.08]	5.14 [3.9–6.78]	4.43 [3.26–6.02]
42	NA	3.85 [3.39–4.37]	2.98 [1.51–5.90]	3.86 [3.40–4.39]	2.03 [1.70–2.42]	4.50 [4.01–5.05]	2.12 [1.37–3.28]	5.44 [4.14–7.14]	4.43 [3.26–6.02]
48	NA	4.12 [3.64–4.66]	3.54 [1.85–6.78]	4.05 [3.57–4.59]	2.2 [1.85–2.61]	4.84 [4.32–5.42]	2.58 [1.71–3.89]	5.76 [4.41–7.53]	4.43 [3.26–6.02]

Time since first cancer diagnosis (months)	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
54	NA	4.39 [3.89–4.96]	3.54 [1.85–6.78]	4.40 [3.90–4.97]	2.49 [2.10–2.95]	5.00 [4.46–5.60]	3.32 [2.26–4.88]	5.76 [4.41–7.53]	4.62 [3.40–6.26]
60	NA	4.70 [4.17–5.30]	4.30 [2.28–8.09]	4.80 [4.26–5.41]	2.57 [2.16–3.05]	5.45 [4.86–6.1]	3.53 [2.42–5.15]	5.98 [4.58–7.80]	4.62 [3.40–6.26]
Lung cancer (N of VTE)	8	1,414	19	890	414	1,277	163	560	187
6	0.99 [0.37–2.63]	2.62 [2.43–2.82]	2.47 [1.25–4.89]	2.80 [2.55–3.06]	1.24 [1.07–1.44]	3.76 [3.50–4.04]	1.57 [1.27–1.96]	2.87 [2.57–3.20]	4.58 [3.82–5.49]
12	1.29 [0.54–3.09]	3.53 [3.31–3.76]	3.40 [1.90–6.07]	3.75 [3.47–4.06]	1.94 [1.72–2.20]	4.79 [4.50–5.10]	1.95 [1.60–2.37]	3.73 [3.39–4.11]	5.57 [4.73–6.56]
18	1.67 [0.74–3.74]	4.15 [3.92–4.40]	4.34 [2.60–7.25]	4.30 [4.00–4.63]	2.31 [2.05–2.60]	5.46 [5.15–5.80]	2.23 [1.86–2.68]	4.33 [3.96–4.74]	6.14 [5.26–7.18]
24	1.67 [0.74–3.74]	4.47 [4.22–4.72]	4.70 [2.86–7.70]	4.67 [4.36–5.01]	2.74 [2.45–3.07]	6.08 [5.74–6.44]	2.51 [2.10–2.99]	4.69 [4.30–5.11]	6.71 [5.79–7.79]
30	2.20 [1.02–4.74]	4.70 [4.45–4.96]	5.08 [3.15–8.20]	4.99 [4.66–5.34]	3.19 [2.87–3.55]	6.41 [6.06–6.78]	2.82 [2.39–3.33]	5.02 [4.62–5.46]	7.21 [6.24–8.32]
36	2.80 [1.35–5.84]	5.00 [4.74–5.27]	5.91 [3.76–9.28]	5.18 [4.84–5.53]	3.63 [3.27–4.04]	6.72 [6.36–7.10]	3.07 [2.62–3.61]	5.21 [4.79–5.66]	7.42 [6.44–8.56]
42	2.80 [1.35–5.84]	5.20 [4.94–5.48]	5.91 [3.76–9.28]	5.38 [5.04–5.74]	3.88 [3.50–4.31]	6.97 [6.60–7.36]	3.22 [2.75–3.77]	5.32 [4.90–5.78]	7.67 [6.66–8.84]
48	2.80 [1.35–5.84]	5.35 [5.08–5.63]	6.43 [4.14–10.01]	5.56 [5.21–5.94]	4.14 [3.73–4.60]	7.15 [6.78–7.55]	3.45 [2.95–4.03]	5.51 [5.08–5.98]	7.74 [6.72–8.91]
54	2.80 [1.35–5.84]	5.50 [5.23–5.79]	6.43 [4.14–10.01]	5.70 [5.35–6.09]	4.47 [4.02–4.97]	7.38 [6.99–7.79]	3.49 [2.99–4.08]	5.60 [5.16–6.08]	7.90 [6.86–9.09]

Time since first cancer diagnosis (months)	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
60	2.80 [1.35–5.84]	5.63 [5.35–5.93]	6.43 [4.14–10.01]	5.81 [5.44–6.19]	4.72 [4.24–5.26]	7.52 [7.12–7.94]	3.67 [3.15–4.29]	5.62 [5.18–6.10]	8.25 [7.17–9.49]
Lymphoma and Leukaemia (N of VTE)	<5	742	25	574	429	716	90	305	94
6	NA	1.01 [0.89–1.15]	0.82 [0.39–1.72]	0.99 [0.86–1.14]	0.41 [0.34–0.49]	1.04 [0.92–1.16]	1.39 [1.03–1.89]	1.34 [1.13–1.59]	1.55 [1.16–2.07]
12	NA	1.44 [1.29–1.59]	1.29 [0.72–2.33]	1.32 [1.17–1.49]	0.63 [0.54–0.74]	1.40 [1.27–1.54]	2.10 [1.64–2.70]	1.82 [1.57–2.11]	1.66 [1.25–2.19]
18	NA	1.72 [1.56–1.89]	1.53 [0.89–2.63]	1.60 [1.44–1.79]	0.87 [0.76–1.00]	1.69 [1.54–1.85]	2.36 [1.87–2.99]	2.16 [1.89–2.48]	2.07 [1.61–2.66]
24	NA	2.02 [1.85–2.20]	1.93 [1.19–3.14]	1.89 [1.71–2.09]	1.08 [0.95–1.22]	1.91 [1.75–2.08]	2.51 [2.00–3.16]	2.45 [2.16–2.79]	2.49 [1.99–3.13]
30	NA	2.25 [2.07–2.45]	2.08 [1.30–3.33]	2.12 [1.93–2.33]	1.30 [1.16–1.45]	2.15 [1.97–2.33]	2.59 [2.07–3.25]	2.65 [2.34–3.01]	2.76 [2.22–3.43]
36	NA	2.51 [2.32–2.71]	2.42 [1.54–3.78]	2.40 [2.19–2.62]	1.54 [1.38–1.71]	2.41 [2.23–2.61]	2.77 [2.22–3.44]	2.82 [2.50–3.18]	3.05 [2.48–3.76]
42	NA	2.75 [2.54–2.96]	2.79 [1.82–4.29]	2.62 [2.41–2.86]	1.69 [1.52–1.87]	2.64 [2.45–2.86]	2.96 [2.39–3.67]	3.01 [2.67–3.39]	3.23 [2.63–3.96]
48	NA	3.03 [2.81–3.26]	3.21 [2.12–4.85]	2.80 [2.57–3.05]	1.94 [1.75–2.15]	2.85 [2.64–3.07]	3.13 [2.53–3.87]	3.26 [2.90–3.66]	3.38 [2.76–4.13]
54	NA	3.20 [2.98–3.44]	3.44 [2.29–5.16]	3.02 [2.78–3.28]	2.15 [1.95–2.38]	3.05 [2.83–3.29]	3.33 [2.70–4.10]	3.50 [3.12–3.92]	3.38 [2.76–4.13]
60	NA	3.38 [3.15–3.63]	3.70 [2.48–5.53]	3.19 [2.94–3.46]	2.39 [2.16–2.64]	3.25 [3.01–3.50]	3.48 [2.82–4.29]	3.76 [3.36–4.22]	3.44 [2.81–4.20]

Time since first cancer diagnosis (months)	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
Melanoma (N of VTE)	0	260	6	148	181	133	35	87	23
6	NA	0.39 [0.31–0.50]	0 [0–0]	0.50 [0.37–0.67]	0.14 [0.10–0.21]	0.37 [0.26–0.53]	0.33 [0.18–0.60]	0.32 [0.20–0.52]	0.28 [0.12–0.68]
12	NA	0.59 [0.48–0.72]	0 [0–0]	0.69 [0.54–0.89]	0.25 [0.19–0.33]	0.62 [0.47–0.83]	0.46 [0.27–0.78]	0.52 [0.35–0.76]	0.56 [0.30–1.05]
18	NA	0.79 [0.67–0.94]	0 [0–0]	0.89 [0.71–1.11]	0.35 [0.28–0.45]	0.85 [0.67–1.09]	0.63 [0.41–0.99]	0.85 [0.63–1.16]	0.79 [0.47–1.33]
24	NA	0.95 [0.82–1.11]	0.66 [0.21–2.03]	1.00 [0.81–1.23]	0.52 [0.43–0.64]	1.08 [0.86–1.35]	0.74 [0.49–1.13]	1.06 [0.80–1.41]	0.91 [0.56–1.48]
30	NA	1.09 [0.94–1.26]	0.66 [0.21–2.03]	1.24 [1.02–1.49]	0.64 [0.54–0.77]	1.30 [1.06–1.60]	0.82 [0.55–1.22]	1.34 [1.04–1.72]	1.03 [0.65–1.64]
36	NA	1.22 [1.07–1.40]	0.66 [0.21–2.03]	1.37 [1.15–1.65]	0.76 [0.64–0.90]	1.42 [1.16–1.74]	0.90 [0.61–1.32]	1.56 [1.23–1.98]	1.10 [0.70–1.72]
42	NA	1.36 [1.19–1.56]	0.91 [0.34–2.43]	1.52 [1.28–1.82]	0.90 [0.76–1.06]	1.71 [1.41–2.07]	1.04 [0.73–1.51]	1.78 [1.42–2.23]	1.17 [0.76–1.82]
48	NA	1.51 [1.33–1.72]	1.23 [0.51–2.96]	1.79 [1.52–2.12]	1.00 [0.86–1.18]	1.96 [1.63–2.35]	1.21 [0.85–1.71]	1.89 [1.51–2.36]	1.17 [0.76–1.82]
54	NA	1.65 [1.46–1.87]	1.23 [0.51–2.96]	1.85 [1.57–2.18]	1.16 [1.00–1.35]	2.19 [1.83–2.62]	1.21 [0.85–1.71]	2.08 [1.68–2.59]	1.34 [0.88–2.04]
60	NA	1.73 [1.53–1.95]	1.63 [0.72–3.71]	1.93 [1.64–2.27]	1.27 [1.09–1.48]	2.42 [2.03–2.89]	1.42 [1.01–1.99]	2.27 [1.83–2.81]	1.44 [0.95–2.17]
Ovarian cancer (N of VTE)	0	213	12	199	117	181	22	148	54

Time since first cancer diagnosis (months)	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
6	NA	3.16 [2.56–3.89]	2.35 [0.99–5.58]	3.19 [2.61–3.90]	1.34 [1.00–1.79]	2.65 [2.14–3.28]	2.15 [1.16–3.97]	5.19 [4.20–6.42]	6.40 [4.64–8.82]
12	NA	4.81 [4.07–5.68]	2.82 [1.28–6.20]	4.22 [3.55–5.02]	2.02 [1.58–2.57]	3.72 [3.11–4.45]	3.50 [2.16–5.66]	6.94 [5.78–8.34]	7.50 [5.58–10.06]
18	NA	5.87 [5.05–6.83]	3.80 [1.93–7.51]	4.60 [3.90–5.43]	2.43 [1.94–3.04]	4.31 [3.64–5.10]	4.44 [2.89–6.81]	7.72 [6.49–9.18]	7.68 [5.74–10.27]
24	NA	6.24 [5.40–7.22]	4.34 [2.29–8.22]	5.19 [4.44–6.06]	2.72 [2.19–3.37]	5.02 [4.29–5.88]	4.68 [3.08–7.11]	8.24 [6.96–9.75]	8.63 [6.56–11.34]
30	NA	6.67 [5.79–7.68]	4.92 [2.68–9.02]	5.41 [4.64–6.30]	3.30 [2.70–4.03]	5.38 [4.61–6.27]	4.68 [3.08–7.11]	8.70 [7.39–10.26]	9.22 [7.08–12.00]
36	NA	7.13 [6.22–8.17]	5.52 [3.10–9.84]	5.73 [4.94–6.64]	3.50 [2.87–4.25]	5.72 [4.92–6.65]	4.95 [3.29–7.44]	9.22 [7.85–10.82]	9.22 [7.08–12.00]
42	NA	7.31 [6.39–8.37]	5.52 [3.10–9.84]	6.24 [5.41–7.20]	3.80 [3.13–4.60]	6.29 [5.43–7.28]	4.95 [3.29–7.44]	9.61 [8.20–11.26]	9.45 [7.27–12.27]
48	NA	7.56 [6.62–8.64]	5.52 [3.10–9.84]	6.58 [5.72–7.57]	4.09 [3.37–4.95]	6.29 [5.43–7.28]	4.95 [3.29–7.44]	9.96 [8.51–11.65]	9.96 [7.71–12.88]
54	NA	8.05 [7.07–9.18]	5.52 [3.10–9.84]	6.74 [5.87–7.76]	4.58 [3.79–5.55]	6.73 [5.81–7.79]	4.95 [3.29–7.44]	10.08 [8.62–11.80]	10.27 [7.96–13.26]
60	NA	8.23 [7.23–9.37]	6.59 [3.72–11.70]	7.47 [6.52–8.56]	4.79 [3.95–5.80]	6.90 [5.96–8.00]	4.95 [3.29–7.44]	10.53 [9.00–12.32]	10.27 [7.96–13.26]
Pancreatic cancer (N of VTE)	0	440	15	499	149	380	52	149	93
6	NA	4.15 [3.68–4.68]	4.95 [2.62–9.35]	3.99 [3.58–4.45]	1.89 [1.52–2.35]	3.54 [3.11–4.03]	2.91 [2.10–4.03]	4.65 [3.84–5.64]	8.22 [6.55–10.33]

Time since first cancer diagnosis (months)	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
12	NA	5.34 [4.80–5.93]	6.04 [3.41–10.72]	5.10 [4.63–5.62]	2.85 [2.36–3.44]	4.54 [4.06–5.09]	3.88 [2.92–5.15]	5.84 [4.91–6.94]	9.92 [8.07–12.18]
18	NA	6.05 [5.48–6.68]	7.20 [4.27–12.16]	5.56 [5.07–6.09]	3.23 [2.69–3.87]	5.12 [4.60–5.70]	4.24 [3.23–5.56]	6.60 [5.61–7.77]	10.40 [8.51–12.70]
24	NA	6.46 [5.87–7.11]	7.20 [4.27–12.16]	5.90 [5.39–6.45]	3.58 [3.00–4.29]	5.77 [5.21–6.39]	4.34 [3.32–5.68]	6.77 [5.77–7.95]	10.66 [8.75–12.99]
30	NA	6.74 [6.14–7.40]	7.91 [4.78–13.10]	6.18 [5.66–6.74]	4.31 [3.62–5.14]	6.03 [5.46–6.67]	4.34 [3.32–5.68]	6.89 [5.87–8.09]	11.14 [9.18–13.53]
36	NA	6.92 [6.32–7.59]	8.69 [5.34–14.15]	6.33 [5.81–6.90]	4.53 [3.80–5.40]	6.11 [5.53–6.75]	4.34 [3.32–5.68]	6.96 [5.93–8.16]	11.32 [9.33–13.74]
42	NA	7.07 [6.45–7.75]	8.69 [5.34–14.15]	6.45 [5.92–7.03]	4.89 [4.09–5.84]	6.30 [5.71–6.96]	4.46 [3.41–5.82]	7.03 [6.00–8.24]	11.32 [9.33–13.74]
48	NA	7.12 [6.50–7.79]	8.69 [5.34–14.15]	6.53 [6.00–7.12]	5.21 [4.35–6.23]	6.36 [5.76–7.03]	4.46 [3.41–5.82]	7.20 [6.14–8.43]	11.32 [9.33–13.74]
54	NA	7.30 [6.67–7.99]	8.69 [5.34–14.15]	6.59 [6.06–7.18]	5.21 [4.35–6.23]	6.48 [5.86–7.15]	4.46 [3.41–5.82]	7.56 [6.46–8.84]	11.32 [9.33–13.74]
60	NA	7.34 [6.70–8.03]	8.69 [5.34–14.15]	6.64 [6.10–7.23]	5.36 [4.46–6.44]	6.56 [5.94–7.25]	4.46 [3.41–5.82]	7.56 [6.46–8.84]	11.63 [9.57–14.12]
Prostate cancer (N of VTE)	8	742	34	764	568	465	110	342	128
6	0.12 [0.02–0.84]	0.61 [0.53–0.71]	0.87 [0.46–1.67]	0.50 [0.43–0.58]	0.36 [0.31–0.42]	0.45 [0.37–0.55]	0.59 [0.40–0.86]	0.51 [0.41–0.64]	0.46 [0.33–0.66]
12	0.25 [0.06–0.99]	0.97 [0.86–1.09]	1.16 [0.66–2.04]	0.71 [0.63–0.80]	0.55 [0.48–0.62]	0.77 [0.66–0.90]	1.03 [0.77–1.38]	0.92 [0.78–1.09]	0.67 [0.50–0.89]

Time since first cancer diagnosis (months)	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
18	0.54 [0.20–1.44]	1.22 [1.10–1.35]	1.47 [0.89–2.42]	0.93 [0.83–1.03]	0.72 [0.64–0.81]	1.03 [0.90–1.17]	1.35 [1.04–1.74]	1.26 [1.09–1.46]	0.89 [0.69–1.14]
24	0.86 [0.39–1.93]	1.47 [1.33–1.61]	2.13 [1.39–3.26]	1.15 [1.05–1.27]	0.84 [0.76–0.94]	1.28 [1.13–1.44]	1.65 [1.31–2.09]	1.49 [1.31–1.71]	1.06 [0.85–1.34]
30	0.86 [0.39–1.93]	1.74 [1.60–1.90]	2.63 [1.78–3.88]	1.39 [1.28–1.52]	0.98 [0.89–1.08]	1.55 [1.38–1.73]	1.89 [1.52–2.35]	1.68 [1.48–1.92]	1.22 [0.99–1.52]
36	1.06 [0.50–2.25]	2.02 [1.86–2.19]	3.03 [2.10–4.38]	1.59 [1.46–1.73]	1.18 [1.08–1.29]	1.79 [1.61–1.99]	2.01 [1.62–2.48]	1.96 [1.74–2.22]	1.43 [1.17–1.75]
42	1.06 [0.50–2.25]	2.30 [2.12–2.48]	3.18 [2.21–4.57]	1.80 [1.66–1.94]	1.33 [1.21–1.45]	2.09 [1.89–2.31]	2.23 [1.82–2.74]	2.22 [1.98–2.49]	1.69 [1.40–2.04]
48	1.32 [0.65–2.68]	2.52 [2.34–2.72]	3.34 [2.34–4.77]	2.02 [1.87–2.18]	1.43 [1.31–1.56]	2.37 [2.15–2.61]	2.44 [2.00–2.98]	2.49 [2.22–2.78]	1.88 [1.57–2.25]
54	1.32 [0.65–2.68]	2.69 [2.50–2.90]	3.54 [2.49–5.05]	2.24 [2.08–2.40]	1.56 [1.43–1.70]	2.57 [2.34–2.83]	2.69 [2.22–3.28]	2.74 [2.46–3.06]	2.03 [1.71–2.42]
60	1.32 [0.65–2.68]	2.90 [2.70–3.12]	4.20 [2.98–5.94]	2.45 [2.28–2.63]	1.74 [1.60–1.89]	2.83 [2.58–3.11]	3.08 [2.54–3.72]	2.93 [2.63–3.27]	2.11 [1.77–2.51]
Stomach cancer (N of VTE)	0	207	15	176	80	258	18	61	42
6	NA	3.06 [2.57–3.65]	3.33 [1.52–7.32]	3.40 [2.83–4.10]	1.13 [0.83–1.56]	3.18 [2.72–3.71]	1.39 [0.73–2.66]	3.69 [2.68–5.08]	5.18 [3.51–7.65]
12	NA	3.89 [3.34–4.54]	4.44 [2.26–8.75]	4.06 [3.43–4.81]	1.64 [1.25–2.15]	4.21 [3.68–4.82]	2.35 [1.43–3.88]	5.30 [4.06–6.93]	6.48 [4.58–9.16]
18	NA	4.32 [3.73–5.00]	6.76 [3.92–11.69]	4.63 [3.96–5.42]	1.93 [1.50–2.50]	4.63 [4.07–5.27]	2.52 [1.55–4.09]	5.87 [4.56–7.58]	7.99 [5.87–10.89]

