BMJ Open

Defining venous thromboembolism and measuring its incidence using Swedish health registries: A nationwide pregnancy cohort study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2015-008864
Article Type:	Research
Date Submitted by the Author:	22-May-2015
Complete List of Authors:	Abdul Sultan, Alyshah; University of Nottingham, Epidemiology & Public Health West, Joe; University of Nottingham, Epidemiology & Public Health Stephansson, Olof; Karolinska Institutet, Department of Medical Epidemiology and Biostatistics Department of Women's and Children's Health Grainge, Matthew; University of Nottingham, Division of Epidemiology and Public Health Tata, Laila; University of Nottingham, Epidemiology & Public Health Fleming, Kate; University of Nottingham, Division of Epidemiology and Public Health humes, david; Nottingham University, Wolfson Digestive Disease Centre Ludvigsson, Jonas; Karolinska Institutet, Department of Medical Epidemiology and Biostatistics
Primary Subject Heading:	Obstetrics and gynaecology
Secondary Subject Heading:	Obstetrics and gynaecology
Keywords:	OBSTETRICS, HAEMATOLOGY, Anticoagulation < HAEMATOLOGY

Defining venous thromboembolism and measuring its incidence using Swedish health registries: A nationwide pregnancy cohort study

Alyshah Abdul Sultan[1,2,] Joe West[1], Olof Stephansson[5,6], Matthew J Grainge, Laila J Tata[1], Kate M Fleming[1], David Humes[1], Jonas F Ludvigsson[2 4]

[1]Division of Epidemiology and Public Health, University of Nottingham, Clinical Sciences Building, City Hospital, Hucknall Road, Nottingham, UK, NG5,1PB

[2]Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden

[4]Dept Paediatrics, Örebro University Hospital, Örebro, Sweden

[5]Department of Medicine, Solna, Clinical Epidemiology Unit, Karolinska Institutet, Stockholm, Sweden

[6]Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden

Address for Correspondence: Alyshah Abdul Sultan, Division of Epidemiology and Public Health, School of Community Health Sciences, University of Nottingham, Clinical Sciences Building, City Hospital, Nottingham, UK, NG5 1PB

Email: alyshah.sultan@hotmail.com

Word count (manuscript) = 3,082

Word count (abstract) = 228

Keywords: Venous thromboembolism; Pregnancy; Postpartum; Define; Swedish data

Reference=19 Tables=3 Figures=2

Abstract

Objective: To accurately define venous thromboembolism (VTE) in the routinely collected Swedish health registers and quantify its incidence in and around pregnancy.

Study design: Cohort study utilising data from the Swedish Medical Birth Registry (MBR) linked to the National Patient Registry (NPR) and the Swedish Prescribed Drug Register (PDR).

Setting: Secondary care centres, Sweden

Participant: 535,055 women aged of 15-44 years experiencing who had one or more pregnancies resulting in a live or still birth between 2005 and 2011.

Main outcome measure: To estimate the incidence rate (IR) of VTE in and around pregnancy using various VTE definitions allowing direct comparison with other countries.

Results: We found that the rate of VTE varied based on the VTE definition and 47% of cases first recorded as outpatient were not accompanied by anticoagulant prescriptions. Using our most inclusive VTE definition we observed higher rates of VTE compared to previously published data using similar methodology. These reduced by 32% (IR=137/100,000 person-years 95% Confidence interval (CI) 127-147) and 23% (IR=321/100,000 person-years; 95% CI 295-350) during the antepartum and postpartum periods respectively, using a restrictive VTE definition that required anticoagulant prescriptions associated with diagnosis, which were more in line with the existing literature.

Conclusion: We found that including VTE codes without treatment confirmation risks the inclusion of false positive cases. When defining VTE using the NPR, anticoagulant prescription information should therefore be considered.

Strength and limitations of this study

- We analysed more than 700,000 pregnancies to assess the robustness of the VTE diagnosis recorded in the Swedish national patient register using anticoagulant prescriptions in a way which has not been done before using these data.

- The use of recent data from the national registries makes our study findings both contemporary and generalizable to the majority of women of childbearing age.

- Our VTE algorithm confirmation relies principally on anticoagulant prescriptions and we were not able to carry out any sort of external validation. However, our absolute rate of VTE in and around pregnancy are in line with the existing literature.

- We may have incorrectly confirmed VTE cases for some women who were being prescribed heparin as thromboprophylaxis rather than therapy which we cannot accurately differentiate from registry data.

Introduction

Venous thromboembolism (VTE) is a serious and potentially fatal condition with an incidence of 1-3 per 1000 per year.[1] Due to its rarity, the use of large administrative heath care datasets is a very cost-effective way to study the epidemiology of VTE in terms of incidence and identification of risk factors. Whilst routinely collected administrative registry data have the advantage of being large, prospectively recorded, population-based and with minimal selection or recall bias, ensuring its accuracy in diagnosis is paramount for the veracity of findings of studies using these data.

The Swedish National Patient Register (NPR) was established in 1964 with complete national coverage from 1987 onwards. To date, various validation studies have been conducted to ensure the quality of data for a wide range of acute and chronic conditions. For instance, a review of NPR validation studies conducted by Ludvigsson et al,[2] concluded that the positive predictive value (PPV) varies based on the diagnosis but generally ranges between 85-95%. However, it is noteworthy that the previous validation studies were only based on inpatient cases (those admitted) and did not specifically validate the outcome of VTE. The more recent inclusion of the Swedish outpatient data (those presenting to hospital but not admitted) from 2001 onwards and the availability of linked data from the Prescribed Drug Register (PDR) now allows for the assessment of different definitions of VTE to determine the most appropriate for use for epidemiological studies.[3]

A particular setting in which VTE has been commonly studied is in and around pregnancy as, although rare, VTE is one of the leading causes of maternal mortality in developed countries.[4] Our previous work using England's linked primary and secondary health care data has shown that the rate of VTE in and around pregnancy varies markedly depending on the data source

and VTE definition used[5]. A separate VTE validation study among pregnant women from Denmark[6] reported the overall PPV to be 75% for inpatient cases but a much lower PPV (31%) for those only presenting to the emergency department. Similarly Vikrus et al[7] in their study reported that 35% of women who received a diagnosis of VTE in the Danish national patient register during pregnancy/postpartum were not validated (i.e. they were false positives).

The need to identify a robust definition of VTE for use in epidemiological studies is clear. Studying this rare event in and around pregnancy gives us the ability to directly compare estimates of incidence arising from the Swedish data to those from other countries where validation of definitions has previously been undertaken. The aim of our study was to use the Swedish NPR to determine the incidence rate of VTE in and around pregnancy using various VTE definitions and compare these to rates from other countries and health care settings.

Method

Data source

The Swedish Medical Birth Registry (MBR) was established in 1973 and collects information on antenatal and perinatal factors (e.g. gestational length, mode of delivery) and their impact on health of the mother and infant. The registry was standardised and computerised from 1982 onwards.[8] It contains information on birth and delivery records from across Sweden with the addition of 100,000 births each year. Apart from delivery and birth information the MBR also contains up to twelve diagnoses which may occur during pregnancy or in the course of the women's stay in the delivery unit. These are coded using International Classification of Diseases (ICD) codes. The MBR has been subjected to numerous quality checks in the past and the data recorded are of a high standard.[8, 9] The personal identity number assigned to each resident in Sweden allows merging of data between registries such as the NPR and PDR at an individual level.[10] The NPR initially only included the information on hospitalised patients (inpatient data), however since 2001 it also includes outpatient consultations (outpatient data). The data contained within the NPR, can mainly be categorised as patient related data (e.g. personal identity number, sex), administrative data (e.g. visit date, admission and discharge date) and medical data (e.g. diagnosis and procedure information). Each diagnosis recorded within the NPR may be categorised as either a primary or secondary diagnosis. The primary diagnosis should be the main condition that was treated or investigated at the end of a hospital episode whereas the secondary diagnosis may or may not be directly relevant to the primary diagnosis. Finally, filled drug prescriptions are recorded in PDR with the date of dispensing, drug name, strength and preparation according to Anatomical Therapeutic Chemical (ATC) classification.[11]

Study population

For the purpose of this study, we included all women aged 15-44 years from the MBR who experienced a pregnancy between 1 July 2005 and 31 December 2011. Utilising data post 2005 allowed us to obtain prescription data for all of our study population.

Defining Venous thromboembolism

Initially, our working VTE definition included all cases of VTE recorded in either the NPR (inpatient/outpatient) or MBR occurring at any time women were between ages 15-44 years, whilst pregnant or not. We used the relevant ICD-10 codes (as previously used,[5] Supplementary table 1) to identify VTE diagnosis regardless of its hierarchal position in the diagnostic field (primary or secondary). These were then stratified by the source where the first VTE case was identified (inpatient versus outpatient data). All VTEs identified in the birth registry were categorised as inpatient VTEs. For the purpose of this study, we only included the first recorded VTE and therefore excluded all subsequent VTE diagnoses. As the complete NPR coverage dates back to 1987, we also used the relevant ICD-9 codes (415, 451, 452, 453, 671 or 673) to exclude anyone with VTE prior to their entry into the study period. For all women with first VTE, we extracted information on anticoagulant therapy along with information on date and cause of death from the cause of death register, if the woman died within the study period.

Algorithm confirmation

A VTE code was considered to be algorithm confirmed if it was accompanied by an anticoagulant prescription within 90 days of the event or if the woman died within 30 days of the event and the underlying cause of death was VTE. This VTE definition has been validated

previously in English health care data with a positive predictive value of 84%.[12] Based on the above information, we considered 2 VTE definitions.

VTE definition A: Any first VTE recorded in either inpatient or outpatient records regardless of algorithm confirmation.

VTE definition B: A First VTE event recorded as an inpatient or outpatient that was confirmed by our algorithm. As such, definition B consists of cases that are a subset of definition A.

Defining exposure time

Women's follow-up time between age 15–44 years was divided into time associated with pregnancy (defined from the date of conception until 12 weeks postpartum for each and every pregnancy identified) and ''time outside pregnancy" (all other available follow-up time as previous described before.[5]) Time associated with pregnancy was subsequently split into antepartum and postpartum time. In order to minimise the potential misclassification[5] between antepartum and postpartum events a third category of "time around delivery" was also considered. Therefore the time associated with pregnancy was divided into the antepartum period (from the estimated date of conception until 2 days before the date of delivery), time around delivery (1 day before until 2 days after delivery) and the postpartum period, which was defined from 3 days after delivery until 12 weeks postpartum. Postpartum time was then further split into early postpartum (up to 6 weeks after delivery) and late postpartum (7-12 weeks after delivery).

Statistical analysis

We calculated the frequency and proportion of first identified VTE stratified by whether it was registered as an inpatient or outpatient visit. For each VTE definition we calculated the rate of VTE per 100,000 person-years in and around pregnancy separately. We compared our absolute rate of VTE from each VTE definition during specific antepartum and postpartum periods to other similar studies from England[5] and Denmark.[13] The VTE incidence estimates in the English data were derived form a proportion of the English representative population. The diagnosis of VTE, whether they first presented as a hospital admission or in primary care, was confirmed using the same algorithm as used in the present study. The Danish VTE estimates were derived from the Danish NPR[14] where the diagnosis of VTE recorded in the hospital discharge data had been previously validated in a small subset of the data. Finally, we compared the rate of VTE during the antepartum and postpartum period to the time outside pregnancy using a Poisson regression model to calculate incidence rate ratios (IRR) adjusted for age and calendar year. All analyses were performed using Stata version12, Stata Corp., College Station, TX. This study was approved by the ethical review board.

Results

Study population

Our cohort consisted of 535,055 women who experienced at least one pregnancy resulting in a live birth or a stillbirth during the study period. Of the 732,021 pregnancies analysed, 2,530 (0.4%) ended in a stillbirth. Our person-years of follow-up associated with different time periods of pregnancy and outside pregnancy are summarised in Table 1.

Confirmation of VTE cases

We identified a total of 3,163 patients with first incident diagnosis of VTE post 2005, 1,979 (63%) of which were first recorded at an outpatient visit (Figure 1). We observed that 88% of VTEs first recorded during the inpatient visit were algorithm confirmed compared with 55% for those recorded as an outpatient. Of the total of 2,116 patients with validated VTEs, 47% (n=967) received heparin/low molecular weight heparin within 90 days of the event whereas 54% were prescribed warfarin (mostly during the postpartum period and time outside pregnancy). When comparing VTE cases between data sources we found that 22% (n=691) of the patients with VTE had a diagnosis recorded in both inpatient and outpatient data within 90 days of the initial event with 20% (n=626) and 58% (n=1,846) solely recorded in inpatient and outpatient data respectively.

Absolute rate of VTE in and around pregnancy using various VTE definitions

Using VTE definition A (any recording of VTE without confirmation), we found the absolute rate of VTE during the time outside pregnancy to be 42 per 100,000 person years (Table 2). This rate decreased to 37 per 100,000 person years when we used our restricted VTE definition (VTE definition B). Figure 2 shows the rate of VTE by weeks of postpartum which shows variation during the first three weeks postpartum based on the VTE definition. The rate of VTE during specific antepartum and postpartum periods using VTE definition A were

much higher than previous studies using the Danish and English data with no overlapping confidence intervals except for the time around delivery and late postpartum period. However, the absolute rates during the using VTE definition B during the third trimester of antepartum, postpartum period and time outside pregnancy were broadly in line with the Danish and/or English data with overlapping confidence intervals.

Incidence rate ratios of VTE in and around pregnancy compared to time outside pregnancy

Using VTE definition A we found a 5-fold (IRR=5.01; 95%CI 4.61-5.44) and 10-fold (IRR=9.99 95%CI 9.10-10.9) higher rate of VTE during the antepartum and postpartum periods respectively compared to time outside pregnancy. Compared to time outside pregnancy, the overall IRRs during the antepartum (IRR=3.80) and postpartum (IRR=8.70) periods using definition B were very similar to the English data (IRR=3.10 and 8.54, respectively). In contrast, using VTE definition B in the Swedish data gave a statistically significant increased risk of VTE during the first trimester, which was not seen in the English or Danish data.

Discussion

Main finding

Using a nationwide cohort of pregnant women we have shown that the majority of first VTEs recorded in the Swedish inpatient data (85%) were accompanied by anticoagulant therapy following the event. This proportion, however; was much lower for those first recorded as outpatients. We also found that the rate of VTE in and around pregnancy and outside pregnancy varied based on the VTE definition used. Our most inclusive VTE definition (VTE definition A) gave us absolute rates of VTE in and around pregnancy which were much higher compared to those obtained from the English and Danish data. However, using a more restrictive definition of VTE which largely relied on confirmation of VTE cases based on anticoagulant therapy from the prescription register, we found absolute and relative rates of VTE to be more comparable with English and/or Danish data which used similar definitions.

