Existing data sources in clinical epidemiology: the Scandinavian Thrombosis and Cancer Cohort

Hilde Jensvoll1,2
Marianne T Severinsen1,4
Jens Hammerstrøm5
Sigrid K Brækkan1,2
Søren R Kristensen6,8
Suzanne C Cannegieter7
Kristine Blix1,2
Anne Tjønneland8
Frits R Rosendaal1,2
Olga Dziewiecka1
Kim Overvad1,10,11
Inger Anne Næs1,2
John-Bjarne Hansen1,2

1Department of Clinical Medicine, KG Jebsen – Thrombosis Research and Expertise Center (TREC), UIT – The Arctic University of Norway, 2Division of Internal Medicine, University Hospital of North Norway, Tromsø, Norway; 3Department of Hematology, Aalborg University Hospital, 4Department of Clinical Medicine, Aalborg University, Aalborg, Denmark; 5Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway; 6Department of Clinical Biochemistry, Aalborg University Hospital, Aalborg, Denmark; 7Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands; 8Department of Medicine, Aarhus University Hospital, Aarhus, Denmark; 9Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands; 10Department of Clinical Medicine, Aarhus University, Aarhus, Denmark; 11Department of Clinical Medicine, Aalborg University Hospital, Aalborg, Denmark; 12Department of Public Health, Section for Epidemiology, Aarhus University, Aarhus, Denmark; 13Department of Hematology, Trondheim University Hospital, Trondheim, Norway

Correspondence: Hilde Jensvoll
Department of Clinical Medicine, KG Jebsen – Thrombosis Research and Expertise Center, UIT – The Arctic University of Norway, PO Box 6050, Langnes, 9037 Tromsø, Norway
Tel +47 7762 0992
Email hilde.jensvoll@uit.no

Background: Although venous thromboembolism (VTE) is a known common complication in cancer patients, there is limited knowledge on patient-related and cancer-specific risk factors in the general population. The Scandinavian Thrombosis and Cancer (STAC) Cohort was established by merging individual data from three large Scandinavian cohorts (The Tromsø Study, the second Nord-Trøndelag Health Study, and the Danish Diet, Cancer and Health Study). Here, we present the profile of the STAC cohort and provide age-specific incidence rates of VTE and cancer.

Methods: The STAC cohort includes 144,952 subjects aged 19–101 years without previous VTE or cancer. Baseline information collected in 1993–1997 included physical examination, self-administered questionnaires, and blood samples. Validated VTE events and cancer diagnoses were registered up to 2007–2012.

Results: There were 2,444 VTE events (1.4 per 1,000 person-years [PY]) during follow-up, and the incidence increased exponentially from 0.3 per 1,000 PY in subjects aged 20–29 years to 6.4 per 1,000 PY in subjects aged 80+. Overall, 51% of the VTE events were provoked, and cancer was the most common provoking factor (19%), followed by immobilization and surgery (both 15%). In total, 19,757 subjects developed cancer during follow-up (9.8 per 1,000 PY), and the 5-year age-specific incidence rates of cancer were coherent with corresponding rates from the Norwegian Cancer Registry.

Conclusion: The STAC cohort will provide a unique opportunity to explore the epidemiology and impact of genetic and environmental patient-related and cancer-specific risk factors for VTE in the general population.

Keywords: venous thromboembolism, incidence rates, person-years, pulmonary embolism, population-based cohort, prospective, cancer

Introduction
Venous thromboembolism (VTE), including deep venous thrombosis (DVT) and pulmonary embolism (PE), is a common complication and a predictor of poor survival in cancer patients.1 The reported annual incidence proportions of VTE in cancer patients vary substantially between 0.5% and 20%,2 and cancer is associated with 20%–25% of all incident VTE events.3–5 Thromboprophylaxis is not recommended to all cancer patients due to the heterogeneity of VTE risk in various cancer types.6 These considerations emphasize the need to identify risk factors for cancer-related VTE in order to recognize high-risk patients who may benefit from thromboprophylaxis.

Discharge registries may represent large and valuable databases when exploring the risk of VTE in cancer patients. However, a recent Danish study reported that the positive predictive value of VTE discharge diagnoses was only 75%, indicating that the
registry data without any further VTE validation should be used with caution in medical research.7 We have established the Scandinavian Thrombosis and Cancer (STAC) Cohort, which is a large population-based cohort study, including individual data from three well-established prospective cohort studies in Norway and Denmark. These cohorts include thorough validation of symptomatic VTE events and the information regarding cancer exposure is obtained from complete national cancer registries.