Time since first cancer diagnosis (months)	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
24	NA	4.67 [4.06–5.38]	8.02 [4.85–13.26]	4.95 [4.25–5.76]	2.12 [1.66–2.72]	5.01 [4.42–5.67]	2.70 [1.69–4.31]	6.23 [4.87–7.98]	8.44 [6.25–11.40]
30	NA	4.92 [4.29–5.64]	8.72 [5.37–14.17]	5.18 [4.46–6.01]	2.33 [1.83–2.97]	5.14 [4.55–5.82]	2.70 [1.69–4.31]	6.36 [4.97–8.12]	8.44 [6.25–11.40]
36	NA	5.01 [4.37–5.73]	8.72 [5.37–14.17]	5.32 [4.59–6.16]	2.45 [1.93–3.11]	5.29 [4.68–5.98]	2.70 [1.69–4.31]	6.36 [4.97–8.12]	8.44 [6.25–11.40]
42	NA	5.10 [4.46–5.84]	8.72 [5.37–14.17]	5.43 [4.70–6.29]	2.60 [2.04–3.30]	5.52 [4.89–6.23]	2.94 [1.86–4.66]	6.36 [4.97–8.12]	9.05 [6.75–12.13]
48	NA	5.14 [4.49–5.88]	8.72 [5.37–14.17]	5.56 [4.81–6.42]	3.22 [2.53–4.09]	5.59 [4.95–6.31]	2.94 [1.86–4.66]	6.53 [5.12–8.33]	9.36 [7.01–12.51]
54	NA	5.26 [4.60–6.02]	8.72 [5.37–14.17]	5.60 [4.85–6.47]	3.33 [2.62–4.23]	5.63 [4.99–6.35]	2.94 [1.86–4.66]	6.53 [5.12–8.33]	9.36 [7.01–12.51]
60	NA	5.35 [4.68–6.12]	8.72 [5.37–14.17]	5.65 [4.90–6.53]	3.44 [2.71–4.38]	5.78 [5.12–6.52]	2.94 [1.86–4.66]	6.53 [5.12–8.33]	9.36 [7.01–12.51]

N=total number of splanchnic vein thrombosis (SVT) events in liver cancer and the total number of events of pulmonary embolism and deep vein thrombosis combined (VTE) in each of the other types of cancer. NA=not applicable (due to 0 or <5 event counts); NC = not captured. IQVIA LPD=IQVIA Longitudinal Patient Database Belgium (IQVIA LPD Belgium); DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; IPCI=Integrated Primary Care Information; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank.

9.2.2. Cumulative incidence of thromboembolic events five years after the cancer diagnosis by sex, age, and study sub-period

The full set of tables and figures for this study stratified by sex, age, and study sub-period is available through an interactive web-application ShinyApp at EUPAS1000000814.

The cumulative incidence estimates of thromboembolic events after the first cancer diagnosis had overlapping 95% CIs between females and males. However, differences in the point estimates of the cumulative incidence of SVT in liver cancer were observed between females and males across data sources. The cumulative incidence of SVT was lower among females than males from six months to five years after the diagnosis of liver cancer in each of the five data sources with at least five event counts, except for IQVIA DA Germany. The 6-month cumulative incidence of SVT after the first diagnosis of liver cancer among females ranged from 0.21% [95% CI: 0.07%–0.66%] in FinOMOP-THL to 2.86% [95% CI: 2.16%–3.80%] in SIDIAP, versus 0.79% [95% CI: 0.52%–1.21%] in FinOMOP-THL and 4.56% [95% CI: 3.92%–5.32%] in SIDIAP among males. The 5-year cumulative incidence was 0.72% [95% CI: 0.39%–1.34%] among females versus 1.40% [95% CI: 1.01%–1.92%] among males in FinOMOP-THL and 5.44% [95% CI: 4.39%–6.74%] among females versus 11.36% [95% CI: 10.22%–12.63%] among males in SIDIAP. The cumulative incidence estimates of thromboembolic events after the cancer diagnosis had overlapping 95% CIs between age groups, but some differences in the point estimates were observed across data sources.

Although no clear pattern was observed in the 6-month cumulative incidence of VTE with age after the first diagnosis of liver cancer, the 5-year cumulative incidence was lower in each consecutive age group in each of the six data sources, except for IQVIA DA Germany, with at least five event counts in at least three consecutive age groups. The 6-month cumulative incidence of VTE after the first diagnosis of liver cancer in the age group 55–64 years ranged from 0.85% [95% CI: 0.32%–2.25%] in IQVIA DA Germany to 7.81% [95% CI: 3.37%–18.13%] in UKBB versus 1.62% [95% CI: 0.87%–3.02%] in IQVIA DA Germany and 2.74% [95% CI: 1.04%–7.20%] in UKBB in the age group 75–84 years. The 5-year cumulative incidence was 4.24% [95% CI: 2.21%–8.12%] in the age group 55–64 years versus 5.66% [95% CI: 2.84%–11.31%] in the age group 75–84 years in IQVIA DA Germany and 9.42% [95% CI: 4.40%–20.17%] in the age group 55–64 years versus 4.11% [95% CI: 1.88%–9.00%] in the age group 75–84 years in UKBB.

The cumulative incidence of VTE from six months to five years after the first diagnosis of pancreatic cancer was lower in each consecutive age group from 45–54 years onwards in each of the seven data sources, except for IQVIA DA Germany, with at least five event counts in at least three consecutive age groups. The 6-month cumulative incidence of VTE after the first diagnosis of pancreatic cancer in the age group 55–64 years ranged from 1.51% [95% CI: 0.88%–2.59%] in IQVIA DA Germany to 10.08% [95% CI: 6.02%–16.87%] in UKBB versus 1.77% [95% CI: 1.20%–2.62%] in IQVIA DA Germany and 7.07% [95% CI: 4.68%–10.68%] in UKBB in the age group 75–84 years. The 5-year cumulative incidence was 4.77% [95% CI: 3.28%–6.94%] in the age group 55–64 years versus 6.33% [95% CI: 4.56%–8.80%] in the age group 75–84 years in IQVIA DA Germany and 12.47% [95% CI: 7.88%–19.73%] in the age group 55–64 years versus 10.45% [95% CI: 7.13%–15.34%] in the age group 75–84 years in UKBB.

Although no clear pattern was observed in the 6-month cumulative incidence of PE with age after the first diagnosis of liver cancer, the 5-year cumulative incidence was lower in each consecutive age group in each of the five data sources, except for IQVIA DA Germany, with at least five event counts in at least three consecutive age groups. The 6-month cumulative incidence of PE after the first diagnosis of liver cancer in the age group 55–64 years ranged from 0.48% [95% CI: 0.21%–1.06%] in SIDIAP to 7.58% [95% CI: 3.26%–17.60%] in UKBB versus 0.58% [95% CI: 0.29%–1.15%] in SIDIAP and 2.03% [95% CI: 0.66%–6.21%] in UKBB in the age group 75–84 years. The 5-year cumulative incidence was 2.59% [95% CI: 1.74%–3.84%] in the age group 55–64 years versus 1.24% [95% CI: 0.74%–2.09%] in the age group 75–84 years in SIDIAP and 9.13% [95% CI: 4.26%–19.58%] in the age group 55–64 years versus 3.38% [95% CI: 1.43%–8.00%] in the age group 75–84 years in UKBB.

The results were generally similar between the two study sub-periods across outcomes and cancer types. For example, in both study sub-periods, the highest 5-year cumulative incidences of VTE after the first cancer diagnosis were observed in pancreatic and ovarian cancers, while the lowest was observed in melanoma.

9.2.3. Median time to thromboembolic event after the first cancer diagnosis

Across thromboembolic outcomes, individuals with pancreatic cancer and liver cancer consistently showed the shortest time from cancer diagnosis to thromboembolic event (**Table 4**).

In individuals with pancreatic cancer, the median [IQR] times to thromboembolic event were:

- VTE (ranging from 42 days [0–116] in UKBB to 134 days [24–260] in IPCI);
- DVT (ranging from 40 days [14–104] in UKBB to 218 days [123–482] in EBB);
- PE (ranging from 23 days [1–131] in EBB to 160 days [46–331] in IPCI);
- SVT (ranging from 3 days [0–126] in CPRD GOLD to 186 days [58–444] in DK-DHR);
- RVT (1 days [0–494] in SIDIAP, 248 days [228–492] in DK-DHR, and 321 days [188–844] in FinOMOP-THL);
- DIC (50 days [10–70] in SIDIAP.

In individuals with liver cancer, the median [IQR] times to thromboembolic event were :

- VTE (ranging from 55 days [1–210] in CPRD GOLD to 184 days [106–331] in IPCI);
- DVT (ranging from 49 days [0–312] in SIDIAP to 233 days [67–929] in IQVIA DA Germany);
- PE (ranging from 40 days [0–170] in CPRD GOLD to 213 days [122–248] in IPCI);
- SVT (ranging from 6 days [0–238] in SIDIAP to 66 days [4–301] in DK-DHR).

In general, median times to event for several cancers varied depending on the outcome and data source. Lung cancer showed relatively short median [IQR] event times for PVT (166 days [80–425] in FinOMOP-THL and 227 days [112–496] in SIDIAP) and moderate times for RVT (ranging from 411 days [226–833] in IQVIA DA Germany to 803 days [432–1,386] in DK-DHR).

Ovarian cancer showed short median [IQR] event time for PVT (82 days [44–196] in FinOMOP-THL). The median days [IQR] event times for DVT were similar between brain (ranging from 106 days [100–302] in IPCI to 251 days [147–460] in DK-DHR), oesophageal (ranging from 93 days [56–448] in FinOMOP-THL to 258 days [91–428] in IPCI), and liver cancers (ranging from 49 days [0–312] in SIDIAP to 233 days [67–929] in IQVIA DA Germany).

Colorectal cancer showed short event time for DIC (median 51 days in SIDIAP), while kidney cancer showed longer median [IQR] event time for RVT (ranging from 663 days [474–1,629] in CPRD GOLD to 1,111 days [642–1,619] in FinOMOP-THL).

In contrast, melanoma consistently showed the longest median [IQR] event times across almost all outcomes and data sources, including:

- VTE (ranging from 435 days [246–1,032] in UKBB to 919 days [696–1,366] in EBB);
- DVT (ranging from 550 days [217–1,180] in SIDIAP to 875 days [229–1,329] in FinOMOP-THL);
- PE (ranging from 432 days [294–1,516] in UKBB to 1,135 days [703–1,443] in EBB);
- SVT (869 days [692–1,253] in DK-DHR and 968 days [290–1,483] SIDIAP);
- PVT (748 days [480–1,293] in FinOMOP-THL and 919 days [528–1,301] in SIDIAP).

Table 4. Median time in days [IQR] to thromboembolic event after first cancer diagnosis.

	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
Bone cancer (N)	0	16	0	79	27	85	NC	15	0
Deep vein thrombosis (DVT)	NA	NA	NA	296 [53–630]	648 [258–1,071]	257 [43–837]	NC	92 [14–206]	NA
PE and DVT combined (VTE)	NA	315 [124–934]	NA	110 [21–446]	434 [55–1,071]	226 [36–830]	NC	47 [5–226]	NA
Pelvic vein thrombosis (PVT)	NA	NA	NA	NA	NA	NA	NC	NA	NA
Pulmonary embolism (PE)	NA	167 [117–530]	NA	104 [22–418]	393 [33–682]	105 [6–805]	NC	NA	NA
Retinal vein thrombosis (RVT)	NA	NA	NA	NA	NA	NA	NC	NA	NA
Splanchnic extrahepatic vein thrombosis (SVT)	NA	NA	NA	NA	NA	NA	NC	NA	NA
Disseminated intravascular coagulation (DIC)	NA	NA	NA	NA	NA	NA	NC	NA	NA
Brain cancer (N)	0	345	14	521	118	531	36	121	73

	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
Deep vein thrombosis (DVT)	NA	251 [147–460]	NA	112 [58–289]	190 [75–887]	176 [67–341]	106 [100–302]	162 [87–361]	108 [35–278]
PE and DVT combined (VTE)	NA	214 [84–461]	68 [39–291]	98 [54–290]	128 [34–494]	142 [62–330]	218 [98–422]	146 [78–302]	122 [46–294]
Pelvic vein thrombosis (PVT)	NA	NA	NA	170 [122–254]	NA	238 [134–723]	NA	NA	NA
Pulmonary embolism (PE)	NA	190 [71–437]	105 [46–302]	93 [49–295]	79 [0–183]	130 [55–322]	316 [120–412]	142 [88–242]	122 [29–280]
Retinal vein thrombosis (RVT)	NA	NA	NA	NA	NA	NA	NA	NA	NA
Splanchnic extrahepatic vein thrombosis (SVT)	NA	NA	NA	NA	NA	NA	NA	NA	NA
Disseminated intravascular coagulation (DIC)	NA	NA	NA	NA	NA	NA	NA	NA	NA
Breast cancer (N)	14	1,495	54	2,207	1,413	1,571	272	1,180	214
Deep vein thrombosis (DVT)	NA	756 [263–1,448]	292 [128–964]	508 [148–1,217]	531 [93–1,113]	369 [125–974]	531 [194–1,220]	261 [112–910]	264 [134–928]
PE and DVT combined (VTE)	14 [0–238]	732 [240–1,369]	279 [164–1,077]	442 [134–1,175]	448 [36–1,113]	348 [113–956]	614 [193–1,058]	345 [131–1,006]	299 [133–884]

	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
Pelvic vein thrombosis (PVT)	NA	NA	1,146 [481–1,958]	512 [164–1,296]	NA	741 [232–1,301]	NA	NA	NA
Pulmonary embolism (PE)	14 [0–238]	716 [217–1,336]	312 [164–949]	420 [118–1,163]	391 [1–1,094]	344 [102–879]	620 [190–1,030]	464 [148–1,130]	310 [114–843]
Retinal vein thrombosis (RVT)	NA	857 [508–1,530]	NA	1,172 [527–1,768]	833 [316–1,361]	508 [230–804]	NA	894 [377–1,416]	578 [81–1,136]
Splanchnic extrahepatic vein thrombosis (SVT)	NA	759 [366–1,742]	NA	778 [557–1,426]	475 [405–998]	482 [139–1,159]	NA	1,380 [868–1,877]	NA
Disseminated intravascular coagulation (DIC)	NA	NA	NA	NA	NA	351 [87–978]	NA	NA	NA
Colorectal cancer (N)	0	2,838	87	2,754	1,058	2,621	333	1,440	486
Deep vein thrombosis (DVT)	NA	462 [150–1,028]	615 [134–795]	308 [122–798]	324 [67–993]	251 [75–667]	416 [155–804]	212 [100–580]	230 [85–544]
PE and DVT combined (VTE)	NA	323 [91–857]	230 [55–674]	211 [67–631]	216 [14–837]	226 [56–641]	327 [131–804]	176 [82–489]	180 [50–424]
Pelvic vein thrombosis (PVT)	NA	908 [292–1,698]	135 [34–896]	282 [77–705]	NA	265 [20–519]	NA	NA	NA
Pulmonary embolism (PE)	NA	296 [83–828]	164 [18–570]	183 [52–576]	176 [0–750]	190 [24–628]	307 [104–906]	167 [69–470]	148 [42–340]

	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
Retinal vein thrombosis (RVT)	NA	935 [459–1,572]	NA	770 [378–1,390]	794 [310–1,351]	845 [396–1,348]	NA	536 [224–1,602]	325 [182–1,085]
Splanchnic extrahepatic vein thrombosis (SVT)	NA	188 [34–453]	343 [218–1,472]	160 [54–470]	454 [57–955]	300 [49–732]	NA	208 [99–744]	346 [68–878]
Disseminated intravascular coagulation (DIC)	NA	NA	NA	NA	NA	51 [0–269]	NA	NA	NA
Corpus uteri cancer (N)	0	420	15	627	189	506	38	225	102
Deep vein thrombosis (DVT)	NA	414 [157–885]	226 [26–818]	602 [176–1,103]	437 [36–1,088]	228 [68–643]	610 [156–1,086]	430 [177–935]	733 [596–1,152]
PE and DVT combined (VTE)	NA	376 [130–916]	45 [22–250]	358 [64–982]	373 [34–1,045]	243 [66–694]	612 [87–1,290]	361 [100–825]	408 [170–860]
Pelvic vein thrombosis (PVT)	NA	NA	NA	616 [161–1,075]	NA	416 [88–848]	NA	NA	NA
Pulmonary embolism (PE)	NA	365 [126–960]	NA	209 [42–810]	338 [6–1,050]	245 [56–685]	615 [28–1,337]	208 [65–778]	355 [142–559]
Retinal vein thrombosis (RVT)	NA	1,039 [725–1,957]	NA	1,223 [389–1,575]	NA	674 [314–757]	NA	285 [259–305]	NA
Splanchnic extrahepatic vein	NA	NA	NA	360 [46–1,379]	NA	446 [428–726]	NA	NA	NA

