HHS Public Access

Author manuscript

Stud Health Technol Inform. Author manuscript; available in PMC 2018 January 09.

Published in final edited form as:

Stud Health Technol Inform. 2017 ; 245: 370–374.

Detecting Signals of Interactions Between Warfarin and Dietary Supplements in Electronic Health Records

Yadan Fan[a], **Terrence J. Adam**[a,b], **Reed McEwan**[c], **Serguei V. Pakhomov**[a,b], **Genevieve B. Melton**[a,d], and **Rui Zhang**[a,d]

[a]Institute for Health Informatics, University of Minnesota, Minneapolis, MN, USA

[b]College of Pharmacy, University of Minnesota, Minneapolis, MN, USA

[c]Academic Health Center-Information Systems, University of Minnesota, Minneapolis, MN, USA

[d]Department of Surgery, University of Minnesota, Minneapolis, MN, USA

Abstract

Drug and supplement interactions (DSIs) have drawn widespread attention due to their potential to affect therapeutic response and adverse event risk. Electronic health records provide a valuable source where the signals of DSIs can be identified and characterized. We detected signals of interactions between warfarin and seven dietary supplements, viz., alfalfa, garlic, ginger, ginkgo, ginseng, St. John's Wort, and Vitamin E by analyzing structured clinical data and unstructured clinical notes from the University of Minnesota Clinical Data Repository. A machine learning-based natural language processing module was further developed to classify supplement use status and applied to filter out irrelevant clinical notes. Cox proportional hazards models were fitted, controlling for a set of confounding factors: age, gender, and Charlson Index of Comorbidity. There was a statistically significant association of warfarin concurrently used with supplements which can potentially increase the risk of adverse events, such as gastrointestinal bleeding.

Keywords

Electronic Health Records; Natural Language Processing; Warfarin

Introduction

Drug and supplement interactions (DSIs) have drawn widespread attention in recent years due to the increased prevalence of dietary supplements worldwide. Patients often take prescribed medications along with dietary supplements to boost the immune system or to mitigate the side effects of a particular treatment. A major safety concern is the potential for dangerous adverse events caused by DSIs, particularly for drugs with narrow therapeutic indexes, such as warfarin. Increasing our knowledge base about DSIs will assist pharmacists and healthcare providers to provide guidance to patients on the safety and efficacy of the

Address for correspondence: Rui Zhang, PhD, University of Minnesota, Minneapolis, MN, USA zhan1386.

concomitant use of prescribed medications and dietary supplements, especially for the elderly, who have increased vulnerability to DSIs. Given the difficulty of testing DSIs in human populations, information on DSIs mostly comes from *in vitro*, animal research, or case reports [1]. Unfortunately, this information is under-reported and can be inconsistent. Also, clinical trials for drug approval may not reveal DSIs since supplements and drug-drug interactions often require large patient populations for adequate study power, especially with rare events. Our prior study identified several known and potential DSIs by mining 23 million biomedical literature abstracts (MEDLINE) [2]. Although the biomedical literature may help us to infer DSI knowledge and potential hypotheses for novel DSIs, we may also leverage electronic health record (EHR) systems to complement DSI understanding and validate DSI hypotheses. EHRs offer a rich source of patient information since they serve as the primary patient care documentation platform for clinical care delivery. Some EHR data of interest to study DSIs include medication information, problem lists, laboratory data, and clinical notes.

Warfarin, as one of the most commonly prescribed anticoagulants, is widely used to treat and prevent thromboembolic events associated with atrial fibrillation, heart valve replacement, myocardial infarction and existing thromboembolic disease. However, warfarin is often involved in interactions with supplements because its metabolism involves multiple active metabolic pathways [3]. Natural products such as garlic, ginger and ginkgo are among the most common supplements implicated in DSIs with warfarin. Garlic has the side effect of platelet inhibition, which can increase the risk of bleeding when used with anticoagulant drugs [4]. Ginger can inhibit thromboxane synthetase and therefore lead to prolonged bleeding times [4]. Ginkgo will increase the International Normalized Ratio (INR) with warfarin, and ginseng might reduce the effect of warfarin [4]. Vitamin E can interact with warfarin due to blood thinning effects, especially in Vitamin K deficient individuals [5]. According to the Natural Medicines Comprehensive Database (NMCD) [6], warfarin also has significant interactions with alfalfa, grapefruit, and St. John’s Wort.

