The association between Non-vitamin K antagonist oral anticoagulants and gastrointestinal bleeding: a meta-analysis of observational studies

Ying HE¹, Ian C K WONG¹,³, Xue LI¹, Shweta ANAND¹, Wai K LEUNG², Chung Wah SIU², Esther W CHAN¹

¹Centre for Safe Medication Practice and Research, Department of Pharmacology and Pharmacy, The University of Hong Kong, Hong Kong; ²Department of Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong; ³Research Department of Practice and Policy, School of Pharmacy, University College London, London, UK.

Running title: NOAC and GIB: meta-analysis of observational studies

Keywords: dabigatran, rivaroxaban, gastrointestinal bleeding, real-life
Correspondence to: Dr. Esther W Chan

Centre for Safe Medication Practice and Research

Department of Pharmacology and Pharmacy

Li Ka Shing Faculty of Medicine

The University of Hong Kong

2/F Laboratory Block FMB, 21 Sassoon Road

Hong Kong SAR, China

Tel: (852) 3917 9029

Fax: (852) 2817 0859

Email: ewchan@hku.hk

Word count: 4,898

No. of Tables: 2

No. of Figures: 4

No. of Supplementary Tables: 2

No. of Supplementary Figures: 10
Summary: (Word count: 250)

Aims: Particular concerns have been raised regarding the association between Non-vitamin K antagonist oral anticoagulants (NOACs) and the risk of gastrointestinal bleeding (GIB); however current findings are still inconclusive. We conducted a systematic review with meta-analysis to examine the association between NOACs and GIB in real-life settings.

Methods: We performed a systematic search of PubMed, EMBASE, and CINAHL Plus up to September 2015. Observational studies that evaluated exposure to NOACs reporting GIB outcomes were included. The inverse variance method with random-effects model was used to calculate the pooled estimates.

Results: Eight cohort studies were included in the primary meta-analysis, enrolling 1,442 GIB cases among 106,626 dabigatran users (49,486 patient-years), and 184 GIB cases among 10,713 rivaroxaban users (4,046 patient-years). The pooled incidence rates of GIB were 4.50 (95%CI 3.17-5.84) and 7.18 (95%CI 2.42-12.0) per 100 patient-years among dabigatran and rivaroxaban users, respectively. The summary risk ratio (RR) was 1.21 (95%CI 1.05-1.39) for dabigatran compared to warfarin, and 1.09 (95%CI 0.92-1.30) for rivaroxaban. Subgroup analyses showed a dose effect of dabigatran, with a significantly higher risk of GIB for 150mg bid (RR=1.51, 95%CI 1.34-1.70) but not for 75mg bid or 110mg bid. In addition, the use of proton pump inhibitors (PPIs)/histamine-2 receptor antagonists (H2RAs) influenced the association in dabigatran users, whereas this effect was modest among rivaroxaban users.

Conclusion: Our meta-analysis of real-life observational studies suggests a slightly higher risk of GIB with dabigatran use compared to warfarin, whereas no significant difference was found between rivaroxaban and warfarin for GIB risk.
What is known about this subject:

- Meta-analyses of randomised controlled trials (RCTs) on the association of non-vitamin K antagonist oral anticoagulants (NOACs) use and the risk of gastrointestinal bleeding (GIB) provide inconsistent results. It has been reported that dabigatran and rivaroxaban are associated with an increased overall GIB risk of approximately 50% compared to warfarin use (regardless of severity). However, it has also been reported that NOACs are not associated with a risk of major GIB compared to standard care.

- Post-marketing pharmacovigilance studies and observational studies have been frequently conducted to investigate this association in the real-world practice, however the findings are conflicting.

What this study adds:

- In real-life settings, dabigatran was associated with a slightly higher risk of GIB compared to warfarin.

- No significant difference was found between rivaroxaban and warfarin in the risk of causing GIB.

- Dabigatran and rivaroxaban do not raise unexpected safety concerns in terms of GIB compared with findings from RCTs.
**Competing Interests Disclosures**

All authors have completed the Unified Competing Interest form at http://www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: EWC has received financial support from Janssen, a division of Johnson and Johnson, Bristol Myers Squibb, Pfizer and Eisai; The Pharmaceutical Society of Hong Kong; The University of Hong Kong; Early Career Scheme and the General Research Fund, Research Grants Council, Hong Kong, all unrelated to the current work. WKL has received honorarium for attending advisory board meeting from Boeringher Ingelheim. Other authors disclose no conflicts.
Introduction

As alternatives to traditional standard anticoagulant warfarin, a new generation of oral anticoagulants has been approved for stroke prevention in atrial fibrillation (AF), prevention and/or treatment of venous thromboembolism (VTE).[1] Unlike vitamin K antagonist (VKA) warfarin, non-vitamin K antagonist oral anticoagulants (NOACs) directly target individual clotting proteins, including direct thrombin inhibitor dabigatran[2] and direct factor Xa inhibitors (rivaroxaban[3], apixaban[4] and edoxaban[5]). Compared to warfarin, NOACs have a rapid onset of action, no food interactions and fewer drug interactions, and do not require bridging with subcutaneous administration of shorter-acting VKA low-molecular-weight-heparins (LMWH). Large randomized controlled trials (RCTs) show superior or non-inferior[2-5] efficacy of NOACs compared to warfarin in stroke prevention in AF and prevention/treatment of VTE[6-8], with lower or non-inferior risk of major bleeding.

Although NOACs are recommended by several guidelines from US[9], Canada[10] and Europe[11], gastrointestinal bleeding (GIB) remains a major concern of NOAC use. NOAC-associated GIB is potentially severe and even fatal[12-14]. The possible non-adherence of NOACs due to GIB may increase the risk of stroke and thromboembolism[15]. However, existing findings from meta-analyses of RCTs are heterogeneous. A meta-analysis of 43 pivotal trials by Holster et al.[16] first reported a 1.6 or 1.5 fold increased GIB risk among dabigatran or rivaroxaban users compared to standard care, respectively. Subsequent meta-analyses have reported varying association between NOACs and GIB, indicating no increase in risk[17-19] or a significant but marginal increased risk[20-22].
In 2012, US Food and Drug Administration (FDA) reported a 1.6 to 2.2 times higher GIB risk for warfarin than dabigatran based on a post-marketing Mini Sentinel Modular Program analysis [23, 24], which contradicted RCT findings[16]. However, in May 2014, the FDA gave an updated dabigatran safety announcement which reported a higher GIB risk (hazard ratio, HR=1.28, 95%CI 1.14-1.44) compared to warfarin[25, 26]. Recently, post-marketing pharmacovigilance studies using available public databases of spontaneous adverse drug reactions (ADRs) from Japan[27], Australia, Canada and US[24, 28-30] also reported signals of GIB risk among NOAC users. Moreover, increasing observational studies[26, 31-45] have been continuously conducted to investigate the association of GIB risk and NOAC use in the real-world clinical practice, with conflicting results. On that notion, clarification about this association is needed, especially to compare results from the real-life setting versus pivotal RCTs.