Strength and limitations

Utilising Swedish national registries, we analysed more than 700,000 pregnancies resulting in a live birth or stillbirth to determine how the incidence of VTE in and around pregnancy varies based on the VTE definition and dataset used. The ability to combine national data from inpatient, outpatient, birth and prescription registers allowed us to directly assess the robustness of the VTE diagnosis occurring in different data sources in a way which has not been done before using these data. Furthermore, the use of recent data from the national registries makes our study findings both contemporary and generalizable to the majority of women of childbearing age.

For our VTE definition B, we used an algorithm that mimics a previously validated VTE definition in the UK's primary care data with a positive predictive value of 84%.[12] This

definition, though itself specifically validated in an outside pregnancy setting, has been shown to produce absolute rates of VTE during both antepartum and postpartum periods which are comparable to pooled estimates from several countries published in the existing literature.[5] It is noteworthy that such a validation does not give us any indication of the negative predictive value and we cannot rule out that, particularly within pregnancy, certain women with VTE may receive anticoagulant prescription from hospital immediately following diagnosis not routinely recorded in the PDR. In our present study, however, we believe that the impact of this will be minimal as women diagnosed with VTE are likely to be on anticoagulant therapy for at least 4 weeks, which should therefore be captured in the PDR.

Another limitation of this study is that our algorithm confirmation relies principally on anticoagulant prescriptions and we were not able to carry out any sort of external validation (review of medical records). Secondly, our algorithm only examines one dimension of case confirmation. The optimal validation study would examine specificity, sensitivity, and positive and negative predictive values for which no other data sources was available. However, the absolute rate using VTE definition B for antepartum and the early postpartum period of 137 and 573 per 100,000 person years respectively are broadly in concordance with the pooled estimates[5] from previous studies conducted after 2005 (118 and 424 per 100,000 person years respectively) with overlapping confidence intervals.

Comparison with other studies

This study is not a traditional validation study as previously conducted in other settings,[6, 7, 12, 15] (i.e. validating VTE cases by reviewing patients medical records. However, we found that 85% of women with an inpatient diagnosis of VTE had a corresponding anticoagulant prescription within 90 days of the event, consistent with the previous validation studies

conducted in the Danish patient register[6, 15] where VTE diagnosis was validated by conducting manual review of medical records for each patient. These studies reported the overall positive predictive value for inpatient cases to be more than 79%. Similarly our lower proportion of cases recorded in outpatient data confirmed through our algorithm (48%) has also been previously demonstrated in the Danish data[6] (PPV=31%) among those who received VTE diagnosis in the emergency department. We found that using our restrictive VTE definition, our absolute rates of antepartum and postpartum VTE were 32% and 23% lower respectively than VTE without confirmation (VTE definition A). This finding is consistent with another Danish validation study which reported that 40% and 21% of the VTE diagnosis during the antepartum and postpartum period respectively are not validated. We do acknowledge that there may be differences between the Danish and Swedish registers in terms of data recording therefore the above comparisons should be interpreted with caution.

Our absolute rates of VTE during the first and second trimester using VTE definition A where slightly higher than those reported by the English and Danish data. This finding is supplemented by the lower maternal VTE related mortality observed in Sweden[16] compared to England[4] and Denmark.[17] This may reflect better management/treatment of VTE particularly during early pregnancy in Sweden. On that note, we cannot ignore the fact that there is a possibility we have incorrectly included as VTE cases some women who were being treated with low molecular weight heparin (LMWH) as preventive thromboprophylaxis rather than as a therapy for a newly occurring VTE event. There is, however, considerable difficulty in knowing the exact dosage of LMWH that a woman would or should receive for thromboprophylaxis as this is dependent on a combination of risk factors and physician judgment which we cannot presume to infer from registry data. Accordingly we cannot

accurately ascribe dosage information as being for therapeutic or thromboprophylactic purposes. The lower rate of VTE during the postpartum period in the Danish data might be due to the potential inclusion of only inpatient VTE cases (although this is not clearly stated in the study). However, even with the inclusion of outpatient data, concerns over its coverage and the use of many different definitions to record this information within the register should not be overlooked.[14] Finally, our absolute rate of VTE during the antepartum and postpartum periods were much higher using VTE definition A compared to the previous studies done on the subject utilising data up to or after 2005.[13, 18, 19]

Implications

Our study has important implications in the way VTE is defined in the Swedish NPR. Whilst a higher proportion of VTEs first recorded in the inpatient data are accompanied by an anticoagulant prescription, this proportion was far lower for those first recorded in the outpatient data. If all outpatient cases were to be included in future research studies this is likely to include false positives, leading to an overestimation of risk if prescription information is not utilised. Our restrictive VTE definition provided absolute rates of VTE during the antepartum and postpartum periods that were more comparable to the estimates from the previously published English data, which utilised similar methodology for case confirmation. Though this study only assessed women of childbearing age (15-44 years) and further comparative studies should be recommended in other patient groups, we believe it is not unreasonable to suggest that prescription confirmation of VTE should be carefully considered when studying the epidemiology of VTE using the Swedish health registries.

Acknowledgements

Competing interest statement

Authors have conflicts of interest to declare.

Details of Contributions

AAS, LJT, JW, KMF, MJG and DH conceived the idea for the study, with OS and JFL also making important contributions to the design of the study. AAS carried out the data management and analysis and wrote the first draft of the manuscript. All authors were involved in the interpretation of the data, contributed towards critical revision of the manuscript and approved the final draft. AAS had full access to all of the data and had final responsibility for the decision to submit for publication.

Data sharing statement

Extra data is available by emailing the corresponding author

Financial disclosure

AAS is jointly funded by the Nottingham Digestive Disease Centre and CORE/Coeliac UK. The collaboration (between the University Of Nottingham and the Karolinska Institute) has been enabled by funding from JW's University of Nottingham/Nottingham University Hospital's NHS Trust Senior Clinical Research Fellowship which also pays his salary. DH is funded by an NIHR post-doctoral fellowship.

Tables

Table 1: Basic characteristics of the study population
All figures are number (%) unless otherwise specified.

Variable	Number (%)
Total number of women aged 15-44	535,055
Total number of pregnancies	732,021
Median follow-up	7 years
Person years of follow-up	
Outside pregnancy	2,999,451
Antepartum	522,922
Around delivery	7,988
Postpartum	164,979
First VTE type (n=VTE events)	
Pulmonary embolism	743 (23)
Deep vein thrombosis	2,420 (77)
Birth outcome (n=pregnancies)	
Live birth	729,491
Stillbirth	2,530

Table 2: Absolute rate of first VTE per 100,000 person-years in and around pregnancy and outside pregnancy

Time period	VTE definition A (Using any code for VTE regardless of confirmation)		VTE definition B (Using only algorithm confirmed VTE)		English data[5]		Danish data[13]	
	N	Rate (95%CI)	N	Rate (95%CI)	N	Rate	N	Rate
Outside pregnancy	1,276	42 (40-44)	1,123	37 (35-39)	1,480	32 (30-33)	2,895	36 (34-37)
Antepartum	1,067	204 (192-216)	718	137 (127-147)	156	99 (85-116)	491	107 (97-116)
Trimester 1	221	134 (118-153)	177	107 (93-124)	23	46 (31-70)	61	41 (32-52)
Trimester 2	296	173 (155-194)	186	109 (94-126)	30	58 (41-83)	75	57 (46-72)
Trimester 3	550	291 (268-316)	355	187 (169-208)	103	182 (150-221)	355	197 (177-219)
Around delivery	131	1640 (1381-1946)	88	1100 (893-1356)	34	1428 (1020-1998)		-
Postpartum	689	417 (387-450)	531	321 (295-350)	135	274 (231-324)	218	175 (153-200)
Early postpartum	620	745 (688-806)	478	573 (524-627)	177	468 (391-561)	199	304 (264-350)
Late postpartum	69	84 (66-106)	53	64 (69-84)	18	73 (46-116)	319	32 (19-50)

Early postpartum (first six weeks after delivery)
Late postpartum (six weeks after delivery)

Table 3: Incidence rate ratio of the rate of VTE using various analysis methods

Time period	VTE definition A IRR[1] (95%CI)	VTE definition B IRR[1] (95%CI)	English data[5] IRR (95%CI)*	Danish data[13] IRR (95%CI)*
Outside pregnancy		Reference		
Antepartum	5.01 (4.61-5.44)	3.80 (3.46-4.18)	3.10 (2.63-3.66)	2.95 (2.68-3.25)
Trimester 1	3.36 (2.91-3.88)	3.02 (2.57-3.54)	1.46 (0.96-2.20)	1.12 (0.86-1.45)
Trimester 2	4.27 (3.76-4.85)	3.02 (2.58-3.53)	1.82 (1.27-2.62)	1.58 (1.24-1.99)
Trimester 3	7.04 (6.37-7.79)	5.13 (4.55-5.78)	5.69 (4.66-6.95)	5.48 (4.89-6.12)
Around delivery	39.4 (32.9-47.2)	29.9 (24.1-37.2)	44.5 (31.68-62.54)	-
Postpartum	9.99 (9.10-10.9)	8.70 (7.84-9.65)	8.54 (7.16-10.19)	4.85 (4.21-5.57)
Early postpartum	17.8 (16.2-19.7)	15.5 (13.9-17.3)	14.61 (12.10-17.67)	8.44 (7.27-9.75)
Late postpartum	2.01 (1.57-2.56)	1.74 (1.32-2.29)	2.29 (1.44-3.65)	0.89 (0.53-1.39)

Early postpartum (first six weeks after delivery)
Late postpartum (second six weeks after delivery)
*Unadjusted ratio calculated based on the data provided
[1]Adjusted for age and calendar year

Figure legends

Figure 1 Flow diagram of the first recorded VTE using both inpatient and outpatient registry

Figure 2 Absolute rate of VTE per 100,000 person years in and around pregnancy using different methods and data sources

References

1. Rashid S, Thursz M, Razvi N, Voller R, Orchard T, Shlebak A. Venous thromboprophylaxis in UK medical inpatients. *J R Soc Med* 2005; **98**(11): 507-12.

2. Ludvigsson JF, Andersson E, Ekbom A, et al. External review and validation of the Swedish national inpatient register. *BMC public health* 2011; **11**(1): 450.

3. Martineau M, Nelson-Piercy C. Venous thromboembolic disease and pregnancy. *Postgrad Med J* 2009; **85**(1007): 489-94.

4. Lewis G, editor. The Confidential Enquiry into Maternal and Child Health. Saving Mothers' Lives: reviewing maternal deaths to make motherhood safer-2003-2005. London; 2007.

5. Abdul Sultan A, Tata LJ, Grainge MJ, West J. The Incidence of First Venous Thromboembolism in and around Pregnancy Using Linked Primary and Secondary Care Data: A Population Based Cohort Study from England and Comparative Meta-Analysis. *Plos One* 2013; **8**(7): e70310.

6. Severinsen MT, Kristensen SR, Overvad K, Dethlefsen C, Tjønneland A, Johnsen SP. Venous thromboembolism discharge diagnoses in the Danish National Patient Registry should be used with caution. *J Clin Epidemiol* 2010; **63**(2): 223-8.

7. Virkus RA, Løkkegaard EC, Lidegaard Ø, et al. Venous thromboembolism in pregnancy and the puerperal period: A study of 1210 events. *Acta Obstet Gynecol Scand* 2013; **92**(10): 1135-42.

8. Centre for Epidemiology The National Board of Health and Welfare. The Swedish Medical Birth Registry - A Summary of Content and Quality2003. http://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/10655/2003-112-3_20031123.pdf (04/07/2014).

9. Cnattingius S, Ericson A, Gunnarskog J, Källén B. A quality study of a medical birth registry. *Scandinavian Journal of Public Health* 1990; **18**(2): 143-8.

10. Ludvigsson JF, Otterblad-Olausson P, Pettersson BU, Ekbom A. The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research. *Eur J Epidemiol* 2009; **24**(11): 659-67.

11. Wettermark B, Hammar N, Fored CM, et al. The new Swedish Prescribed Drug Register--opportunities for pharmacoepidemiological research and experience from the first six months. *Pharmacoepidemiol Drug Saf* 2007; **16**(7): 726-35.

12. Lawrenson R, Todd JC, Leydon GM, Williams TJ, Farmer RD. Validation of the diagnosis of venous thromboembolism in general practice database studies. *Br J Clin Pharmacol* 2000; **49**(6): 591-6.

13. Virkus RA, Lokkegaard ECL, Bergholt T, Mogensen U, Langhoff-Roos J, Lidegaard O. Venous thromboembolism in pregnant and puerperal women in Denmark 1995-2005. A national cohort study. *Thromb Haemost* 2011; **106**(2): 304-9.

14. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scandinavian Journal of Public Health* 2011; **39**(7 suppl): 30-3.

15. Larsen TB, Johnsen SP, Møller CI, Larsen H, Sørensen HT. A review of medical records and discharge summary data found moderate to high predictive values of discharge diagnoses of venous thromboembolism during pregnancy and postpartum. *J Clin Epidemiol* 2005; **58**(3): 316-9.