	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
thrombosis (SVT)									
Disseminated intravascular coagulation (DIC)	NA	NA	NA	NA	NA	NA	NA	NA	NA
Oesophageal cancer (N)	<5	264	5	264	149	192	61	456	120
Deep vein thrombosis (DVT)	NA	175 [77–695]	NA	93 [56–448]	180 [78–904]	106 [39–298]	258 [91–428]	145 [78–288]	188 [46–317]
PE and DVT combined (VTE)	NA	149 [77–419]	146 [105–256]	118 [51–436]	148 [57–862]	101 [18–288]	179 [106–274]	132 [78–252]	112 [54–264]
Pelvic vein thrombosis (PVT)	NA	NA	NA	NA	NA	NA	NA	NA	NA
Pulmonary embolism (PE)	NA	148 [76–418]	NA	126 [50–438]	142 [48–821]	86 [0–225]	172 [109–223]	126 [82–208]	108 [65–187]
Retinal vein thrombosis (RVT)	NA	870 [548–1,110]	NA	NA	NA	NA	NA	NA	NA
Splanchnic extrahepatic vein thrombosis (SVT)	NA	NA	NA	NA	NA	255 [55–427]	NA	NA	NA
Disseminated intravascular coagulation (DIC)	NA	NA	NA	NA	NA	418 [140–541]	NA	NA	NA

	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
Kidney cancer (N)	0	588	22	604	303	754	58	117	86
Deep vein thrombosis (DVT)	NA	547 [89–1,176]	364 [98–816]	361 [109–912]	510 [44–1,073]	61 [0–394]	959 [204–1,319]	195 [28–536]	0 [0–37]
PE and DVT combined (VTE)	NA	456 [98–1,051]	123 [23–493]	264 [62–895]	186 [0–867]	86 [0–532]	356 [107–1,300]	193 [42–749]	36 [0–181]
Pelvic vein thrombosis (PVT)	NA	NA	NA	321 [116–1,008]	NA	362 [76–637]	NA	NA	NA
Pulmonary embolism (PE)	NA	478 [111–1,069]	62 [42–308]	255 [61–872]	83 [0–551]	182 [0–697]	320 [104–1,042]	219 [39–882]	131 [58–318]
Retinal vein thrombosis (RVT)	NA	1,061 [714–1,426]	NA	1,111 [642–1,619]	1,014 [597–1,615]	740 [459–1,116]	NA	663 [474–1,629]	NA
Splanchnic extrahepatic vein thrombosis (SVT)	NA	212 [188–680]	NA	874 [376–1,509]	NA	59 [0–440]	NA	NA	NA
Disseminated intravascular coagulation (DIC)	NA	NA	NA	NA	NA	85 [0–545]	NA	NA	NA
Liver cancer (N)	0	297	11	328	145	747	24	78	78
Deep vein thrombosis (DVT)	NA	203 [50–599]	NA	150 [20–404]	233 [67–929]	49 [0–312]	154 [80–544]	146 [56–363]	124 [80–204]

	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
PE and DVT combined (VTE)	NA	124 [20–393]	65 [24–85]	68 [2–247]	162 [21–353]	72 [0–386]	184 [106–331]	55 [1–210]	96 [47–170]
Pelvic vein thrombosis (PVT)	NA	NA	NA	NA	NA	NA	NA	NA	NA
Pulmonary embolism (PE)	NA	82 [19–330]	NA	48 [0–193]	119 [0–333]	63 [0–384]	213 [122–248]	40 [0–170]	92 [47–210]
Retinal vein thrombosis (RVT)	NA	770 [439–1,242]	NA	580 [316–784]	NA	NA	NA	NA	NA
Splanchnic extrahepatic vein thrombosis (SVT)	NA	66 [4–301]	52 [2–136]	9 [0–216]	15 [0–253]	6 [0–238]	NA	16 [0–240]	49 [0–361]
Disseminated intravascular coagulation (DIC)	NA	NA	NA	NA	NA	240 [26–532]	NA	NA	NA
Lung cancer (N)	16	2,955	40	1,930	877	2,833	335	1,163	396
Deep vein thrombosis (DVT)	NA	218 [77–572]	390 [44–729]	194 [50–497]	238 [40–787]	80 [9–307]	238 [72–797]	102 [38–380]	172 [42–492]
PE and DVT combined (VTE)	100 [15–337]	188 [64–529]	163 [15–421]	114 [20–420]	140 [4–597]	81 [2–302]	182 [64–736]	118 [31–376]	83 [18–324]
Pelvic vein thrombosis (PVT)	NA	NA	NA	166 [80–425]	NA	227 [112–496]	NA	NA	NA

	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
Pulmonary embolism (PE)	100 [15–337]	186 [62–525]	110 [9–358]	109 [14–404]	121 [0–551]	81 [0–325]	160 [49–720]	138 [25–404]	80 [15–321]
Retinal vein thrombosis (RVT)	NA	803 [432–1,386]	NA	464 [256–872]	411 [226–833]	446 [121–1,136]	NA	704 [399–1,316]	NA
Splanchnic extrahepatic vein thrombosis (SVT)	NA	564 [445–1,196]	NA	209 [32–341]	265 [51–434]	89 [0–602]	NA	NA	NA
Disseminated intravascular coagulation (DIC)	NA	49 [20–283]	NA	NA	NA	106 [16–264]	NA	NA	NA
Lymphoma and Leukaemia (N)	<5	1,684	61	1,374	949	1,714	182	658	213
Deep vein thrombosis (DVT)	NA	680 [196–1,296]	255 [110–690]	565 [120–1,150]	444 [60–1,011]	188 [25–734]	224 [109–892]	247 [57–841]	296 [26–774]
PE and DVT combined (VTE)	NA	487 [121–1,146]	422 [102–1,058]	461 [96–1,057]	414 [22–1,071]	165 [9–735]	214 [76–747]	258 [64–914]	159 [17–765]
Pelvic vein thrombosis (PVT)	NA	588 [200–905]	532 [400–688]	432 [116–976]	NA	447 [116–1,236]	NA	NA	NA
Pulmonary embolism (PE)	NA	398 [90–1,078]	478 [142–1,062]	434 [96–990]	392 [0–1,202]	142 [0–782]	207 [46–458]	290 [78–1,065]	144 [16–675]
Retinal vein thrombosis (RVT)	NA	964 [258–1,596]	NA	706 [306–1,128]	557 [188–1,166]	677 [36–1,257]	NA	461 [100–1,184]	888 [244–1,263]

	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
Splanchnic extrahepatic vein thrombosis (SVT)	NA	514 [104–1,042]	NA	216 [62–861]	384 [0–889]	218 [0–858]	NA	NA	429 [112–893]
Disseminated intravascular coagulation (DIC)	NA	221 [30–1,056]	NA	360 [96–531]	NA	12 [0–281]	NA	NA	NA
Melanoma (N)	0	580	11	350	397	319	70	193	47
Deep vein thrombosis (DVT)	NA	771 [282–1,412]	NA	875 [229–1,329]	723 [198–1,467]	550 [217–1,180]	783 [65–1,678]	755 [364–1,190]	554 [198–778]
PE and DVT combined (VTE)	NA	632 [202–1,290]	919 [696–1,366]	643 [147–1,294]	684 [195–1,301]	569 [192–1,184]	694 [180–1,496]	689 [354–1,185]	435 [246–1,032]
Pelvic vein thrombosis (PVT)	NA	NA	NA	748 [480–1,293]	NA	919 [528–1,301]	NA	NA	NA
Pulmonary embolism (PE)	NA	570 [157–1,244]	1,135 [703–1,443]	529 [114–1,260]	647 [130–1,205]	547 [171–1,192]	694 [214–1,507]	742 [360–1,185]	432 [294–1,516]
Retinal vein thrombosis (RVT)	NA	898 [450–1,737]	NA	1,122 [712–1,684]	700 [187–1,142]	710 [294–1,478]	NA	964 [394–1,686]	NA
Splanchnic extrahepatic vein thrombosis (SVT)	NA	869 [692–1,253]	NA	NA	NA	968 [290–1,483]	NA	NA	NA
Disseminated intravascular	NA	NA	NA	NA	NA	NA	NA	NA	NA

	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
coagulation (DIC)									
Ovarian cancer (N)	0	456	26	430	241	404	44	317	116
Deep vein thrombosis (DVT)	NA	380 [189–823]	411 [207–626]	396 [66–1,061]	272 [55–877]	183 [26–665]	261 [108–390]	161 [41–483]	80 [5–292]
PE and DVT combined (VTE)	NA	250 [72–608]	281 [18–758]	113 [9–670]	148 [0–705]	108 [5–512]	216 [76–380]	149 [39–378]	77 [18–300]
Pelvic vein thrombosis (PVT)	NA	NA	NA	82 [44–196]	NA	NA	NA	NA	NA
Pulmonary embolism (PE)	NA	195 [65–484]	128 [14–621]	71 [7–446]	113 [0–431]	76 [0–412]	172 [73–363]	147 [42–466]	74 [5–368]
Retinal vein thrombosis (RVT)	NA	596 [148–975]	NA	NA	NA	NA	NA	672 [670–702]	NA
Splanchnic extrahepatic vein thrombosis (SVT)	NA	164 [82–1,241]	NA	NA	NA	110 [4–659]	NA	NA	NA
Disseminated intravascular coagulation (DIC)	NA	NA	NA	NA	NA	NA	NA	NA	NA
Pancreatic cancer (N)	0	965	40	1,150	334	1,066	107	318	220

	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
Deep vein thrombosis (DVT)	NA	147 [42–386]	218 [123–482]	129 [33–486]	136 [24–315]	76 [3–317]	76 [26–165]	104 [34–221]	40 [14–104]
PE and DVT combined (VTE)	NA	102 [31–344]	111 [18–341]	67 [5–225]	94 [3–284]	54 [0–221]	134 [24–260]	66 [6–198]	42 [0–116]
Pelvic vein thrombosis (PVT)	NA	NA	NA	109 [20–290]	NA	306 [78–875]	NA	NA	NA
Pulmonary embolism (PE)	NA	94 [26–334]	23 [1–131]	51 [0–190]	84 [0–260]	27 [0–178]	160 [46–331]	40 [0–140]	45 [0–149]
Retinal vein thrombosis (RVT)	NA	248 [228–492]	NA	321 [188–844]	NA	1 [0–494]	NA	NA	NA
Splanchnic extrahepatic vein thrombosis (SVT)	NA	186 [58–444]	96 [16–190]	104 [16–284]	66 [0–391]	41 [0–283]	NA	3 [0–126]	39 [1–216]
Disseminated intravascular coagulation (DIC)	NA	NA	NA	NA	NA	50 [10–70]	NA	NA	NA
Prostate cancer (N)	16	1,678	84	1,910	1,237	1,085	225	773	274
Deep vein thrombosis (DVT)	NA	692 [253–1,246]	758 [400–1,645]	748 [240–1,398]	463 [97–1,174]	559 [164–1,140]	625 [276–1,214]	645 [243–1,207]	438 [88–974]
PE and DVT combined (VTE)	666 [361–1,486]	716 [230–1,289]	695 [286–1,301]	760 [212–1,434]	385 [48–1,090]	568 [124–1,202]	544 [201–1,200]	595 [224–1,274]	568 [136–1,134]

	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
Pelvic vein thrombosis (PVT)	NA	688 [440–1,144]	314 [180–508]	670 [242–1,165]	NA	326 [168–963]	NA	NA	NA
Pulmonary embolism (PE)	666 [361–1,486]	732 [224–1,306]	689 [231–1,062]	751 [196–1,439]	387 [11–1,032]	507 [60–1,202]	553 [139–1,201]	594 [226–1,370]	602 [166–1,186]
Retinal vein thrombosis (RVT)	NA	987 [363–1,562]	NA	897 [404–1,504]	744 [285–1,570]	416 [139–848]	NA	628 [314–1,272]	1,092 [489–1,443]
Splanchnic extrahepatic vein thrombosis (SVT)	NA	718 [443–1,416]	NA	1,092 [758–1,545]	819 [233–1,583]	940 [281–1,664]	NA	924 [713–1,808]	NA
Disseminated intravascular coagulation (DIC)	NA	NA	NA	NA	NA	226 [3–758]	NA	NA	NA
Stomach cancer (N)	0	438	32	384	167	621	38	125	91
Deep vein thrombosis (DVT)	NA	210 [107–783]	400 [258–486]	157 [40–505]	318 [144–1,268]	97 [27–247]	211 [128–646]	177 [83–292]	84 [8–424]
PE and DVT combined (VTE)	NA	141 [65–375]	205 [111–486]	110 [29–373]	145 [7–544]	92 [14–260]	149 [91–246]	123 [54–264]	108 [30–357]
Pelvic vein thrombosis (PVT)	NA	NA	NA	119 [40–432]	NA	290 [123–1,062]	NA	NA	NA
Pulmonary embolism (PE)	NA	132 [58–360]	300 [132–595]	106 [24–344]	84 [0–338]	74 [0–250]	124 [82–218]	80 [20–214]	96 [30–338]

	IQVIA LPD Belgium	DK-DHR	EBB	FinOMOP-THL	IQVIA DA Germany	SIDIAP	IPCI	CPRD GOLD	UKBB
Retinal vein thrombosis (RVT)	NA	822 [736–1,331]	NA	881 [442–1,349]	NA	418 [93–517]	NA	NA	NA
Splanchnic extrahepatic vein thrombosis (SVT)	NA	252 [186–663]	NA	80 [16–1,043]	NA	19 [0–263]	NA	NA	137 [48–149]
Disseminated intravascular coagulation (DIC)	NA	NA	NA	NA	NA	56 [0–312]	NA	NA	NA

N=total number of thromboembolic events, of any of the seven specific outcomes, per type of cancer. NA=not applicable (due to 0 or <5 event counts); NC=not captured.
IQVIA LPD=IQVIA Longitudinal Patient Database Belgium (IQVIA LPD Belgium); DK-DHR=Danish Data Health Registries; EBB=Estonian Biobank; FinOMOP-THL=Finnish National Health Registers; IQVIA DA=IQVIA Disease Analyzer Germany (IQVIA DA Germany); SIDIAP=The Information System for Research in Primary Care; IPCI=Integrated Primary Care Information; CPRD GOLD=Clinical Practice Research Datalink GOLD; UKBB=UK BioBank.

9.2.4. Median time to thromboembolic event after the first cancer diagnosis by sex, age, and study sub-period

The full set of tables and figures for this study stratified by sex, age, and study sub-period is available through an interactive web-application ShinyApp at EUPAS1000000814.

No clear differences were observed in the median time to thromboembolic event between females and males across data sources, except for a possible pattern towards longer event times among females than males for SVT in liver cancer. Although the median [IQR] time from the first diagnosis of liver cancer until the occurrence of SVT was shorter among females than males in SIDIAP (0 days [0–164] versus 9 days [0–286]), it was longer among females than males in DK-DHR (162 days [30–514] versus 48 days [0–255]), FinOMOP-THL (30 days [0–274] versus 6 days [0–168]), IQVIA DA Germany (161 days [0–858] versus 2 days [0–223]), and UKBB (133 days [42–248] versus 42 days [0–383]).

The events of SVT in liver cancer were recorded predominantly among males, with the number of events ranging from 6 among females and 15 among males in UKBB to 83 among females and 329 among males in SIDIAP.

Although the median [IQR] time from the diagnosis of pancreatic cancer until the occurrence of VTE increased with increasing age in IQVIA DA Germany (from 18 days [11–139] in the age group 45–54 years to 113 days [4–201] in the age group ≥85 years), a different pattern was observed in other data sources. In DK-DHR, FinOMOP-THL, IPCI, SIDIAP, CPRD GOLD, and UKBB, the median time to VTE was highest in younger age groups (35–64 years) and lower in the higher age groups. For example, in DK-DHR it declined from 224 days [81–560] in the age group 45–54 years to 15 days [8–51] in the age group ≥85 years. In FinOMOP-THL, it decreased from 524 days [201–727] in the age group 35–44 years to 6 days [0–36] in the age group ≥85 years. Similar patterns were observed in IPCI (from 184 days [110–304] in the age group 55–64 years to 62 days [16–229] in the age group 75–84 years), CPRD GOLD (from 111 days [40–262] in the age group 45–54 years to 30 days [0–122] in the age group 75–84 years), and UKBB (from 50 days [18–106] in the age group 55–64 years to 25 days [0–130] in the age group 75–84 years). The highest number of events of VTE was recorded in the age group 65–74 years in all data sources, except for IQVIA DA Germany. In IQVIA DA Germany, the highest number of events of VTE in any of the age groups was recorded in the age group 75–84 years (50 events), followed by the age group 65–74 years (49 events).

In both study sub-periods, i.e. 2016–2019 and 2020–2022, individuals with pancreatic and liver cancers had the shortest median time to VTE, whereas melanoma had the longest. The median event times were shorter in the second study sub-period than in the first, due to the shorter maximum possible follow-up duration in the second study sub-period (i.e. <5 years).

10. DISCUSSION

10.1. Key results

Overall, we included 975,962 individuals across nine data sources from eight European countries. The total number of individuals across all data sources per cancer type was: bone cancer (n=3,328), brain cancer (n=16,019), breast cancer (n=191,151), colorectal cancer (n=128,431), corpus uteri cancer (n=25,296), kidney cancer (n=32,881), liver cancer (n=18,767), lung cancer (n=105,527), lymphoma and leukaemia (n=126,499), melanoma (n=73,649), oesophageal cancer (n=15,547), ovarian cancer (n=16,794), pancreatic cancer (n=32,647), prostate cancer (n=169,739), and stomach cancer (n=19,687). 64,650 thromboembolic events were used in the calculation of cumulative incidence.

The steepest increase in the cumulative incidence of thromboembolic events was within the first six months after cancer diagnosis for VTE, DVT, PE, SVT, PVT, and DIC across all cancer types. It varied between cancer types only for RVT.

The highest 5-year cumulative incidence of VTE, DVT, and PE was observed in ovarian cancer (VTE: ranging from 4.79% [95% CI: 3.95%–5.80%] in IQVIA DA Germany to 10.53% [95% CI: 9.00%–12.32%] in CPRD GOLD) and pancreatic cancer (VTE: ranging from 4.46% [95% CI: 3.41%–5.82%] in IPCI to 11.63% [95% CI: 9.57%–14.12%] in UKBB), while the lowest was observed in melanoma (VTE: ranging from 1.27% [95% CI: 1.09%–1.48%] in IQVIA DA Germany to 2.42% [95% CI: 2.03%–2.89%] in SIDIAP).

The highest 5-year cumulative incidence estimates of SVT after the cancer diagnosis were observed in liver cancer (ranging from 1.16% [95% CI: 0.87%–1.54%] in FinOMOP-THL to 9.44% [95% CI: 8.58%–10.38%] in SIDIAP) and pancreatic cancer (ranging from 0.46% [95% CI: 0.25%–0.85%] in CPRD GOLD to 4.41% [95% CI: 2.22%–8.73%] in EBB), while the lowest was observed in breast cancer (ranging from 0.02% [95% CI: 0.01%–0.04%] in IQVIA DA Germany to 0.11% [95% CI: 0.07%–0.16%] in SIDIAP).

The cumulative incidence of PVT, RVT, and DIC up to five years after the first cancer diagnosis remained below 1% across cancer types and data sources.