The overall aim of the STAC cohort is to investigate the impact of cancer on the risk of VTE by exploring genetic and environmental patient-related and cancer-specific risk factors in order to identify high-risk patients who may benefit from thromboprophylaxis. Age-specific absolute and relative risks of VTE and related complications (mortality and recurrent VTE) in cancer patients will be calculated and compared with the general population. The risk will be assessed for overt and occult cancer, within various cancer sites, across cancer stages, across histologic grades, and at different time intervals after cancer diagnosis. The impact of established and novel genetic variants will be investigated by targeted gene sequencing in a large case–control study derived from STAC.

Here, we present the overall design and methods of the STAC cohort. We will also provide age-specific incidence rates (IRs) of VTE and cancer in the three cohorts and in the STAC cohort. The age-specific IRs of cancer in the STAC cohort will be compared with expected IRs based on data from the Cancer Registry of Norway.

Methods

Study population
The STAC cohort includes individual data from three large population-based cohort studies with enrollment in 1993–1997, ie, the fourth survey of the Tromsø Study (Tromsø 4) in Norway, the second Nord-Trøndelag Health (HUNT 2) Study in Norway, and the Diet, Cancer and Health (DCH) Study in Denmark.

The Tromsø Study8 consists of six health surveys conducted in Tromsø, which is the largest city in Northern Norway (about 70,000 inhabitants). The Tromsø Study was established to investigate the epidemiology of cardiovascular diseases and a broad spectrum of other chronic diseases. In the Tromsø 4 Study, conducted in 1994–1995, all inhabitants ≥25 years in the municipality of Tromsø were invited, of whom 27,158 attended (77% of the eligible population).

The HUNT Study9 has conducted three health surveys in the Nord-Trøndelag County, which is located geographically in the middle of Norway and includes six small towns (between 12,000 and 21,000 inhabitants) combined with sparsely populated rural areas. The HUNT 2 Study in 1995–1997 was also designed to investigate cardiovascular disease and other chronic diseases in accordance with national health priorities. All residents ≥20 years of Nord-Trøndelag County were invited of which 65,237 participated (69% of the eligible population).

The DCH Study10 was initiated in 1993–1997 and designed to investigate the associations between diet, lifestyle factors, cancer, and other chronic diseases. In the DCH Study, all inhabitants aged 50–64 years, who were born in Denmark, living in the urban areas of Copenhagen and Aarhus, and without a previous diagnosis of cancer registered in the Danish Cancer Registry were invited, of whom 57,054 participated (35% of those invited).

In the two Norwegian cohorts, subjects with a cancer diagnosis (registered in the Norwegian Cancer Registry) prior to enrollment were excluded before merging the three cohorts. Moreover, subjects with known VTE prior to baseline were excluded in all cohorts prior to merging. In the DCH Study, discharge diagnoses were used as basis for exclusions of subjects with prior VTE, whereas in the two Norwegian cohorts, subjects with prior VTE were detected through the medical records review. After merging, the study population counted 144,952 subjects aged 19–101 years who were followed from the date of inclusion (1993–1997) to the end of follow-up (2007–2012).

Figure 1 summarizes the number of invited and the number of subjects who attended in the three cohorts and the total number of participants in the STAC cohort.

All included study participants gave their informed written consent, and respective regional committees for research ethics in Norway and Denmark approved the cohort studies and the collaboration study.

Baseline measurements
Baseline information was obtained by physical examination, self-administered questionnaires, and nonfasting blood samples. Height and weight were measured with light clothing and without shoes, and body mass index was calculated as weight in kilograms divided by the height in meters squared (kg/m²). Trained personnel recorded blood pressure using an automatic device. Three measurements were performed after 2 minutes at rest in the Tromsø 4 Study and the HUNT 2 Study, and the average of the two last readings were chosen. In the DCH Study, blood pressure was measured twice after 5 minutes at rest, and the lowest reading was used.