To resolve this issue, we conducted a systematic review with meta-analysis of published observational studies to clarify the association between NOAC use and GIB, and also investigate the effects of various factors that may affect GIB risk.

**Materials and Methods**

This systematic review was conducted following guidance provided by the Cochrane Handbook [46] and is reported in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement[47] for the flowchart of study inclusion and exclusion; and the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) statement[48] for overall reporting.
Study definitions

The exposure of interest was defined as exposure to NOAC or warfarin in clinical setting. The different doses of NOACs studied in this meta-analysis were indication-specific daily doses based on recommendations by the FDA or the European Medicines Agency (EMA).[16] The outcome of this meta-analysis was the risk of GIB. In observational studies, GIB was defined as any bleeding in the GI tract that was identified through medical records or by International Classification of Diseases, Ninth or Tenth Revision, Clinical Modification (ICD-9-CM or ICD-10-CM) codes, as described in the original literature. The classification of the severity of GIB was based on the description in the original studies[26, 38]. Major GIB was defined as a fatal GI hemorrhagic event, or a severe GIB event resulting in hospitalization or even requiring transfusion[26, 38], and the remaining were defined as non-major GIB.

Data sources and search strategy

A systematic literature search was conducted using PubMed, EMBASE and CINAHL Plus with the search strategy: (Gastrointestinal ulcer OR Peptic ulcer OR gastric ulcer OR duodenal ulcer OR gastrojejunal ulcer OR stomach ulcer OR peptic ulcer disease OR Gastrointestinal bleeding OR Gastrointestinal Hemorrhage OR Peptic Ulcer Hemorrhage OR GI hemorrhage OR GI bleeding OR GI bleed) AND (dabigatran OR rivaroxaban OR apixaban OR edoxaban OR pradaxa OR xarelto OR eliquis OR lixiana OR new oral anticoagulant OR novel oral anticoagulant OR direct oral anticoagulant OR oral anticoagulant OR TSOAC). Key words, MeSH and Emtree terms were used where appropriate. All databases were searched up to 28 September 2015. English titles and abstracts were screened and full texts of relevant articles
were retrieved for further review to identify relevant studies. The bibliographies of review articles were also searched to identify any pertinent studies.

**Study selection**

Studies included in this meta-analysis were comparative observational studies that investigated the association between NOAC use and the risk of GIB. Studies were included if they: (i) clearly defined exposure to NOACs and other comparative exposure groups; (ii) clearly defined the outcome of GIB; (iii) reported HRs, relative risks (RR) or odds ratios (OR) or provided data for calculation of HR/RR/OR. There were no restrictions on study size. Conference proceedings were excluded as we were unable to assess the quality of these studies and some may be preliminary results. Single-arm observational studies including case series/case reports/medical chart review studies without comparisons were excluded. Studies were excluded if there was insufficient data for determining the HR/RR/OR with 95% confidence intervals (CI).

**Quality assessment**

The methodological quality of the included cohort and cross-sectional studies were assessed using the Newcastle-Ottawa scale[49] as recommended by the Cochrane Collaboration. This scale provides specific criteria to assess the study selection (four items), comparability (two items) and ascertainment of exposure/outcome (three items). The ‘high’ quality items were scored with a ‘star’ and the maximum score would be nine, the scale was used previously in many published meta-analyses[50-52]. A final score ≥7 was considered high quality. Two researchers (YH and SA) independently reviewed and scored each of the studies. Any discrepancies were addressed by re-evaluation to reach consensus.
Data abstraction

YH and XL performed the initial searches, screened abstracts and full texts for eligibility. Study characteristics and measure of effect were extracted using a standardized data collection form by YH, XL and SA independently. For each study, the risk estimates that were adjusted for the largest number of confounding variables were extracted for analysis. The incidence rates of GIB among specific NOAC users were also extracted. For studies that did not report an adjusted result [53, 54], the unadjusted result was extracted for analysis. If the unadjusted estimates were not directly reported, patient and event number, patient-years or other available data were used to compute the HR/RR/OR manually. For the Danish studies that only reported dose-specific estimates of dabigatran (110mg bid and 150mg bid separately) [34, 43], the HR of high dose dabigatran was used. This is because all other cohort studies were from the US, where 150mg bid dabigatran was the most commonly used dose, while 110mg bid was not available. The dose-specific association between dabigatran and risk of GIB was examined in the subgroup analysis.

Statistical methods

Using data available from original studies, unreported incidence rate and crude incidence rate ratio were calculated manually based on Rothman/Greenland’s formula[55] or computed using Review Manager 5.3. Meta-analyses with forest plots of comparative groups were generated using Review Manager 5.3.[56] The inverse variance method with random-effects model was used to compute the pooled estimates and 95% CI. Heterogeneity was assessed using Cochran’s Q statistical test and $P < 0.10$ was considered significant. $I^2$ statistic was also calculated to estimate the proportion of total variation among studies, where a value of 25%, 50% and 75% was regarded as low, moderate and high heterogeneous, retrospectively[57]. The point estimate
of the pooled incident rates (IR) of GIB among dabigatran and rivaroxaban users were analyzed respectively with random-effects model, using an established and validated module published by Neyeloff et al.[58] Forest plots of point estimate meta-analysis were generated using SAS version 9.3 (SAS Institute Inc., Cary, NC, United States). P-values (two-tailed) <0.05 was regarded as statistically significant, except for heterogeneity test. Publication bias test was not performed, as funnel plots are more informative in cases where ten or more studies are included in the meta-analysis[59]. Otherwise, the power of the test is considered too low to distinguish between chance and real asymmetry.