The shortest median times to VTE, DVT, PE, and SVT were observed in pancreatic cancer and liver cancer, while the longest was observed in melanoma. Out of these four outcomes, the median [IQR] time was shortest to SVT (in liver cancer: ranging from 6 days [0–238] in SIDIAP to 66 days [4–301] in DK-DHR), followed by PE (in pancreatic cancer: ranging from 23 days [1–131] in EBB to 160 days [46–331] in IPCI), DVT (in pancreatic cancer: ranging from 40 days [14–104] in UKBB to 218 days [123–482] in EBB), and VTE (in pancreatic cancer: ranging from 42 days [0–116] in UKBB to 134 days [24–260] in IPCI). The median [IQR] time was longest to SVT (869 days [692–1,253] in DK-DHR and 968 days [290–1,483] in SIDIAP), followed by DVT (ranging from 550 days [217–1,180] in SIDIAP to 875 days [229–1,329] in FinOMOP-THL), VTE (ranging from 435 days [246–1,032] in UKBB to 919 days [696–1,366] in EBB), and PE (ranging from 432 days [294–1,516] in UKBB to 1,135 days [703–1,443] in EBB).

The shortest median [IQR] times to PVT after the first cancer diagnosis were observed in ovarian cancer (82 days [44–196] in FinOMOP-THL) and lung cancer (166 days [80–425] in FinOMOP-THL and 227 days [112–496] in SIDIAP), while the longest was observed in melanoma (748 days [480–1,293] in FinOMOP-THL and 919 days [528–1,301] in SIDIAP). The shortest median [IQR] times to RVT after the first cancer diagnosis were observed in pancreatic cancer (1 days [0–494] in SIDIAP, 248 days [228–492] in DK-DHR, and 321 days [188–844] in FinOMOP-THL) and lung cancer (ranging from 411 days [226–833] in IQVIA DA Germany to 803 days [432–1,386] in DK-DHR), while the longest was observed in kidney cancer (from 663 days [474–1,629] in CPRD GOLD to 1,111 days [642–1,619] in FinOMOP-THL). The shortest median [IQR] times to DIC after the first cancer diagnosis were observed in pancreatic cancer (50 days [10–70] in SIDIAP) and colorectal cancer (51 days [0–269] in SIDIAP), while the longest was observed in oesophageal cancer (418 days [140–541] in SIDIAP).

The point estimates of the cumulative incidence in five years were lower among females than males for SVT in liver cancer. A possible pattern was observed towards longer event times among females than males from the first diagnosis of liver cancer until the occurrence of SVT.

While confidence intervals overlapped, the point estimates of the 5-year cumulative incidence were lower in older age groups for VTE in liver cancer and pancreatic cancer, and for PE in liver cancer. A possible pattern was observed towards shorter event times with increasing age for VTE in pancreatic cancer.

The cumulative incidence of VTE, DVT, PE and SVT from six months to five years was similar between those diagnosed in the two different study sub-periods: 2016–2019 and 2020–2022. The relatively low event counts by study sub-period complicated the comparison of the cumulative incidence of PVT, RVT, and DIC between the two study sub-periods.

10.2. Strengths and limitations of the research methods

A strength of this study was that we included nine data sources across eight different European countries, ensuing some degree of representativeness of the findings across Europe.

One limitation of this study was that event counts were low for some types of cancer, especially bone cancer, and for some outcomes. The cumulative incidence of PVT, RVT, and DIC up to five years after the first cancer diagnosis remained below 1% across cancer types and data sources. The relatively low event counts by sex, age group, and study sub-period complicated the comparison of results by strata. A third limitation was that we could not consider the cancer stage or treatment, both of which may be associated with a higher risk of venous thromboembolic events. The effect of cancer stage (especially metastatic versus non-metastatic) and treatment on the risk of thromboembolic events among individuals with cancer should be investigated in future studies, for which VTE would be the most likely outcome to allow for relatively high precision and sufficient event counts per subgroup for all cancer types (and SVT for liver cancer).

One data source-specific limitation of this study was that death data were incomplete in IQVIA LPD Belgium and IQVIA DA Germany. Another data source-specific limitation was that survival bias may be present in EBB, as three-quarters of the EBB population joined the data source in 2018–2019 and had to be alive to join. Among individuals with a first cancer diagnosis prior to 2018, this may have led to the selection of those individuals who survived long enough to be included in the data source, which could explain the lower cumulative incidence of death in EBB compared to the other data sources.

10.3. Interpretation

Overall, we included 975,962 individuals with selected cancers across nine data sources, among whom a total of 64,650 thromboembolic events occurred.

The steepest increase in the cumulative incidence of thromboembolic events occurred within the first six months after cancer diagnosis for most thromboembolic events and cancer types. This was expected, as it has been reported previously that the risk of these events is highest during the first year after the cancer diagnosis and can also be linked to active treatment phase.[7] In a cohort study on the risk of VTE post-diagnosis of cancer, the cumulative risks of VTE, DVT, and PE increased most steeply during the first ninety days, with more gradual increases afterwards up to the first year.[8] In line with this, the shortest median event times of PE, DVT, and VTE (all in pancreatic cancer) ranged from approximately one month to approximately seven months.

The cancer types that showed the highest and lowest cumulative incidence of venous thromboembolic events, and correspondingly with the shortest and longest events, were the same as those showing the highest and lowest incidence rates in the DARWIN EU® study Incidence rates of venous thromboembolic events in patients with selected cancers. The cancer types with the highest rates of venous thromboembolic events were pancreatic cancer, ovarian cancer, brain cancer, stomach cancer, and lung cancer, while the cancer types with the lowest rates were breast cancer, prostate cancer, and melanoma.

The cancer types that showed the highest and lowest rates of venous thromboembolic events were generally consistent with the literature. In a recent meta-analysis, the cancer types with the highest incidence of VTE were pancreatic cancer, brain cancer, and lung cancer; the incidence of DVT and PE was highest in lung cancer.[3] In a recent review, however, the cancer type with the highest incidence of VTE during the first six months after the cancer diagnosis was pancreatic cancer, followed by liver cancer, biliary cancer (not included in our study), non-small cell lung cancer (with small-cell lung cancer ranking ten cancer types lower than non-small cell lung cancer), ovarian cancer, and stomach cancer.[1] In this review, the cancer types associated with a higher risk of venous thrombosis were pancreatic cancer, brain cancer, lung cancer, ovarian cancer, lymphoma, myeloma (not included in our study), kidney cancer, stomach cancer,

and bone cancer; and the cancer types associated with a lower risk of venous thrombosis were breast cancer, prostate cancer, melanoma, and testicular cancer (not included in our study).

In a cohort study specifically on the risk of VTE in breast cancer, the risk was highest in the first year after the cancer diagnosis.[9] In this study, the cumulative incidence of VTE one year after the diagnosis of breast cancer was 2.0%, increasing to 4.0% five years after the diagnosis of breast cancer. Although the corresponding absolute estimates were slightly lower in our study, an approximate doubling was seen across data sources in the cumulative incidence of VTE from one year to four years after the first diagnosis of breast cancer.

The point estimates of the cumulative incidence of SVT five years after the liver cancer diagnosis were lower among females than males. This was reflected in a possible pattern towards longer event times among females than males from the initial cancer diagnosis. Apart from this finding, no clear differences were observed across data sources in estimates of cumulative incidence or time-to-event between females and males.

These results are in line with a recent review, which reported similar overall proportions of DVT and PE (and their composite, VTE in our study) in females and males with cancer. The review also cited a study in which portal vein thrombosis was less common among females than males with hepatocellular carcinoma, although this study was conducted in an Asian population.

In pancreatic cancer, the estimated cumulative incidence of VTE decreased with increasing age, although with shorter event times observed in older individuals. Similarly, a recent cohort study found that increasing age was possibly associated with a slightly lower risk of VTE among individuals with pancreatic cancer (hazard ratio in univariable analysis 0.98 [95%CI: 0.95–1.00] among 361 individuals with pancreatic cancer, of whom 64 (18%) experienced a VTE).[10]

The results for cumulative incidence and event times were generally similar among those diagnosed with selected cancers in the two study sub-periods: 2016–2019 and 2020–2022, although the observed event times were longer among those diagnosed in 2016–2019 than among those diagnosed in 2020–2022. This was likely due to the shorter follow-up in the latter study sub-period.

Death was considered a competing event in the time-to-event analysis of thromboembolic events. The cumulative incidence of death was higher than the cumulative incidence of VTE across cancer types and data sources, except for IQVIA DA Germany, in which the death records were incomplete. This is in line with a cohort study of the risk of VTE in individuals with incident cancer, in which 7.3% of individuals with cancer experienced VTE and 37.1% died during follow-up up to two years.[11] The authors found that the risk of VTE is overestimated when death is not treated as a competing risk for VTE (unless mortality is low) and that the extent of the overestimation increases with increasing mortality. Accounting for death as an event that precluded the subsequent occurrence of a thromboembolic event thus made the cumulative incidence estimates of the outcomes more accurate than they would have been if death had not been accounted for as a competing risk.

10.4. Generalisability

We utilised electronic healthcare data from nine data sources across eight European countries, thereby ensuring a certain level of generalisability of our findings to the European region.

These data sources included primary care records, national health registries, and claims data sources, or a combination of those. The observed heterogeneity in our results suggests differences in follow-up duration, data collection practices, and cancer diagnostic and management practices. While these differences may limit comparability within Europe, they can further constrain generalisability to non-European settings, where healthcare systems, population characteristics, and data infrastructure may differ substantially.

11. CONCLUSION

In this large multinational cohort study, we found the steepest increase in the cumulative incidence of thromboembolic events within the first six months after the first cancer diagnosis. This was observed for all thromboembolic events, except for RVT, for which the six-month interval with the steepest increase in the cumulative incidence varied by cancer type. The magnitude and timing of the increase in cumulative incidence varied by cancer type. Ovarian and pancreatic cancers had the highest percentages of cumulative incidence of VTE, DVT, and PE up to five years after the first cancer diagnosis, while melanoma had the lowest. Similarly, SVT had the highest cumulative incidence in five years after the first cancer diagnosis in liver and pancreatic cancers, while the lowest was observed in breast cancer.

These differences were reflected in time-to-event analyses, with pancreatic cancer and liver cancer showing the shortest median times from cancer diagnosis to VTE, DVT, PE, and SVT, while the longest were observed in melanoma. The cumulative incidence of PVT, RVT, and DIC up to five years after the first cancer diagnosis remained below 1% across cancer types and data sources. Overall, these findings highlight a critical early time period for the incidence of thromboembolic events after cancer diagnosis and demonstrate consistent patterns across multiple European populations, supporting their use for clinical and regulatory considerations.

12. REFERENCES

1. Wan T, Song J, Zhu D. Cancer-associated venous thromboembolism: a comprehensive review. Thromb J. 2025 Apr 16;23(1):35.
2. Vladić N, Englisch C, Berger JM, Moik F, Berghoff AS, Preusser M, et al. Validation of risk assessment models for venous thromboembolism in patients with cancer receiving systemic therapies. Blood Adv. 2025 June 30;9(13):3340–9.
3. Deng HW, Li J, Zhai YS, Mei WY, Lin XX, Xu Q, et al. Incidence of arterial and venous thromboembolism in cancer patients- insights from more than 5,000,000 patients. J Thromb Thrombolysis. 2025 Mar 1;58(3):370–9.
4. López-Güell K, Burn E, Catala M, Li X, Newby D, Mercade-Besora N (2025). *CohortSurvival: Estimate Survival from Common Data Model Cohorts*. R package version 1.0.3, https://darwin-eu-dev.github.io/CohortSurvival/
5. Therneau T, Crowson C, Atkinson E. Multi-state models and competing risks. 2024.
6. Catala M, Guo Y, Lopez-Guell K, Burn E, Mercade-Besora N, Alcalde M (2025). *CohortCharacteristics: Summarise and Visualise Characteristics of Patients in the OMOP CDM*. R package version 1.1.0, https://darwin-eu.github.io/CohortCharacteristics/.
7. Blom JW, Doggen CJM, Osanto S, Rosendaal FR. Malignancies, Prothrombotic Mutations, and the Risk of Venous Thrombosis. JAMA. 2005 Feb 9;293(6):715–22.
8. Ording AG, Skjøth F, Søgaard M, Højen AA, Overvad TF, Noble S, et al. Increasing Incidence and Declining Mortality After Cancer-Associated Venous Thromboembolism: A Nationwide Cohort Study. Am J Med. 2021 July 1;134(7):868-876.e5.
9. Brand JS, Hedayati E, Bhoo-Pathy N, Bergh J, Hall P, Humphreys K, et al. Time-dependent risk and predictors of venous thromboembolism in breast cancer patients: A population-based cohort study. Cancer. 2017;123(3):468–75.
10. Hanna-Sawires RG, Groen JV, Hamming A, Tollenaar RAEM, Mesker WE, Luelmo SAC, et al. Incidence, timing and risk factors of venous thromboembolic events in patients with pancreatic cancer. Thromb Res. 2021 Nov 1;207:134–9.
11. Ay C, Posch F, Kaider A, Zielinski C, Pabinger I. Estimating Risk of Venous Thromboembolism in Patients with Cancer in the Presence of Competing Mortality. J Thromb Haemost JTH. 2015 Mar 1;13(3):390–7.

13. ANNEXES

ANNEX I. Description of data sources

IQVIA Longitudinal Patient Database Belgium (IQVIA LPD Belgium)

#	Section	Description
1	Database Identification and country	IQVIA LPD Belgium (IQVIA Longitudinal Patient Database Belgium) Belgium
2	Data partner information section	IQVIA IQVIA Europe
3	Coverage and timespan	Data collection since: 2005 Extent: Nation-wide. Panel of 300 GPs in Belgium. The panel is maintained as a representative sample of the primary care physician population in Belgium, according to three criteria known to influence prescribing: age, sex, and geographical distribution.
4	Healthcare setting / type of data	Primary care – gps. Ambulatory visits, with diagnosis, prescriptions, procedures, and laboratory tests.
5	Data collection process	Outpatient electronic health records. Records are entered by GPs at the healthcare encounter.
6	General representativeness	The panel of contributing physicians (a stable 300 GPs) is maintained as a representative sample of the primary care physician population in Belgium, according to three criteria known to influence prescribing: age, sex, and geographical distribution. The panel consists of a stable 300 GPs that are geographically well spread. The total number of active GPs in Belgium is 15,602. The regional geographical spread of physicians in the LPD data is also representative of the distribution across the country: 57% GPs in the North (compared to 54% nationally), 31% in the South (33% nationally), and 12% in Brussels (13%).The provider of the data has more than 2,250 GPs under contract so in case of a drop out a replacement is easily found.
7	Data content /source coding	No information on source coding.
8	Data Harmonisation	The data has been mapped to the OMOP CDM v5.4 and the OMOP standard vocabularies (SNOMED, RxNorm, LOINC). The format, structural and semantic conformance has been verified upon onboarding into the DARWIN EU® data network. The patient ID is per practice. So a patient can have different IDs in the DB, one per practice. In Belgium, patients are typically registered at only one GP practice, so duplication should be minimal.
9	Quality control (database specific)	No QC. Integrity constraints only.
10	Linkage	No linkage.
11	Vital status	Death information is derived from healthcare events.
12	Limitations	No database-specific limitations documented. General limitations for the data type applicable.
13	Main references	No main reference provided.
14	Link to HMA-EMA catalogue and database webpage	HMA-EMA Catalogue entry: https://catalogues.ema.europa.eu/data-source/1111116 Website: https://iqvia.com

Danish Data Health Registries (DK-DHR)

#	Section	Description
1	Database Identification and country	DK-DHR (Danish Data Health Registries) Denmark
2	Data partner information section	Danish Medicines Agency (DKMA), Data Analytics Centre (DAC)
3	Coverage and timespan	Data collection since: 1995, Extent: Nationwide. The data is representative of the entire Danish population.
4	Healthcare setting / type of data	Community pharmacists, and secondary care – specialists (ambulatory or hospital outpatient care), and hospital inpatient care. The following data elements are collected: diagnosis (including rare diseases and pregnancy data), hospital admissions, discharge and ICU data, Cause of death, Drug prescription retrievals, vaccination and contraception, Procedures, and Sociodemographic information (sex and age but no information on income, education, occupation).
5	Data collection process	Outpatient electronic health records, and Inpatient hospital electronic health records, and Registries, and Other. All causes of deaths, all retrieved drug prescriptions, all records of vaccinations, all hospital inpatient and outpatients contacts including disease diagnoses and hospital surgical and non-surgical procedures, histologically confirmed incident cancers, laboratory test results for the entire Danish population from 1/1/1995 onwards.
6	General representativeness	The data is representative of the entire Danish population. Healthcare is free in Denmark, so we do not expect any bias in data collection based on socio-economic status.
7	Data content /source coding	Diagnoses and causes of death are collected using the ICD-10 vocabulary. ATC and RxNorm are used for Drugs. SNOMED codes are used for Procedures.
8	Data Harmonisation	The data has been mapped to the OMOP CDM v5.4 and the OMOP standard vocabularies (SNOMED, RxNorm, LOINC, ICDO3, cancer Modifier). The format, structural and semantic conformance has been verified upon onboarding into the DARWIN EU® data network. .
9	Quality control (database specific)	The data we have received relating to nationwide Danish Health Data registries offer an opportunity for large-scale, population-based studies with several advantages 1) Their large size improves the precision of estimates and enables the study of rare exposures and outcomes with long-term latency, 2) Inclusion of nearly all individuals in the target population ensures that the data reflect routine clinical care and all clinical segments of the source population, 3) Data are collected independently of each research study, thus minimising certain types of bias, e.g. non-response, and the influence from attention to the research question on the diagnostic process. Before the source data is sent to us, the Danish Health Data Authority does running and comprehensive checks of the registry table data validity of the variables, breaks in data, changes in variable coding, missingness, etc. We perform checks of missingness/completeness in relation to requested variables. In essence, we are receiving a dump of a mirror of the data that is controlled by the SDS. The documentation performed by SDS is available online, in Danish primarily https://www.esundhed.dk/Dokumentation (all variables), but also in English https://sundhedsdatastyrelsen.dk/da/english/health_data_and_registers/national_health_registers
10	Linkage	There is no linkage in this data source.
11	Vital status	The Cause of Death registry (DAR) is used, the cause of death is collected using ICD-10 codes.
12	Limitations	There are no clinical measurements in the data. DK-DHR has the following limitations, which may be relevant confounders for certain complex Darwin EU studies: • We lack information on key socio-economic status (SES) factors, such as occupation, education, and income. These variables may be important for analysis in some studies. • We only have complete data on lifestyle factors (such as smoking status and weight) for pregnant women.