Information regarding education level, smoking habits (never/former/current and duration in years), alcohol...
consumption (units per week), physical activity, diabetes mellitus, and use of antihypertensives was collected from the self-administered questionnaires. The level of education was categorized into basic schooling (7–10 years), upper secondary education (high school/vocational school), and tertiary education (college/university). In the STAC cohort, hard physical activity was defined as equal to or more than 1 hour of sports/physical activity per week that caused sweating or breathlessness. Detailed information about the baseline variables in each cohort is described elsewhere.8–10

Identification and validation of cancer diagnosis
Cancer registration is mandatory by law in both Norway and Denmark. Individual data from the three cohorts were linked to the Norwegian and Danish cancer registries by the use of the unique national civil registration number, which is assigned to every Norwegian and Danish citizen at birth. The obtained information included date of cancer diagnosis, location of the disease (ICD10 codes C00-96), histological grade (ICO-3), and cancer stage (localized, regional, distant, or unknown). Subjects with nonmelanoma skin cancers (ICD10 C44) were regarded as cancer-free. Both cancer registries are considered complete and valid, and evaluations have reported 98.8% completeness in Norway and 95%–98% in Denmark, with 94% and 93% microscopically confirmed diagnoses, respectively.11,12

Identification and validation of venous thromboembolic events
Only first time, objectively confirmed symptomatic VTE events have been included in the STAC cohort. Trained personnel have reviewed the medical records for each potential case of VTE in all three included cohorts, and both inpatients and outpatients have been included.

In the Tromsø 4 Study,13 VTE events during follow-up were identified by searching the hospital discharge diagnosis registry, the radiology procedure registry, and the autopsy registry at the University Hospital of North Norway, which is the only hospital serving the municipality of Tromsø. The four validation criteria were symptomatic VTE combined with objective confirmatory tests (ultrasound, venography, computed tomography (CT) scan, perfusion–ventilation scan, pulmonary angiography, or autopsy), which resulted in a VTE diagnosis that required treatment.

In the HUNT 2 Study, a VTE endpoint registry, the TROL Study was established.14 VTE events during follow-up were identified by the use of the hospital discharge diagnosis
registry and the radiology procedure registry at the two local hospitals (Levanger and Namsos in Nord-Trøndelag County, Norway) and at the tertiary-care center of the region, St Olav’s Hospital. The validation criteria included symptomatic VTE events that required treatment and confirmation by objective diagnostic tests (ultrasound, venography, CT scan, or perfusion–ventilation scan).

In the DCH Study,15 incident VTE events were found by the use of the participants’ civil registration numbers and linkage to the Danish National Patient Registry and the Danish National Death Registry. A verified VTE diagnosis required typical clinical symptoms combined with a confirmatory diagnostic test (ultrasound, venography, CT scan, perfusion–ventilation scan, echocardiography, or autopsy). Detailed information on the identification and validation of the VTE events in each cohort study has been previously described.13–15

The VTE events in all three cohorts were classified as DVT or PE, and concurrent DVT and PE were registered as PE. The VTE events were further classified as provoked and unprovoked VTE. A provoked event in the STAC cohort required that one or more of the following provoking factors were present: overt cancer at the time of the VTE event, recent trauma or surgery, marked immobilization, travel, acute medical condition, or other risk factors described by the physician in the medical record such as intravascular catheters. There were some differences in the definitions of provoking factors in the three included cohorts.7,13,14 Provoking factors were registered up to 8 weeks prior to the VTE event in the Tromsø Study, while it was registered up to 3 months ahead of the event in the other two studies. Additionally, the minimum duration of the travel variable defined to be a provoking factor varied between 4 and 8 hours in the three cohorts. Furthermore, the HUNT 2 Study did not include acute medical conditions as individual factors, but instead registered prolonged bed rest (immobility) due to acute medical conditions. The acute medical conditions included in both the Tromsø 4 Study and the DCH Study were stroke, acute myocardial infarction, and infection, while the DCH Study also included exacerbation of chronic lung disease and activity in collagenous disease.

Study design

In our prospective cohort, all participants were followed up from the date of inclusion in 1993–1997 to the date of death, migration, or end of follow-up for VTE, and cancer registration in the respective cohorts. Incident VTE events were identified until December 31, 2012 in Tromsø 4, December 31, 2007 in HUNT 2, and April 30, 2008 in DCH. Cancer diagnoses were registered up to December 31, 2012 in Tromso 4 and DCH and to December 31, 2009 in HUNT 2. Information regarding date of migration and death were obtained by linkage of the national civil registration numbers to the Central Population Registers in Norway and Denmark.