In our previous study, we found that clinical notes contain some supplement mentions that do not exist in the medication list [7]. Much information about supplement use is embedded in clinical notes, and thus in this study we demonstrate that informatics techniques, especially natural language processing (NLP) methods, are effective in extracting supplement use status information from clinical notes. Specifically, we conducted survival analysis to test the significance of the concomitant use of warfarin and supplements associated with the appearance of adverse events based on the EHR data from University of Minnesota Clinical Data Repository (UMN-CDR). We focused our assessment on the adverse interactions of warfarin with seven dietary supplements: alfalfa, garlic, ginger, ginkgo, ginseng, Vitamin E, and St. John’s Wort, with potential interactions indicated in the NMCD knowledge base. Adverse events were limited to embolic stroke and thromboembolism, which are defined as warfarin treatment failure events. In addition, bleeding (including GI bleeding) and the subset of patients with GI bleeding were evaluated and defined as side effects of warfarin in the context of this study.

Background

Warfarin potential interactions

Warfarin is a medication with a long history of clinical use due to its effect on the human coagulation system, but also as a poisoning agent/rodenticide due to these same pharmacologic characteristics. This dichotomy is carefully balanced clinically by having regular testing of the therapeutic response with the use of warfarin. Patient monitoring is managed by assessing blood coagulation with prothrombin times and INR testing which provide standard clotting measures. The testing is typically done on a monthly basis or even more frequently with substantial dosage changes or use of medications associated with drug-drug interactions. This intensive follow-up therapy helps to reduce the clinical risk of excess therapeutic effect (anticoagulation) which can result in bleeding. In addition, the monitoring allows dosages to be constantly adjusted to maintain adequate levels of anticoagulation to prevent thromboembolic events. Given this frequent follow-up by patients, the shifts in levels of anticoagulation are typically noticed before clinically significant events occur. However, if patients initiate new medications or supplements shortly after regular testing, they may be at risk for several weeks before routine testing can detect drug responses outside of the usual therapeutic range.

BioMedICUS and NLP-PIER

Both BioMedICUS and NLP-PIER are tools developed by the NLP/Information Extraction group at the UMN. BioMedICUS (BioMedical Information Collection and Understanding System) [8] is an open-source NLP system based on the Unstructured Information Management Architecture – Asynchronous Scaleout (UIMA-AS) architecture [9] specializing in NLP-related information extraction and understanding of clinical notes. NLP-PIER (Patient Information Extraction for Researchers) is a web-based search engine for clinical notes from the EHR [10]. Clinical notes in the CDR are run through a BioMedICUS NLP pipeline and indexed for use in NLP-PIER. BioMedICUS identifies UMLS Metathesaurus concepts (concept unique identifiers, or CUIs) from lexical variants expressed in the notes, and whether the identified concepts were used in a negated context. These negation-qualified CUIs are added to a set of 15 patient-related and encounter-related note attributes from the CDR, including five attributes from the HL7-LOINC document ontology [11]. Attributes and CUIs are stored in an Elasticsearch cluster along with the clinical note itself, which is run through an Elasticsearch snowball analyzer when it is indexed. This setup enables full text searches to be run on research-related note sets within NLP-PIER. Search terms can be expanded by specifying UMLS CUIs as part of the search query and results can be filtered using the attributes.

Methods

The method of this study consists of five steps: 1) data collection: search for patients taking warfarin and collect information about patients' demographics, warfarin usage, diagnosis and clinical notes; 2) NLP for supplement information extraction: apply NLP module to extract information about supplement use in clinical notes; 3) structured data query: query medication table for supplement use and diagnosis table for adverse events; 4) data

combination: combine information from structured and unstructured data to generate a comprehensive data set for each patient; 5) statistical analysis: conduct survival analysis to detect the significance of adverse events caused by concurrent use of warfarin and supplements.

Data collection

Patient cohort data in the Epic EHR were extracted from the UMN-CDR hosted by the Academic Health Center-Information Services (AHC-IS) exchange platform and supported through the Clinical Translational Science Institute (CTSI) at the UMN. The data in the CDR comes from the EHR of more than 2 million patients who sought health services at eight hospitals and over 40 clinics. Data are available for hospital visits starting from 2011. IRB approval was obtained for accessing the clinical notes.