#	Section	Description
		We have no information on patient contacts in primary care (visits to the GP). Consequently, the incidence of chronic diseases like Type 2 Diabetes might be underestimated.
13	Main references	Schmidt M, Schmidt SAJ,Adelborg K,Sundbøll J,Laugesen K,Ehrenstein V,Sørensen HT "The Danish health care system and epidemiological research: from health care contacts to database records." Clinical epidemiology (2019): 31372058
14	Link to HMA-EMA catalogue and database webpage	Website: https://sundhedsdatastyrelsen.dk/da/english/health_data_and_registers/healthdatadenmark HMA-EMA Catalogue entry: https://catalogues.ema.europa.eu/data-source/1111217

Estonian Biobank (EBB)

#	Section	Description
1	Database Identification and country	EBB (Estonian Biobank) Estonia
2	Data partner information section	University of Tartu Institute of Computer Science
3	Coverage and timespan	Data collection since: 2004 Extent: Nation-wide. EBB is a nation-wide database containing records from 2004 onwards. Estonian population-based cohort size of 211,800 participants (01/01/2024) aged 18 years and older recruited at GP offices, private practices, and hospitals or in the recruitment offices of the Estonian Genome Center.
4	Healthcare setting / type of data	Primary care – GPs, and community pharmacists, and primary care specialists (e.g. paediatricians), and secondary care – specialists (ambulatory or hospital outpatient care), and hospital inpatient care. Registry which collects electronic records from the biobank and cohort study.
5	Data collection process	Data is retrieved by Estonian Biobank once a year from national registries. The insurance claims are requested from Estonian Health Insurance Fund. The inpatient and outpatient electronic health records are requested from National Health Information System. The cancer registry and cause of death registry information is requested from The National Institute for Health Development. The data is sent to the national registry by the healthcare providers.
6	General representativeness	The age, sex, and geographical distribution closely reflect those of the Estonian adult population and encompass close to 5approximately 20% of adult population. Female participants are over-represented in EBB. Overall, 3.4% of Estonian men and 5.5% of Estonian women are represented in EBB. Older people tend to participate less frequently, however, all age groups are well represented.
7	Data content /source coding	All participants have undergone a standardized health assessment, including provision of blood samples for purification of DNA, white blood cells, and plasma, and completed a questionnaire covering various health-related topics, such as lifestyle, diet, and clinical diagnoses. Diseases and health problems are recorded as ICD-10 codes and prescribed medicine according to the ATC classification and local package codes. Procedures and services are coded with NOMESCO classifier and local service codes.
8	Data Harmonisation	The data has been mapped to the OMOP CDM v5.4 and the OMOP standard vocabularies (SNOMED, RxNorm, LOINC). The format, structural and semantic conformance has been verified upon onboarding into the DARWIN EU® data network. There is one national identifier that allows linking together all encounters across databases.
9	Quality control (database specific)	The quality control procedures in the Estonian Biobank aim to remove the most obvious mistakes in the data, misspellings, impossible dates, duplicates. Before performing the ETL, several problems are fixed on the source data. Since the ETL

#	Section	Description
		procedures are used for a number of different datasets (from the same national sources), we have a growing number of pre-processing steps that correspond to the issues we have discovered previously in the data, such as checking for the presence of critical values, harmonizing date and unit of measurement formats, checking the validity of certain entries against classifiers, etc.
10	Linkage	Follow-up data are available via linkage with national health-related registries and via re-examination of participants. Furthermore, electronic health records are updated for phenotypic outcome information every year. The EBB database is regularly linked with national registries, hospital databases, and the databases of the Estonian Health Insurance Fund (EHIF) and the National Health Information System (NHIS)
11	Vital status	Vital status (death date and causes of death) are obtained from the Causes of Death Registry.
12	Limitations	Participation in EBB cohort is voluntary, therefore the biobank does not represent a random sample and could be subject to recruitment bias. Although recruitment was open to everyone, there is a disproportion between ethnic Estonians and ethnic Russians in the biobank, with Estonians being overrepresented.
13	Main references	Milani, L., Alver, M., Laur, S. *et al.* The Estonian Biobank's journey from biobanking to personalized medicine. *Nat Commun* **16**, 3270 (2025). https://doi.org/10.1038/s41467-025-58465-3
14	Link to HMA-EMA catalogue and database webpage	HMA-EMA Catalogue entry: https://catalogues.ema.europa.eu/data-source/1111114 Website: https://genomics.ut.ee/en/content/estonian-biobank

Finnish National Health Registers (FinOMOP-THL)

#	Section	Description
1	Database Identification and country	FinOMOP-THL (Population Registers at the Finnish Institute for Health and Welfare) Finland
2	Data partner information section	Finnish Institute for Health and Welfare (THL) Department of Knowledge Brokers
3	Coverage and timespan	Data collection since: 1998 Extent: Nation-wide. The current CDM population comprises all persons having been alive and residing in Finland since the beginning of 2011.
4	Healthcare setting / type of data	Primary care – gps, and primary care specialists (e.g. paediatricians), and secondary care – specialists (ambulatory or hospital outpatient care), and hospital inpatient care. THL maintains health registers that cover both public and private, primary, and specialised inpatient and outpatient health care encounters in Finland, starting from 2011. The entire public sector and private inpatient encounters have been included since 2011, while private outpatient encounters, including occupational care, are included since 2020. Since 1998, the Care Register for Health Care (TerveysHilmo) has covered both public outpatient and inpatient specialized care and private inpatient care. Since 2011, the Register of Primary Health Care Visits (AvoHilmo) has covered public primary care. Since 2020, the register has also covered private outpatient care and occupational care. In addition, the CDM also contains positive COVID-19 test results from the Finnish National Infectious Diseases Register. The register itself covers all laboratory confirmations for around 70 specific microbes from 1995 onwards, but only COVID-19 has currently been mapped to the CDM.

#	Section	Description
		The CDM also includes prescription records from multiple different sources. Both Care Register for Health Care and Register of Primary Health Care Visits contain very basic prescription data recorded during health care encounters that include just the ATC-code and trade name of medication. More comprehensive prescription data from Kanta Prescription Centre, maintained by Social Insurance Institution of Finland (Kela), has been integrated into the CDM since its 2024 release. The Kanta Prescription records are based electronic prescriptions, which were adopted by most public health care providers in 2010 and by most private providers by 2017.
5	Data collection process	Outpatient electronic health records, and Inpatient hospital electronic health records, and Registries. Data is entered by clinicians upon healthcare contact and processed by THL (Kela in the case of Kanta Prescription Centre).
6	General representativeness	The THL data has national coverage and is therefore well representative of the Finnish population. Using the complete population as a basis for the person table also serves to facilitate calculations on a population level, e.g. incidence rates.
7	Data content /source coding	The following coding systems have been OMOP-mapped, typically to a good level of completeness: ICD10fi Finnish Extension, ATC, Toimenpideluokitus (procedure classification adapted from the Nordic Classification of Surgical Procedures (NCSP)), Terveydenhuollon erikoisalat (Hilmo specific provider speciality), Rokotustapa (AR/YDIN National classification for vaccine administration), Tupakointistatus (AR/YDIN National classification for smoking status). Vaccinations are identified on product level based on batch number, trade name, vaccine title, and ATC-code. This is mapped on brand and type in the OMOP CDM.
8	Data Harmonisation	The data has been mapped to the OMOP CDM v5.4 and the OMOP standard vocabularies (SNOMED, RxNorm, LOINC). The format, structural and semantic conformance has been verified upon onboarding into the DARWIN EU® data network. Each patient in THL has a unique identifier.
9	Quality control (database specific)	The source data collection undergoes a structural and semantic validation before entry into the source database. Additionally, some coded variables undergo quality assessment against the respective code systems post entry into the database. The source registers are also assessed for completeness and coverage, with the aim of improving future collection in the areas where data is lacking.
10	Linkage	THL is already a linkage of multiple Finnish registries (see above).
11	Vital status	The National Population registry data forms the basis for forming the patient population. This ensures an up-to-date location (municipality of residence) of patients, as well as complete death occurrences (although not the cause of death).
12	Limitations	Not all private health care records are covered for the CDM's entire follow-up time from 2011 onwards. For Register of Primary Health Care Visits and Finnish National Vaccination Register, records from private health care have been available from 2020 onwards. For Kanta Prescription Centre, the coverage of private health care records has been good from 2017 onwards. The inclusion of private health care mainly presents itself as an increase in the number of observations, meaning that it has to be accounted for when interpreting any time series data from the CDM. All drug records in the CDM are currently based on prescriptions. Kanta Prescription Centre also includes information on drug dispensings, but these have not currently been converted into OMOP CDM. Depending on the type of the medication, a single prescription can be for up to two years dosage of the drug. This means a patient can have up to two-year breaks between observations while actively using the drug. Observation of a prescription also does not mean that the patient necessarily bought or used the medication. The CDM does not currently have meaningful end dates or days of supply for drug exposure. This information is not available for Care Register for Health Care or Register of Primary

#	Section	Description
		Health Care Visits, and for Kanta Prescription Centre it is only available in unstructured, free-form format that has not been converted into OMOP CDM.
13	Main references	Häkkinen, Pirjo; Mölläri, Kaisa; Saukkonen, Sanna-Mari; Väyrynen, Riikka; Mielikäinen, Lasse; Järvelin, Jutta "Hilmo–Sosiaali- ja terveydenhuollon hoitoilmoitus 2020 : Määrittelyt ja ohjeistus : Voimassa 1.1.2020 alkaen" Terveyden ja hyvinvoinnin laitos (2019):
14	Link to HMA-EMA catalogue and database webpage	HMA-EMA Catalogue entries: https://catalogues.ema.europa.eu/node/1090/ https://catalogues.ema.europa.eu/node/1094/ Websites: https://thl.fi/en/statistics-and-data/data-and-services/register-descriptions/care-register-for-health-care https://thl.fi/en/statistics-and-data/data-and-services/register-descriptions/register-of-primary-health-care-visits https://thl.fi/en/topics/infectious-diseases-and-vaccinations/surveillance-and-registers/finnish-national-vaccination-register-and-monitoring-of-the-vaccination-programme https://www.kanta.fi/en/research-and-knowledge-management

IQVIA Disease Analyzer Germany (IQVIA DA Germany)

#	Section	Description
1	Database Identification and country	IQVIA DA Germany (IQVIA Disease Analyzer Germany) Germany
2	Data partner information section	IQVIA
3	Coverage and timespan	Data collection since: 1989 Extent: Nation-wide. GP and specialists in Germany using specific patient management software.
4	Healthcare setting / type of data	Primary care – gps, and primary care specialists (e.g. paediatricians). Diagnoses, medication, and procedures from an ambulatory setting. Medications are recorded as prescriptions of marketed products.
5	Data collection process	Outpatient electronic health records. By clinicians at healthcare contact.
6	General representativeness	No specific details on general representativeness given.
7	Data content /source coding	Prescription is on product code level (German PZN), ICD10, NFC, Local lab coding.
8	Data Harmonisation	The data has been mapped to the OMOP CDM v5.4 and the OMOP standard vocabularies (SNOMED, RxNorm, LOINC). The format, structural and semantic conformance has been verified upon onboarding into the DARWIN EU® data network. There can be patients registered under different ID numbers, because there is no linkage between different GPs.
9	Quality control (database specific)	Data is quality checked on plausibility.

#	Section	Description
10	Linkage	No.
11	Vital status	Death information is derived from medical events.
12	Limitations	No database-specific limitations documented. General limitations for the data type applicable.
13	Main references	No main reference provided.
14	Link to HMA-EMA catalogue and database webpage	HMA-EMA Catalogue entry: https://catalogues.ema.europa.eu/data-source/104282 Website: https://www.iqvia.com/

The Information System for Research in Primary Care (SIDIAP)

#	Section	Description
1	Database Identification and country	SIDIAP (The Information System for Research in Primary Care) Catalunya, Spain
2	Data partner information section	IDIAPJGol
3	Coverage and timespan	Data collection since: 2006 Extent: Regional. The SIDIAP database contains records of around 6 million people residing in Catalonia, estimated to be representing around 76% of the Catalan population.
4	Healthcare setting / type of data	Primary care – gps, and hospital inpatient care. SIDIAP captured data includes routine visits, sociodemographic information, diagnoses, laboratory tests, drugs (prescribed and dispensed), referrals, and lifestyle information.
5	Data collection process	Outpatient electronic health records, and Inpatient hospital electronic health records, and Other. Data is entered by primary care physicians upon healthcare contact, supplemented with hospital discharge records. The Institut Catala de la Salut is the owner of the data and acts as the data controller.
6	General representativeness	It was previously shown that the captured SIDIAP population is highly representative of the entire Catalan region in terms of geographic, age, and sex distributions.
7	Data content /source coding	SIDIAP data covers all services that occur at the Primary Care Centres, as well as support services, such as sexual and reproductive health or home end-of-life care. Drugs are coded in ATC-WHO terminology in the source data. Health outcomes are captured in ICD-10CM codes. The SIDIAP contains all laboratory tests and results performed in primary health centres. Demographics, geographical, as well as socio-economic factors are recorded for each patient.
8	Data Harmonisation	The data has been mapped to the OMOP CDM v5.4 and the OMOP standard vocabularies (SNOMED, RxNorm, LOINC). The format, structural and semantic conformance has been verified upon onboarding into the DARWIN EU® data network. No.
9	Quality control (database specific)	Internal and external validation processes are carried out to determine the data quality of the SIDIAP information at each data update. These include stratifying the data by geographical regions and year in order to identify differences in data collection that need to be harmonized (e.g. recording of specific information under different codes). The measurement units of variables measuring one characteristic are also homogenized (e.g. transformation of the data from every laboratory that measures haemoglobin to grams per

#	Section	Description
		decilitre). Visual inspection of all data included in the database by week is also conducted, allowing one to see temporal patterns in the registry of a certain variable. With this information, the SIDIAP team can issue recommendations to researchers about the most common variable(s) where certain information is recorded (e.g. there are several variables with information concerning the women's menopausal status and with these visual inspection tools the SIDIAP team can inform the researchers about which related variables have the largest number of records and could be more helpful to capture menopause). Data availability (longitudinally and reliability), plausibility (range checks and unusual values), and consistency are inspected through visualisation tools. In addition, before accessing the data for a requested project, research teams have access to a quality-control report. This document contains counts, years, percentiles, maximums and minimums, incidences, and prevalence of the data requested for the project, allowing detection of inconsistencies in the data extraction prior to data delivery. External validation processes of the SIDIAP database mainly include assessing the data recorded in SIDIAP through linkage to external gold standard data sources, by analysing free text, or by sending questionnaires to health professionals.
10	Linkage	SIDIAP is linked to a hospital discharge database, pharmacy dispensation, and primary care laboratories. It can also be linked to other registries in Catalonia on a project-by-project basis.
11	Vital status	Mortality is fully captured in SIDIAP. The cause of death is not available but can be linked to the Spanish death registry on a project-by-project basis.
12	Limitations	The SIDIAP data is representative of individuals using public primary care. Conditions that are usually followed by specialist care or in private practices might not be properly captured. Patients are followed until death or when transferring to another primary health care centre that does not contribute to SIDIAP.
13	Main references	Recalde M, Rodríguez C,Burn E,Far M,García D,Carrere-Molina J,Benítez M,Moleras A,Pistillo A,Bolíbar B,Aragón M,Duarte-Salles T "Data Resource Profile: The Information System for Research in Primary Care (SIDIAP)." International journal of epidemiology (2022): 35415748
14	Link to HMA-EMA catalogue and database webpage	HMA-EMA Catalogue entry: https://catalogues.ema.europa.eu/data-source/50190 Website: https://www.sidiap.org/index.php/en

Integrated Primary Care Information (IPCI)

#	Section	Description
1	Database Identification and country	IPCI (Integrated Primary Care Information) The Netherlands
2	Data partner information section	Erasmus University Medical Center Department of Medical Informatics
3	Coverage and timespan	Data collection since: 2006 Extent: Nation-wide. IPCI is a Dutch database that contains patient records from 2006 onwards. However, it mainly covers the central part of the country, including the most densely populated area (the 'Randstad') and non-urban areas. IPCI contains information on all patients registered with GPs responsible for non-emergency care and referrals. A patient is registered at birth or at first encounter with the GP.
4	Healthcare setting / type of data	Primary care – gps. Data is collected from primary care EHR. This includes demographic information, complaints and symptoms, diagnoses, laboratory test results, lifestyle factors (in limited amount), and correspondence with secondary care, such as referral and discharge letters.