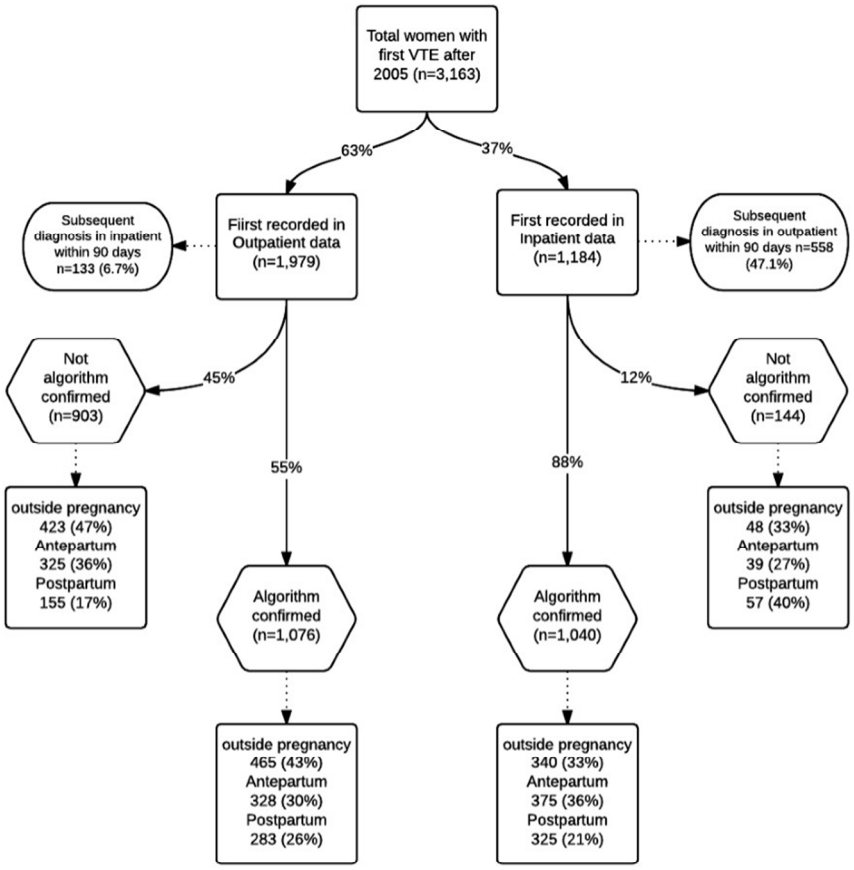
16. The National Board of Health and Welfare. Stastical Database. http://www.socialstyrelsen.se/statistics/statisticaldatabase (accessed 02/01/2015 2014).

17. BØDker B, Hvidman L, Weber TOM, et al. Maternal deaths in Denmark 2002–2006. *Acta Obstet Gynecol Scand* 2009; **88**(5): 556-62.

18. Lyall H, Myers B. Incidence and management of 82 cases of pregnancy-associated venous thromboembolism occurring at a single centre--comparison with Voke et al. *Br J Haematol* 2008; **142**(2): 309-11; author reply 11-2.

19. O'Connor DJ, Scher LA, Gargiulo NJ, 3rd, Jang J, Suggs WD, Lipsitz EC. Incidence and characteristics of venous thromboembolic disease during pregnancy and the postnatal period: a contemporary series. *Ann Vasc Surg* 2011; **25**(1): 9-14.

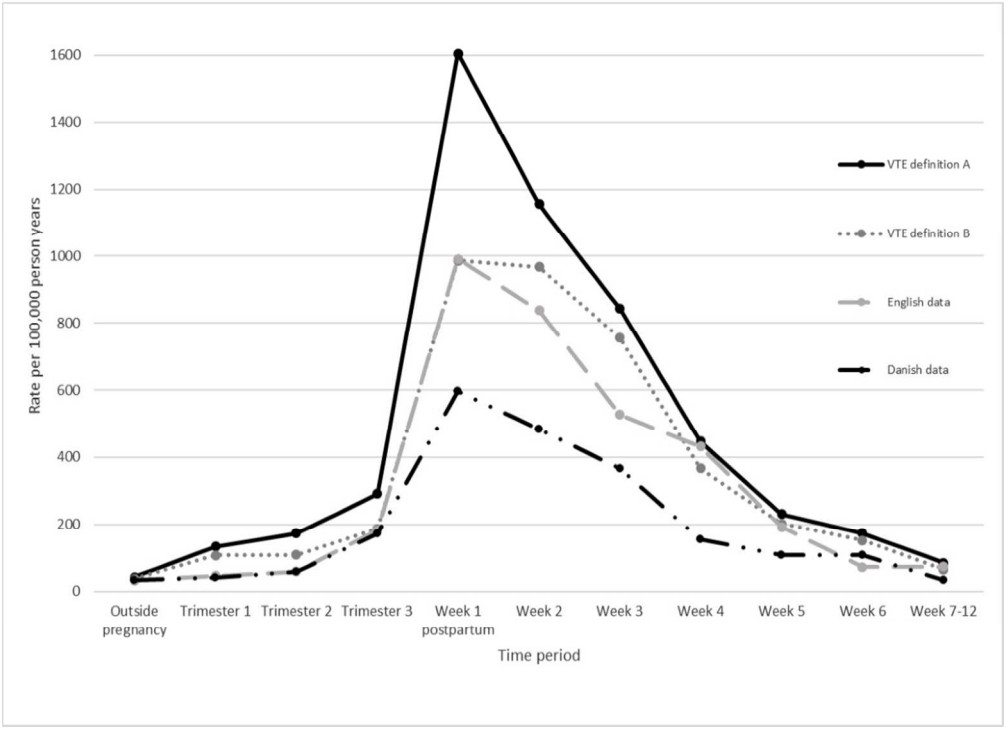
Figure 1: Flow diagram of the first recorded VTE using both inpatient and outpatient registry



Algorithm confirmed
Anticoagulant prescription within 90 days or death within 30 days of the VTE event.
Postpartum time period also includes time around delivery.

175x183mm (150 x 150 DPI)

Figure 2: Absolute rate of VTE per 100,000 person years in and around pregnancy using different methods and data sources



202x159mm (150 x 150 DPI)

Supplementary table 1:ICD-10 codelist for VTE

I260	Pulmonary embolism with mention of acute cor pulmonale
I269	Pulmonary embolism without mention of acute cor pulmonale
I801	Phlebitis and thrombophlebitis of femoral vein
I802	Phlebitis and thrombophlebitis of other deep vessels of lower extremities
I803	Phlebitis and thrombophlebitis of lower extremities, unspecific
I808	Phlebitis and thrombophlebitis of other sites
I809	Phlebitis and thrombophlebitis of unspecified site
I820	Budd-Chiari syndrome
I821	Thrombophlebitis migrans
I822	Embolism and thrombosis of vena cava
I823	Embolism and thrombosis of renal vein
I828	Embolism and thrombosis of other specified veins
I829	Embolism and thrombosis of unspecified vein
O082	Embolism following abortion and ectopic and molar pregnancy
O223	Deep phlebothrombosis in pregnancy
O871	Deep phlebothrombosis in the puerperium
O882	Obstetric blood-clot embolism

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page no.
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	6-8
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-7
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6-7
Bias	9	Describe any efforts to address potential sources of bias	6-8
Study size	10	Explain how the study size was arrived at	
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	N/A
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	8
		(*b*) Describe any methods used to examine subgroups and interactions	
		(*c*) Explain how missing data were addressed	
		(*d*) If applicable, explain how loss to follow-up was addressed	
		(*e*) Describe any sensitivity analyses	
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	9
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	9-10
		(b) Indicate number of participants with missing data for each variable of interest	
		(c) Summarise follow-up time (eg, average and total amount)	
Outcome data	15*	Report numbers of outcome events or summary measures over time	9-10
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear	10-11

		which confounders were adjusted for and why they were included	
		(*b*) Report category boundaries when continuous variables were categorized	
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	11
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	11-12
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	11-12
Generalisability	21	Discuss the generalisability (external validity) of the study results	11-12
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	15

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

BMJ Open

Defining venous thromboembolism and measuring its incidence using Swedish health registries: A nationwide pregnancy cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-008864.R1
Article Type:	Research
Date Submitted by the Author:	10-Aug-2015
Complete List of Authors:	Abdul Sultan, Alyshah; University of Nottingham, Epidemiology & Public Health West, Joe; University of Nottingham, Epidemiology & Public Health Stephansson, Olof; Karolinska Institutet, Department of Medical Epidemiology and Biostatistics Department of Women's and Children's Health Grainge, Matthew; University of Nottingham, Division of Epidemiology and Public Health Tata, Laila; University of Nottingham, Epidemiology & Public Health Fleming, Kate; University of Nottingham, Division of Epidemiology and Public Health humes, david; Nottingham University, Wolfson Digestive Disease Centre Ludvigsson, Jonas; Karolinska Institutet, Department of Medical Epidemiology and Biostatistics
Primary Subject Heading:	Obstetrics and gynaecology
Secondary Subject Heading:	Obstetrics and gynaecology
Keywords:	OBSTETRICS, HAEMATOLOGY, Anticoagulation < HAEMATOLOGY

Defining venous thromboembolism and measuring its incidence using Swedish health registries: A nationwide pregnancy cohort study

Alyshah Abdul Sultan[1,2,] Joe West[1], Olof Stephansson[5,6], Matthew J Grainge, Laila J Tata[1], Kate M Fleming[1], David Humes[1], Jonas F Ludvigsson[2 4]

[1]Division of Epidemiology and Public Health, University of Nottingham, Clinical Sciences Building, City Hospital, Hucknall Road, Nottingham, UK, NG5,1PB

[2]Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden

[4]Dept Paediatrics, Örebro University Hospital, Örebro, Sweden

[5]Department of Medicine, Solna, Clinical Epidemiology Unit, Karolinska Institutet, Stockholm, Sweden

[6]Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden

Address for Correspondence: Alyshah Abdul Sultan, Division of Epidemiology and Public Health, School of Community Health Sciences, University of Nottingham, Clinical Sciences Building, City Hospital, Nottingham, UK, NG5 1PB

Email: alyshah.sultan@hotmail.com

Word count (manuscript) = 3,082

Word count (abstract) = 239

Keywords: Venous thromboembolism; Pregnancy; Postpartum; Define; Swedish data

Reference=19 Tables=3 Figures=2

Abstract

Objective: To accurately define venous thromboembolism (VTE) in the routinely collected Swedish health registers and quantify its incidence in and around pregnancy.

Study design: Cohort study utilising data from the Swedish Medical Birth Registry (MBR) linked to the National Patient Registry (NPR) and the Swedish Prescribed Drug Register (PDR).

Setting: Secondary care centres, Sweden

Participant: 509,198 women aged of 15-44 years who had one or more pregnancies resulting in a live or still birth between 2005 and 2011.

Main outcome measure: To estimate the incidence rate (IR) of VTE in and around pregnancy using various VTE definitions allowing direct comparison with other countries.

Results: The rate of VTE varied based on the VTE definition. We found that 43% of cases first recorded as outpatient were not accompanied by anticoagulant prescriptions whereas this proportion was much lower those cases first recorded in the inpatient (9%). Using our most inclusive VTE definition we observed higher rates of VTE compared to previously published data using similar methodology. These reduced by 31% (IR=142/100,000 person-years 95% Confidence interval (CI) 132-153) and 22% (IR=331/100,000 person-years; 95% CI 304-361) during the antepartum and postpartum periods respectively, using a restrictive VTE definition that required anticoagulant prescriptions associated with diagnosis, which were more in line with the existing literature.

Conclusion: We found that including VTE codes without treatment confirmation risks the inclusion of false positive cases. When defining VTE using the NPR, anticoagulant prescription information should therefore be considered.

Strength and limitations of this study

- We analysed more than 680,000 pregnancies to assess the robustness of the VTE diagnosis recorded in the Swedish national patient register using anticoagulant prescriptions in a way which has not been done before using these data.

- The use of recent data from the national registries makes our study findings both contemporary and generalizable to the majority of women of childbearing age.

- Our VTE algorithm confirmation relies principally on anticoagulant prescriptions and we were not able to carry out any sort of external validation. However, our absolute rate of VTE in and around pregnancy are in line with the existing literature.

- We may have incorrectly confirmed VTE cases for some women who were being prescribed heparin as thromboprophylaxis rather than therapy which we cannot accurately differentiate from registry data.

Introduction

Venous thromboembolism (VTE) is a serious and potentially fatal condition with an incidence of 1-3 per 1000 per year.[1] Due to its rarity, the use of large administrative heath care datasets is a very cost-effective way to study the epidemiology of VTE in terms of incidence and identification of risk factors. Whilst routinely collected administrative registry data have the advantage of being large, prospectively recorded, population-based and with minimal selection or recall bias, ensuring its accuracy in diagnosis is paramount for the veracity of findings of studies using these data.

The Swedish National Patient Register (NPR) was established in 1964 with complete national coverage from 1987 onwards. To date, various validation studies have been conducted to ensure the quality of data for a wide range of acute and chronic conditions. For instance, a review of NPR validation studies conducted by Ludvigsson et al,[2] concluded that the positive predictive value (PPV) varies based on the diagnosis but generally ranges between 85-95%. However, it is noteworthy that the previous validation studies were only based on inpatient cases (those admitted) and did not specifically validate the outcome of VTE. The more recent inclusion of the Swedish outpatient data (those presenting to hospital but not admitted) from 2001 onwards and the availability of linked data from the Prescribed Drug Register (PDR) now allows for the assessment of different definitions of VTE to determine the most appropriate for use for epidemiological studies.[3]

A particular setting in which VTE has been commonly studied is in and around pregnancy as, although rare, VTE is one of the leading causes of maternal mortality in developed countries.[4] Our previous work using England's linked primary and secondary health care data has shown that the rate of VTE in and around pregnancy varies markedly depending on the data source

and VTE definition used[5]. A separate VTE validation study among pregnant women from Denmark[6] reported the overall PPV to be 75% for inpatient cases but a much lower PPV (31%) for those only presenting to the emergency department. Similarly Virkus et al[7] in their study reported that 35% of women who received a diagnosis of VTE in the Danish national patient register during pregnancy/postpartum were not validated (i.e. they were false positives).

The need to identify a robust definition of VTE for use in epidemiological studies is clear. Studying this rare event in and around pregnancy gives us the ability to directly compare estimates of incidence arising from the Swedish data to those from other countries where validation of definitions has previously been undertaken. The aim of our study was to use the Swedish NPR to determine the incidence rate of VTE in and around pregnancy using various VTE definitions and compare these to rates from other countries and health care settings.

Method

Data source

The Swedish Medical Birth Registry (MBR) was established in 1973 and collects information on antenatal and perinatal factors (e.g. gestational length, mode of delivery) and their impact on health of the mother and infant. The registry was standardised and computerised from 1982 onwards.[8] It contains information on birth and delivery records from across Sweden with the addition of 100,000 births each year. Apart from delivery and birth information the MBR also contains up to twelve diagnoses which may occur during pregnancy or in the course of the women's stay in the delivery unit. These are coded using International Classification of Diseases (ICD) codes. The MBR has been subjected to numerous quality checks in the past and the data recorded are of a high standard.[8, 9] The personal identity number assigned to each resident in Sweden allows merging of data between registries such as the NPR and PDR at an individual level.[10] The NPR initially only included the information on hospitalised patients (inpatient data), however since 2001 it also includes outpatient consultations (outpatient data). The data contained within the NPR, can mainly be categorised as patient related data (e.g. personal identity number, sex), administrative data (e.g. visit date, admission and discharge date) and medical data (e.g. diagnosis and procedure information). Each diagnosis recorded within the NPR may be categorised as either a primary or secondary diagnosis. The primary diagnosis should be the main condition that was treated or investigated at the end of a hospital episode whereas the secondary diagnosis may or may not be directly relevant to the primary diagnosis. Finally, filled drug prescriptions are recorded in PDR with the date of dispensing, drug name, strength and preparation according to Anatomical Therapeutic Chemical (ATC) classification.[11]

Study population

For the purpose of this study, we included all women aged 15-44 years from the MBR who experienced a pregnancy between 1 July 2005 and 31 December 2011. Utilising data post 2005 allowed us to obtain prescription data for all of our study population.