The primary analysis focused on assessing the risk of GIB among NOAC users in cohort studies. HR, RR or OR with their 95% CIs in the observational studies were extracted and log-transformed for the meta-analysis. The pooled estimates were reported as RRs. Analyses were performed on both crude and adjusted estimates from the studies.

Due to the heterogeneity of study design and the study quality, three cross-sectional studies[31, 42, 60] which also investigated the GIB outcome of NOAC users were included in the secondary meta-analyses. The cross-sectional studies did not adjust for confounding factors and no follow up was allowed, and therefore were of a lower quality of evidence.

Post-hoc sensitivity analyses were performed to assess the robustness of the main results. Of the included studies, Larsen et al.[32, 34] and Staerk et al.[43] used Danish registry database, while Graham et al.[26] and Hernandez et al.[40] used US Medicare database with overlapping study period. To rule out the influence of possible repeated subjects, studies with better representative population, better study design, and longer follow-up time or with most recent publication date were included in the primary analysis. We performed a sensitivity analysis by substituting the
results of Larsen 2014[34] and Graham 2015[26] with those of Staerk 2015[43]/Larsen 2013[32] and Hernandez 2015[40], respectively to test the robustness of this inclusion criteria. Moreover, another sensitivity analysis was performed by removing Abraham et al.[37] with obvious heterogeneous result to examine the influence of this study on the pooled estimate. Lastly, a sensitivity analysis was carried out by removing Chan et al.[38] to investigate the effect of the inclusion of study specifically on patients with renal dysfunction.

Pre-defined subgroup analyses stratified according to study characteristics were also carried out to investigate the source of heterogeneity, which included different comparison groups of NOAC type, study location (US and Denmark), indication of NOAC use, dabigatran doses (75mg bid, 110mg bid and 150mg bid), patient age group (≥18y, 18-64y and ≥65y), GIB severity (major vs non-major). Studies were further stratified based on whether there was adjustment of the history of warfarin use, and use of drugs that were associated with GIB including non-steroidal anti-inflammatory drugs (NSAIDs), aspirin/antiplatelet agents, steroids, selective serotonin reuptake inhibitors (SSRI) and gastroprotective agents including proton pump inhibitors (PPIs) and histamine type-2 antagonists (H2RAs).

Results

Search results

A total of 1,634 records were retrieved from the literature search, and 51 records were identified through bibliographies and other sources. After removal of duplicates, 1,427 titles and abstracts were screened and full texts of 81 articles were further reviewed. Sixteen studies met the criteria for this systematic review (Figure 1), including 11 cohort studies and 5 cross-sectional studies. Among them, eight cohort studies[26, 33, 34, 36-39, 41] investigating the GIB risk of NOAC
compared to warfarin were included in the primary meta-analysis and the remaining three[32, 40, 43] were included for sensitivity analysis to test the influence of possible repeated subjects. Three cross sectional studies[31, 42, 60] were included in the secondary meta-analysis and the remaining two[35, 45] were narratively reviewed.

**Characteristics and quality of included studies**

As shown in Supplementary Table 1, the quality scores of the cohort studies ranged from 6 to 9, which are of modest to high quality (Newcastle-Ottawa score ≥7). The cross-sectional studies were of lower quality with scores 2 to 7. Study characteristics and results from individual studies are summarized in Table 1 and Supplementary Table 2. The cohort studies were conducted in US (n=8) and Denmark (n=3). Five studies used propensity score analysis[26, 33, 37, 39, 41] either by matching or adjusted in statistical model to minimize the influence of confounding factors and heterogeneity of patient characteristics between comparison groups. Studies were matched for demographic variables, and several also accounted for other potential confounders, including the use of drugs that are associated with GIB (NSAIDs, aspirin, steroid and SSRI), gastroprotective agents (PPI or H2RA), as well as comorbidities (cardiovascular and cerebrovascular diseases, diabetes and GIB history).

**Risk of gastrointestinal bleeding**

*Primary analysis: Cohort studies*

Eight cohort studies were included in the primary analysis, enrolling 106,626 patients on dabigatran and 10,713 on rivaroxaban. Abraham et al.[37], Chan et al.[38] and Chang et al.[39] compared both dabigatran and rivaroxaban with warfarin, and the subgroup by different NOACs
were also stratified in the primary result. The crude analysis showed a trend of increased GIB risk in both the dabigatran and rivaroxaban group compared to warfarin use, albeit not significant (Figure 2A). The adjusted estimate showed that dabigatran was associated with a 21% increase of GIB risk compared to warfarin, with an RR of 1.21 (1.05-1.39) with a moderate to high heterogeneity ($I^2=69\%, p_{heterogeneity}=0.004$; Figure 2B), whereas rivaroxaban show no higher risk of GIB than warfarin (RR=1.09, 95%CI 0.92-1.30) with no substantial heterogeneity ($I^2=0\%, p_{heterogeneity}=0.42$; Figure 2B).

**Secondary analysis: Cross-sectional studies**

Of the five cross-sectional studies with comparison arms of NOAC users, three studies[31, 42, 60] compared the use of dabigatran vs warfarin, while Nagao et al.[45] specifically compared the GIB outcome in apixaban versus warfarin. Sherid et al.[35] firstly conducted a direct head-to-head comparison of rivaroxaban versus dabigatran in the GIB risk. Meta-analysis showed a trend of increased GIB risk among dabigatran compared to warfarin users, albeit not significant (RR=1.47, 95%CI 0.45-4.85), with a high heterogeneity ($I^2=88\%, p_{heterogeneity}<0.001$; Supplementary Figure 1). In Nagal et al.[45], there was one GIB event among 105 apixaban users but none in 105 warfarin users, which led to a wide CI. Sherid et al.[35] reported 7 GIB events (4.8%) among 147 rivaroxaban users compared to 12 GIB events (5.3%) among 227 dabigatran users, of which the crude analysis showed no significant difference between rivaroxaban and dabigatran (RR=0.90, 95%CI 0.36-2.24, data not shown) (Supplementary Table 2).

*Incidence rate of GIB*
In total, our meta-analysis of eight cohort studies enrolled 1,442 GIB cases among 106,626 dabigatran users (49,486 patient-years), and 184 GIB cases among 10,713 rivaroxaban users (4,046 patient-years) (Supplementary Table 2). The pooled incidence rate of GIB was 4.50 (95%CI 3.17-5.84) and 7.18 (95%CI 2.42-12.0) per 100 patient-years for dabigatran and rivaroxaban, respectively (Figure 3).