When comparing studies regarding the risk of VTE in cancer patients, heterogeneity in the follow-up time is a methodological challenge, and it is recommended that follow-up should preferably start at the time of cancer diagnosis.1 In the STAC cohort, subjects with cancer or VTE prior to enrollment were excluded. Consequently, all individuals started to accrue nonexposed person–time in the beginning of the study period. Individuals who developed cancer during follow-up changed their exposure status and started to accrue cancer-exposed observational time from the date of cancer diagnosis. Thus, individuals who developed cancer contributed to both nonexposed and cancer-exposed person-years (PY) at risk during the study period. The design of the STAC cohort also facilitated an observational time with occult cancer exposure, which is particularly interesting as VTE risk during this period is not influenced by cancer treatment, a well-known risk factor for cancer-related VTE.1 Figure 2 illustrates the potential observational times.

Statistical analyses

Descriptive characteristics of the study population and the VTE events were presented within each cohort as means with
5–95% ranges for normally distributed continuous variables, and as medians with 25–75% ranges for continuous variables with skewed distribution. Ten-year age-specific IRs of VTE and cancer with 95% confidence intervals (CIs) were calculated as numbers of events per 1,000 PY at risk. In these analyses, the subjects contributed with PY into different age groups, as they were aging during follow-up.

For overall cancer in the STAC cohort, we calculated 5-year age-specific IRs per 1,000 PY in women and men for the period 1998–2009 and compared these figures graphically with corresponding IRs in the Norwegian Cancer Registry. This period was chosen because it included complete follow-up of cancer for all the three included cohorts. Statistical analyses were performed with STATA version 13.0 (Stata Corporation, College Station, Texas, USA).

### Results

The distribution of baseline characteristics in the three included cohorts and in the merged STAC cohort is presented in Table 1. Median age at inclusion was highest in the DCH Study, followed by the HUNT 2 Study and the Tromsø 4 Study (56 years, 48 years, and 45 years, respectively). The proportion of women was slightly higher in the STAC cohort (52.6% vs 47.4%, respectively) as well as in the three cohorts separately. In the DCH Study, the proportion of subjects with low education (basic school) was lower than in the Tromsø 4 Study and HUNT 2 Study (14.9%, 33.9%, and 34.2%, respectively). It has been reported that subjects with basic school as highest completed education are underrepresented in the DCH Study.\(^\text{10}\) However, subjects with basic school as highest completed education represent 26.7% of the