Patients who have warfarin prescriptions from 2011 to September 2015 were included by using both generic name and brand names (*i.e.,* Coumadin, Panwarfin, Sofarin) of warfarin. Patients with medication records showing at least one warfarin prescription and complete information about the warfarin start date and end date were included in this study. The data from a total of 48,426 patients were stored in AHC-IS data shelter, which included patients' demographic information, diagnosis, and medications. Clinical comorbidities were calculated using the Charlson Index of Comorbidity. Their corresponding clinical notes were processed by BioMedICUS and indexed by NLP-PIER for further information extraction.

NLP for supplement information extraction

Since much of the information about supplement use was embedded in clinical notes, we retrieved the related clinical notes using PIER for further information extraction. Selected supplements and all their lexical variants were used for retrieving clinical notes. For instance, "ginkgo", "Vitamin E" and "St. John's Wort" have their lexical variants including "gingko", "ginko" and "ginkoba", "Vit E", "St. Johns Wort", "St. John Wort", "St John's wort", "St Johns wort", "St John Wort", respectively. However, we found instances of negative mentions (such as discontinuation of supplements) of supplements in the notes, such as "she may try ginkgo biloba for her memory issues" or "Denies using St John's Wort". Therefore, we applied a NLP module to classify the use status of the supplements, especially the active ones, such as "started" and "continuing", and also filtered the irrelevant clinical notes, such as "discontinuing" and negative mentions.

In our prior study [12], we used machine learning-based methods to automatically classify the use status of the supplements into four categories (Continuing (C), Discontinued (D), Started (S), Unclassified (U)). A total of 1,300 sentences on 25 most commonly consumed supplements were randomly selected and annotated. The training set consisting of 1000 sentences of 10 supplements was used to select the optimal algorithm with the identified feature sets. The test set included 300 sentences on the remaining 15 supplements. We trained four algorithms with seven different feature sets in the study. The best model (*i.e.,* Support Vector Machine (SVM) with the feature set of unigram, bigram and indicator words within window size of four tokens on both sides of supplement mention) achieved F-measure of 0.906, 0.913, 0.914, 0.715 for status C, D, S, U on the test set, respectively. We

further applied the trained SVM model on the notes retrieved in this study. We only consider the “Continuing” and “Started” categories since they are the active status for supplement use. “Discontinued” category may hold important information about the past use of supplements, however, the start date of the supplements remains unclear, therefore, “Discontinued” was considered negative case in this study.

Structured data query

Supplements and warfarin—Both structured and unstructured data were used in the search for supplement use. Since some supplement products were also included in the medication tables, we queried them using the common names and the lexical variants of supplements as above. Warfarin information was also retrieved from the medication table.

Adverse events—A list of ICD9-CM codes (International Classification of Disease, Ninth Revision, Clinical Modification) was used to identify patients having adverse events, which included gastrointestinal (GI) bleeding, general bleeding (GI bleeding included), embolic stroke and thromboembolism (see Table 1).

Data combination

Structured and unstructured data were combined to generate a comprehensive data set for each patient. Patients were further divided into two groups based on their use of seven supplements. The supplements-reported group include patients who have at least one prescription record showing that they take warfarin and at least one of the seven supplements concurrently. The warfarin-only group include patients taking warfarin only and do not have exposure to any of the seven supplements during the time period six months before the warfarin initiation to the first occurrence of an adverse event of interest based on the EHR data. We used the note date as the supplement’s start date if the start date was not specifically mentioned in the clinical notes and medication table. For example, “Pt has started taking ginseng” and “Patient has been taking garlic” indicate that the patient has already started taking the supplements before the visit; however, detailed information about the start date is unavailable in the notes. Additionally, it usually takes weeks for warfarin to demonstrate a stable therapeutic level. The initial titration phase may be a time of increased risk of adverse events until a stable warfarin dose is reached. To reduce the drug titration effect bias, the first 30 days of warfarin use were eliminated for both groups [13]. For the supplements-reported group, day 1 was the first day when any of the supplements were first noted in the EHR after eliminating the first 30 days of warfarin use. For the warfarin-only group, day 1 was actually the day 31 for the warfarin use.