Defining Venous thromboembolism

Initially, our working VTE definition included all cases of VTE recorded in either the NPR (inpatient/outpatient) or MBR occurring at any time women were between ages 15-44 years, whilst pregnant or not. We used the relevant ICD-10 codes (as previously used,[5] Supplementary table 1) to identify VTE diagnosis regardless of its hierarchal position in the diagnostic field (primary or secondary). These were then stratified by the source where the first VTE case was identified (inpatient versus outpatient data). All VTEs identified in the birth registry were categorised as inpatient VTEs. For the purpose of this study, we only included the first recorded VTE and therefore excluded all subsequent VTE diagnoses. As the complete NPR coverage dates back to 1987, we also used the relevant ICD-9 codes (415, 451, 452, 453, 671 or 673) to exclude anyone with VTE prior to their entry into the study period. For all women with first VTE, we extracted information on anticoagulant therapy along with information on date and cause of death from the cause of death register, if the woman died within the study period.

Algorithm confirmation

A VTE code was considered to be algorithm confirmed if it was accompanied by an anticoagulant prescription (based on ATC codes; Supplementary table 2) within 90 days of the event or if the woman died within 30 days of the event and the underlying cause of death

was VTE. This VTE definition has been validated previously in English health care data with a positive predictive value of 84%.[12] Based on the above information, we considered 2 VTE definitions.

VTE definition A: Any first VTE recorded in either inpatient or outpatient records regardless of algorithm confirmation.

VTE definition B: A First VTE event recorded as an inpatient or outpatient that was confirmed by our algorithm. As such, definition B consists of cases that are a subset of definition A.

Defining exposure time

Women's follow-up time between age 15–44 years was divided into time associated with pregnancy (defined from the date of conception until 12 weeks postpartum for each and every pregnancy identified) and ''time outside pregnancy" (all other available follow-up time as previous described before.[5]) Time associated with pregnancy was subsequently split into antepartum and postpartum time. In order to minimise the potential misclassification[5] between antepartum and postpartum events a third category of "time around delivery" was also considered. Therefore the time associated with pregnancy was divided into the antepartum period (from the estimated date of conception until 2 days before the date of delivery), time around delivery (1 day before until 2 days after delivery) and the postpartum period, which was defined from 3 days after delivery until 12 weeks postpartum. Postpartum time was then further split into early postpartum (up to 6 weeks after delivery) and late postpartum (7-12 weeks after delivery).

Statistical analysis

We calculated the frequency and proportion of first identified VTE stratified by whether it was registered as an inpatient or outpatient visit. For each VTE definition we calculated the rate of VTE per 100,000 person-years in and around pregnancy separately. We compared our absolute rate of VTE from each VTE definition during specific antepartum and postpartum periods to other similar studies from England[5] and Denmark.[13] The VTE incidence estimates in the English data were derived form a proportion of the English representative population. The diagnosis of VTE, whether they first presented as a hospital admission or in primary care, was confirmed using the same algorithm as used in the present study. The Danish VTE estimates were derived from the Danish NPR[13] where the diagnosis of VTE recorded in the hospital discharge data had been previously validated in a small subset of the data. Finally, we compared the rate of VTE during the antepartum and postpartum period to the time outside pregnancy using a Poisson regression model to calculate incidence rate ratios (IRR) adjusted for age and calendar year. All analyses were performed using Stata version12, Stata Corp., College Station, TX. This study was approved by the ethical review board.

Results

Study population

Our cohort consisted of 509,198 women who experienced at least one pregnancy resulting in a live birth or a stillbirth during the study period. Of the 681,756 pregnancies analysed, 2,389 (0.4%) ended in a stillbirth. Our person-years of follow-up associated with different time periods of pregnancy and outside pregnancy are summarised in Table 1.

Confirmation of VTE cases

We identified a total of 2,864 patients with first incident diagnosis of VTE post 2005, 1,789 (62%) of which were first recorded at an outpatient visit (Figure 1). We observed that 91% of VTEs first recorded during the inpatient visit were algorithm confirmed compared with 57% for those recorded as an outpatient. Of the total of 2,000 patients with validated VTEs, 46% (n=928) received heparin/low molecular weight heparin within 90 days of the event whereas 54% were prescribed warfarin (mostly during the postpartum period and time outside pregnancy). When comparing VTE cases between data sources we found that 22% (n=627) of the patients with VTE had a diagnosis recorded in both inpatient and outpatient data within 90 days of the initial event with 20% (n=569) and 58% (n=1,668) solely recorded in inpatient and outpatient data respectively. Confirmation of VTE cases based on VTE type and position of diagnosis (primary versus secondary) is presented as supplementary table 3. Of 757 cases of non-specific thrombophlebitis diagnoses, 25% (n=190) had evidence of Hirudiod (C05BA01) prescriptions within 90 days of the event of which 44% (n=83) were confirmed based on our algorithm.

Absolute rate of VTE in and around pregnancy using various VTE definitions

Using VTE definition A (any recording of VTE without confirmation), we found the absolute rate of VTE during the time outside pregnancy to be 42 per 100,000 person years (Table 2).

This rate decreased to 38 per 100,000 person years when we used our restricted VTE definition (VTE definition B). Figure 2 shows the rate of VTE by weeks of postpartum which shows variation during the first three weeks postpartum based on the VTE definition. The rate of VTE during specific antepartum and postpartum periods using VTE definition A were much higher than previous studies using the Danish and English data with no overlapping confidence intervals except for the time around delivery and late postpartum period. However, the absolute rates during the using VTE definition B during the third trimester of antepartum, postpartum period and time outside pregnancy were broadly in line with the Danish and/or English data with overlapping confidence intervals.

Incidence rate ratios of VTE in and around pregnancy compared to time outside pregnancy

Using VTE definition A we found a 5-fold (IRR=5.08; 95%CI 4.66-5.54) and 10-fold (IRR=10.2 95%CI 9.27-11.25) higher rate of VTE during the antepartum and postpartum periods respectively compared to time outside pregnancy. Compared to time outside pregnancy, the overall IRRs during the antepartum (IRR=3.80) and postpartum (IRR=8.72) periods using definition B were very similar to the English data (IRR=3.10 and 8.54, respectively). In contrast, using VTE definition B in the Swedish data gave a statistically significant increased risk of VTE during the first trimester, which was not seen in the English or Danish data.

Discussion

Main finding

Using a nationwide cohort of pregnant women we have shown that the majority of first VTEs recorded in the Swedish inpatient data (91%) were accompanied by anticoagulant therapy following the event. This proportion, however; was much lower for those first recorded as outpatients. We also found that the rate of VTE in and around pregnancy and outside pregnancy varied based on the VTE definition used. Our most inclusive VTE definition (VTE definition A) gave us absolute rates of VTE in and around pregnancy which were much higher compared to those obtained from the English and Danish data. However, using a more restrictive definition of VTE which largely relied on confirmation of VTE cases based on anticoagulant therapy from the prescription register, we found absolute and relative rates of VTE to be more comparable with English and/or Danish data which used similar definitions.

Strength and limitations

Utilising Swedish national registries, we analysed more than 680,000 pregnancies resulting in a live birth or stillbirth to determine how the incidence of VTE in and around pregnancy varies based on the VTE definition and dataset used. The ability to combine national data from inpatient, outpatient, birth and prescription registers allowed us to directly assess the robustness of the VTE diagnosis occurring in different data sources in a way which has not been done before using these data. Furthermore, the use of recent data from the national registries makes our study findings both contemporary and generalizable to the majority of women of childbearing age.

For our VTE definition B, we used an algorithm that mimics a previously validated VTE definition in the UK's primary care data with a positive predictive value of 84%.[12] This

definition, though itself specifically validated in an outside pregnancy setting, has been shown to produce absolute rates of VTE during both antepartum and postpartum periods which are comparable to pooled estimates from several countries published in the existing literature.[5] Whilst we did not have the data to externally validate our VTE confirmation algorithm in the Swedish data, a similar VTE definition have been externally validated in the Danish patient register with the positive predictive value of 99%.[14] The study also highlighted that cross linkage with national register of medicinal products provide reliable validation of the VTE event. It is noteworthy that such a validation does not give us any indication of the negative predictive value and we cannot rule out that, particularly within pregnancy, certain women with VTE may receive anticoagulant prescription from hospital immediately following diagnosis not routinely recorded in the PDR. In our present study, however, we believe that the impact of this will be minimal as women diagnosed with VTE are likely to be on anticoagulant therapy for at least 4 weeks, which should therefore be captured in the PDR.

Another limitation of this study is that our algorithm confirmation relies principally on anticoagulant prescriptions and we were not able to carry out any sort of external validation (review of medical records). Secondly, our algorithm only examines one dimension of case confirmation. The optimal validation study would examine specificity, sensitivity, and positive and negative predictive values for which no other data sources was available. However, the absolute rate using VTE definition B for antepartum and the early postpartum period of 142 and 593 per 100,000 person years respectively are broadly in concordance with the pooled estimates[5] from previous studies conducted after 2005 (118 and 424 per 100,000 person years respectively) with overlapping confidence intervals.

Comparison with other studies

This study is not a traditional validation study as previously conducted in other settings,[6, 7, 12, 15] (i.e. validating VTE cases by reviewing patients medical records). However, Rosengren et[16] al carried out external validation of VTE cases identified in a random sample of men born between 1915 and 1922 and between 1924 and 1925 in a particular Swedish county in their study to quantify the impact of psychosocial factors on VTE using NRP. The study found that 85% and 70% of their inpatient recorded PE and DVT cases respectively were externally validated. Our algorithm was able to confirm 94% and 88% of inpatient recorded PE and DVT respectively which may be due to changes in the diagnostic modalities and improvements in the recording of VTE from 1998 onwards. Our higher and lower number of cases confirmed in the outpatient and inpatient respectively have also been previously demonstrated in the Danish data[6] where VTE diagnosis was validated by conducting manual review of medical records for each patient. One possible explanation of the lower number of outpatient cases confirmed by our algorithm could be the potential inclusion of superficial thrombophlebitis treated with hirudoid. However only 26% (n=98) of non-confirmed cases (based on our algorithm) of non-specific thrombophlebitis (ICD-10: I809, I821, I808, I803, ICD-9: 451, 671) recorded in the outpatient were prescribed hirudoids within 90 days of the event. Therefore it may be possible that those with suspected VTE seen in the outpatient are erroneously given VTE diagnosis before case confirmation which

We found that using our restrictive VTE definition, our absolute rates of antepartum and postpartum VTE were 31% and 22% lower respectively than VTE without confirmation (VTE definition A). This finding is consistent with another Danish validation study[7] which reported that 40% and 21% of the VTE diagnosis during the antepartum and postpartum period respectively are not validated. We do acknowledge that there may be differences

between the Danish and Swedish registers in terms of data recording therefore the above comparisons should be interpreted with caution.

Our absolute rates of VTE during the first and second trimester using VTE definition A where slightly higher than those reported by the English and Danish data. This finding is supplemented by the lower maternal VTE related mortality observed in Sweden[17] compared to England[4] and Denmark.[18] This may reflect better management/treatment of VTE particularly during early pregnancy in Sweden. On that note, we cannot ignore the fact that there is a possibility we have incorrectly included as VTE cases some women who were being treated with low molecular weight heparin (LMWH) as preventive thromboprophylaxis rather than as a therapy for a newly occurring VTE event. There is, however, considerable difficulty in knowing the exact dosage of LMWH that a woman would or should receive for thromboprophylaxis as this is dependent on a combination of risk factors and physician judgment which we cannot presume to infer from registry data. Accordingly we cannot accurately ascribe dosage information as being for therapeutic or thromboprophylactic purposes. The lower rate of VTE during the postpartum period in the Danish data might be due to the potential inclusion of only inpatient VTE cases (although this is not clearly stated in the study). However, even with the inclusion of outpatient data, concerns over its coverage and the use of many different definitions to record this information within the register should not be overlooked.[19] Finally, our absolute rate of VTE during the antepartum and postpartum periods were much higher using VTE definition A compared to the previous studies done on the subject utilising data up to or after 2005.[13, 20, 21] Whilst our incidence rate of VTE in and around pregnancy varied based on VTE definition used, our incidence rate ratios during the antepartum and postpartum periods compared to time outside pregnancy remained broadly

similar. This however needs to be further investigated for other risk factors which may be more sensitive to VTE definition in these data.

Implications

Our study has important implications in the way VTE is defined in the Swedish NPR. Whilst a higher proportion of VTEs first recorded in the inpatient data are accompanied by an anticoagulant prescription, this proportion was far lower for those first recorded in the outpatient data. If all outpatient cases were to be included in future research studies this is likely to include false positives, leading to an overestimation of risk if prescription information is not utilised. Our restrictive VTE definition provided absolute rates of VTE during the antepartum and postpartum periods that were more comparable to the estimates from the previously published English data, which utilised similar methodology for case confirmation. Though this study only assessed women of childbearing age (15-44 years) and further comparative studies should be recommended in other patient groups, we believe it is not unreasonable to suggest that prescription confirmation of VTE should be carefully considered when studying the epidemiology of VTE using the Swedish health registries.

Acknowledgements

Competing interest statement

Authors have conflicts of interest to declare.

Details of Contributions

AAS, LJT, JW, KMF, MJG and DH conceived the idea for the study, with OS and JFL also making important contributions to the design of the study. AAS carried out the data management and analysis and wrote the first draft of the manuscript. All authors were involved in the interpretation of the data, contributed towards critical revision of the manuscript and approved the final draft. AAS had full access to all of the data and had final responsibility for the decision to submit for publication.

Data sharing statement

Extra data is available by emailing the corresponding author

Financial disclosure

AAS is jointly funded by the Nottingham Digestive Disease Centre and CORE/Coeliac UK. The collaboration (between the University Of Nottingham and the Karolinska Institute) has been enabled by funding from JW's University of Nottingham/Nottingham University Hospital's NHS Trust Senior Clinical Research Fellowship which also pays his salary. DH is funded by an NIHR post-doctoral fellowship.

Tables

Table 1: Basic characteristics of the study population
All figures are number (%) unless otherwise specified.