#	Section	Description
5	Data collection process	Outpatient electronic health records. Data is entered into the EHR system by the GPs, during or after the visit. Data is aggregated by Erasmus MC data managers and combined in one harmonized database. Several checks are done on this database to ensure correct data processing. Persons are mostly uniquely identified, with the exception of when persons change GP practice (when the same individual can receive several different identifiers).
6	General representativeness	More than 99% of the Dutch population has health insurance, and almost all citizens are registered with a general practitioner. Over 12 months, around 78% of the population has at least one contact with their GP. IPCI included around 350 GP practices out of around 5000 in the country (~ 7%). The demographic composition of the IPCI population mirrors that of the general Dutch population in terms of age and sex.
7	Data content /source coding	Dutch GPs use mainly Dutch standard codes, like ICPC-1 and Diagnostische Bepalingen maintained by NHG. And for therapy the G-Standard is used, maintained by ZIndex.
8	Data Harmonisation	The data has been mapped to the OMOP CDM v5.4 and the OMOP standard vocabularies (SNOMED, RxNorm, LOINC). The format, structural and semantic conformance has been verified upon onboarding into the DARWIN EU® data network. Patients can be registered under different IDs. However, in the Netherlands, patients typically have one GP and changing practice is uncommon.
9	Quality control (database specific)	Prior to each data release, extensive quality control steps are performed, e.g. comparison of patient characteristics between practices, and checks to identify abnormal temporal data patterns in practices. For each practice, around 200 quality indicators are obtained. Of these indicators, a quarter refer to population characteristics, e.g. number of birth and mortalities relative to practice size, temporal consistency. The other indicators are based on medical data, e.g. distribution of measurement values, frequencies of diagnoses and procedures relative to age, completeness of data. The indicators are combined in a couple of quality scores for each practice. For these scores, cut-off values for acceptable quality have been defined. Practices with a score below a cut-off are excluded for research. This approach has shown to be very important, for example to check if data from practices that just joined the database are at an acceptable level of quality. The details of the approach, like the cut-off values for acceptance, are based on years of experience. In addition, trends are compared with the previous database release. Extensive quality control steps are performed before each data release. These include comparing patient characteristics between practices and checks to identify abnormal temporal data patterns in practices. Additional checks include over 200 indicators related to population characteristics (e.g. reliability of birth and mortality rates) and medical data (e.g. availability of durations of prescriptions and completeness of laboratory results). Records of low quality are excluded from the database.
10	Linkage	Linkage requires additional approval steps and needs to be assessed on a case-by-case basis. IPCI is not routinely linked with other databases.
11	Vital status	Vital status (death date and cause) is collected based on GP records.
12	Limitations	The main limitation comes with the fact that IPCI is limited to GP records, and although it contains information on referrals and discharge letters, it may not fully capture specific hospital information. IPCI does not include coded/detailed data about medications/procedures/test results from the hospital or other care-providers.
13	Main references	de Ridder MAJ, de Wilde M,de Ben C,Leyba AR,Mosseveld BMT,Verhamme KMC,van der Lei J,Rijnbeek PR "Data Resource Profile: The Integrated Primary Care Information (IPCI) database, The Netherlands." International journal of epidemiology (2022): 35182143
14	Link to HMA-EMA catalogue and database webpage	HMA-EMA Catalogue entry: https://catalogues.ema.europa.eu/data-source/42618 Website: http://www.ipci.nl

Clinical Practice Research Datalink GOLD (Oxford) (CPRD GOLD)

#	Section	Description
1	Database Identification and country	CPRD GOLD (Clinical Practice Research Datalink GOLD) The United Kingdom
2	Data partner information section	University of Oxford NDORMS
3	Coverage and timespan	Data collection since: 1987 Extent: Nation-wide. CPRD GOLD consists of patients in contributing practices using Vision software. Historically this covered the whole of the UK, but the number of contributing practices in the England is dropping. In January 2025 only 3 practices from England were a part of CPRD GOLD, while historical patient data were from the whole of the UK, and will continue to be so. In the future, no practices from England will be present, only practices from Scotland, Wales, and Northern Ireland.
4	Healthcare setting / type of data	Primary care – gps, and primary care specialists (e.g. paediatricians), and secondary care – specialists (ambulatory or hospital outpatient care), and hospital inpatient care. CPRD GOLD data include patient demographics, biological measurements, clinical symptoms and diagnoses, referrals to specialist/hospital and their outcome, laboratory tests/results, and prescribed medications.
5	Data collection process	Outpatient electronic health records. Data is entered by clinicians into the EHR. Data is processed by CPRD and provides data releases for research.
6	General representativeness	CPRD GOLD has been assessed and found to be broadly representative of the UK general population in terms of age, gender, and ethnicity. In CPRD GOLD in January 2025 there were 2,730,707 current acceptable patients (i.e. registered at currently contributing practices that use Vision software, excluding transferred out, deceased patients, and those flagged by CPRD as not acceptable for clinical research for data quality issues). This equals to 4.07%, based on the UK population estimates of 67,026,300 from the Office of National Statistics (mid-2023). Current patients are only from Scotland, Wales, and Northern Ireland. Historically, GOLD does contain data from England as well.
7	Data content /source coding	Gemscript, Read, dm+d
8	Data Harmonisation	The data has been mapped to the OMOP CDM v5.4 and the OMOP standard vocabularies (SNOMED, RxNorm, LOINC). The format, structural and semantic conformance has been verified upon onboarding into the DARWIN EU® data network. In GOLD, a patient can be registered under different ID numbers upon changing practice or re-registration. Researchers are not able to identify these patients, as the data are anonymised. However, GOLD covers less than 5% of the current UK GP practices and it is unlikely that an individual who does change GP practice ends up in another GP practice which uses the Vision software and accepts the CPRD data collection agreement. The very small number of duplicated IDs will have different observation periods and should not have an impact on the data analyses.
9	Quality control (database specific)	CPRD GOLD only includes practices whose data quality is assessed to be up-to-standard (uts). Each practice is associated to an uts date set when the data quality standards become satisfactory, and CPRD recommend using only longitudinal data starting from this uts date. Every time CPRD collect the EHR from a practice, checks are run for the data quality standards and if they are not adequate, the EHR is not accepted. When the data quality becomes acceptable again, CPRD updates the practice uts date. CPRD also check data quality standards at the patient level and associate each patient to a flag, reporting if its data is acceptable for clinical research. Only patients with acceptable data quality are included in the population to be mapped to CDM.

#	Section	Description
10	Linkage	CPRD GOLD can be linked to several sources, however our Oxford OMOP CDM is only linked to the CPRD GOLD Ethnicity Record and to the CPRD Townsend Deprivation Index at Practice Level
11	Vital status	Vital status is retrieved from the GP records. Population registry (ONS) data can be requested on a study-by-study basis and linked. This data only covers England and is planned to be mapped to OMOP in the future. The cause of death is not captured.
12	Limitations	The main limitation is due to the fact that CPRD GOLD is limited to GP records, and although it contains information on referrals and discharge letters, it may not fully capture specific hospital information. Events from hospital and specialist care are not covered.
13	Main references	Sanchez-Santos MT, Axson EL,Dedman D,Delmestri A "Data Resource Profile Update: CPRD GOLD." International journal of epidemiology (2025): 40499193
14	Link to HMA-EMA catalogue and database webpage	HMA-EMA Catalogue entry: https://catalogues.ema.europa.eu/data-source/1111113 Website: https://cprd.com

UK BioBank (UKBB)

#	Section	Description
1	Database Identification and country	UKBB (UK BioBank) The United Kingdom
2	Data partner information section	Oxford University NDORMS
3	Coverage and timespan	Data collection since: 2006 Extent: Nation-wide. People recruited from whole of the UK.
4	Healthcare setting / type of data	Primary care – gps, and primary care specialists (e.g. paediatricians), and secondary care – specialists (ambulatory or hospital outpatient care), and hospital inpatient care, and other (specify). UK Biobank is made by a rich variety of data sources, which include genetic data, primary care data, hospital inpatient data, death data, and cancer registry.
5	Data collection process	Inpatient hospital electronic health records, and Registries, and Biobank. The baseline assessment is consisting of both patient reported data (questionnaire) and physical measurements. GP and hospital data, as well as death and cancer registry records, are linked afterwards and are subject to data validation: https://biobank.ctsu.ox.ac.uk/~bbdatan/Data_cleaning_overall_doc_showcase_v1.pdf
6	General representativeness	The database population consists of volunteers aged 40-69years. We can expect that volunteers might have been more willing to participate if living closer to one of the 22 recruitment centres or if more interested in health issues compared to the general population. These aspects might have introduced an unavoidable bias in the cohort.
7	Data content /source coding	READ2, READ3, DM+D, ICD9, ICD10, OPCS3, OPCS4, ICD-O-3 are used.
8	Data Harmonisation	The data has been mapped to the OMOP CDM v5.4 and the OMOP standard vocabularies (SNOMED, RxNorm, LOINC). The format, structural and semantic conformance has been verified upon onboarding into the DARWIN EU® data network. No.
9	Quality control (database specific)	All UK BioBank data are provided already curated and each of the many datasets have specific curation algorithms and procedures. As always, primary care and hospital data, which come from real-world setting, need special attention regarding data quality. Please refer to the link below for

#	Section	Description
		specific details https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/primary_care_data.pdf https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/HospitalEpisodeStatistics.pdf
10	Linkage	The database contains liked data from death and cancer registry, GP and hospital for the participants.
11	Vital status	Linked to the national death registry.
12	Limitations	The UKBB source data will not be updated anymore and have no new records after December 2022. There is no day's supply information captured in the source. GP prescription data are available for 45% of the cohort. There is no information on dispensed medicines. GP laboratory tests and results are not available.
13	Main references	Hewitt J, Walters M,Padmanabhan S,Dawson J "Cohort profile of the UK Biobank: diagnosis and characteristics of cerebrovascular disease." BMJ open (2016): 27006341
14	Link to HMA-EMA catalogue and database webpage	HMA-EMA Catalogue entry: https://catalogues.ema.europa.eu/data-source/1111233 Website: https://www.ukbiobank.ac.uk/

ANNEX II. Fitness for use assessment

Data source justification for inclusion and key characteristics

The selected data sources met the criteria required to capture outcomes of interest and relevant data, enabling a patient-level characterisation of newly diagnosed individuals with cancer across different European settings and regions. The main criterion was a meaningful number of person counts for the population of interest (individuals with cancer) and outcomes (thromboembolic events) assessed at the feasibility stage for all data sources included in the study. Data sources were also selected based on European representativeness. Not all data sources had records of all outcomes of interest.

Additional criteria reflect other data quality domains assessed at the DARWIN EU® data partners onboarding stage. With every new release of the data partners OMOP CDM, the DARWIN EU® coordination centre also receives new results of the *CdmOnboarding*, *DashboardExport,* and *DataQualityDashboard* packages and assesses the quality of the data. No open quality issues related to the study population and outcomes were present for any of the data sources selected for the study.

Relevance was also assessed based on the previous research related to the study population of the outcome. Previously, IPCI, EBB, SIDIAP, CPRD GOLD, IQVIA DA Germany, UKBB, and EBB were used in studies with thromboembolic events as an outcome. (Ali et al., 2020, Mercadé-Besora et al., 2024, Li et al., 2022, Voss et al, 2023). CPRD GOLD, FinOMOP-THL, UKBB, SIDIAP, IPCI, and UKBB were used in studies repeated on individuals with cancer (Hagberg et al., 2023; Corby et al., 2024, Leinonen et al., 2017, Smith et al., 2024; Recalde et al, 2019; van Soest 2008; Chen et al., 2024). In addition to that, DK-DHR is a nationwide, fully representative data source that includes information from the National Patient Registry and the National Cancer Register.

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element, including justification where relevant
Study population	Objective 1 Inclusion criteria • First diagnosis of a selected cancer (index date) between 01/01/2016 and 31/12/2022 • Age ≥18 years at index date • Minimum 365 days of available history before index date • Index date ≥365 days prior to end of data availability of the data source Exclusion criteria • Diagnoses of multiple primary tumours at index date • History of cancer diagnosis ever before index date	• First diagnosis of a selected cancer Diagnosis of: • Deep vein thrombosis (DVT) • Pulmonary embolisms (PE) • Venous thromboembolism (VTE, composite of DVT and PE) • Pelvic venous thrombosis (PVT) • Splanchnic vein thrombosis (SVT), including hepatic and extra-hepatic vein thrombosis	Low/Medium/**High**

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element, including justification where relevant
	Objective 2 Inclusion criteria • Included in the study population of objective 1 • Occurrence of the outcome during follow-up	• Retinal vein thrombosis (RVT), including retinal central vein thrombosis • Disseminated intravascular coagulation (DIC)	
Treatment/ exposure	Not applicable.		Low/Medium/High
Comparator group (if relevant)	Not applicable.		Low/Medium/High
Outcomes (if relevant)	<u>All objectives</u> The specific thromboembolic events of interest are: • Deep vein thrombosis (DVT) • Pulmonary embolisms (PE) • Venous thromboembolism (VTE, composite of DVT and PE) • Pelvic venous thrombosis (PVT) • Splanchnic vein thrombosis (SVT), including hepatic and extra-hepatic vein thrombosis • Retinal vein thrombosis (RVT), including retinal central vein thrombosis • Disseminated intravascular coagulation (DIC) Each of the specific thromboembolic events is a binary outcome. Each outcome is assessed at any diagnosis position in the electronic health record and from any of the care settings in each data source. The list of concepts, from EUPAS1000000440 and based on SNOMED codes and aligned with previous studies that used OMOP CDM and VTE as an outcome (Burn et al., 2022), is provided in **Annex IV**.	Diagnosis of: • Deep vein thrombosis (DVT) • Pulmonary embolisms (PE) • Venous thromboembolism (VTE, composite of DVT and PE) • Pelvic venous thrombosis (PVT) • Splanchnic vein thrombosis (SVT), including hepatic and extra-hepatic vein thrombosis • Retinal vein thrombosis (RVT), including retinal central vein thrombosis • Disseminated intravascular coagulation (DIC)	Low/Medium/**High**
Covariates (including confounders if relevant)	<u>All Objectives</u> • Sex ○ Female/male	• Date of first diagnosis of selected cancer (index date)	Low/Medium/**High**

Design elements	Operational definition	Data elements for valid capture	Criticality of the quality of the element, including justification where relevant
	• Age (years) at date of the first cancer diagnosis, categorised into age groups: ○ 18–34 ○ 35–44 ○ 45–54 ○ 55–64 ○ 65–74 ○ 75–84 ○ ≥85 Age at index date will be calculated using January 1st of the year of birth as proxy for the actual birthday. Date/month is either not present or cannot be made available for governance reasons. If available, date is often set to first of the month for personal privacy. • Calendar year of the first cancer diagnosis, categorised into subperiods of the study period: ○ 2016–2019 ○ 2020-2022 The status of each covariate will be assessed at index date. Each covariate will be used to stratify the results.	• Sex • Age at index date	
Follow-up time (if relevant)	For both objectives, follow-up will start on the date of cancer diagnosis (index date) and end on the earliest of occurrence of the outcome, loss to follow-up, end of data availability, death, or end of the study period (31/12/2024).	• Date of first diagnosis of selected cancer (index date) • Date of thromboembolic event • Data of end of data availability • Death date	Low/Medium/**High**

EMA Data Quality Framework for EU medicines regulation: application to Real-World Data for more information (https://www.ema.europa.eu/system/files/documents/other/data-quality-framework-eu-medicines-regulation-application-real-world-data_en.pdf).

ANNEX III. Operational and reporting considerations

DATA MANAGEMENT

Data management

All data sources have previously mapped their data to the OMOP common data model. This enables the use of standardised analytics and using DARWIN EU® tools across the network, since the structure of the data and the terminology system is harmonised. The OMOP CDM was developed and maintained by the Observational Health Data Sciences and Informatics (OHDSI) initiative and is described in detail on the wiki page of the CDM: https://ohdsi.github.io/CommonDataModel and in The Book of OHDSI: http://book.ohdsi.org.

The analytic code for this study will be written in R and will use standardized analytics wherever possible. Each data partner will execute the study code against their data source containing patient-level data and then return the results (csv files), which will only contain aggregated data. The results from each of the contributing data sites will then be combined in tables and figures for the study report.

Data storage and protection

For this study, participants from various EU member states will process personal data from individuals that is collected in national/regional electronic health record data sources. Due to the sensitive nature of this personal medical data, it is important to be fully aware of ethical and regulatory aspects and to strive to take all reasonable measures to ensure compliance with ethical and regulatory issues on privacy.

All data sources used in this study are already used for pharmaco-epidemiological research and have a well-developed mechanism to ensure that European and local regulations dealing with ethical use of the data and adequate privacy control are adhered to. In agreement with these regulations, rather than combining person level data and performing only a central analysis, local analyses will be run, which generate non-identifiable aggregate summary results.

The output files are stored in the DARWIN EU® Remote Research Environment (RRE). These output files do not contain any data that allow identification of subjects included in the study. The RRE implements further security measures to ensure a high level of stored data protection to comply with the local implementation of the General Data Protection Regulation (GDPR) (EU) 679/20161 in the various member states.

QUALITY CONTROL

Data source quality control

When defining drug cohorts, non-systemic products will be excluded from the list of included codes summarised on the ingredient level.

When defining cohorts for indications, a systematic search of possible codes for inclusion will be identified using the *CodelistGenerator* R package (https://github.com/darwin-eu/CodelistGenerator). This package allows the user to define a search strategy and will use this to query the vocabulary tables of the OMOP common data model so as to find potentially relevant codes. In addition, the *CohortDiagnostics* (https://github.com/OHDSI/CohortDiagnostics) and *DrugExposureDiagnostics* (https://cran.r-project.org/web/packages/DrugExposureDiagnostics/index.html) R packages will be run, if needed, to assess the use of different codes across the data sources contributing to the study and identify any codes potentially omitted in error. The *DrugExposureDiagnostics* package evaluates ingredient-specific attributes and patterns in drug exposure records.

The study code will be based on DARWIN EU® R packages: *IncidencePrevalence* to estimate Incidence and Prevalence, *DrugUtilisation* to characterise the drug use, and *CohortCharacteristics* to characterise the cohort by indication. These packages will include numerous automated unit tests to ensure the validity of

the codes, alongside software peer review and user testing. The R package will be made publicly available via GitHub.

PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

A PDF report including an executive summary, and the specified tables and/or figures will be submitted to EMA by the DARWIN EU® CC upon completion of the study.

An interactive dashboard incorporating all the results (tables and figures) will be provided alongside the PDF report. The full set of underlying aggregated data used in the dashboard will also be made available, if requested.

ANNEX IV. List of stand-alone documents

Concepts to define individuals with cancer (populations of interest) are available in a stand-alone document (DARWIN_EU_P3_C3_005 _Cancer_phenotypes.xlsx). **Tables S1 to S7** include concepts used to define outcomes (thromboembolic events).

Table S1. List of concepts used to define deep vein thrombosis (DVT).