Survival analysis

Cox proportional hazards models were fitted to compare the hazard of adverse events between two groups, controlling for a set of confounding factors including age, gender, and comorbidities. All the patients were followed for one year for the first occurrence of adverse events. Follow-up ended with the first adverse event, or the end of the warfarin therapy. Kaplan-Meier survival curves were also plotted.

Results

A total of 41,257 patients were included in the study, among which 2,640 subjects were in the supplements-reported group who took warfarin and at least one of the seven supplements concurrently. The control group included 38,617 subjects in the warfarin-only exposures.

The number of patients taking each of the seven supplements were counted based on the information from both structured and unstructured data. The results in Table 2 indicate that the identification of supplement use was much larger with the use of the combination of structured and unstructured data approach, especially for garlic and ginger, since much of the information about dietary supplements related to food such as “garlic bread” and “ginger tea” were detected by our NLP module.

The hazard ratio, 95% CI, and p-value for the four adverse events are listed in Table 3. The results show that the hazard ratio of the four adverse events in the supplements-reported group are statistically significant and higher in the supplement exposed patients when compared with the warfarin-only group. The results indicate taking warfarin concurrently with supplements is associated with side effects such as bleeding, or therapeutic failure events like embolic stroke.

The Kaplan-Meier survival curves for four adverse events were shown in Figure 1. The results of the log-rank test indicate the survival curves for the supplements-reported group and the warfarin-only group are significantly different ($P<0.01$) in GI bleeding, general bleeding, and embolic stroke, however, for thromboembolism, there is no significant difference in the curves between the supplements-reported group and warfarin-only group.

Discussion

The literature has shown there may be potential adverse interactions between warfarin and supplements, however, in many cases, the limited available data impedes the assessment of the potential risk associated with concurrent use. Additionally, the data is hard to detect in clinical trials due to limited sample sizes and high costs for the evaluations.

Due to its blood thinning effect, patients on warfarin are warned to be careful taking other supplements, such as ginkgo, ginger, Vitamin E, which can potentially increase the risk of bleeding events. Our confirmation of these potential adverse interactions provides evidence to support the current clinical guidance and provides data to assess drug safety with DSIs. For example, alfalfa contains a large amount of Vitamin K, which can reduce the anticoagulant activity of warfarin [14]. Alfalfa was part of the original research on Vitamin K metabolism and was one of the first substances on which Vitamin K was synthesized. This finding is consistent with the expected response with Vitamin K directly reversing the effects of warfarin. Taking St. John’s Wort induces cytochrome P450 2C19 which may clinically affect warfarin [15]. Ginkgo also affects the CYP3AY path by inducing the enzyme which may affect the R-enantiomer of warfarin [16].

This study demonstrates the feasibility of using clinical data from EHR to detect the signals for adverse events associated with drug and supplement interactions. The results of the study

as noted in the hazard ratio indicate a higher risk of adverse events and therapeutic failure well beyond typical screening triggers to assess the signal for potential adverse events. Additional assessment of the clinical cases will be needed to confirm the temporal and pharmacological patterns with the results to better assess the risk of exposure to supplements. However, from a medication safety perspective, the approach substantially reduces the assessment effort by patient safety officers or clinician managers to identify potential drug safety issues.

Though we used structured data for our outcomes assessment, our results provide support for the use of unstructured data to assess clinical exposures and outcomes. In addition, the combination of structured data (i.e., structured medication table) with unstructured data (i.e., clinical notes) in identifying supplements use has shown that clinical notes contain valuable information related to supplements which can complement structured data for DSI detection in the EHR. It is noted from our study that very little information about supplements is stored in the medication table since dietary supplements are regulated as food and can be obtained over the counter without a prescription, consequently, much of the information about supplement consumption is documented in clinical notes during the medical encounter. Therefore, the combination of text information with a structured medication module is necessary for the supplements use identification, where NLP is essential for extracting supplements use related information from clinical notes.

One limitation of the study is that for some patients, the actual start date of supplements is before the date of the clinical note because we found some patterns like ‘she has started ginger two months ago’, which could lead to misclassification of the exposure in the statistical analysis. The data has limitations on both the medication and supplement orders, which limits the ability to directly assess the association between exposures and clinical outcomes. Correlation of these results with other peripheral data sources such as retail data, if available, could help better identify the acquisition of supplements for presumed use by patients. Patient diaries, medication adherence apps and other sources could also be considered as part of usual care processes to better identify supplement use.