Variable	Number (%)
Total number of women aged 15-44	509,198
Total number of pregnancies	681,756
Median follow-up	6.5 years
Person years of follow-up	
Outside pregnancy	2,625,033
Antepartum	486,170
Around delivery	7,438
Postpartum	153,465
First VTE type (n=VTE events)	
Pulmonary embolism	678 (24)
Deep vein thrombosis	2,186 (76)
Birth outcome (n=pregnancies)	
Live birth	681,756 (99.6)
Stillbirth	2,389 (0.4)

Table 2: Absolute rate of first VTE per 100,000 person-years in and around pregnancy and outside pregnancy

Time period	VTE definition A (Using any code for VTE regardless of confirmation)		VTE definition B (Using only algorithm confirmed VTE)		English data[5]		Danish data[13]	
	N	Rate (95%CI)	N	Rate (95%CI)	N	Rate	N	Rate
Outside pregnancy	1105	42 (40-44)	1015	38	1,480	32 (30-33)	2,895	36 (34-37)
Antepartum	995	205 (192-218)	690	142 (132-153)	156	99 (85-116)	491	107 (97-116)
Trimester 1	207	136 (118-155)	172	113 (97-131)	23	46 (31-70)	61	41 (32-52)
Trimester 2	275	174 (154-196)	178	112 (97-130)	30	58 (41-83)	75	57 (46-72)
Trimester 3	513	297 (268-319)	340	194 (174-216)	103	182 (150-221)	355	197 (177-219)
Around delivery	115	1546 (1288-1856)	79	1061 (851-1323)	34	1428 (1020-1998)		-
Postpartum	649	423 (392-457)	509	331 (304-361)	135	274 (231-324)	218	175 (153-200)
Early postpartum	584	754 (696-818)	460	593 (541-650)	177	468 (391-561)	199	304 (264-350)
Late postpartum	65	85 (70-109)	49	64 (49-85)	18	73 (46-116)	319	32 (19-50)

Early postpartum (first six weeks after delivery)
Late postpartum (six weeks after delivery)

Table 3: Incidence rate ratio of the rate of VTE using various analysis methods

Time period	VTE definition A IRR[1] (95%CI)	VTE definition B IRR[1] (95%CI)	English data[5] IRR (95%CI)*	Danish data[13] IRR (95%CI)*
Outside pregnancy		Reference		
Antepartum	5.08 (4.66-5.54)	3.80 (3.44-4.19)	3.10 (2.63-3.66)	2.95 (2.68-3.25)
Trimester 1	3.42 (2.95-3.98)	3.04 (2.58-3.56)	1.46 (0.96-2.20)	1.12 (0.86-1.45)
Trimester 2	4.31 (3.78-4.93)	3.01 (2.56-3.53	1.82 (1.27-2.62)	1.58 (1.24-1.99)
Trimester 3	7.14 (6.43-7.94)	5.12 (4.53-5.80)	5.69 (4.66-6.95)	5.48 (4.89-6.12)
Around delivery	37.5 (30.9-44.45)	27.97 (22.24-35.17)	44.5 (31.68-62.54)	-
Postpartum	10.21 (9.27-11.25)	8.72 (7.83-9.70)	8.54 (7.16-10.19)	4.85 (4.21-5.57)
Early postpartum	19.27 (16.53-20.21)	15.62 (14.00-17.45)	14.61 (12.10-17.67)	8.44 (7.27-9.75)
Late postpartum	2.06 (1.60-2.64)	1.69 (1.26-2.25)	2.29 (1.44-3.65)	0.89 (0.53-1.39)

Early postpartum (first six weeks after delivery)
Late postpartum (second six weeks after delivery)
*Unadjusted ratio calculated based on the data provided
[1]Adjusted for age and calendar year

Figure legends

Figure 1 Flow diagram of the first recorded VTE using both inpatient and outpatient registry

Figure 2 Absolute rate of VTE per 100,000 person years in and around pregnancy using different methods and data sources compared to the English[5] and Danish data[13]

References

1. Rashid S, Thursz M, Razvi N, Voller R, Orchard T, Shlebak A. Venous thromboprophylaxis in UK medical inpatients. *J R Soc Med* 2005; **98**(11): 507-12.

2. Ludvigsson JF, Andersson E, Ekbom A, et al. External review and validation of the Swedish national inpatient register. *BMC public health* 2011; **11**(1): 450.

3. Martineau M, Nelson-Piercy C. Venous thromboembolic disease and pregnancy. *Postgrad Med J* 2009; **85**(1007): 489-94.

4. Lewis G, editor. The Confidential Enquiry into Maternal and Child Health. Saving Mothers' Lives: reviewing maternal deaths to make motherhood safer-2003-2005. London; 2007.

5. Abdul Sultan A, Tata LJ, Grainge MJ, West J. The Incidence of First Venous Thromboembolism in and around Pregnancy Using Linked Primary and Secondary Care Data: A Population Based Cohort Study from England and Comparative Meta-Analysis. *Plos One* 2013; **8**(7): e70310.

6. Severinsen MT, Kristensen SR, Overvad K, Dethlefsen C, Tjønneland A, Johnsen SP. Venous thromboembolism discharge diagnoses in the Danish National Patient Registry should be used with caution. *J Clin Epidemiol* 2010; **63**(2): 223-8.

7. Virkus RA, Løkkegaard EC, Lidegaard Ø, et al. Venous thromboembolism in pregnancy and the puerperal period: A study of 1210 events. *Acta Obstet Gynecol Scand* 2013; **92**(10): 1135-42.

8. Centre for Epidemiology The National Board of Health and Welfare. The Swedish Medical Birth Registry - A Summary of Content and Quality2003. http://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/10655/2003-112-3_20031123.pdf (04/07/2014).

9. Cnattingius S, Ericson A, Gunnarskog J, Källén B. A quality study of a medical birth registry. *Scandinavian Journal of Public Health* 1990; **18**(2): 143-8.

10. Ludvigsson JF, Otterblad-Olausson P, Pettersson BU, Ekbom A. The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research. *Eur J Epidemiol* 2009; **24**(11): 659-67.

11. Wettermark B, Hammar N, Fored CM, et al. The new Swedish Prescribed Drug Register--opportunities for pharmacoepidemiological research and experience from the first six months. *Pharmacoepidemiol Drug Saf* 2007; **16**(7): 726-35.

12. Lawrenson R, Todd JC, Leydon GM, Williams TJ, Farmer RD. Validation of the diagnosis of venous thromboembolism in general practice database studies. *Br J Clin Pharmacol* 2000; **49**(6): 591-6.

13. Virkus RA, Lokkegaard ECL, Bergholt T, Mogensen U, Langhoff-Roos J, Lidegaard O. Venous thromboembolism in pregnant and puerperal women in Denmark 1995-2005. A national cohort study. *Thromb Haemost* 2011; **106**(2): 304-9.

14. Lidegaard Ø, Nielsen LH, Skovlund CW, Skjeldestad FE, Løkkegaard E. Risk of venous thromboembolism from use of oral contraceptives containing different progestogens and oestrogen doses: Danish cohort study, 2001-9; 2011.

15. Larsen TB, Johnsen SP, Møller CI, Larsen H, Sørensen HT. A review of medical records and discharge summary data found moderate to high predictive values of discharge diagnoses of venous thromboembolism during pregnancy and postpartum. *J Clin Epidemiol* 2005; **58**(3): 316-9.

16. Rosengren A, Freden M, Hansson PO, Wilhelmsen L, Wedel H, Eriksson H. Psychosocial factors and venous thromboembolism: a long-term follow-up study of Swedish men. *J Thromb Haemost* 2008; **6**(4): 558-64.

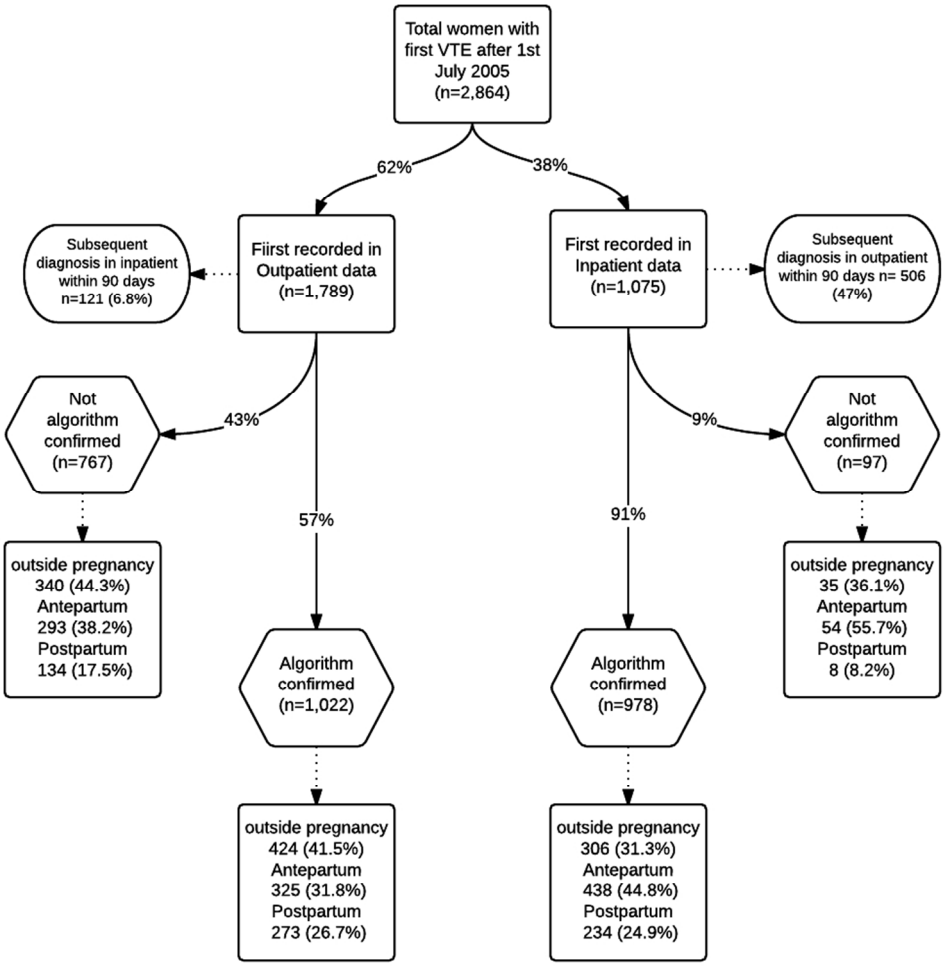
17. The National Board of Health and Welfare. Stastical Database. http://www.socialstyrelsen.se/statistics/statisticaldatabase (accessed 02/01/2015 2014).

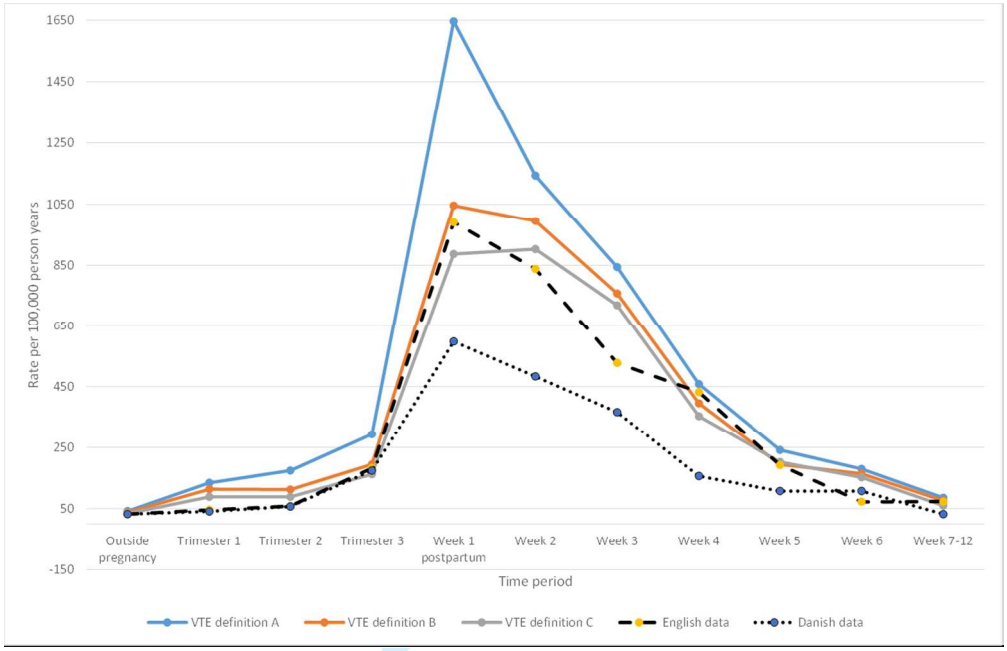
18. BØDker B, Hvidman L, Weber TOM, et al. Maternal deaths in Denmark 2002–2006. *Acta Obstet Gynecol Scand* 2009; **88**(5): 556-62.

19. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scandinavian Journal of Public Health* 2011; **39**(7 suppl): 30-3.

20. Lyall H, Myers B. Incidence and management of 82 cases of pregnancy-associated venous thromboembolism occurring at a single centre--comparison with Voke et al. *Br J Haematol* 2008; **142**(2): 309-11; author reply 11-2.

21. O'Connor DJ, Scher LA, Gargiulo NJ, 3rd, Jang J, Suggs WD, Lipsitz EC. Incidence and characteristics of venous thromboembolic disease during pregnancy and the postnatal period: a contemporary series. *Ann Vasc Surg* 2011; **25**(1): 9-14.