Sensitivity analyses

We performed four sensitivity analyses (Table 2). First, sensitivity analysis by substituting the results of Larsen et al.[34] and Graham et al.[26] with those of Staerk et al.[43] and Hernandez et al. [40] respectively yielded similar results to the primary analysis (Figure 2), with an RR of 1.30 (95%CI 1.01-1.66) for dabigatran compared to warfarin. Substituting Larsen 2014 results with Larsen 2013 yielded a similar but non-significant result (RR 1.24, 95%CI 0.98-1.59). Moreover, The I² reduced from 69% to 48% with omission of Abraham et al.[37], which led to a similar RR of 1.27 (95% 1.13-1.41) compared to the main result of 1.21 (95%CI 1.05-1.39). The fourth sensitivity analysis conducted by removing Chan et al. [38] also show similar results with the main analysis. Results of our sensitivity analyses support the robustness of the main results.

Subgroup analyses

A series of subgroup analyses were performed on the eight cohort studies included in the primary analysis to examine how the different factors may affect the risk of developing GIB in NOAC users (Table 2, Supplementary Figure 2-10). Results of dabigatran and rivaroxaban groups were stratified accordingly. Subgroup analyses based on dabigatran dose showed that dabigatran 150mg bid was associated with a significantly higher risk of GIB, with a RR of 1.51 (95%CI 1.34-1.70). However, this association was marginal for 75mg bid (RR=1.01, 95%CI 0.78-1.31).
and was not reflected for the 110mg bid dose (RR=0.53, 95%CI 0.29-0.98). However, it is important to note that the data for 75mg bid and 110mg bid were from separate single studies, respectively (Figure 4). Subgroup analysis by different indications for dabigatran or rivaroxaban use yielded similar results between subgroups (Table 2, Supplementary Figure 2).

By stratifying major and non-major GIB, this showed that dabigatran was associated with a significantly increased risk of major GIB (RR=1.30, 95%CI 1.17-1.46), however no significant association was observed with non-major GIB (RR=0.85, 95%CI 0.40-1.79) (Table 2, Supplementary Figure 3).

In the cohort studies, the majority (60%-97%) of patients were 65 years or above (Supplementary Table 2). Specifically, subjects from Graham et al.[26] were all 65 years or above. Stratified analysis of the older NOAC users (≥ 65y) showed a higher GIB risk when on dabigatran or rivaroxaban compared to that of adult patients in general (Table 2, Supplementary Figure 4), although it was not statistically significant.

Subgroup analysis also showed that dabigatran was associated with a trend of non-significantly higher GIB risk among warfarin-experienced patients than warfarin-naïve patients (Table 2, Supplementary Figure 5). When considering the adjustment of different confounders, this showed a high impact of concomitant PPI/H2 use ($I^2=83.8\%, p_{\text{heterogeneity}}=0.01$) on dabigatran-associated GIB, whereas the effect on rivaroxaban-associated GIB was relatively modest ($I^2=46.7\%, p_{\text{heterogeneity}}=0.17$). A 1.3-fold increase of GIB risk with dabigatran use was only observed in studies that did not adjust for the use of antiplatelet agents, steroid or SSRI, however the difference between subgroups was not significant. There were no significant differences
between subgroups that had adjusted or not adjusted for NSAID use (Table 2, Supplementary Figure 6-10).

**Discussion**

To our knowledge, this is the first meta-analysis of observational studies investigating the association of NOAC use and the risk of GIB in the real-world setting. We undertook a rigorous systematic review and meta-analysis with data extraction and statistical analysis by independent reviewers. Based on our results, a slightly higher risk of GIB in dabigatran use compared to warfarin was observed, with an RR of 1.21 (95% CI 1.05-1.39). No significant difference was found between rivaroxaban and warfarin for GIB risk. (RR=1.09, 95%CI 0.92-1.30).

**Heterogeneity**

Notably, there was substantial heterogeneity between studies on dabigatran for both meta-analyses on cohort studies (I²=69%, p heterogeneity=0.004) and cross-sectional studies (I²=88%, p heterogeneity<0.001). However, there seemed low or no substantial heterogeneity between included studies for rivaroxaban.

Our sensitivity analysis showed that this substantial heterogeneity was mainly attributed to Abraham et al.[37]. However, this did not significantly change the pooled result, with an even slightly higher RR (1.27 vs 1.21 in primary analysis) with overlapping confidence intervals. This finding suggests no substantial influence of the heterogeneity and the possible association between dabigatran use and an increased risk of GIB cannot be ruled out.

For cross-sectional studies, Choi et al.[60] was survey-based and the other two studies were based on hospital medical records. Additionally, the sample size in Sherid et al.[42] and Choi et
were limited, which may affect the robustness of the result. Therefore, findings from this secondary analysis should be interpreted cautiously due to the intrinsic heterogeneity of included studies and limited number of available studies.

**Risk of GIB: Comparison of our meta-analysis to findings from RCTs**

Four landmark phase III RCTs, the RE-LY trial for dabigatran[2], the ROCKET-AF trial for rivaroxaban[3], the ARISTOTLE trial for apixaban[4], the ENGAGE AF-TIMI 48 for edoxaban[5] were conducted which supported the approval of each NOACs. Various meta-analyses of RCTs have been published to investigate the association of NOACs and GIB. A meta-analysis by Holster *et al.*[16] reported that NOACs were associated with a significantly higher risk of GIB, with a pooled OR of 1.45 (95%CI 1.07-1.97), compared with standard care (defined as the use of any of the following: LMWH, a VKA, antiplatelet agent, placebo, or no additional therapy). The GIB risk varied among different NOACs, with ORs of 1.58 (95% CI 1.29-1.93) for dabigatran, 1.48 (95%CI 1.21-1.82) for rivaroxaban[16]. Subsequently, two meta-analyses including more updated trials reported a significant but modest association of all NOACs and GIB, with RRs of 1.25 (95%CI 1.01-1.55) from Ruff *et al.*[20] and 1.23 (95%CI 1.03-1.46) from Loffredo *et al.*[21]. However, these two studies did not report results of specific NOACs separately. On the contrary, findings from three recent meta-analyses [17-19, 61] did not support an association between NOACs and GIB, of which separate risks of specific NOACs were not reported. Additionally, a recent meta-analysis by Caldeira *et al.*[62], which included the most recent RCTs, reported that NOACs were not associated with risk of major GIB.