Concept ID	Concept Name	Domain	Vocabulary
762047	Acute bilateral thrombosis of subclavian veins	Condition	SNOMED
762148	Acute deep vein thrombosis of bilateral iliac veins	Condition	SNOMED
37169261	Acute deep vein thrombosis of bilateral lower limbs following procedure	Condition	SNOMED
37169249	Acute deep vein thrombosis of bilateral upper limbs following procedure	Condition	SNOMED
35616028	Acute deep vein thrombosis of left iliac vein	Condition	SNOMED
35615035	Acute deep vein thrombosis of left lower limb following procedure	Condition	SNOMED
35615031	Acute deep vein thrombosis of left upper limb following procedure	Condition	SNOMED
43531681	Acute deep vein thrombosis of lower limb	Condition	SNOMED
35616027	Acute deep vein thrombosis of right iliac vein	Condition	SNOMED
35615034	Acute deep vein thrombosis of right lower limb following procedure	Condition	SNOMED
35615030	Acute deep vein thrombosis of right upper limb following procedure	Condition	SNOMED
44782746	Acute deep venous thrombosis	Condition	SNOMED
44782751	Acute deep venous thrombosis of axillary vein	Condition	SNOMED
762008	Acute deep venous thrombosis of bilateral axillary veins	Condition	SNOMED
760875	Acute deep venous thrombosis of bilateral calves	Condition	SNOMED
765155	Acute deep venous thrombosis of bilateral ileofemoral veins	Condition	SNOMED
762017	Acute deep venous thrombosis of bilateral internal jugular veins	Condition	SNOMED
762417	Acute deep venous thrombosis of bilateral legs	Condition	SNOMED
761461	Acute deep venous thrombosis of bilateral pelvic veins	Condition	SNOMED
762020	Acute deep venous thrombosis of bilateral popliteal veins	Condition	SNOMED
765546	Acute deep venous thrombosis of bilateral tibial veins	Condition	SNOMED
762004	Acute deep venous thrombosis of both upper extremities	Condition	SNOMED
44782742	Acute deep venous thrombosis of calf	Condition	SNOMED
44782747	Acute deep venous thrombosis of femoral vein	Condition	SNOMED
762015	Acute deep venous thrombosis of ileofemoral vein of left leg	Condition	SNOMED
765541	Acute deep venous thrombosis of ileofemoral vein of right lower extremity	Condition	SNOMED
44782748	Acute deep venous thrombosis of iliofemoral vein	Condition	SNOMED
44782752	Acute deep venous thrombosis of internal jugular vein	Condition	SNOMED
762009	Acute deep venous thrombosis of left axillary vein	Condition	SNOMED
760876	Acute deep venous thrombosis of left calf	Condition	SNOMED
765540	Acute deep venous thrombosis of left femoral vein	Condition	SNOMED
765922	Acute deep venous thrombosis of left internal jugular vein	Condition	SNOMED

Concept ID	Concept Name	Domain	Vocabulary
762418	Acute deep venous thrombosis of left lower extremity	Condition	SNOMED
761462	Acute deep venous thrombosis of left pelvic vein	Condition	SNOMED
618482	Acute deep venous thrombosis of left peroneal vein	Condition	SNOMED
765537	Acute deep venous thrombosis of left upper extremity	Condition	SNOMED
44782767	Acute deep venous thrombosis of lower extremity as complication of procedure	Condition	SNOMED
44782761	Acute deep venous thrombosis of pelvic vein	Condition	SNOMED
762022	Acute deep venous thrombosis of politeal vein of right leg	Condition	SNOMED
44782743	Acute deep venous thrombosis of popliteal vein	Condition	SNOMED
762021	Acute deep venous thrombosis of popliteal vein of left leg	Condition	SNOMED
762010	Acute deep venous thrombosis of right axillary vein	Condition	SNOMED
760877	Acute deep venous thrombosis of right calf	Condition	SNOMED
762013	Acute deep venous thrombosis of right femoral vein	Condition	SNOMED
762018	Acute deep venous thrombosis of right internal jugular vein	Condition	SNOMED
762419	Acute deep venous thrombosis of right lower extremity	Condition	SNOMED
765229	Acute deep venous thrombosis of right pelvic vein	Condition	SNOMED
618681	Acute deep venous thrombosis of right peroneal vein	Condition	SNOMED
762005	Acute deep venous thrombosis of right upper extremity	Condition	SNOMED
44782745	Acute deep venous thrombosis of thigh	Condition	SNOMED
44782744	Acute deep venous thrombosis of tibial vein	Condition	SNOMED
762026	Acute deep venous thrombosis of tibial vein of left leg	Condition	SNOMED
765156	Acute deep venous thrombosis of tibial vein of right leg	Condition	SNOMED
44782421	Acute deep venous thrombosis of upper extremity	Condition	SNOMED
44782766	Acute deep venous thrombosis of upper extremity as complication of procedure	Condition	SNOMED
37171353	Acute ischemia of colon due to thrombosis of mesenteric vein	Condition	SNOMED
37170675	Acute ischemia of small intestine due to thrombosis of mesenteric vein	Condition	SNOMED
762048	Acute thrombosis of left subclavian vein	Condition	SNOMED
45757410	Acute thrombosis of mesenteric vein	Condition	SNOMED
762049	Acute thrombosis of right subclavian vein	Condition	SNOMED
36712892	Acute thrombosis of splenic vein	Condition	SNOMED
44782762	Acute thrombosis of subclavian vein	Condition	SNOMED
4179911	Axillary vein thrombosis	Condition	SNOMED
37109253	Bilateral acute deep vein thrombosis of femoral veins	Condition	SNOMED
618678	Bilateral acute deep venous thrombosis of peroneal veins	Condition	SNOMED
609003	Bilateral deep femoral vein thrombophlebitis	Condition	SNOMED
3179900	Bilateral deep vein thromboses	Condition	Nebraska Lexicon
40478951	Bilateral deep vein thrombosis of lower extremities	Condition	SNOMED
609002	Bilateral femoral vein thrombophlebitis	Condition	SNOMED

Concept ID	Concept Name	Domain	Vocabulary
608965	Bilateral iliac vein thrombophlebitis	Condition	SNOMED
1245776	Bilateral popliteal vein thrombophlebitis	Condition	SNOMED
609006	Bilateral tibial vein thrombophlebitis	Condition	SNOMED
4042396	Deep thrombophlebitis	Condition	SNOMED
4046884	Deep vein thrombosis of leg related to air travel	Condition	SNOMED
3655221	Deep vein thrombosis of lower extremity due to intravenous drug use	Condition	SNOMED
4133004	Deep venous thrombosis	Condition	SNOMED
761013	Deep venous thrombosis of bilateral pelvic veins	Condition	SNOMED
37163011	Deep venous thrombosis of calf	Condition	SNOMED
45773536	Deep venous thrombosis of femoropopliteal vein	Condition	SNOMED
763942	Deep venous thrombosis of left lower extremity	Condition	SNOMED
1075379	Deep venous thrombosis of left posterior tibial vein	Condition	SNOMED
761980	Deep venous thrombosis of left upper extremity	Condition	SNOMED
443537	Deep venous thrombosis of lower extremity	Condition	SNOMED
4133975	Deep venous thrombosis of pelvic vein	Condition	SNOMED
40480555	Deep venous thrombosis of peroneal vein	Condition	SNOMED
1075377	Deep venous thrombosis of posterior tibial vein	Condition	SNOMED
4322565	Deep venous thrombosis of profunda femoris vein	Condition	SNOMED
763941	Deep venous thrombosis of right lower extremity	Condition	SNOMED
1075378	Deep venous thrombosis of right posterior tibial vein	Condition	SNOMED
761928	Deep venous thrombosis of right upper extremity	Condition	SNOMED
4207899	Deep venous thrombosis of tibial vein	Condition	SNOMED
4028057	Deep venous thrombosis of upper extremity	Condition	SNOMED
193512	Embolism and thrombosis of the renal vein	Condition	SNOMED
435565	Embolism and thrombosis of the vena cava	Condition	SNOMED
4258295	Embolism from thrombosis of vein of distal lower extremity	Condition	SNOMED
40481089	Embolism from thrombosis of vein of lower extremity	Condition	SNOMED
40479840	Embolism from thrombosis of vein of thigh	Condition	SNOMED
4119760	Iliofemoral deep vein thrombosis	Condition	SNOMED
4124856	Inferior mesenteric vein thrombosis	Condition	SNOMED
608964	Left iliac vein thrombophlebitis	Condition	SNOMED
602592	Left peroneal vein thrombophlebitis	Condition	SNOMED
600938	Left subclavian vein thrombophlebitis	Condition	SNOMED
37164448	Lemierre syndrome	Condition	SNOMED
4281689	Phlegmasia alba dolens	Condition	SNOMED
4284538	Phlegmasia cerulea dolens	Condition	SNOMED
3185768	Popliteal vein thrombosis	Condition	Nebraska Lexicon

Concept ID	Concept Name	Domain	Vocabulary
4309333	Postoperative deep vein thrombosis	Condition	SNOMED
1245858	Postpartum acute deep vein thrombosis	Condition	SNOMED
46285905	Provoked deep vein thrombosis	Condition	SNOMED
608963	Right iliac vein thrombophlebitis	Condition	SNOMED
602583	Right peroneal vein thrombophlebitis	Condition	SNOMED
600939	Right subclavian vein thrombophlebitis	Condition	SNOMED
4033521	Splenic vein thrombosis	Condition	SNOMED
4055089	Superior mesenteric vein thrombosis	Condition	SNOMED
4230403	Thrombophlebitis of axillary vein	Condition	SNOMED
4069561	Thrombophlebitis of deep femoral vein	Condition	SNOMED
761831	Thrombophlebitis of deep vein of bilateral lower limbs	Condition	SNOMED
761830	Thrombophlebitis of deep vein of left lower limb	Condition	SNOMED
761808	Thrombophlebitis of deep vein of left upper limb	Condition	SNOMED
761832	Thrombophlebitis of deep vein of right lower limb	Condition	SNOMED
761809	Thrombophlebitis of deep vein of right upper limb	Condition	SNOMED
4221821	Thrombophlebitis of deep veins of lower extremity	Condition	SNOMED
440750	Thrombophlebitis of deep veins of upper extremities	Condition	SNOMED
4203618	Thrombophlebitis of femoropopliteal vein	Condition	SNOMED
4176614	Thrombophlebitis of iliac vein	Condition	SNOMED
764715	Thrombophlebitis of internal jugular vein	Condition	SNOMED
608904	Thrombophlebitis of left axillary vein	Condition	SNOMED
761821	Thrombophlebitis of left deep femoral vein	Condition	SNOMED
761819	Thrombophlebitis of left femoral vein	Condition	SNOMED
609000	Thrombophlebitis of left popliteal vein	Condition	SNOMED
609005	Thrombophlebitis of left tibial vein	Condition	SNOMED
4318407	Thrombophlebitis of mesenteric vein	Condition	SNOMED
608903	Thrombophlebitis of right axillary vein	Condition	SNOMED
761820	Thrombophlebitis of right deep femoral vein	Condition	SNOMED
761818	Thrombophlebitis of right femoral vein	Condition	SNOMED
609001	Thrombophlebitis of right popliteal vein	Condition	SNOMED
609004	Thrombophlebitis of right tibial vein	Condition	SNOMED
4205652	Thrombophlebitis of subclavian vein	Condition	SNOMED
4110339	Thrombophlebitis of the anterior tibial vein	Condition	SNOMED
4111868	Thrombophlebitis of the common iliac vein	Condition	SNOMED
4110343	Thrombophlebitis of the external iliac vein	Condition	SNOMED
439314	Thrombophlebitis of the femoral vein	Condition	SNOMED
4109877	Thrombophlebitis of the internal iliac vein	Condition	SNOMED

Concept ID	Concept Name	Domain	Vocabulary
4112171	Thrombophlebitis of the popliteal vein	Condition	SNOMED
4112172	Thrombophlebitis of the posterior tibial vein	Condition	SNOMED
4250765	Thrombophlebitis of tibial vein	Condition	SNOMED
42538533	Thrombosis of iliac vein	Condition	SNOMED
44811347	Thrombosis of internal jugular vein	Condition	SNOMED
765049	Thrombosis of left peroneal vein	Condition	SNOMED
4317289	Thrombosis of mesenteric vein	Condition	SNOMED
4203836	Thrombosis of subclavian vein	Condition	SNOMED
4175649	Thrombosis of the popliteal vein	Condition	SNOMED
4153353	Traumatic thrombosis of axillary vein	Condition	SNOMED
46285904	Unprovoked deep vein thrombosis	Condition	SNOMED
37163265	Venous thromboembolism due to thrombosis of vein of lower limb	Condition	SNOMED

Table S2. List of concepts used to define pulmonary embolism (PE).

Concept ID	Concept Name	Domain	Vocabulary
608954	Acute cor pulmonale due to septic pulmonary embolism	Condition	SNOMED
4120091	Acute massive pulmonary embolism	Condition	SNOMED
45768439	Acute pulmonary embolism	Condition	SNOMED
45768888	Acute pulmonary thromboembolism	Condition	SNOMED
762808	Infarction of lung due to embolus	Condition	SNOMED
40480461	Infarction of lung due to iatrogenic pulmonary embolism	Condition	SNOMED
4108681	Postoperative pulmonary embolus	Condition	SNOMED
37160752	Postoperative pulmonary thromboembolism	Condition	SNOMED
1244882	Pulmonary artery embolism due to foreign body	Condition	SNOMED
440417	Pulmonary embolism	Condition	SNOMED
37109911	Pulmonary embolism due to and following acute myocardial infarction	Condition	SNOMED
37016922	Pulmonary embolism on long-term anticoagulation therapy	Condition	SNOMED
43530605	Pulmonary embolism with pulmonary infarction	Condition	SNOMED
4253796	Pulmonary microemboli	Condition	SNOMED
4121618	Pulmonary thromboembolism	Condition	SNOMED
36713113	Saddle embolus of pulmonary artery	Condition	SNOMED
35615055	Saddle embolus of pulmonary artery with acute cor pulmonale	Condition	SNOMED
40479606	Septic pulmonary embolism	Condition	SNOMED
4119607	Subacute massive pulmonary embolism	Condition	SNOMED

Table S3. List of concepts used to define venous thromboembolism (VTE).

Concept ID	Concept Name	Domain	Vocabulary
762047	Acute bilateral thrombosis of subclavian veins	Condition	SNOMED
608954	Acute cor pulmonale due to septic pulmonary embolism	Condition	SNOMED
762148	Acute deep vein thrombosis of bilateral iliac veins	Condition	SNOMED
37169261	Acute deep vein thrombosis of bilateral lower limbs following procedure	Condition	SNOMED
37169249	Acute deep vein thrombosis of bilateral upper limbs following procedure	Condition	SNOMED
35616028	Acute deep vein thrombosis of left iliac vein	Condition	SNOMED
35615035	Acute deep vein thrombosis of left lower limb following procedure	Condition	SNOMED
35615031	Acute deep vein thrombosis of left upper limb following procedure	Condition	SNOMED
43531681	Acute deep vein thrombosis of lower limb	Condition	SNOMED
35616027	Acute deep vein thrombosis of right iliac vein	Condition	SNOMED
35615034	Acute deep vein thrombosis of right lower limb following procedure	Condition	SNOMED
35615030	Acute deep vein thrombosis of right upper limb following procedure	Condition	SNOMED
44782746	Acute deep venous thrombosis	Condition	SNOMED
44782751	Acute deep venous thrombosis of axillary vein	Condition	SNOMED
762008	Acute deep venous thrombosis of bilateral axillary veins	Condition	SNOMED
760875	Acute deep venous thrombosis of bilateral calves	Condition	SNOMED
765155	Acute deep venous thrombosis of bilateral ileofemoral veins	Condition	SNOMED
762017	Acute deep venous thrombosis of bilateral internal jugular veins	Condition	SNOMED
762417	Acute deep venous thrombosis of bilateral legs	Condition	SNOMED
761461	Acute deep venous thrombosis of bilateral pelvic veins	Condition	SNOMED
762020	Acute deep venous thrombosis of bilateral popliteal veins	Condition	SNOMED
765546	Acute deep venous thrombosis of bilateral tibial veins	Condition	SNOMED
762004	Acute deep venous thrombosis of both upper extremities	Condition	SNOMED
44782742	Acute deep venous thrombosis of calf	Condition	SNOMED
44782747	Acute deep venous thrombosis of femoral vein	Condition	SNOMED
762015	Acute deep venous thrombosis of ileofemoral vein of left leg	Condition	SNOMED
765541	Acute deep venous thrombosis of ileofemoral vein of right lower extremity	Condition	SNOMED
44782748	Acute deep venous thrombosis of iliofemoral vein	Condition	SNOMED
44782752	Acute deep venous thrombosis of internal jugular vein	Condition	SNOMED
762009	Acute deep venous thrombosis of left axillary vein	Condition	SNOMED
760876	Acute deep venous thrombosis of left calf	Condition	SNOMED
765540	Acute deep venous thrombosis of left femoral vein	Condition	SNOMED
765922	Acute deep venous thrombosis of left internal jugular vein	Condition	SNOMED
762418	Acute deep venous thrombosis of left lower extremity	Condition	SNOMED
761462	Acute deep venous thrombosis of left pelvic vein	Condition	SNOMED

Concept ID	Concept Name	Domain	Vocabulary
618482	Acute deep venous thrombosis of left peroneal vein	Condition	SNOMED
765537	Acute deep venous thrombosis of left upper extremity	Condition	SNOMED
44782767	Acute deep venous thrombosis of lower extremity as complication of procedure	Condition	SNOMED
44782761	Acute deep venous thrombosis of pelvic vein	Condition	SNOMED
762022	Acute deep venous thrombosis of politeal vein of right leg	Condition	SNOMED
44782743	Acute deep venous thrombosis of popliteal vein	Condition	SNOMED
762021	Acute deep venous thrombosis of popliteal vein of left leg	Condition	SNOMED
762010	Acute deep venous thrombosis of right axillary vein	Condition	SNOMED
760877	Acute deep venous thrombosis of right calf	Condition	SNOMED
762013	Acute deep venous thrombosis of right femoral vein	Condition	SNOMED
762018	Acute deep venous thrombosis of right internal jugular vein	Condition	SNOMED
762419	Acute deep venous thrombosis of right lower extremity	Condition	SNOMED
765229	Acute deep venous thrombosis of right pelvic vein	Condition	SNOMED
618681	Acute deep venous thrombosis of right peroneal vein	Condition	SNOMED
762005	Acute deep venous thrombosis of right upper extremity	Condition	SNOMED
44782745	Acute deep venous thrombosis of thigh	Condition	SNOMED
44782744	Acute deep venous thrombosis of tibial vein	Condition	SNOMED
762026	Acute deep venous thrombosis of tibial vein of left leg	Condition	SNOMED
765156	Acute deep venous thrombosis of tibial vein of right leg	Condition	SNOMED
44782421	Acute deep venous thrombosis of upper extremity	Condition	SNOMED
44782766	Acute deep venous thrombosis of upper extremity as complication of procedure	Condition	SNOMED
37171353	Acute ischemia of colon due to thrombosis of mesenteric vein	Condition	SNOMED
37170675	Acute ischemia of small intestine due to thrombosis of mesenteric vein	Condition	SNOMED
4120091	Acute massive pulmonary embolism	Condition	SNOMED
45768439	Acute pulmonary embolism	Condition	SNOMED
45768888	Acute pulmonary thromboembolism	Condition	SNOMED
762048	Acute thrombosis of left subclavian vein	Condition	SNOMED
45757410	Acute thrombosis of mesenteric vein	Condition	SNOMED
762049	Acute thrombosis of right subclavian vein	Condition	SNOMED
36712892	Acute thrombosis of splenic vein	Condition	SNOMED
44782762	Acute thrombosis of subclavian vein	Condition	SNOMED
4179911	Axillary vein thrombosis	Condition	SNOMED
37109253	Bilateral acute deep vein thrombosis of femoral veins	Condition	SNOMED
618678	Bilateral acute deep venous thrombosis of peroneal veins	Condition	SNOMED
609003	Bilateral deep femoral vein thrombophlebitis	Condition	SNOMED
3179900	Bilateral deep vein thromboses	Condition	Nebraska Lexicon