### Table 1 Baseline characteristics obtained in 1993–1997 within the separate cohorts and in the merged STAC cohort

<table>
<thead>
<tr>
<th></th>
<th>Tromsø 4</th>
<th>HUNT 2</th>
<th>DCH</th>
<th>STAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects (n)</td>
<td>26,093</td>
<td>62,846</td>
<td>56,013</td>
<td>144,952</td>
</tr>
<tr>
<td>Age (years)*</td>
<td>45 (35–57)</td>
<td>48 (36–63)</td>
<td>56 (53–60)</td>
<td>53 (43–61)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>47.7 (12,438)</td>
<td>47.0 (29,538)</td>
<td>47.6 (26,674)</td>
<td>47.4 (68,650)</td>
</tr>
<tr>
<td>Women</td>
<td>52.3 (13,655)</td>
<td>53.0 (33,308)</td>
<td>52.4 (29,339)</td>
<td>52.6 (76,302)</td>
</tr>
<tr>
<td><strong>Questionnaire data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic school</td>
<td>33.9 (8,852)</td>
<td>34.2 (21,471)</td>
<td>14.9 (8,318)</td>
<td>26.7 (38,641)</td>
</tr>
<tr>
<td>High school</td>
<td>35.9 (9,366)</td>
<td>41.6 (26,153)</td>
<td>62.7 (35,130)</td>
<td>48.7 (70,649)</td>
</tr>
<tr>
<td>College/University</td>
<td>29.8 (7,780)</td>
<td>19.1 (11,995)</td>
<td>22.1 (12,375)</td>
<td>22.2 (32,150)</td>
</tr>
<tr>
<td>Missing</td>
<td>0.4 (95)</td>
<td>5.1 (3,227)</td>
<td>0.3 (190)</td>
<td>2.4 (3,512)</td>
</tr>
<tr>
<td>Hard physical activity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 1 hour per week</td>
<td>30.5 (7,948)</td>
<td>24.5 (15,389)</td>
<td>31.7 (17,736)</td>
<td>28.3 (41,073)</td>
</tr>
<tr>
<td>Missing</td>
<td>&lt;0.1 (22)</td>
<td>10.3 (6,462)</td>
<td>0.1 (75)</td>
<td>4.5 (6,559)</td>
</tr>
<tr>
<td><strong>Smoking</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>36.0 (9,399)</td>
<td>44.6 (28,009)</td>
<td>35.1 (19,633)</td>
<td>39.4 (57,041)</td>
</tr>
<tr>
<td>Former</td>
<td>26.6 (6,947)</td>
<td>20.4 (12,827)</td>
<td>28.8 (16,116)</td>
<td>24.8 (35,890)</td>
</tr>
<tr>
<td>Current &lt; 15 cigarettes</td>
<td>24.3 (6,401)</td>
<td>20.4 (12,825)</td>
<td>13.1 (7,342)</td>
<td>18.3 (26,568)</td>
</tr>
<tr>
<td>Current 15–25 cigarettes</td>
<td>10.7 (2,789)</td>
<td>6.7 (4,231)</td>
<td>16.1 (9,030)</td>
<td>11.1 (16,050)</td>
</tr>
<tr>
<td>Current 25+ cigarettes</td>
<td>2.0 (509)</td>
<td>2.2 (1,357)</td>
<td>6.9 (3,857)</td>
<td>3.9 (5,723)</td>
</tr>
<tr>
<td>Missing</td>
<td>0.2 (48)</td>
<td>5.7 (3,597)</td>
<td>0.6 (35)</td>
<td>2.5 (3,680)</td>
</tr>
<tr>
<td>Smoke time (years)*</td>
<td>20 (12–30)</td>
<td>20 (11–30)</td>
<td>32 (20–39)</td>
<td>25 (14–35)</td>
</tr>
<tr>
<td>Missing</td>
<td>0.5 (88)</td>
<td>11.1 (3,483)</td>
<td>0.7 (272)</td>
<td>4.4 (3,843)</td>
</tr>
<tr>
<td>Alcohol units/week*</td>
<td>1 (0–3)</td>
<td>1 (0–2)</td>
<td>8 (3–18)</td>
<td>2 (1–7)</td>
</tr>
<tr>
<td>Missing</td>
<td>0.8 (197)</td>
<td>12.8 (8,027)</td>
<td>&lt;1.0 (3)</td>
<td>5.7 (8,227)</td>
</tr>
<tr>
<td>Self-reported diabetes</td>
<td>1.7 (452)</td>
<td>2.9 (1,794)</td>
<td>2.1 (1,175)</td>
<td>2.4 (3,421)</td>
</tr>
<tr>
<td>Use of antihypertensives</td>
<td>6.0 (1,575)</td>
<td>10.7 (6,718)</td>
<td>12.3 (6,868)</td>
<td>10.5 (15,161)</td>
</tr>
<tr>
<td><strong>Health examination</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systolic blood pressure (mmHg)</td>
<td>134 (109–175)</td>
<td>137 (109–179)</td>
<td>140 (109–176)</td>
<td>137 (109–177)</td>
</tr>
<tr>
<td>Diastolic blood pressure (mmHg)</td>
<td>78 (60–100)</td>
<td>80 (62–102)</td>
<td>83 (67–101)</td>
<td>81 (63–101)</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>170 (155–186)</td>
<td>170 (156–186)</td>
<td>170 (156–185)</td>
<td>170 (156–185)</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>73 (53–97)</td>
<td>76 (56–101)</td>
<td>76 (55–100)</td>
<td>75 (53–100)</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>25.2 (19.9–32)</td>
<td>25.8 (20.6–33.7)</td>
<td>26.0 (20.4–33.3)</td>
<td>26.0 (20.4–33.3)</td>
</tr>
</tbody>
</table>

**Notes:** Values are percentages with absolute numbers in parentheses or means with 5%–95% ranges except for*. Values are medians with 25–75% ranges due to skewed distribution of data. The proportion of missing is not reported for variables that contain less than 5% missing. \(^{\text{a}}\) Smoke time is reported for smokers only (previous and current), and the missing proportion is among smokers.

**Abbreviations:** STAC, Scandinavian Thrombosis and Cancer; DCH, Danish Diet, Cancer and Health Study; HUNT 2, second Nord-Trøndelag Health Study.
participants in the STAC cohort due to the higher proportion of subjects with low education in the Norwegian studies. In coherence with reported differences in alcohol consumption between Norway and Denmark, study participants in the DCH cohort had substantially higher weekly alcohol consumption than the Norwegian participants. Participants in the HUNT 2 Study had the highest proportion of nonsmokers. The systolic and diastolic blood pressure and the use of antihypertensives increased with advancing median age in the cohorts.