Our meta-analysis found an association between dabigatran use and risk of GIB and a non-significant association in rivaroxaban use, however both were slightly lower than that of
RCTs[16]. This suggests that in real-life practice, dabigatran and rivaroxaban do not cause more harm nor raise unexpected safety concerns compared with findings from RCTs.

A potential explanation for the lower risk of GIB in this meta-analysis of observational studies may be due to patient selection. In RCTs, only patients at high risk of stroke were recruited, resulting in an older group of patients than in real-life practice. In the RE-LY trial, the mean age was 71.4 (SD 8.6) and 71.5 (SD 8.8) for dabigatran 110mg bid and 150mg bid users, respectively[2]. However, NOAC users in five studies included in this meta-analysis had a mean age below 70 years (Supplementary Table 2) [36-39, 41]. Moreover, we cannot exclude the possibility of a ‘healthy subject effect’. Over 40% of subjects from some included observational studies had a low CHADS2 score of 0 to 1[33, 37]. However, only approximately a third of the RE-LY cohort and none of the ROCKET-AF cohort had a CHADS2 score of 0 to 1. With better awareness of the potential GIB-inducing safety issue with NOACs, clinicians may tend to prescribe NOACs to lower-risk patients, preventing exposure of NOACs to many high-risk patients. However, most of the studies have already controlled for various factors using propensity score adjustment[26, 33, 37, 39, 41], therefore the “healthy subject effect” is unlikely to be a significant issue in this analysis.

Notably, among the included studies, the outcome was specifically defined as major GIB in Graham et al.[26], while Chan et al.[38] stratified the outcomes and reported the results of minor and major GIB separately. It was reported in the meta-analysis of RCTs that dabigatran and rivaroxaban were associated with an approximately 50% increase of overall GIB risk compared to warfarin [16]; however they were not associated with the risk of major GIB[62]. Thus, the lower risk of GIB observed in our meta-analysis may also be partially attributed to the effect of combining overall GIB and major GIB in the outcome. Interestingly, in contrast to the results
reported by Caldeira et al.[62], our subgroup analysis found that dabigatran use was mainly associated with major GIB rather than non-major GIB (RR 1.30, 95%CI 1.17-1.46). However, due to the limited number of studies (n=2), we cannot draw firm conclusions (Supplementary Figure 2).

It is important to note that although results from our meta-analysis showed a significant association for dabigatran but non-significant for rivaroxaban, indirect comparison of the risk of GIB associated with dabigatran and rivaroxaban use should be evaluated with caution. Rivaroxaban is a newer agent compared to dabigatran, with fewer available real-life studies and shorter follow-up duration. Further, in included studies, patients receiving rivaroxaban were of a younger age compared to those on dabigatran. This might be another possible explanation of the lower risk of GIB found. Moreover, subgroup analysis showed that among patients with age ≥65 years, a 50% non-significant increased risk of GIB was observed both in dabigatran and rivaroxaban use compared to warfarin, respectively (Table 2, Supplementary Figure 3). Therefore, an increased risk of GIB in rivaroxaban users cannot be ruled out.

Risk factors of GIB

Subgroup analyses show that there was a dose effect with respect to the safety of dabigatran, with a higher risk of GIB in dabigatran 150mg bid group compared to the low-dose groups, which is consistent with findings from the RE-LY trial. However, there was only one study available for the dabigatran 75mg bid and 110mg bid subgroups, respectively. Data of different doses of dabigatran are still needed from further studies to confirm this finding. Furthermore, the subgroup analysis by different indications (AF and non-AF) showed similar results, probably due to the limited data available (only one study, Abraham et al. [37] reported data from non-AF
patients). Notably, the patient’s age would be a main risk factor for an increased GIB risk among NOAC users, thus specific caution is needed for prescription of NOAC among elderly patients. In the RE-LY trial, the increased risk of GIB was only observed in patients with age ≥75yr for both dabigatran 110mg bid (RR=1.39, 95%CI 1.03-1.98) and 150mg bid (RR=1.79, 95%CI 1.35-2.37) groups, but not in patients <75yr.[63] However, for rivaroxaban, a secondary analysis of the ROCKET-AF trial comparing outcomes in patients aged≥75 and <75 years show no significant difference between older and younger patients. This suggests that rivaroxaban may be an alternative for older patients.[64] Our meta-analysis investigated the age group ≥ 65yr and show a non-significant trend of higher risk of GIB among these elderly patients, possibly due to the scanty number of studies with available data.

In addition, normal renal function is crucial in the elimination of dabigatran[65]. A population pharmacokinetic analysis of the AF patients from the RE-LY trial suggested that patients with renal dysfunction warrants dose adjustment, which highlights the importance of renal function monitoring in patients already on or being considered for dabigatran[66]. In our meta-analysis, only Chan et al.[38] specifically focused on AF patients with end-stage renal disease who were on hemodialysis. It may partially explain the higher incidence of GIB among NOAC users in this study. However, sensitivity analysis by removing Chan et al. [38] show similar results with the main analysis (Table 2) and support the robustness of main analysis. Further studies are needed to investigate the efficacy and safety of NOACs in patients with renal dysfunction.

Moreover, prior history of warfarin use was also noted to be a factor that may influence the risk of GIB in NOAC users. Ezekowitz et al.[67] conducted a secondary analysis of RE-LY trial and reported a similarly higher GIB rate of dabigatran 150mg bid in both VKA-naïve and – experienced cohorts, whereas similar non-significant findings were observed for dabigatran
110mg bid in both cohorts. In our meta-analysis, Larsen et al. [34] specifically investigated both VKA-naïve and –experienced cohorts and reported similar GIB risk in both groups. Our subgroup analysis by the history of warfarin use suggested that previous warfarin exposure did not influence the GIB risk of dabigatran compared with warfarin, which is consistent with the RE-LY study findings.

Furthermore, discrepancies between the findings from subgroups of adjustment for the use of GPA (PPI/H2RA), antiplatelet agents, steroid and SSRI implied that concomitant use with these agents would affect the safety profile of NOACs. Further studies to investigate the potential protective role of GPA (PPI/H2RA) would be of great clinical impact[68], as there is currently no well-developed specific antidote for NOACs for the management of their ADRs. In addition, we were unable to investigate the drug-drug interaction directly as the original studies did not aim to evaluate drug-drug interaction hence such data were not reported in the original studies.