Supplementary table 1:ICD-10 codelist for VTE

I260	Pulmonary embolism with mention of acute cor pulmonale
I269	Pulmonary embolism without mention of acute cor pulmonale
I801	Phlebitis and thrombophlebitis of femoral vein
I802	Phlebitis and thrombophlebitis of other deep vessels of lower extremities
I803	Phlebitis and thrombophlebitis of lower extremities, unspecific
I808	Phlebitis and thrombophlebitis of other sites
I809	Phlebitis and thrombophlebitis of unspecified site
I820	Budd-Chiari syndrome
I821	Thrombophlebitis migrans
I822	Embolism and thrombosis of vena cava
I823	Embolism and thrombosis of renal vein
I828	Embolism and thrombosis of other specified veins
I829	Embolism and thrombosis of unspecified vein
O082	Embolism following abortion and ectopic and molar pregnancy
O223	Deep phlebothrombosis in pregnancy
O871	Deep phlebothrombosis in the puerperium
O882	Obstetric blood-clot embolism
I81	Portal vein thrombosis

Supplementary table 2: Anticoagulants ATC codes

ATC codes	Description
B01AA03	Warfarin
B01AA04	Fenprokumon
B01AA07	Acenokumarol
B01AB01	Heparin
B01AB02	Antitrombin III
B01AB04	Dalteparin
B01AB05	Enoxaparin
B01AB09	Danaparoid
B01AB10	Tinzaparin
B01AC	Platlet aggregation inhibitor
B01AC04	Clopidogrel
B01AC05	Ticlopidine
B01AC06	Aspirin
B01AC07	Dipyridamole
B01AC09	Epoprostenol
B01AC11	Iloprost
B01AC21	Treprostinil
B01AC22	Prasugrel
B01AC24	Tikagrelor
B01AC30	Antiplatelet agents
B01AD01	Streptokinase
B01AD02	Alteplase
B01AD10	Drotrekogin
B01AE05	Ximelagatran
B01AE07	Dabigatran
B01AF01	Rivaroxaban
B01AX05	Fondaparinux

Supplementary table 3: **Confirmation of cases based on VTE type and position of diagnosis**

Parameters	**Anticoagulant prescriptions confirmed diagnosis by source of diagnosis**	
	Inpatient Numbers (%)	**Outpatient Numbers (%)**
VTE diagnostic code		
Pulmonary embolism (ICD 10: 0882, I269, I260) (ICD 9: 415, 673)	470 (94.4)	93 (52.0)
Deep vein thrombosis (ICD 10: I801, I828, I829, O223, I822, I820, I802, I81, O082, I823, O871) (ICD 9:452, 453)	467 (94.0)	622 (66.7)
Thrombophlebitis (ICD 10: I809, I821, I808, I803) (ICD 9: 451, 671)	41 (51.2)	307 (45.4)
Position of diagnosis		
Main diagnosis	738 (94.9)	93 (52.0)
First secondary diagnosis	123 (88.5)	622 (66.7)
Second diagnosis	117 (74.1)	307 (45.4)

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page no.
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	6-8
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-7
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6-7
Bias	9	Describe any efforts to address potential sources of bias	6-8
Study size	10	Explain how the study size was arrived at	
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	N/A
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	8
		(*b*) Describe any methods used to examine subgroups and interactions	
		(*c*) Explain how missing data were addressed	
		(*d*) If applicable, explain how loss to follow-up was addressed	
		(*e*) Describe any sensitivity analyses	
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	9
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	9-10
		(b) Indicate number of participants with missing data for each variable of interest	
		(c) Summarise follow-up time (eg, average and total amount)	
Outcome data	15*	Report numbers of outcome events or summary measures over time	9-10
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear	10-11

		which confounders were adjusted for and why they were included	
		(*b*) Report category boundaries when continuous variables were categorized	
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	11
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	11-12
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	11-12
Generalisability	21	Discuss the generalisability (external validity) of the study results	11-12
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	15

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

BMJ Open

Defining venous thromboembolism and measuring its incidence using Swedish health registries: A nationwide pregnancy cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2015-008864.R2
Article Type:	Research
Date Submitted by the Author:	25-Sep-2015
Complete List of Authors:	Abdul Sultan, Alyshah; University of Nottingham, Epidemiology & Public Health West, Joe; University of Nottingham, Epidemiology & Public Health Stephansson, Olof; Karolinska Institutet, Department of Medical Epidemiology and Biostatistics Department of Women's and Children's Health Grainge, Matthew; University of Nottingham, Division of Epidemiology and Public Health Tata, Laila; University of Nottingham, Epidemiology & Public Health Fleming, Kate; University of Nottingham, Division of Epidemiology and Public Health humes, david; Nottingham University, Wolfson Digestive Disease Centre Ludvigsson, Jonas; Karolinska Institutet, Department of Medical Epidemiology and Biostatistics
Primary Subject Heading:	Obstetrics and gynaecology
Secondary Subject Heading:	Obstetrics and gynaecology
Keywords:	OBSTETRICS, HAEMATOLOGY, Anticoagulation < HAEMATOLOGY

Defining venous thromboembolism and measuring its incidence using Swedish health registries: A nationwide pregnancy cohort study

Alyshah Abdul Sultan[1,2,3,] Joe West[1], Olof Stephansson[5,6], Matthew J Grainge[1], Laila J Tata[1], Kate M Fleming[1], David Humes[1], Jonas F Ludvigsson[2 4]

[1]Division of Epidemiology and Public Health, University of Nottingham, Clinical Sciences Building, City Hospital, Hucknall Road, Nottingham, UK, NG5,1PB

[2]Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden

[3]Nottingham Digestive Diseases Biomedical Research Unit, University of Nottingham, Queens Medical Center, UK, NG7 2UH

[4]Dept Paediatrics, Örebro University Hospital, Örebro, Sweden

[5]Department of Medicine, Solna, Clinical Epidemiology Unit, Karolinska Institutet, Stockholm, Sweden

[6]Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden

Address for Correspondence: Alyshah Abdul Sultan, Division of Epidemiology and Public Health, School of Community Health Sciences, University of Nottingham, Clinical Sciences Building, City Hospital, Nottingham, UK, NG5 1PB

Email: alyshah.sultan@hotmail.com

Word count (manuscript) = 3,082

Word count (abstract) = 249

Keywords: Venous thromboembolism; Pregnancy; Postpartum; Define; Swedish data

Reference=19 Tables=3 Figures=2

Abstract

Objective: To accurately define venous thromboembolism (VTE) in the routinely collected Swedish health registers and quantify its incidence in and around pregnancy.

Study design: Cohort study utilising data from the Swedish Medical Birth Registry (MBR) linked to the National Patient Registry (NPR) and the Swedish Prescribed Drug Register (PDR).

Setting: Secondary care centres, Sweden

Participant: 509,198 women aged of 15-44 years who had one or more pregnancies resulting in a live or still birth between 2005 and 2011.

Main outcome measure: To estimate the incidence rate (IR) of VTE in and around pregnancy using various VTE definitions allowing direct comparison with other countries.

Results: The rate of VTE varied based on the VTE definition. We found that 43% of cases first recorded as outpatient were not accompanied by anticoagulant prescriptions whereas this proportion was much lower than those cases first recorded in the inpatient registry (9%). Using our most inclusive VTE definition we observed higher rates of VTE compared to previously published data using similar methodology. These reduced by 31% (IR=142/100,000 person-years 95% Confidence interval (CI) 132-153) and 22% (IR=331/100,000 person-years; 95% CI 304-361) during the antepartum and postpartum periods respectively, using a restrictive VTE definition that required anticoagulant prescriptions associated with diagnosis, which were more in line with the existing literature.

Conclusion: We found that including VTE codes without treatment confirmation risks the inclusion of false positive cases. When defining VTE using the NPR, anticoagulant prescription information should therefore be considered particularly for cases recorded in the outpatient.

Strength and limitations of this study

- We analysed more than 680,000 pregnancies to assess the robustness of the VTE diagnosis recorded in the Swedish national patient register using anticoagulant prescriptions in a way which has not been done before using these data.

- The use of recent data from the national registries makes our study findings both contemporary and generalizable to the majority of women of childbearing age.

- Our VTE algorithm confirmation relies principally on anticoagulant prescriptions and we were not able to carry out any sort of external validation. However, our absolute rate of VTE in and around pregnancy are in line with the existing literature.

- We may have incorrectly confirmed VTE cases for some women who were being prescribed heparin as thromboprophylaxis rather than therapy which we cannot accurately differentiate from registry data.

Introduction

Venous thromboembolism (VTE) is a serious and potentially fatal condition with an incidence of 1-3 per 1000 per year.[1] Due to its rarity, the use of large administrative heath care datasets is a very cost-effective way to study the epidemiology of VTE in terms of incidence and identification of risk factors. Whilst routinely collected administrative registry data have the advantage of being large, prospectively recorded, population-based and with minimal selection or recall bias, ensuring its accuracy in diagnosis is paramount for the veracity of findings of studies using these data.

The Swedish National Patient Register (NPR) was established in 1964 with complete national coverage from 1987 onwards. To date, various validation studies have been conducted to ensure the quality of data for a wide range of acute and chronic conditions. For instance, a review of NPR validation studies conducted by Ludvigsson et al,[2] concluded that the positive predictive value (PPV) varies based on the diagnosis but generally ranges between 85-95%. However, it is noteworthy that the previous validation studies were only based on inpatient cases (those admitted) and did not specifically validate the outcome of VTE. The more recent inclusion of the Swedish outpatient data (those presenting to hospital but not admitted) from 2001 onwards and the availability of linked data from the Prescribed Drug Register (PDR) now allows for the assessment of different definitions of VTE to determine the most appropriate for use for epidemiological studies.[3]

A particular setting in which VTE has been commonly studied is in and around pregnancy as, although rare, VTE is one of the leading causes of maternal mortality in developed countries.[4] Our previous work using England's linked primary and secondary health care data has shown that the rate of VTE in and around pregnancy varies markedly depending on the data source

and VTE definition used[5]. A separate VTE validation study among pregnant women from Denmark[6] reported the overall PPV to be 75% for inpatient cases but a much lower PPV (31%) for those only presenting to the emergency department. Similarly Virkus et al[7] in their study reported that 35% of women who received a diagnosis of VTE in the Danish national patient register during pregnancy/postpartum were not validated (i.e. they were false positives).

The need to identify a robust definition of VTE for use in epidemiological studies is clear. Studying this rare event in and around pregnancy gives us the ability to directly compare estimates of incidence arising from the Swedish data to those from other countries where validation of definitions has previously been undertaken. The aim of our study was to use the Swedish NPR to determine the incidence rate of VTE in and around pregnancy using various VTE definitions and compare these to rates from other countries and health care settings.

Method

Data source

The Swedish Medical Birth Registry (MBR) was established in 1973 and collects information on antenatal and perinatal factors (e.g. gestational length, mode of delivery) and their impact on health of the mother and infant. The registry was standardised and computerised from 1982 onwards.[8] It contains information on birth and delivery records from across Sweden with the addition of 100,000 births each year. Apart from delivery and birth information the MBR also contains up to twelve diagnoses which may occur during pregnancy or in the course of the women's stay in the delivery unit. These are coded using International Classification of Diseases (ICD) codes. The MBR has been subjected to numerous quality checks in the past and the data recorded are of a high standard.[8, 9] The personal identity number assigned to each resident in Sweden allows merging of data between registries such as the NPR and PDR at an individual level.[10] The NPR initially only included the information on hospitalised patients (inpatient data), however since 2001 it also includes outpatient consultations (outpatient data). The data contained within the NPR, can mainly be categorised as patient related data (e.g. personal identity number, sex), administrative data (e.g. visit date, admission and discharge date) and medical data (e.g. diagnosis and procedure information). Each diagnosis recorded within the NPR may be categorised as either a primary or secondary diagnosis. The primary diagnosis should be the main condition that was treated or investigated at the end of a hospital episode whereas the secondary diagnosis may or may not be directly relevant to the primary diagnosis. Finally, filled drug prescriptions are recorded in PDR with the date of dispensing, drug name, strength and preparation according to Anatomical Therapeutic Chemical (ATC) classification.[11]

Study population

For the purpose of this study, we included all women aged 15-44 years from the MBR who experienced a pregnancy between 1 July 2005 and 31 December 2011. Utilising data post 2005 allowed us to obtain prescription data for all of our study population.

Defining Venous thromboembolism

Initially, our working VTE definition included all cases of VTE recorded in either the NPR (inpatient/outpatient) or MBR occurring at any time women were between ages 15-44 years, whilst pregnant or not. We used the relevant ICD-10 codes (as previously used,[5] Supplementary table 1) to identify VTE diagnosis regardless of its hierarchal position in the diagnostic field (primary or secondary). These were then stratified by the source where the first VTE case was identified (inpatient versus outpatient data). All VTEs identified in the birth registry were categorised as inpatient VTEs. For the purpose of this study, we only included the first recorded VTE and therefore excluded all subsequent VTE diagnoses. As the complete NPR coverage dates back to 1987, we also used the relevant ICD-9 codes (415, 451, 452, 453, 671 or 673) to exclude anyone with VTE prior to their entry into the study period. For all women with first VTE, we extracted information on anticoagulant therapy along with information on date and cause of death from the cause of death register, if the woman died within the study period.

Algorithm confirmation

A VTE code was considered to be algorithm confirmed if it was accompanied by an anticoagulant prescription (based on ATC codes; Supplementary table 2) within 90 days of the event or if the woman died within 30 days of the event and the underlying cause of death

was VTE. This VTE definition has been validated previously in English health care data with a positive predictive value of 84%.[12] Based on the above information, we considered 2 VTE definitions.

VTE definition A: Any first VTE recorded in either inpatient or outpatient records regardless of algorithm confirmation.

VTE definition B: A First VTE event recorded as an inpatient or outpatient that was confirmed by our algorithm. As such, definition B consists of cases that are a subset of definition A.

Defining exposure time

Women's follow-up time between age 15–44 years was divided into time associated with pregnancy (defined from the date of conception based on woman's gestational age until 12 weeks postpartum for each and every pregnancy identified) and ''time outside pregnancy" (all other available follow-up time as previous described before.[5]) Time associated with pregnancy was subsequently split into antepartum and postpartum time. In order to minimise the potential misclassification[5] between antepartum and postpartum events a third category of "time around delivery" was also considered. Therefore the time associated with pregnancy was divided into the antepartum period (from the estimated date of conception until 2 days before the date of delivery), time around delivery (1 day before until 2 days after delivery) and the postpartum period, which was defined from 3 days after delivery until 12 weeks postpartum. Postpartum time was then further split into early postpartum (up to 6 weeks after delivery) and late postpartum (7-12 weeks after delivery).

Statistical analysis

We calculated the frequency and proportion of first identified VTE stratified by whether it was registered as an inpatient or outpatient visit. For each VTE definition we calculated the rate of VTE per 100,000 person-years in and around pregnancy separately. We compared our absolute rate of VTE from each VTE definition during specific antepartum and postpartum periods to other similar studies from England[5] and Denmark.[13] The VTE incidence estimates in the English data were derived form a proportion of the English representative population. The diagnosis of VTE, whether they first presented as a hospital admission or in primary care, was confirmed using the same algorithm as used in the present study. The Danish VTE estimates were derived from the Danish NPR[13] where the diagnosis of VTE recorded in the hospital discharge data had been previously validated in a small subset of the data. Finally, we compared the rate of VTE during the antepartum and postpartum period to the time outside pregnancy using a Poisson regression model to calculate incidence rate ratios (IRR) adjusted for age and calendar year. All analyses were performed using Stata version12, Stata Corp., College Station, TX. This study was approved by the ethical review board.