Characteristics of patients with VTE by study origin are shown in Table 2. The proportion of provoked VTE was slightly higher in the Tromsø 4 Study (55.1%) than in the DCH Study and HUNT 2 Study (50.1% and 48.0%, respectively). In the STAC cohort, 51% of the VTE events were provoked, whereas 49% were unprovoked. Cancer was the most common provoking factor in the STAC cohort and is associated with 18.9% of all the VTE events, followed by immobilization (15.1%) and surgery (14.6%).

Because the age distributions differed between the included cohorts, we estimated 10-year age-specific IRs of VTE per 1,000 PY in order to enable comparisons between the cohorts (Table 3). There were 2,444 VTE events during a median follow-up of 11.7 years and the incidence increased with age for all cohorts. In the STAC cohort, the overall incidence of VTE was 1.4 per 1,000 PY and the incidence increased from 0.3 per 1,000 PY in subjects aged 20–29 years to 6.4 per 1,000 PY in subjects above 80 years. In general, overall and age-specific IRs of VTE were highest in the Tromsø 4 Study, intermediate in the HUNT 2 Study, and lowest in the STAC cohort.

Table 2 Characteristics of patients with VTE events at the time of VTE diagnosis within the separate cohorts and in the merged STAC cohort

<table>
<thead>
<tr>
<th>VTE type</th>
<th>Tromsø 4</th>
<th>HUNT 2</th>
<th>DCH</th>
<th>STAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE</td>
<td>42.8 (284)</td>
<td>38.7 (383)</td>
<td>41.2 (326)</td>
<td>40.6 (993)</td>
</tr>
<tr>
<td>DVT</td>
<td>57.2 (380)</td>
<td>61.3 (606)</td>
<td>54.5 (431)</td>
<td>58.0 (1,417)</td>
</tr>
<tr>
<td>Unprovoked VTE</td>
<td>44.9 (298)</td>
<td>52.0 (514)</td>
<td>49.9 (367)</td>
<td>49.4 (1,179)</td>
</tr>
<tr>
<td>Provoked VTE</td>
<td>55.1 (366)</td>
<td>48.0 (475)</td>
<td>50.1 (369)</td>
<td>50.6 (1,210)</td>
</tr>
</tbody>
</table>

**Provoking factors**

- Overt cancer: 22.9 (152)
- Surgery: 14.3 (95)
- Trauma: 7.7 (51)
- Acute medical condition: 14.3 (95)
- Immobilization: 17.6 (117)
- Travel: 0.9 (6)
- Other factor: 5.1 (34)

**Notes:** Values are percentages with absolute numbers in parentheses or mean with 5–95% ranges in parentheses. There are 34 subjects in the DCH Study with missing information for PE and DVT; several subjects have more than one provoking factor. The given percentages of unprovoked VTE, provoked VTE, and each provoking factor are among total VTE events in each study with exclusion of 55 events with missing information on provoking factors in the DCH Study.

**Abbreviations:** STAC, Scandinavian Thrombosis and Cancer; DCH, Danish Diet, Cancer and Health Study; HUNT, second nord-Trøndelag Health Study; VTE, venous thromboembolism; DVT, deep venous thrombosis; PE, pulmonary embolism; N/A, not applicable.

Table 3 Age-specific IRs of VTE per 1,000 PY with 95% confidence intervals within the separate cohorts and in the merged STAC cohort