**Strengths and limitations**

Our study is the first meta-analysis to include all relevant, observational studies to provide a thorough summarized analysis and systematic review of NOAC safety in respect to GIB risk in the real-world setting. The eight cohort studies included in the primary analysis were of high quality in study design, as indicated by the Newcastle-Ottawa Scale quality assessment. Sensitivity analysis show that there was no significant difference between different studies using the same data source, which support the robustness of the main result.

This review however has several limitations. First, non-English studies were not included in the meta-analysis and potential language bias may not be ruled out. However this effect may be minimal and have diminished due to the shift of the publication tendency towards English
journals in recent decades [69]. Second, our meta-analysis attempted to investigate the influence of different dabigatran dose, patient’s age and GIB severity. However, the studies reporting corresponding outcomes were scant and therefore the study conclusions should be interpreted with caution. Further studies may be helpful to strengthen the evidence. Third, this meta-analysis compiles results from available non-randomised epidemiological studies. The risk of our results being influenced by bias and confounding factors is generally higher than that of meta-analyses of RCTs. In an attempt to overcome some of the confounder issues, we performed subgroup and sensitivity analyses to evaluate confounding factors including patient characteristics, concomitant diseases and medications, albeit some of the above mentioned factors were not analysed due to unavailability of the data. Some heterogeneity can be partially overcome through subgroup analysis by confounder adjustments. Sensitivity analyses by omitting certain specific studies significantly reduce the level of heterogeneity, and similar pooled results support the robustness of the primary analysis. Fourth, to date, Sherid et al.[35] is the only study available which reported a direct head-to-head comparison of rivaroxaban versus dabigatran. Due to the small sample size and relatively low power of this study, it is difficult to draw any firm conclusion. Likewise, apixaban and edoxaban are also relatively new[27, 45], hence there are few studies and small sample sizes. Long-term monitoring is required to assess their association with GIB.

Conclusion

In conclusion, our meta-analysis suggests a slightly higher risk of GIB in dabigatran use compared to warfarin, whereas no significant difference was found between rivaroxaban and
warfarin for GIB risk. Dabigatran and rivaroxaban do not raise unexpected safety concerns in terms of GIB compared with findings from RCTs.
Notes

Funding

This work was not supported by any funding.

Acknowledgements

We thank our colleagues in the Department of Pharmacology and Pharmacy of the University of Hong Kong – Dr. Martijn Schuemie and Mr. Kenneth K. C. Man for statistical advice. We thank Ms. Lisa Wong for proofreading the manuscript.

Author Contributions

YH, ICKW and EWC had the original idea for this study and contributed to the development of the idea and the study design. YH and XL independently conducted a systematic review and reviewed the literature for relevance. YH, SA and XL undertook the analysis. YH, ICKW, SA, XL, WKL, CWS and EWC contributed to the interpretation of the analysis. YH wrote the first draft of the paper. EWC, WKL, CWS and ICKW critically reviewed the results and the manuscript. ICKW and EWC provided oversight over all aspects of this project. YH and EWC are the guarantors. All authors had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of data analysis.
Figure 1 PRISMA flowchart summarizing study identification and selection.

Identification
1,634 records identified through database searching:
- PubMed: 447
- EMBASE: 1108
- CINAHL: 79

Screening
258 duplicates excluded, 1,427 records remained

Eligibility
1,427 records screened (titles/abstracts)
- 1,346 records excluded:
  - Basic science/animal studies
  - Case reports, case series
  - Reviews and commentary
  - Not related to GIB or NOAC use
- 65 records excluded:
  - Non-comparative studies
  - Pharmacovigilance studies
  - Conference abstracts
  - Detailed GIB results not reported

Included
81 full-text articles assessed for eligibility
- 16 studies included in the systematic review:
  - 11 Cohort studies (8 for primary meta-analysis, 3 for sensitivity analysis)
  - 5 Cross sectional studies (3 for secondary meta-analysis, 2 narrative review)
Figure 2 Primary analysis of cohort studies: summarised estimates (crude and adjusted) of GIB risk among users of NOAC vs warfarin

A. Crude estimate

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>log(RR)</th>
<th>SE</th>
<th>Total</th>
<th>Total</th>
<th>Weight</th>
<th>IV, Random, 95% CI</th>
<th>Risk Ratio</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 dabigatran vs warfarin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abraham 2015</td>
<td>0.2231</td>
<td>0.1339</td>
<td>7749</td>
<td>7749</td>
<td>14.4%</td>
<td>0.60 [0.52, 1.04]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chan 2015</td>
<td>0.3888</td>
<td>0.1603</td>
<td>281</td>
<td>8094</td>
<td>13.7%</td>
<td>1.49 [1.03, 2.10]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chang 2015</td>
<td>0.1823</td>
<td>0.1206</td>
<td>4197</td>
<td>39637</td>
<td>14.8%</td>
<td>1.20 [0.95, 1.52]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graham 2015</td>
<td>0.2548</td>
<td>0.0597</td>
<td>67207</td>
<td>67207</td>
<td>15.7%</td>
<td>1.29 [1.15, 1.45]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Larsen 2014</td>
<td>-0.1744</td>
<td>0.2606</td>
<td>4018</td>
<td>14126</td>
<td>11.2%</td>
<td>0.84 [0.50, 1.40]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lautenburger 2015</td>
<td>-0.3857</td>
<td>0.05</td>
<td>21070</td>
<td>43856</td>
<td>15.8%</td>
<td>0.68 [0.62, 0.75]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vaughan 2015</td>
<td>0.2465</td>
<td>0.1192</td>
<td>1394</td>
<td>93950</td>
<td>14.7%</td>
<td>1.71 [1.35, 2.16]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td></td>
<td></td>
<td>100626</td>
<td>284508</td>
<td>100.0%</td>
<td>1.05 [0.90, 1.26]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: $I^2 = 0.16$, $Chi^2 = 103.54$, df = 6 ($P < 0.00001$); $I^2 = 24$
Test for overall effect: $Z = 0.55$ ($P = 0.58$)