The second limitation is that we only applied a limited set of common names of supplements in the search for notes related to supplement use which may have missed some supplement information. The recall of supplement information might be increased when using more complete supplement terms.

Another limitation of the study is that we did not take into consideration the end date of the supplements since this information was often missing. Such information may also be contained in the clinical notes but requires additional analysis. Future work including the development of the NLP system to accurately extract temporal supplement information from the clinical notes could better assess the relationships of supplement exposures to medication use and clinical outcomes.

Conclusion

This study indicates that it is possible to use existing EHR data to detect signals of DSIs. The current findings also demonstrate the feasibility of applying NLP methods to extract supplement usage information from clinical notes. Furthermore, these methods can likely be extended to detect other potential drug and supplement interactions providing an important approach for post-market surveillance for DSI as well as drug-drug interactions.

Acknowledgments

The project was partly supported by the National Center for Complementary & Integrative Health Award (R01AT009457) (Zhang), the Agency for Healthcare Research & Quality grant (#1R01HS022085-01) (Melton), the University of Minnesota Clinical and Translational Science Award (8UL1TR000114) (Blazer), and the UMN grant-in-aid award (Zhang). The clinical data was provided by the University of Minnesota's Clinical Translational Science Institute (CTSI) Informatics Consulting Service. The authors thank Fairview Health Services for their data support of this research.

References

1. Levy I, Attias S, Ben-Arye E, Goldstein L, Schiff E. Adverse events associated with interactions with dietary and herbal supplements among inpatients. Br J Clin Pharmacol. 2016
2. Zhang R, Simon G, Cairelli MJ, Rindflesch T, Pakhomov S, Melton GB. Mining Biomedical Literature to Explore Interactions between Cancer Drugs and Dietary Supplements. AMIA Jt Summits Transl Sci Proc. 2015; 2015:69–73. [PubMed: 26306241]
3. Vaes LPJ, Chyka PA. Interactions of warfarin with garlic, ginger, ginkgo, or ginseng: Nature of the evidence. Annals of Pharmacotherapy. 2000; 34:1478–1482. [PubMed: 11144706]
4. Ulbricht C, Chao W, Costa D, Rusie-Seamon E, Weissner W, Woods J. Clinical Evidence of Herb-Drug Interactions: A Systematic Review by the Natural Standard Research Collaboration. Current Drug Metabolism. 2008; 9:1063–1120. [PubMed: 19075623]
5. Podszun M, Frank J. Vitamin E-drug interactions: molecular basis and clinical relevance. Nutr Res Rev. 2014; 27:215–231. [PubMed: 25225959]
6. [Accessed. December 5, 2016] Natural Medicines Comprehensive Databases [Internet]. http://naturaldatabase.therapeuticresearch.com
7. Zhang, R., Manohar, N., Arsoniadis, E., Wang, Y., Adam, TJ., Pahkhomov, S., Melton, GB. Evaluating Term Coverage of Herbal and Dietary Supplements in Electronic Health Records. AMIA Annu Symp Proc; 2015; 2015. p. 1361-1370.
8. [Accessed September 20, 2015] BioMedICUS [Internet]. Available from: https://github.com/nlpie/biomedicus
9. [Accessed September 20, 2015] UIMA [Internet]. Available from: https://uima.apache.org/
10. [Accessed September 20, 2015] NLP-PIER [Internet]. http://athena.ahc.umn.edu/nlppier/
11. McEwan, R., Melton, GB., Knoll, BC., Wang, Y., Hultman, G., Dale, JL., Meyer, T., Pakhomov, S. NLP-PIER: A Scalable Natural Language Processing, Indexing, and Searching Architecture for Clinical Notes. AMIA Joint Summits Proceedings; 2016; p. 150-9.
12. Fan YD, He L, Pakhomov S, Melton GB, Zhang R. Classifying Supplement Use Status in Clinical Notes. AMIA Jt Summits Transl Sci Proc. 2017; 2017:493–501. [PubMed: 28815149]
13. Stoddard, GJ., Shane-McWhorter, L., et al. Ginkgo and Warfarin Interaction in a Large Veterans Administration Population. AMIA Annu Symp Proc; 2015; p. 1174-1183.
14. Newall, CA., Anderson, LA., Philpson, JD. Herbal Medicine: A Guide for Healthcare Professionals. London, UK: The Pharmaceutical Press; 1996.
15. Wang LS, Zhu B, Abd El-Aty A, et al. The influence of St. John's wort on CYP2C19 activity with respect to genotype. J Clin Pharmacol. 2004; 44:577–81. [PubMed: 15145964]