Results

Study population

Our cohort consisted of 509,198 women who experienced at least one pregnancy resulting in a live birth or a stillbirth during the study period. Of the 681,756 pregnancies analysed, 2,389 (0.4%) ended in a stillbirth. Our person-years of follow-up associated with different time periods of pregnancy and outside pregnancy are summarised in Table 1.

Confirmation of VTE cases

We identified a total of 2,864 patients with first incident diagnosis of VTE post 2005, 1,789 (62%) of which were first recorded at an outpatient visit (Figure 1). We observed that 91% of VTEs first recorded during the inpatient visit were algorithm confirmed compared with 57% for those recorded as an outpatient. Of the total of 2,000 patients with validated VTEs, 46% (n=928) received heparin/low molecular weight heparin within 90 days of the event whereas 54% were prescribed warfarin (mostly during the postpartum period and time outside pregnancy). When comparing VTE cases between data sources we found that 22% (n=627) of the patients with VTE had a diagnosis recorded in both inpatient and outpatient data within 90 days of the initial event with 20% (n=569) and 58% (n=1,668) solely recorded in inpatient and outpatient data respectively. Confirmation of VTE cases based on VTE type and position of diagnosis (primary versus secondary) is presented as supplementary table 3. Of 757 cases of non-specific thrombophlebitis diagnoses, 25% (n=190) had evidence of Hirudiod (C05BA01) prescriptions within 90 days of the event of which 44% (n=83) were confirmed based on our algorithm.

Absolute rate of VTE in and around pregnancy using various VTE definitions

Using VTE definition A (any recording of VTE without confirmation), we found the absolute rate of VTE during the time outside pregnancy to be 42 per 100,000 person years (Table 2).

This rate decreased to 38 per 100,000 person years when we used our restricted VTE definition (VTE definition B). Figure 2 shows the rate of VTE by weeks of postpartum which shows variation during the first three weeks postpartum based on the VTE definition. The rate of VTE during specific antepartum and postpartum periods using VTE definition A were much higher than previous studies using the Danish and English data with no overlapping confidence intervals except for the time around delivery and late postpartum period. However, the absolute rates during the using VTE definition B during the third trimester of antepartum, postpartum period and time outside pregnancy were broadly in line with the Danish and/or English data with overlapping confidence intervals.

Incidence rate ratios of VTE in and around pregnancy compared to time outside pregnancy

Using VTE definition A we found a 5-fold (IRR=5.08; 95%CI 4.66-5.54) and 10-fold (IRR=10.2 95%CI 9.27-11.25) higher rate of VTE during the antepartum and postpartum periods respectively compared to time outside pregnancy (Table 3). Compared to time outside pregnancy, the overall IRRs during the antepartum (IRR=3.80) and postpartum (IRR=8.72) periods using definition B were very similar to the English data (IRR=3.10 and 8.54, respectively). In contrast, using VTE definition B in the Swedish data gave a statistically significant increased risk of VTE during the first trimester, which was not seen in the English or Danish data.

Discussion

Main finding

Using a nationwide cohort of pregnant women we have shown that the majority of first VTEs recorded in the Swedish inpatient data (91%) were accompanied by anticoagulant therapy following the event. This proportion, however; was much lower for those first recorded as outpatients. We also found that the rate of VTE in and around pregnancy and outside pregnancy varied based on the VTE definition used. Our most inclusive VTE definition (VTE definition A) gave us absolute rates of VTE in and around pregnancy which were much higher compared to those obtained from the English and Danish data. However, using a more restrictive definition of VTE which largely relied on confirmation of VTE cases based on anticoagulant therapy from the prescription register, we found absolute and relative rates of VTE to be more comparable with English and/or Danish data which used similar definitions.

Strength and limitations

Utilising Swedish national registries, we analysed more than 680,000 pregnancies resulting in a live birth or stillbirth to determine how the incidence of VTE in and around pregnancy varies based on the VTE definition and dataset used. The ability to combine national data from inpatient, outpatient, birth and prescription registers allowed us to directly assess the robustness of the VTE diagnosis occurring in different data sources in a way which has not been done before using these data. Furthermore, the use of recent data from the national registries makes our study findings both contemporary and generalizable to the majority of women of childbearing age.

For our VTE definition B, we used an algorithm that mimics a previously validated VTE definition in the UK's primary care data with a positive predictive value of 84%.[12] This

definition, though itself specifically validated in an outside pregnancy setting, has been shown to produce absolute rates of VTE during both antepartum and postpartum periods which are comparable to pooled estimates from several countries published in the existing literature.[5] Whilst we did not have the data to externally validate our VTE confirmation algorithm in the Swedish data, a similar VTE definition have been externally validated in the Danish patient register with the positive predictive value of 99%.[14] The study also highlighted that cross linkage with national register of medicinal products provide reliable validation of the VTE event. It is noteworthy that such a validation does not give us any indication of the negative predictive value and we cannot rule out that, particularly within pregnancy, certain women with VTE may receive anticoagulant prescription from hospital immediately following diagnosis not routinely recorded in the PDR. In our present study, however, we believe that the impact of this will be minimal as women diagnosed with VTE are likely to be on anticoagulant therapy for at least 4 weeks, which should therefore be captured in the PDR.

Another limitation of this study is that our algorithm confirmation relies principally on anticoagulant prescriptions and we were not able to carry out any sort of external validation (review of medical records). Secondly, our algorithm only examines one dimension of case confirmation. The optimal validation study would examine specificity, sensitivity, and positive and negative predictive values for which no other data sources was available. However, the absolute rate using VTE definition B for antepartum and the early postpartum period of 142 and 593 per 100,000 person years respectively are broadly in concordance with the pooled estimates[5] from previous studies conducted after 2005 (118 and 424 per 100,000 person years respectively) with overlapping confidence intervals.

Comparison with other studies

This study is not a traditional validation study as previously conducted in other settings,[6, 7, 12, 15] (i.e. validating VTE cases by reviewing patients medical records). However, Rosengren et[16] al carried out external validation of VTE cases identified in a random sample of men born between 1915 and 1922 and between 1924 and 1925 in a particular Swedish county in their study to quantify the impact of psychosocial factors on VTE using NRP. The study found that 85% and 70% of their inpatient recorded PE and DVT cases respectively were externally validated. Our algorithm was able to confirm 94% and 88% of inpatient recorded PE and DVT respectively which may be due to changes in the diagnostic modalities and improvements in the recording of VTE from 1998 onwards. Our higher and lower number of cases confirmed in the outpatient and inpatient respectively have also been previously demonstrated in the Danish data[6] where VTE diagnosis was validated by conducting manual review of medical records for each patient. One possible explanation of the lower number of outpatient cases confirmed by our algorithm could be the potential inclusion of superficial thrombophlebitis treated with hirudoid. However only 26% (n=98) of non-confirmed cases (based on our algorithm) of non-specific thrombophlebitis (ICD-10: I809, I821, I808, I803, ICD-9: 451, 671) recorded in the outpatient were prescribed hirudoids within 90 days of the event. Therefore it may be possible that those with suspected VTE seen in the outpatient are erroneously given VTE diagnosis before case confirmation which

We found that using our restrictive VTE definition, our absolute rates of antepartum and postpartum VTE were 31% and 22% lower respectively than VTE without confirmation (VTE definition A). This finding is consistent with another Danish validation study[7] which reported that 40% and 21% of the VTE diagnosis during the antepartum and postpartum period respectively are not validated. We do acknowledge that there may be differences

between the Danish and Swedish registers in terms of data recording therefore the above comparisons should be interpreted with caution.

Our absolute rates of VTE during the first and second trimester using VTE definition A where slightly higher than those reported by the English and Danish data. This finding is supplemented by the lower maternal VTE related mortality observed in Sweden[17] compared to England[4] and Denmark.[18] This may reflect better management/treatment of VTE particularly during early pregnancy in Sweden. On that note, we cannot ignore the fact that there is a possibility we have incorrectly included as VTE cases some women who were being treated with low molecular weight heparin (LMWH) as preventive thromboprophylaxis rather than as a therapy for a newly occurring VTE event. There is, however, considerable difficulty in knowing the exact dosage of LMWH that a woman would or should receive for thromboprophylaxis as this is dependent on a combination of risk factors and physician judgment which we cannot presume to infer from registry data. Accordingly we cannot accurately ascribe dosage information as being for therapeutic or thromboprophylactic purposes. The lower rate of VTE during the postpartum period in the Danish data might be due to the potential inclusion of only inpatient VTE cases (although this is not clearly stated in the study). However, even with the inclusion of outpatient data, concerns over its coverage and the use of many different definitions to record this information within the register should not be overlooked.[19] Finally, our absolute rate of VTE during the antepartum and postpartum periods were much higher using VTE definition A compared to the previous studies done on the subject utilising data up to or after 2005.[13, 20, 21] Whilst our incidence rate of VTE in and around pregnancy varied based on VTE definition used, our incidence rate ratios during the antepartum and postpartum periods compared to time outside pregnancy remained broadly

similar. This however needs to be further investigated for other risk factors which may be more sensitive to VTE definition in these data.

Implications

Our study has important implications in the way VTE is defined in the Swedish NPR. Whilst a higher proportion of VTEs first recorded in the inpatient data are accompanied by an anticoagulant prescription, this proportion was far lower for those first recorded in the outpatient data. If all outpatient cases were to be included in future research studies this is likely to include false positives, leading to an overestimation of risk if prescription information is not utilised. Our restrictive VTE definition provided absolute rates of VTE during the antepartum and postpartum periods that were more comparable to the estimates from the previously published English data, which utilised similar methodology for case confirmation. Though this study only assessed women of childbearing age (15-44 years) and further comparative studies should be recommended in other patient groups, we believe it is not unreasonable to suggest that prescription confirmation of VTE should be carefully considered when studying the epidemiology of VTE using the Swedish health registries.

Acknowledgements

Competing interest statement

Authors have conflicts of interest to declare.

Details of Contributions

AAS, LJT, JW, KMF, MJG and DH conceived the idea for the study, with OS and JFL also making important contributions to the design of the study. AAS carried out the data management and analysis and wrote the first draft of the manuscript. All authors were involved in the interpretation of the data, contributed towards critical revision of the manuscript and approved the final draft. AAS had full access to all of the data and had final responsibility for the decision to submit for publication.

Data sharing statement

Extra data is available by emailing the corresponding author

Financial disclosure

AAS is jointly funded by the Nottingham Digestive Disease Centre and CORE/Coeliac UK. The collaboration (between the University Of Nottingham and the Karolinska Institute) has been enabled by funding from JW's University of Nottingham/Nottingham University Hospital's NHS Trust Senior Clinical Research Fellowship which also pays his salary. DH is funded by an NIHR post-doctoral fellowship.

Tables

Table 1: Basic characteristics of the study population
All figures are number (%) unless otherwise specified.

Variable	Number (%)
Total number of women aged 15-44	509,198
Total number of pregnancies	681,756
Median follow-up	6.5 years
Person years of follow-up	
Outside pregnancy	2,625,033
Antepartum	486,170
Around delivery	7,438
Postpartum	153,465
First VTE type (n=VTE events)	
Pulmonary embolism	678 (24)
Deep vein thrombosis	2,186 (76)
Birth outcome (n=pregnancies)	
Live birth	681,756 (99.6)
Stillbirth	2,389 (0.4)

Table 2: Absolute rate of first VTE per 100,000 person-years in and around pregnancy and outside pregnancy

Time period	VTE definition A (Using any code for VTE regardless of confirmation)		VTE definition B (Using only algorithm confirmed VTE)		English data[5]		Danish data[13]	
	N	Rate (95%CI)	N	Rate (95%CI)	N	Rate	N	Rate
Outside pregnancy	1105	42 (40-44)	1015	38	1,480	32 (30-33)	2,895	36 (34-37)
Antepartum	995	205 (192-218)	690	142 (132-153)	156	99 (85-116)	491	107 (97-116)
Trimester 1	207	136 (118-155)	172	113 (97-131)	23	46 (31-70)	61	41 (32-52)
Trimester 2	275	174 (154-196)	178	112 (97-130)	30	58 (41-83)	75	57 (46-72)
Trimester 3	513	292 (268-319)	340	194 (174-216)	103	182 (150-221)	355	197 (177-219)
Around delivery	115	1546 (1288-1856)	79	1061 (851-1323)	34	1428 (1020-1998)		-
Postpartum	649	423 (392-457)	509	331 (304-361)	135	274 (231-324)	218	175 (153-200)
Early postpartum	584	754 (696-818)	460	593 (541-650)	177	468 (391-561)	199	304 (264-350)
Late postpartum	65	85 (70-109)	49	64 (49-85)	18	73 (46-116)	319	32 (19-50)

Early postpartum (first six weeks after delivery)
Late postpartum (six weeks after delivery)

Table 3: Incidence rate ratio of the rate of VTE using various analysis methods

Time period	VTE definition A IRR[1] (95%CI)	VTE definition B IRR[1] (95%CI)	English data[5] IRR (95%CI)*	Danish data[13] IRR (95%CI)*
Outside pregnancy		Reference		
Antepartum	5.08 (4.66-5.54)	3.80 (3.44-4.19)	3.10 (2.63-3.66)	2.95 (2.68-3.25)
Trimester 1	3.42 (2.95-3.98)	3.04 (2.58-3.56)	1.46 (0.96-2.20)	1.12 (0.86-1.45)
Trimester 2	4.31 (3.78-4.93)	3.01 (2.56-3.53	1.82 (1.27-2.62)	1.58 (1.24-1.99)
Trimester 3	7.14 (6.43-7.94)	5.12 (4.53-5.80)	5.69 (4.66-6.95)	5.48 (4.89-6.12)
Around delivery	37.5 (30.9-44.45)	27.97 (22.24-35.17)	44.5 (31.68-62.54)	-
Postpartum	10.21 (9.27-11.25)	8.72 (7.83-9.70)	8.54 (7.16-10.19)	4.85 (4.21-5.57)
Early postpartum	19.27 (16.53-20.21)	15.62 (14.00-17.45)	14.61 (12.10-17.67)	8.44 (7.27-9.75)
Late postpartum	2.06 (1.60-2.64)	1.69 (1.26-2.25)	2.29 (1.44-3.65)	0.89 (0.53-1.39)

Early postpartum (first six weeks after delivery)
Late postpartum (second six weeks after delivery)
*Unadjusted ratio calculated based on the data provided
[1]Adjusted for age and calendar year

Figure legends

Figure 1 Flow diagram of the first recorded VTE using both inpatient and outpatient register

Figure 2 Absolute rate of VTE per 100,000 person years in and around pregnancy using different methods and data sources compared to the English[5] and Danish data[13]

References

1. Rashid S, Thursz M, Razvi N, Voller R, Orchard T, Shlebak A. Venous thromboprophylaxis in UK medical inpatients. *J R Soc Med* 2005; **98**(11): 507-12.