1.2 rivaroxaban vs warfarin

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>log(RR)</th>
<th>SE</th>
<th>Total</th>
<th>Total</th>
<th>Weight</th>
<th>IV, Random, 95% CI</th>
<th>Risk Ratio</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abraham 2015</td>
<td>-0.0726</td>
<td>0.1508</td>
<td>5186</td>
<td>5186</td>
<td>35.6%</td>
<td>0.93 [0.86, 1.02]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chan 2015</td>
<td>-0.0408</td>
<td>0.2574</td>
<td>244</td>
<td>8084</td>
<td>17.1%</td>
<td>0.96 [0.88, 1.06]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chang 2015</td>
<td>-0.0513</td>
<td>0.5784</td>
<td>1549</td>
<td>39827</td>
<td>4.1%</td>
<td>0.95 [0.91, 0.99]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laiwai 2014</td>
<td>0.3675</td>
<td>0.1257</td>
<td>3654</td>
<td>14616</td>
<td>43.1%</td>
<td>1.24 [1.05, 1.47]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td></td>
<td></td>
<td>10713</td>
<td>67453</td>
<td>100.0%</td>
<td>1.10 [0.97, 1.33]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: $I^2 = 0.02$, $Chi^2 = 4.33$, df = 3 ($P = 0.23$); $I^2 = 11$
Test for overall effect: $Z = 0.81$ ($P = 0.42$)

Test for subgroup differences: $Chi^2 = 0.00$, df = 1 ($P = 0.96$), $P = 0$

B. Adjusted estimate

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>log(RR)</th>
<th>SE</th>
<th>Total</th>
<th>Total</th>
<th>Weight</th>
<th>IV, Random, 95% CI</th>
<th>Risk Ratio</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 dabigatran vs warfarin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abraham 2015</td>
<td>-0.2357</td>
<td>0.1354</td>
<td>7749</td>
<td>7749</td>
<td>12.8%</td>
<td>0.79 [0.61, 1.03]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chan 2015</td>
<td>0.3985</td>
<td>0.1603</td>
<td>281</td>
<td>8094</td>
<td>10.8%</td>
<td>1.49 [1.09, 2.04]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chang 2015</td>
<td>0.1906</td>
<td>0.1197</td>
<td>4907</td>
<td>39637</td>
<td>14.3%</td>
<td>1.21 [0.90, 1.63]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graham 2015</td>
<td>0.2465</td>
<td>0.0801</td>
<td>67207</td>
<td>67207</td>
<td>20.8%</td>
<td>1.28 [1.14, 1.44]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lautenburger 2015</td>
<td>0.3149</td>
<td>0.2666</td>
<td>4018</td>
<td>14126</td>
<td>5.4%</td>
<td>1.37 [1.01, 1.88]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vaughan Sarazin 2014</td>
<td>0.4319</td>
<td>0.1256</td>
<td>1394</td>
<td>83950</td>
<td>13.8%</td>
<td>1.54 [2.00, 1.37]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td></td>
<td></td>
<td>100626</td>
<td>284508</td>
<td>100.0%</td>
<td>1.21 [1.05, 1.39]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: $I^2 = 0.02$, $Chi^2 = 19.17$, df = 6 ($P = 0.0034$); $I^2 = 69$
Test for overall effect: $Z = 2.72$ ($P = 0.007$)

1.2 rivaroxaban vs warfarin

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>log(RR)</th>
<th>SE</th>
<th>Total</th>
<th>Total</th>
<th>Weight</th>
<th>IV, Random, 95% CI</th>
<th>Risk Ratio</th>
<th>IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abraham 2015</td>
<td>-0.0726</td>
<td>0.1508</td>
<td>5186</td>
<td>5186</td>
<td>35.3%</td>
<td>0.93 [0.89, 1.25]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chan 2015</td>
<td>-0.3409</td>
<td>0.2574</td>
<td>244</td>
<td>8084</td>
<td>12.1%</td>
<td>0.96 [0.88, 1.05]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chang 2015</td>
<td>-0.2020</td>
<td>0.5152</td>
<td>1849</td>
<td>39637</td>
<td>3.0%</td>
<td>0.99 [0.95, 1.06]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laiwai 2014</td>
<td>0.2381</td>
<td>0.1273</td>
<td>3654</td>
<td>14616</td>
<td>49.8%</td>
<td>1.27 [1.09, 1.48]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td></td>
<td></td>
<td>10713</td>
<td>67453</td>
<td>100.0%</td>
<td>1.09 [0.92, 1.20]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: $I^2 = 0.00$, $Chi^2 = 2.83$, df = 3 ($P = 0.42$); $I^2 = 0$
Test for overall effect: $Z = 0.77$ ($P = 0.33$)

Test for subgroup differences: $Chi^2 = 0.83$, df = 1 ($P = 0.36$), $P = 0$
Figure 3 Summarised estimates of incidence rate of GIB in NOAC users

### A. Dabigatran

<table>
<thead>
<tr>
<th>Study</th>
<th>IR (per 100 patient-years)</th>
<th>95% CI (per 100 patient-years)</th>
<th>Weight</th>
<th>GI case</th>
<th>Patient-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abraham 2015</td>
<td>2.29</td>
<td>1.88 - 2.79</td>
<td>7.59%</td>
<td>99</td>
<td>4323</td>
</tr>
<tr>
<td>Chan KE 2015</td>
<td>33.3</td>
<td>23.1 - 43.5</td>
<td>0.02%</td>
<td>41</td>
<td>123</td>
</tr>
<tr>
<td>Chang 2015</td>
<td>9.01</td>
<td>7.41 - 10.6</td>
<td>0.60%</td>
<td>122</td>
<td>1354</td>
</tr>
<tr>
<td>Graham 2015</td>
<td>3.42</td>
<td>3.15 - 3.69</td>
<td>21.4%</td>
<td>623</td>
<td>18205</td>
</tr>
<tr>
<td>Larsen 2014</td>
<td>0.49</td>
<td>0.27 - 0.71</td>
<td>31.8%</td>
<td>19</td>
<td>3878</td>
</tr>
<tr>
<td>Lauffenburger 2015</td>
<td>2.16</td>
<td>1.58 - 2.83</td>
<td>38.1%</td>
<td>450</td>
<td>20652</td>
</tr>
<tr>
<td>Vaughan Sarrazin 2014</td>
<td>9.25</td>
<td>7.32 - 11.2</td>
<td>0.41%</td>
<td>88</td>
<td>951</td>
</tr>
<tr>
<td>Overall (Random effects Model)</td>
<td>4.51</td>
<td>3.17 - 5.84</td>
<td>100%</td>
<td>1442</td>
<td>49486</td>
</tr>
</tbody>
</table>