16. Deng Y, Bi HC, Zhao LZ, Wang XD, Chen J, Ou ZM, Ding L, Xu LJ, Guan S, Chen X, Zhou SF, Huang M. Induction of Cytochrome P450 3A by the Ginkgo biloba Extract and Bilobalides in Human and Rat Primary Hepatocytes. Drug Metab Lett. 2:6.

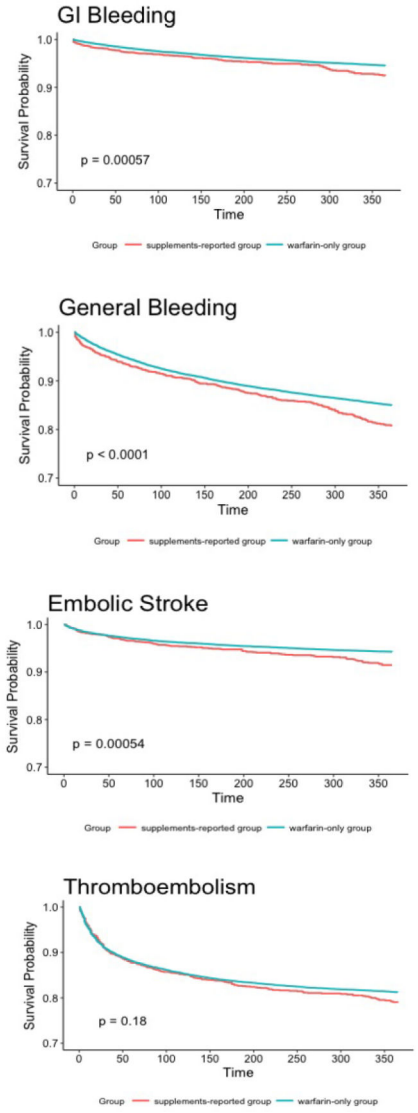

Figure 1.
Kaplan-Meier Curves for Adverse Events

Table 1

ICD-9 Codes for Adverse Events

Diagnosis	ICD-9 Codes
GI bleeding	530.7, 530.82, 531.2, 531.4, 531.6, 532.2, 532.4, 532.6, 533.2, 533.4, 533.6, 534.2, 534.4, 534.6, 535.×1, 537.83, 562.02, 562.03, 562.12, 562.13, 569.3, 578.x
general bleeding	964.2, 964.5, E934.2, E934.5, 459.0, 285.1, 286.59, 362.81, 596.7, 599.70, 599.71, 719.1x, 782.7, 784.7, 784.8, 786.30, 786.39, 423.0, 423.1, 423.9, 568.81, 530.7, 530.82, 531.2, 531.4, 531.6, 532.2, 532.4, 532.6, 533.2, 533.4, 533.6, 534.2, 534.4, 534.6, 535.×1, 537.83, 562.02, 562.03, 562.12, 562.13, 569.3,e 578.x, 430, 431
embolic stroke	346.6x, 433.01, 433.11, 433.21, 433.31, 433.81, 433.91, 434, 434.01, 434.1, 434.10, 434.11, 434.9, 434.90, 434.91, 436
thromboembolism	451.1x, 453.4x, 453.5x, 453.8, 453.9, 415.1x

Table 2

The Number of Patients with Mention of Supplements Use in Structured and Unstructured Data

Supplements	Structured	Structured and Unstructured
alfalfa	30	68
garlic	329	925
ginger	100	1296
ginkgo	141	276
ginseng	42	109
Vitamin E	2273	4145
St. John’s Wort	22	44

Table 3

Multivariable Cox Proportional Hazards Regression for Adverse Events (supplements-reported group VS. warfarin-only group)

Adverse events	HR (95% CI)	p-value
GI Bleeding	1.30 (1.08, 1.57)	**0.005**
General Bleeding	1.20 (1.07, 1.34)	**0.002**
Embolic Stroke	1.27 (1.06, 1.51)	**0.008**
Thromboembolism	1.13 (1.02, 1.25)	**0.021**