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Tromsø 4 PY</th>
<th>Events</th>
<th>IR (95% CI)</th>
<th>HUNT 2 PY</th>
<th>Events</th>
<th>IR (95% CI)</th>
<th>DCH PY</th>
<th>Events</th>
<th>IR (95% CI)</th>
<th>STAC PY</th>
<th>Events</th>
<th>IR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20–29</td>
<td>6,648</td>
<td>1</td>
<td>0.2 (0.02–1.1)</td>
<td>39,515</td>
<td>12</td>
<td>0.3 (0.2–0.5)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>46,163</td>
<td>13</td>
<td>0.3 (0.2–0.5)</td>
</tr>
<tr>
<td>30–39</td>
<td>55,087</td>
<td>17</td>
<td>0.3 (0.2–0.5)</td>
<td>105,397</td>
<td>35</td>
<td>0.3 (0.2–0.5)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>160,484</td>
<td>52</td>
<td>0.3 (0.2–0.4)</td>
</tr>
<tr>
<td>40–49</td>
<td>99,147</td>
<td>46</td>
<td>0.5 (0.3–0.6)</td>
<td>137,592</td>
<td>78</td>
<td>0.6 (0.5–0.7)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>236,739</td>
<td>124</td>
<td>0.5 (0.4–0.6)</td>
</tr>
<tr>
<td>50–59</td>
<td>96,481</td>
<td>118</td>
<td>1.2 (1.0–1.5)</td>
<td>141,081</td>
<td>127</td>
<td>0.9 (0.8–1.1)</td>
<td>221,503</td>
<td>153</td>
<td>0.7 (0.6–0.8)</td>
<td>459,064</td>
<td>398</td>
<td>0.9 (0.8–1.0)</td>
</tr>
<tr>
<td>60–69</td>
<td>65,749</td>
<td>139</td>
<td>2.1 (1.8–2.5)</td>
<td>101,471</td>
<td>176</td>
<td>1.7 (1.5–2.0)</td>
<td>359,362</td>
<td>480</td>
<td>1.3 (1.2–1.5)</td>
<td>526,582</td>
<td>795</td>
<td>1.5 (1.4–1.6)</td>
</tr>
<tr>
<td>70–79</td>
<td>40,241</td>
<td>179</td>
<td>4.4 (3.8–5.2)</td>
<td>83,773</td>
<td>299</td>
<td>3.6 (3.2–4.0)</td>
<td>69,888</td>
<td>158</td>
<td>2.3 (1.9–2.6)</td>
<td>193,903</td>
<td>636</td>
<td>3.3 (3.0–3.5)</td>
</tr>
<tr>
<td>80+</td>
<td>20,342</td>
<td>164</td>
<td>8.1 (6.9–9.4)</td>
<td>45,779</td>
<td>262</td>
<td>5.7 (5.1–6.5)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>66,121</td>
<td>426</td>
<td>6.4 (5.9–7.1)</td>
</tr>
<tr>
<td>Total</td>
<td>383,695</td>
<td>664</td>
<td>1.7 (1.6–1.9)</td>
<td>654,609</td>
<td>989</td>
<td>1.5 (1.4–1.6)</td>
<td>650,753</td>
<td>791</td>
<td>1.2 (1.1–1.3)</td>
<td>1,689,056</td>
<td>2,444</td>
<td>1.4 (1.4–1.5)</td>
</tr>
</tbody>
</table>

**Notes:** Subjects contributed PYs within different age groups while aging during follow-up. Age at the time of VTE diagnosis was registered for the VTE events.

**Abbreviations:** IRs, incidence rates; PY, person-years; STAC, Scandinavian Thrombosis and Cancer; VTE, venous thromboembolism; DCH, Danish Diet, Cancer and Health Study; HUNT, second nord-Trøndelag Health Study; PE, pulmonary embolism; N/A, not applicable.
in the Tromsø 4 Study, followed by the HUNT 2 Study and the DCH Study.

Table 4 presents 10-year age-specific IRs of cancer per 1,000 PY. In total, 19,757 incident cancers were detected during a median follow-up of 14.1 years (9.8 per 1,000 PY), and the incidence increased with age, as expected, for all cohorts. The age-specific cancer rates were approximately similar in the three included cohorts, except for a higher incidence among subjects aged 50–59 years in the DCH Study.

The distribution of selected cancer sites within each cohort is shown in Figure 3. Prostate, breast, colorectal, and lung were the most common cancers in all cohorts, accounting for 53%–61% of the cancer diagnoses. Breast cancer constituted 17.8% of all cancer diagnoses in the DCH Study, while the corresponding percentages in the Tromsø 4 Study and the HUNT 2 Study were 12.2% and 11.6%, respectively. The 10-year age-specific IRs of breast cancer were higher in the DCH Study than in the Norwegian studies (data not shown).

Figure 4 shows 5-year age-specific IRs of overall cancer per 1,000 PY for women and men in the STAC cohort between 1998 and 2009 compared to the corresponding rates given by the Norwegian Cancer Registry in the same time period. The age-specific IRs of cancers in women and men in the STAC cohort are similar to national figures.

**Discussion**

We have established a large population-based cohort study with validated VTE and cancer events to investigate the risk of VTE in cancer patients.