### B. Rivaroxaban

<table>
<thead>
<tr>
<th>Study</th>
<th>IR (per 100 patient-years)</th>
<th>95% CI (per 100 patient-years)</th>
<th>Weight</th>
<th>GI case</th>
<th>Patient-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abraham 2015</td>
<td>2.84</td>
<td>2.30 - 3.52</td>
<td>89.5%</td>
<td>86</td>
<td>3028</td>
</tr>
<tr>
<td>Chan 2015</td>
<td>20.8</td>
<td>12.6 - 34.6</td>
<td>0.29%</td>
<td>15</td>
<td>72</td>
</tr>
<tr>
<td>Chang 2015</td>
<td>3.41</td>
<td>0.07 - 6.75</td>
<td>2.69%</td>
<td>4</td>
<td>117</td>
</tr>
<tr>
<td>Laliberte 2014</td>
<td>9.51</td>
<td>7.64 - 11.9</td>
<td>7.32%</td>
<td>79</td>
<td>830</td>
</tr>
<tr>
<td>Overall (Random effects Model)</td>
<td>7.18</td>
<td>2.42 - 12.0</td>
<td>100%</td>
<td>184</td>
<td>4048</td>
</tr>
</tbody>
</table>

**Abbreviations:** IR=Incidence rate per 100 patient-years; LCL=95% lower confidence limit; UCL=95% upper confidence limit.
**Figure 4 Subgroup analysis:** summarised estimates of GIB risk by different dabigatran doses

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Log(Risk Ratio)</th>
<th>SE</th>
<th>Weight</th>
<th>IV, Random, 95% CI</th>
<th>Risk Ratio IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1.6.1 75mg bid</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graham 2015</td>
<td>0.01</td>
<td>0.1327</td>
<td>100.0%</td>
<td>1.01 [0.78, 1.31]</td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>100.0%</td>
<td>1.01 [0.78, 1.31]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: Not applicable Test for overall effect: Z = 0.03 (P = 0.94)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| **1.6.2 110mg bid** |                |    |        |                     |                               |
| Larson 2014        | -0.6349        | 0.3136 | 100.0% | 0.53 [0.29, 0.98]   |                               |
| Subtotal (95% CI)  | 100.0%         | 0.53 [0.29, 0.98] |
| Heterogeneity: Not applicable Test for overall effect: Z = 2.02 (P = 0.04) |

| **1.6.3 150mg bid** |                |    |        |                     |                               |
| Graham 2015        | 0.4121         | 0.0684 | 72.8%  | 1.51 [1.32, 1.73]   |                               |
| Larson 2014        | 0.3143         | 0.2666 | 4.9%   | 1.37 [0.81, 2.21]   |                               |
| Vaughan Sarrazin 2014 | 0.4318        | 0.1256 | 22.2%  | 1.54 [1.20, 1.97]   |                               |
| Subtotal (95% CI)  | 100.0%         | 1.54 [1.20, 1.97] |
| Heterogeneity: Tau^2 = 0.39, Ch^2 = 13.6, df = 4 (P = 0.03); *P* = 0.9% Test for overall effect: Z = 6.95 (P < 0.00001) |

| **1.6.4 Combined doses** |                |    |        |                     |                               |
| Abraham 2015       | -0.2357        | 0.1364 | 15.6%  | 0.79 [0.61, 1.03]   |                               |
| Chan 2015          | 0.3988         | 0.1603 | 13.0%  | 1.49 [1.09, 2.04]   |                               |
| Cheong 2015        | 0.1903         | 0.1197 | 17.5%  | 1.21 [0.96, 1.53]   |                               |
| Graham 2015        | 0.2489         | 0.0801 | 26.1%  | 1.28 [1.14, 1.44]   |                               |
| Lauffenburger 2015 | 0.1044         | 0.0482 | 27.6%  | 1.11 [1.01, 1.22]   |                               |
| Subtotal (95% CI)  | 100.0%         | 1.11 [1.01, 1.22] |
| Heterogeneity: Tau^2 = 0.02, Ch^2 = 14.1, df = 4 (P = 0.007), *P* = 72% Test for overall effect: Z = 1.88 (P = 0.06) |

Test for subgroup differences: Ch^2 = 20.70, df = 3 (P = 0.0001), *P* = 95.5%
Table 1 Characteristics of included observational studies

**Abbreviations:** ACH= Auckland City Hospital; AF=atrial fibrillation; C=cohort; CS= cross-sectional; FDA=Food and Drug Administration; GFR= glomerular filtration rate; GI=gastrointestinal; GIB=gastrointestinal bleeding; ICD-9=International Classification of Diseases, Ninth Revision; P=prospective; R=retrospective; VTE=venous thromboembolism; OAC=oral anticoagulation; Rx=prescription; IMS=Intercontinental Marketing Services; ESRD=end-stage renal disease; Y=years; CHADS$_2$=congestive heart failure, hypertension, age of 75 years or older, diabetes mellitus, prior stroke or transient ischemic attack or thromboembolism; PPI=proton pump inhibitor
Table 2 Subgroup and sensitivity analyses of the GIB risk among users of NOAC vs warfarin

a substitution of Larsen 2014’s results [34] with that of Staerk 2015[43], and Graham 2015’s results [26] with that of Hernandez 2015[40], respectively

b substitution of Larsen 2014’s results [34] with that of Larsen 2013[32], and Graham 2015’s results [26] with that of Hernandez 2015[40], respectively

c exclusion of the result of Abraham 2015 [37]

d exclusion of the result of Chan 2015 [38]

Abbreviations: NOAC= non-vitamin K antagonist oral anticoagulants; GIB=gastrointestinal bleeding; PPI=proton pump inhibitor; H2RA=histamine type-2 antagonists; NSAID=non-steroidal anti-inflammatory drug; Riva=rivaroxaban; Dabi=dabigatran; Warf=warfarin; RR=risk ratio; CI=confidence interval; SSRI=selective serotonin reuptake inhibitor
References


29. McDonald CJ, Kalisch Ellett LM, Barratt JD, Caughey GE. A cross-country comparison of rivaroxaban spontaneous adverse event reports and concomitant medicine use with the potential to increase the risk of harm. Drug safety. 2014;37(12):1029-35.


Author names in bold designate shared co-first authors.