Concept ID	Concept Name	Domain	Vocabulary
40478951	Bilateral deep vein thrombosis of lower extremities	Condition	SNOMED
609002	Bilateral femoral vein thrombophlebitis	Condition	SNOMED
608965	Bilateral iliac vein thrombophlebitis	Condition	SNOMED
1245776	Bilateral popliteal vein thrombophlebitis	Condition	SNOMED
609006	Bilateral tibial vein thrombophlebitis	Condition	SNOMED
44782732	Chronic pulmonary embolism	Condition	SNOMED
45768887	Chronic pulmonary thromboembolism	Condition	SNOMED
45771016	Chronic pulmonary thromboembolism without pulmonary hypertension	Condition	SNOMED
4042396	Deep thrombophlebitis	Condition	SNOMED
4046884	Deep vein thrombosis of leg related to air travel	Condition	SNOMED
3655221	Deep vein thrombosis of lower extremity due to intravenous drug use	Condition	SNOMED
4133004	Deep venous thrombosis	Condition	SNOMED
761013	Deep venous thrombosis of bilateral pelvic veins	Condition	SNOMED
37163011	Deep venous thrombosis of calf	Condition	SNOMED
45773536	Deep venous thrombosis of femoropopliteal vein	Condition	SNOMED
763942	Deep venous thrombosis of left lower extremity	Condition	SNOMED
1075379	Deep venous thrombosis of left posterior tibial vein	Condition	SNOMED
761980	Deep venous thrombosis of left upper extremity	Condition	SNOMED
443537	Deep venous thrombosis of lower extremity	Condition	SNOMED
4133975	Deep venous thrombosis of pelvic vein	Condition	SNOMED
40480555	Deep venous thrombosis of peroneal vein	Condition	SNOMED
1075377	Deep venous thrombosis of posterior tibial vein	Condition	SNOMED
4322565	Deep venous thrombosis of profunda femoris vein	Condition	SNOMED
763941	Deep venous thrombosis of right lower extremity	Condition	SNOMED
1075378	Deep venous thrombosis of right posterior tibial vein	Condition	SNOMED
761928	Deep venous thrombosis of right upper extremity	Condition	SNOMED
4207899	Deep venous thrombosis of tibial vein	Condition	SNOMED
4028057	Deep venous thrombosis of upper extremity	Condition	SNOMED
193512	Embolism and thrombosis of the renal vein	Condition	SNOMED
435565	Embolism and thrombosis of the vena cava	Condition	SNOMED
4258295	Embolism from thrombosis of vein of distal lower extremity	Condition	SNOMED
40481089	Embolism from thrombosis of vein of lower extremity	Condition	SNOMED
40479840	Embolism from thrombosis of vein of thigh	Condition	SNOMED
4119760	Iliofemoral deep vein thrombosis	Condition	SNOMED
43530934	Induced termination of pregnancy complicated by pulmonary embolism	Condition	SNOMED
762808	Infarction of lung due to embolus	Condition	SNOMED

Concept ID	Concept Name	Domain	Vocabulary
40480461	Infarction of lung due to iatrogenic pulmonary embolism	Condition	SNOMED
4124856	Inferior mesenteric vein thrombosis	Condition	SNOMED
608964	Left iliac vein thrombophlebitis	Condition	SNOMED
602592	Left peroneal vein thrombophlebitis	Condition	SNOMED
600938	Left subclavian vein thrombophlebitis	Condition	SNOMED
37164448	Lemierre syndrome	Condition	SNOMED
4281689	Phlegmasia alba dolens	Condition	SNOMED
4284538	Phlegmasia cerulea dolens	Condition	SNOMED
3185768	Popliteal vein thrombosis	Condition	Nebraska Lexicon
4309333	Postoperative deep vein thrombosis	Condition	SNOMED
4108681	Postoperative pulmonary embolus	Condition	SNOMED
37160752	Postoperative pulmonary thromboembolism	Condition	SNOMED
1245858	Postpartum acute deep vein thrombosis	Condition	SNOMED
46285905	Provoked deep vein thrombosis	Condition	SNOMED
1244882	Pulmonary artery embolism due to foreign body	Condition	SNOMED
440417	Pulmonary embolism	Condition	SNOMED
37109911	Pulmonary embolism due to and following acute myocardial infarction	Condition	SNOMED
3655209	Pulmonary embolism due to and following ectopic pregnancy	Condition	SNOMED
3655210	Pulmonary embolism due to and following molar pregnancy	Condition	SNOMED
37016922	Pulmonary embolism on long-term anticoagulation therapy	Condition	SNOMED
43530605	Pulmonary embolism with pulmonary infarction	Condition	SNOMED
4253796	Pulmonary microemboli	Condition	SNOMED
4121618	Pulmonary thromboembolism	Condition	SNOMED
4236271	Recurrent pulmonary embolism	Condition	SNOMED
608963	Right iliac vein thrombophlebitis	Condition	SNOMED
602583	Right peroneal vein thrombophlebitis	Condition	SNOMED
600939	Right subclavian vein thrombophlebitis	Condition	SNOMED
36713113	Saddle embolus of pulmonary artery	Condition	SNOMED
35615055	Saddle embolus of pulmonary artery with acute cor pulmonale	Condition	SNOMED
40479606	Septic pulmonary embolism	Condition	SNOMED
4033521	Splenic vein thrombosis	Condition	SNOMED
4119607	Subacute massive pulmonary embolism	Condition	SNOMED
4055089	Superior mesenteric vein thrombosis	Condition	SNOMED
4230403	Thrombophlebitis of axillary vein	Condition	SNOMED
4069561	Thrombophlebitis of deep femoral vein	Condition	SNOMED
761831	Thrombophlebitis of deep vein of bilateral lower limbs	Condition	SNOMED

Concept ID	Concept Name	Domain	Vocabulary
761830	Thrombophlebitis of deep vein of left lower limb	Condition	SNOMED
761808	Thrombophlebitis of deep vein of left upper limb	Condition	SNOMED
761832	Thrombophlebitis of deep vein of right lower limb	Condition	SNOMED
761809	Thrombophlebitis of deep vein of right upper limb	Condition	SNOMED
4221821	Thrombophlebitis of deep veins of lower extremity	Condition	SNOMED
440750	Thrombophlebitis of deep veins of upper extremities	Condition	SNOMED
4203618	Thrombophlebitis of femoropopliteal vein	Condition	SNOMED
4176614	Thrombophlebitis of iliac vein	Condition	SNOMED
764715	Thrombophlebitis of internal jugular vein	Condition	SNOMED
608904	Thrombophlebitis of left axillary vein	Condition	SNOMED
761821	Thrombophlebitis of left deep femoral vein	Condition	SNOMED
761819	Thrombophlebitis of left femoral vein	Condition	SNOMED
609000	Thrombophlebitis of left popliteal vein	Condition	SNOMED
609005	Thrombophlebitis of left tibial vein	Condition	SNOMED
4318407	Thrombophlebitis of mesenteric vein	Condition	SNOMED
608903	Thrombophlebitis of right axillary vein	Condition	SNOMED
761820	Thrombophlebitis of right deep femoral vein	Condition	SNOMED
761818	Thrombophlebitis of right femoral vein	Condition	SNOMED
609001	Thrombophlebitis of right popliteal vein	Condition	SNOMED
609004	Thrombophlebitis of right tibial vein	Condition	SNOMED
4205652	Thrombophlebitis of subclavian vein	Condition	SNOMED
4110339	Thrombophlebitis of the anterior tibial vein	Condition	SNOMED
4111868	Thrombophlebitis of the common iliac vein	Condition	SNOMED
4110343	Thrombophlebitis of the external iliac vein	Condition	SNOMED
439314	Thrombophlebitis of the femoral vein	Condition	SNOMED
4109877	Thrombophlebitis of the internal iliac vein	Condition	SNOMED
4112171	Thrombophlebitis of the popliteal vein	Condition	SNOMED
4112172	Thrombophlebitis of the posterior tibial vein	Condition	SNOMED
4250765	Thrombophlebitis of tibial vein	Condition	SNOMED
42538533	Thrombosis of iliac vein	Condition	SNOMED
44811347	Thrombosis of internal jugular vein	Condition	SNOMED
765049	Thrombosis of left peroneal vein	Condition	SNOMED
4317289	Thrombosis of mesenteric vein	Condition	SNOMED
4203836	Thrombosis of subclavian vein	Condition	SNOMED
4175649	Thrombosis of the popliteal vein	Condition	SNOMED
4153353	Traumatic thrombosis of axillary vein	Condition	SNOMED

Concept ID	Concept Name	Domain	Vocabulary
46285904	Unprovoked deep vein thrombosis	Condition	SNOMED
37163265	Venous thromboembolism due to thrombosis of vein of lower limb	Condition	SNOMED

Table S4. List of concepts used to define pelvic vein thrombosis (PVT) (concept sets included all descendants of listed concepts).

Concept ID	Concept Name	Domain	Vocabulary
762148	Acute deep vein thrombosis of bilateral iliac veins	Condition	SNOMED
35616028	Acute deep vein thrombosis of left iliac vein	Condition	SNOMED
35616027	Acute deep vein thrombosis of right iliac vein	Condition	SNOMED
765155	Acute deep venous thrombosis of bilateral ileofemoral veins	Condition	SNOMED
761461	Acute deep venous thrombosis of bilateral pelvic veins	Condition	SNOMED
762015	Acute deep venous thrombosis of ileofemoral vein of left leg	Condition	SNOMED
765541	Acute deep venous thrombosis of ileofemoral vein of right lower extremity	Condition	SNOMED
761462	Acute deep venous thrombosis of left pelvic vein	Condition	SNOMED
44782761	Acute deep venous thrombosis of pelvic vein	Condition	SNOMED
765229	Acute deep venous thrombosis of right pelvic vein	Condition	SNOMED
608965	Bilateral iliac vein thrombophlebitis	Condition	SNOMED
765152	Chronic deep vein thrombosis of bilateral iliac veins	Condition	SNOMED
35616026	Chronic deep vein thrombosis of left iliac vein	Condition	SNOMED
761439	Chronic deep vein thrombosis of left pelvic vein	Condition	SNOMED
46271548	Chronic deep vein thrombosis of pelvic vein	Condition	SNOMED
35616025	Chronic deep vein thrombosis of right iliac vein	Condition	SNOMED
761441	Chronic deep vein thrombosis of right pelvic vein	Condition	SNOMED
765542	Chronic deep venous thrombosis of bilateral ileofemoral veins	Condition	SNOMED
761440	Chronic deep venous thrombosis of bilateral pelvic veins	Condition	SNOMED
765543	Chronic deep venous thrombosis of left ileofemoral vein	Condition	SNOMED
762016	Chronic deep venous thrombosis of right ileofemoral vein	Condition	SNOMED
761013	Deep venous thrombosis of bilateral pelvic veins	Condition	SNOMED
4133975	Deep venous thrombosis of pelvic vein	Condition	SNOMED
608964	Left iliac vein thrombophlebitis	Condition	SNOMED
4285751	Pelvic thrombophlebitis in puerperium	Condition	SNOMED
608963	Right iliac vein thrombophlebitis	Condition	SNOMED
4176614	Thrombophlebitis of iliac vein	Condition	SNOMED
4317290	Thrombophlebitis of pelvic vein	Condition	SNOMED
4111868	Thrombophlebitis of the common iliac vein	Condition	SNOMED
4110343	Thrombophlebitis of the external iliac vein	Condition	SNOMED
4109877	Thrombophlebitis of the internal iliac vein	Condition	SNOMED

Concept ID	Concept Name	Domain	Vocabulary
42538533	Thrombosis of iliac vein	Condition	SNOMED
4319327	Thrombosis of pelvic vein	Condition	SNOMED

Table S5. List of concepts used to splanchnic vein thrombosis (SVT).

Concept ID	Concept Name	Domain	Vocabulary
37171353	Acute ischemia of colon due to thrombosis of mesenteric vein	Condition	SNOMED
37170675	Acute ischemia of small intestine due to thrombosis of mesenteric vein	Condition	SNOMED
45757410	Acute thrombosis of mesenteric vein	Condition	SNOMED
36712892	Acute thrombosis of splenic vein	Condition	SNOMED
196715	Budd-Chiari syndrome	Condition	SNOMED
4301208	Hepatic vein thrombosis	Condition	SNOMED
4124856	Inferior mesenteric vein thrombosis	Condition	SNOMED
4092406	Portal thrombophlebitis	Condition	SNOMED
199837	Portal vein thrombosis	Condition	SNOMED
4033521	Splenic vein thrombosis	Condition	SNOMED
4055089	Superior mesenteric vein thrombosis	Condition	SNOMED
4318407	Thrombophlebitis of mesenteric vein	Condition	SNOMED
4317289	Thrombosis of mesenteric vein	Condition	SNOMED

Table S6. List of concepts used to define retinal vein thrombosis (RVT).

Concept ID	Concept Name	Domain	Vocabulary
437544	Arterial retinal branch occlusion	Condition	SNOMED
3657106	Bilateral occlusion of branch retinal arteries	Condition	SNOMED
37310623	Bilateral occlusion of central retinal arteries	Condition	SNOMED
37169454	Bilateral vascular occlusion of retina of eyes	Condition	SNOMED
4336004	Branch macular artery occlusion	Condition	SNOMED
4339013	Branch retinal vein occlusion with macular edema	Condition	SNOMED
4334248	Branch retinal vein occlusion with neovascularization	Condition	SNOMED
4199035	Branch retinal vein occlusion with no neovascularization	Condition	SNOMED
437540	Central retinal artery occlusion	Condition	SNOMED
313761	Central retinal vein occlusion	Condition	SNOMED
4208221	Central retinal vein occlusion–ischemic	Condition	SNOMED
4208222	Central retinal vein occlusion–non-ischemic	Condition	SNOMED
4339010	Central retinal vein occlusion with macular edema	Condition	SNOMED
4334246	Central retinal vein occlusion with neovascularization	Condition	SNOMED
4338905	Cilioretinal artery occlusion	Condition	SNOMED
42535735	Combined occlusion by thrombus of retinal artery and retinal vein	Condition	SNOMED

Concept ID	Concept Name	Domain	Vocabulary
4102317	Incipient occlusion of retinal vein	Condition	SNOMED
4083482	Macular branch retinal vein occlusion	Condition	SNOMED
37206377	Occlusion of branch of retinal vein of left eye	Condition	SNOMED
37206378	Occlusion of branch of retinal vein of right eye	Condition	SNOMED
37206381	Occlusion of central retinal vein of left eye	Condition	SNOMED
37206380	Occlusion of central retinal vein of right eye	Condition	SNOMED
36713329	Occlusion of left branch retinal artery	Condition	SNOMED
37207955	Occlusion of left central retinal artery	Condition	SNOMED
3657873	Occlusion of left cilioretinal artery	Condition	SNOMED
36713330	Occlusion of right branch retinal artery	Condition	SNOMED
37207895	Occlusion of right central retinal artery	Condition	SNOMED
3657872	Occlusion of right cilioretinal artery	Condition	SNOMED
4334245	Retinal artery occlusion	Condition	SNOMED
4324290	Retinal phlebitis	Condition	SNOMED
440392	Retinal vascular occlusion	Condition	SNOMED
3183076	Right branch retinal artery occlusion	Condition	Nebraska Lexicon
4216561	Thrombophlebitis of retinal vein	Condition	SNOMED
4187790	Thrombosis of retinal vein	Condition	SNOMED
3657847	Vascular occlusion of retina of left eye	Condition	SNOMED
3657848	Vascular occlusion of retina of right eye	Condition	SNOMED
312622	Venous retinal branch occlusion	Condition	SNOMED

Table S7. List of concepts used to define disseminated intravascular coagulation (DIC).

Concept ID	Concept Name	Domain	Vocabulary
37117819	Acquired purpura fulminans	Condition	SNOMED
436093	Disseminated intravascular coagulation	Condition	SNOMED
4028488	Purpura fulminans	Condition	SNOMED

ANNEX V. Glossary

Additional definitions are available in the EMA Glossary of terms https://www.ema.europa.eu/en/about-us/glossaries.

Aggregated Data

Data collected and combined from multiple sources to generate summary information, typically anonymised.

Benefit-Risk Assessment

Evaluation of the positive therapeutic effects of a medicine compared to its risks (e.g. side effects).

Common Data Model (CDM)

A standardized data structure that enables data from multiple sources to be harmonized, making analysis consistent and reproducible. DARWIN EU® utilises the OMOP CDM maintained by the OHDSI community.

Complex Studies (C3)

Studies requiring the development or customisation of specific study designs, protocols, and Statistical Analysis Plans (SAPs), with extensive collection or extraction of data. Examples include etiological studies measuring the strength and determinants of an association between an exposure and the occurrence of a health outcome in a defined population considering sources of bias, potential confounding factors, and effect modifiers.

Coordination Centre (CC)

The central hub responsible for managing and overseeing the activities within DARWIN EU®. It is based at Erasmus University Medical Centre in Rotterdam, the Netherlands.

Data Access

The process of obtaining permission to use specific datasets for regulatory or scientific studies.

Data Quality Framework

A set of standards and procedures to ensure accuracy, completeness, timeliness, and consistency of data used in DARWIN EU®.

Data Source

A database or repository of structured health-related data, such as electronic health records (EHRs), insurance claims, or registries.

DARWIN EU®

The European Medicines Agency's (EMA) federated network of real-world data sources designed to generate evidence to support regulatory decision-making.

EMA (European Medicines Agency)

The regulatory body responsible for the evaluation and supervision of medicinal products in the EU, overseeing DARWIN EU®.

Evidence Generation

The process of analysing real-world data to produce scientific information that can inform healthcare or regulatory decisions.

Federated Network

A data infrastructure where data remain at their original location but can be analysed in a harmonised way across multiple partners using a common model and tools.

GDPR (General Data Protection Regulation)

The EU regulation governing the protection of personal data and privacy, crucial to how DARWIN EU® handles health data.

Health Technology Assessment (HTA)

A systematic evaluation of properties and impacts of health technology, often using DARWIN EU® data to support assessments.

Metadata

Descriptive information about a data source (e.g. its content, quality, and structure), essential for identifying relevant databases in DARWIN EU® studies.

Off-the-Shelf Studies (OTS)

Studies for which a standard protocol per study/analysis type and standardised analytics may be developed and applied or adapted, typically relating to a descriptive research question. This includes studies on disease epidemiology, for example, the estimation of the prevalence or incidence of health outcomes in defined time periods and population groups, or drug utilisation studies at the population or patient level.

OHDSI (Observational Health Data Sciences and Informatics)

An open-science collaborative community that develops tools and standards (including the OMOP CDM) to enable large-scale analytics of observational health data. OHDSI provides the technical and scientific foundation for DARWIN EU®'s analytical ecosystem.

Patient-Level Data

Data related to individuals, de-identified, used for longitudinal or detailed analyses.

OMOP (Observational Medical Outcomes Partnership)

A common data model (CDM) that standardises the structure and content of observational healthcare data, enabling systematic analysis across disparate datasets. DARWIN EU® uses the OMOP CDM to ensure interoperability and consistency in real-world evidence generation.

Real-World Data (RWD)

Data relating to individual health status or healthcare delivery that is collected from routine clinical practice rather than from randomised controlled trials.

Real-World Evidence (RWE)

Clinical evidence derived from the analysis of RWD, used to inform decisions by regulators, payers, or clinicians.

Regulatory Decision-Making

The process by which authorities like EMA assess data to authorise, monitor, or modify the use of medicines in the EU.

Routine Repeated Studies (RR)

Studies that are either Off-the-Shelf or Complex studies repeated on a regular basis, following the same protocol and study code, but with updated data and/or different data partners.

Study Protocol

A detailed plan describing how a specific real-world study will be conducted, including objectives, design, data sources, and analyses.

Very Complex Studies (C4)

Studies which cannot rely only on electronic health care databases, or which would require complex methodological work, for example, due to the occurrence of events that cannot be defined by existing diagnosis codes, including events that do not yet have a diagnosis code, where it may be necessary to combine a diagnosis code with other data such as results of laboratory investigations. These studies might require the collection of data prospectively, or the inclusion of new (not previously onboarded) data sources.