In accordance with previously reported IRs in adult populations in the developed countries, and the reported impact of age, we found an overall IR of VTE in the STAC cohort of 1.4 per 1,000 PY and an exponential increase at higher ages. The overall and 10-year age-specific IRs of VTE were highest in the Tromsø 4 Study, followed by the HUNT 2 Study, and finally the DCH Study. This may be explained by the 5-year longer follow-up time of VTE in the Tromsø Study, along with the most comprehensive procedure to identify VTE events.

Definitions of provoking factors varied somewhat between the three cohorts, but in spite of this, the proportion of provoked VTE events was approximately 50% in all studies, which is in accordance with previous population-based studies. The proportion of provoked VTEs in a given study will clearly depend on the definition of included provoking factors, and a definite definition does not exist. For instance, the Tromsø 4
Study had a slightly higher proportion of provoked VTEs (55%), presumably because both acute medical conditions and trauma were registered as provoking factors. Conversely, the HUNT 2 Study did not include acute medical conditions and the DCH Study registered trauma-related VTEs, but found no events. Nevertheless, since the overall proportion of provoked VTEs was quite similar in the included cohorts, we regard the STAC cohort suitable for future analyses differentiating between provoked and unprovoked VTE events.

The incidence of cancer increased similarly with age in all cohorts. The proportion of breast cancer, however, was higher in the DCH Study than in the Norwegian cohorts. The IR of breast cancer is higher in Denmark than in Norway and Copenhagen, where the majority of DCH participants are recruited from, has particularly high IRs. The DCH population has higher alcohol consumption and a lower proportion of subjects with basic school as their highest education level than the Norwegian cohorts. Interestingly, higher risk of breast cancer in women from the Copenhagen area was recently reported to be related to higher education level and suggested to be attributed to differences in alcohol consumption, reproductive patterns, and use of hormone therapy. Similar results have been found in a Norwegian study including 93,638 women.

Age-specific IRs of cancer in the STAC cohort were similar to corresponding IRs from the Cancer Registry of Norway. The attendance rates in the Tromsø 4 Study and HUNT 2 Study were high (77% and 69%, respectively), while the DCH had a low rate of 35%. Even though it has been reported that subjects with lower socioeconomic positions were under represented in the DCH Study, the age-standardized IRs of cancer were not significantly different from the general Danish population.

The main strengths of the STAC cohort are the prospective design with data from three large, well-established...
cohort studies, characterized by wide age distribution, long follow-up and detailed information on cancer exposure, VTE events, and confounders. The cancer events were obtained by linkage to complete cancer registries, and trained personnel confirmed all the VTE events according to objective criteria after a comprehensive identification process where different patient registries were examined. Some limitations merit consideration. We did not have verified information on previous history of VTE among all participants in the Norwegian cohorts. In these cohorts, individuals who were excluded owing to prior VTE were detected through the medical records review, and the majority of these had a recurrent episode during the study period. Hence, some of the subjects who were treated as healthy participants during follow-up could be prevalent VTE cases who should have been excluded. In the DCH Study, discharge diagnoses were used as basis for exclusions of subjects with prior VTE and may have led to some misclassifications. However, this would only lead to a small change in overall number of PY in the STAC cohort, and thus would presumably have a negligible influence on the IRs of VTE. The DCH Study did not search the radiology procedure registry for VTE events, which may have slightly underestimated the incidence of VTE in the DCH and hence in the STAC cohort. Differences in the follow-up time may also have influenced the overall risk of VTE and cancer in the three included cohorts. Harmonization of baseline variables, such as physical activity and education, had some methodological challenges due to different questionnaires in Norway and Denmark, which may have resulted in some imprecisions in these variables. Another limitation is that we do not have access to measured biomarkers in blood (hemoglobin, platelet count, leukocyte count, etc) from all cohorts at baseline or during follow-up. Information about cancer treatment (ie, radiation and chemotherapy) is also limited and restricted to data from the national cancer registries. Finally, information on lifestyle factors was only assessed at baseline and may have changed over time, and hence resulted in some nondifferential misclassifications.

We have established a large Scandinavian cohort, including individual data from three established population-based cohorts, the STAC cohort that will provide a unique opportunity to explore the impact of genetic and environmental risk factors, gene–environmental interactions, and cancer-specific characteristics on the risk of cancer-related VTE in the general population.

Acknowledgments

The study has used data from the Cancer Registry of Norway and the Danish Cancer Registry. The interpretation and reporting of this data was the sole responsibility of the authors, and no endorsement by these registries is intended nor should be inferred.

Disclosure

The authors declare that they have no conflicts of interests in this work.

References