2. Ludvigsson JF, Andersson E, Ekbom A, et al. External review and validation of the Swedish national inpatient register. *BMC public health* 2011; **11**(1): 450.

3. Martineau M, Nelson-Piercy C. Venous thromboembolic disease and pregnancy. *Postgrad Med J* 2009; **85**(1007): 489-94.

4. Lewis G, editor. The Confidential Enquiry into Maternal and Child Health. Saving Mothers' Lives: reviewing maternal deaths to make motherhood safer-2003-2005. London; 2007.

5. Abdul Sultan A, Tata LJ, Grainge MJ, West J. The Incidence of First Venous Thromboembolism in and around Pregnancy Using Linked Primary and Secondary Care Data: A Population Based Cohort Study from England and Comparative Meta-Analysis. *Plos One* 2013; **8**(7): e70310.

6. Severinsen MT, Kristensen SR, Overvad K, Dethlefsen C, Tjønneland A, Johnsen SP. Venous thromboembolism discharge diagnoses in the Danish National Patient Registry should be used with caution. *J Clin Epidemiol* 2010; **63**(2): 223-8.

7. Virkus RA, Løkkegaard EC, Lidegaard Ø, et al. Venous thromboembolism in pregnancy and the puerperal period: A study of 1210 events. *Acta Obstet Gynecol Scand* 2013; **92**(10): 1135-42.

8. Centre for Epidemiology The National Board of Health and Welfare. The Swedish Medical Birth Registry - A Summary of Content and Quality2003. http://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/10655/2003-112-3_20031123.pdf (04/07/2014).

9. Cnattingius S, Ericson A, Gunnarskog J, Källén B. A quality study of a medical birth registry. *Scandinavian Journal of Public Health* 1990; **18**(2): 143-8.

10. Ludvigsson JF, Otterblad-Olausson P, Pettersson BU, Ekbom A. The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research. *Eur J Epidemiol* 2009; **24**(11): 659-67.

11. Wettermark B, Hammar N, Fored CM, et al. The new Swedish Prescribed Drug Register--opportunities for pharmacoepidemiological research and experience from the first six months. *Pharmacoepidemiol Drug Saf* 2007; **16**(7): 726-35.

12. Lawrenson R, Todd JC, Leydon GM, Williams TJ, Farmer RD. Validation of the diagnosis of venous thromboembolism in general practice database studies. *Br J Clin Pharmacol* 2000; **49**(6): 591-6.

13. Virkus RA, Lokkegaard ECL, Bergholt T, Mogensen U, Langhoff-Roos J, Lidegaard O. Venous thromboembolism in pregnant and puerperal women in Denmark 1995-2005. A national cohort study. *Thromb Haemost* 2011; **106**(2): 304-9.

14. Lidegaard Ø, Nielsen LH, Skovlund CW, Skjeldestad FE, Løkkegaard E. Risk of venous thromboembolism from use of oral contraceptives containing different progestogens and oestrogen doses: Danish cohort study, 2001-9; 2011.

15. Larsen TB, Johnsen SP, Møller CI, Larsen H, Sørensen HT. A review of medical records and discharge summary data found moderate to high predictive values of discharge diagnoses of venous thromboembolism during pregnancy and postpartum. *J Clin Epidemiol* 2005; **58**(3): 316-9.

16. Rosengren A, Freden M, Hansson PO, Wilhelmsen L, Wedel H, Eriksson H. Psychosocial factors and venous thromboembolism: a long-term follow-up study of Swedish men. *J Thromb Haemost* 2008; **6**(4): 558-64.

17. The National Board of Health and Welfare. Stastical Database. http://www.socialstyrelsen.se/statistics/statisticaldatabase (accessed 02/01/2015 2014).

18. BØDker B, Hvidman L, Weber TOM, et al. Maternal deaths in Denmark 2002–2006. *Acta Obstet Gynecol Scand* 2009; **88**(5): 556-62.

19. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scandinavian Journal of Public Health* 2011; **39**(7 suppl): 30-3.

20. Lyall H, Myers B. Incidence and management of 82 cases of pregnancy-associated venous thromboembolism occurring at a single centre--comparison with Voke et al. *Br J Haematol* 2008; **142**(2): 309-11; author reply 11-2.

21. O'Connor DJ, Scher LA, Gargiulo NJ, 3rd, Jang J, Suggs WD, Lipsitz EC. Incidence and characteristics of venous thromboembolic disease during pregnancy and the postnatal period: a contemporary series. *Ann Vasc Surg* 2011; **25**(1): 9-14.

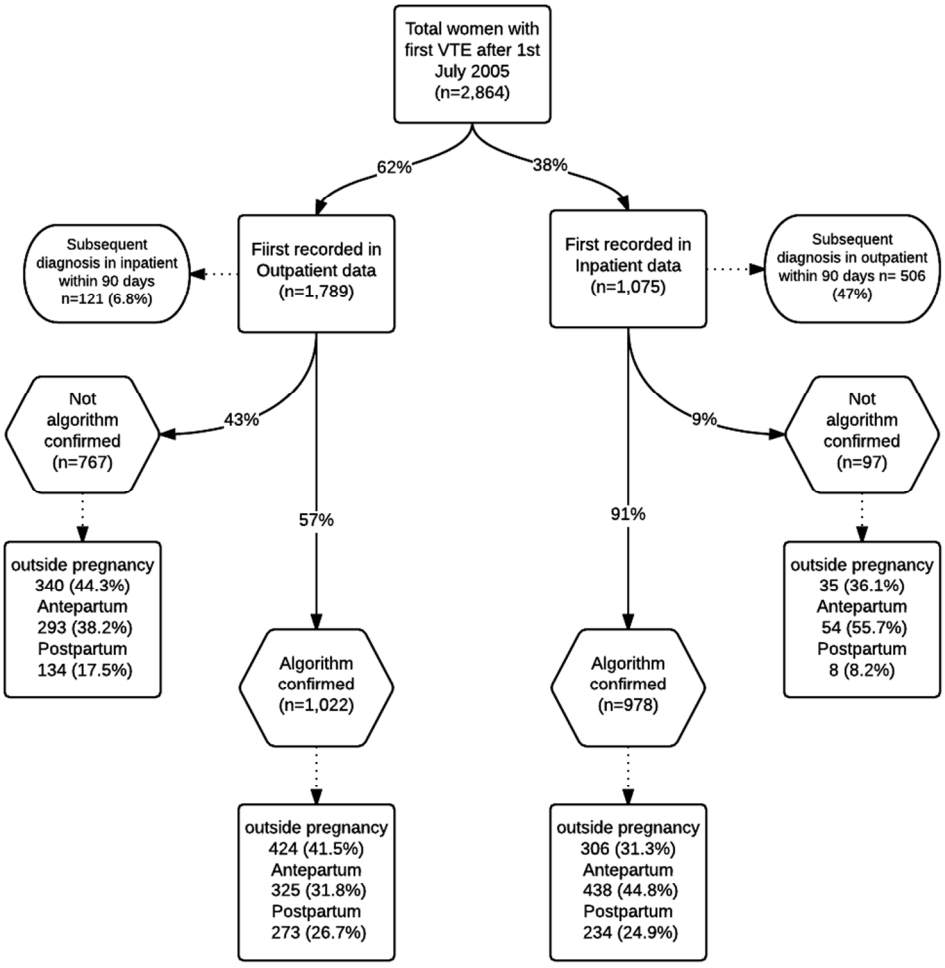


Figure 1 Flow diagram of the first recorded VTE using both inpatient and outpatient registr

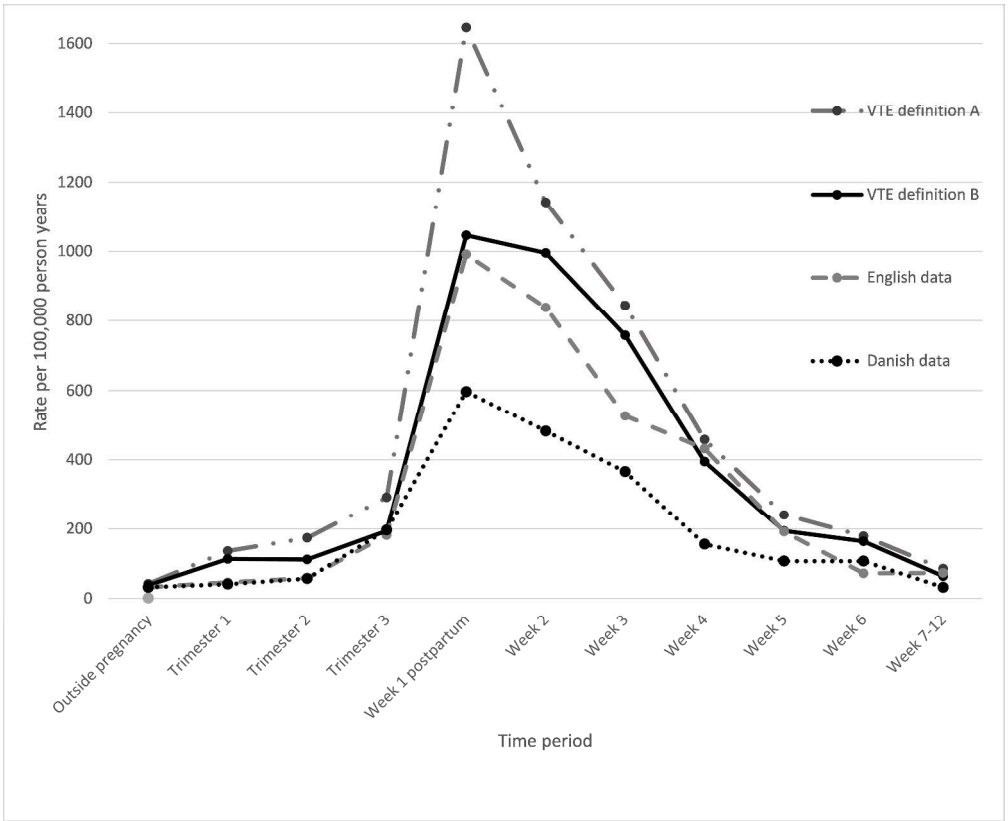


Figure 2 Absolute rate of VTE per 100,000 person years in and around pregnancy using different methods and data sources compared to the English5 and Danish data13

Supplementary table 1:ICD-10 codelist for VTE

I260	Pulmonary embolism with mention of acute cor pulmonale
I269	Pulmonary embolism without mention of acute cor pulmonale
I801	Phlebitis and thrombophlebitis of femoral vein
I802	Phlebitis and thrombophlebitis of other deep vessels of lower extremities
I803	Phlebitis and thrombophlebitis of lower extremities, unspecific
I808	Phlebitis and thrombophlebitis of other sites
I809	Phlebitis and thrombophlebitis of unspecified site
I820	Budd-Chiari syndrome
I821	Thrombophlebitis migrans
I822	Embolism and thrombosis of vena cava
I823	Embolism and thrombosis of renal vein
I828	Embolism and thrombosis of other specified veins
I829	Embolism and thrombosis of unspecified vein
O082	Embolism following abortion and ectopic and molar pregnancy
O223	Deep phlebothrombosis in pregnancy
O871	Deep phlebothrombosis in the puerperium
O882	Obstetric blood-clot embolism
I81	Portal vein thrombosis

Supplementary table 2: Anticoagulants ATC codes

ATC codes	Description
B01AA03	Warfarin
B01AA04	Fenprokumon
B01AA07	Acenokumarol
B01AB01	Heparin
B01AB02	Antitrombin III
B01AB04	Dalteparin
B01AB05	Enoxaparin
B01AB09	Danaparoid
B01AB10	Tinzaparin
B01AC	Platlet aggregation inhibitor
B01AC04	Clopidogrel
B01AC05	Ticlopidine
B01AC06	Aspirin
B01AC07	Dipyridamole
B01AC09	Epoprostenol
B01AC11	Iloprost
B01AC21	Treprostinil
B01AC22	Prasugrel
B01AC24	Tikagrelor
B01AC30	Antiplatelet agents
B01AD01	Streptokinase
B01AD02	Alteplase
B01AD10	Drotrekogin
B01AE05	Ximelagatran
B01AE07	Dabigatran
B01AF01	Rivaroxaban
B01AX05	Fondaparinux

Supplementary table 3: **Confirmation of cases based on VTE type and position of diagnosis**

Parameters	**Anticoagulant prescriptions confirmed diagnosis by source of diagnosis**	
	Inpatient Numbers (%)	**Outpatient Numbers (%)**
VTE diagnostic code		
Pulmonary embolism (ICD 10: 0882, I269, I260) (ICD 9: 415, 673)	470 (94.4)	93 (52.0)
Deep vein thrombosis (ICD 10: I801, I828, I829, O223, I822, I820, I802, I81, O082, I823, O871) (ICD 9:452, 453)	467 (94.0)	622 (66.7)
Thrombophlebitis (ICD 10: I809, I821, I808, I803) (ICD 9: 451, 671)	41 (51.2)	307 (45.4)
Position of diagnosis		
Main diagnosis	738 (94.9)	93 (52.0)
First secondary diagnosis	123 (88.5)	622 (66.7)
Second diagnosis	117 (74.1)	307 (45.4)

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation	Page no.
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	2
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	6-8
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-7
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6-7
Bias	9	Describe any efforts to address potential sources of bias	6-8
Study size	10	Explain how the study size was arrived at	
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	N/A
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	8
		(*b*) Describe any methods used to examine subgroups and interactions	
		(*c*) Explain how missing data were addressed	
		(*d*) If applicable, explain how loss to follow-up was addressed	
		(*e*) Describe any sensitivity analyses	
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	9
		(b) Give reasons for non-participation at each stage	
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	9-10
		(b) Indicate number of participants with missing data for each variable of interest	
		(c) Summarise follow-up time (eg, average and total amount)	
Outcome data	15*	Report numbers of outcome events or summary measures over time	9-10
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear	10-11

		which confounders were adjusted for and why they were included	
		(*b*) Report category boundaries when continuous variables were categorized	
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	N/A
Discussion			
Key results	18	Summarise key results with reference to study objectives	11
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	11-12
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	11-12
Generalisability	21	Discuss the generalisability (external validity) of the study results	11-12
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	15

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.