SEARCH STRATEGY & STUDY QUALITY

Text S1. Full Electronic Search Strategy

Search Strategy Overview A systematic literature search was executed across three major electronic databases: MEDLINE (via PubMed), Embase (via Elsevier), and the Cochrane Central Register of Controlled Trials (CENTRAL). The search strategy was designed to identify randomized controlled trials (RCTs) and their pre-specified secondary sub-analyses evaluating direct oral anticoagulants (DOACs) versus vitamin K antagonists (VKAs) in patients with atrial fibrillation and concurrent heart failure. The syntax utilized a combination of Medical Subject Headings (MeSH), Emtree terms, and free-text keywords covering three primary domains: (1) atrial fibrillation, (2) heart failure or ventricular dysfunction, and (3) specific DOAC agents. No language restrictions were applied. The initial literature search was completed on March 15, 2026, with a final updated search conducted on May 25, 2026, to ensure no recent landmark trials were omitted prior to submission.

1. **MEDLINE (via PubMed)**
   ("atrial fibrillation"[MeSH Terms] OR "atrial fibrillation"[Title/Abstract] OR "nonvalvular atrial fibrillation"[Title/Abstract] OR "AF"[Title/Abstract]) AND ("heart failure"[MeSH Terms] OR "heart failure"[Title/Abstract] OR "ventricular dysfunction, left"[MeSH Terms] OR "left ventricular dysfunction"[Title/Abstract] OR "HFrEF"[Title/Abstract] OR "HFpEF"[Title/Abstract] OR "reduced ejection fraction"[Title/Abstract] OR "preserved ejection fraction"[Title/Abstract]) AND ("administration, oral"[MeSH Terms] AND "anticoagulants"[MeSH Terms] OR "direct oral anticoagulants"[Title/Abstract] OR "DOAC"[Title/Abstract] OR "non-vitamin K antagonist oral anticoagulants"[Title/Abstract] OR "NOAC"[Title/Abstract] OR "apixaban"[Title/Abstract] OR "dabigatran"[Title/Abstract] OR "rivaroxaban"[Title/Abstract] OR "edoxaban"[Title/Abstract]) AND ("randomized controlled trial"[Publication Type] OR "randomized"[Title/Abstract] OR "placebo"[Title/Abstract] OR "clinical trials as topic"[MeSH Terms] OR "subgroup analysis"[Title/Abstract] OR "post hoc"[Title/Abstract])
2. **Embase (via Elsevier)**
   #1 ('heart atrium fibrillation'/exp OR 'atrial fibrillation':ab,ti OR 'nonvalvular atrial fibrillation':ab,ti)
   #2 ('heart failure'/exp OR 'heart ventricle dysfunction'/exp OR 'heart failure':ab,ti OR 'ventricular dysfunction':ab,ti OR 'hfref':ab,ti OR 'hfpef':ab,ti)
   #3 ('direct oral anticoagulant'/exp OR 'apixaban'/exp OR 'dabigatran'/exp OR 'rivaroxaban'/exp OR 'edoxaban'/exp OR 'doac':ab,ti OR 'noac':ab,ti OR 'apixaban':ab,ti OR 'dabigatran':ab,ti)
   #4 ('randomized controlled trial'/exp OR 'random':ab,ti OR 'subgroup':ab,ti)**
   #5 #1 AND #2 AND #3 AND #4
3. **Cochrane Central Register of Controlled Trials (CENTRAL)**
   #1 MeSH descriptor: [Atrial Fibrillation] explode all trees
   #2 ("atrial fibrillation"):ti,ab,kw
   #3 #1 OR #2
   #4 MeSH descriptor: [Heart Failure] explode all trees
   #5 MeSH descriptor: [Ventricular Dysfunction] explode all trees
   #6 ("heart failure" OR "ventricular dysfunction" OR "HFrEF" OR "HFpEF"):ti,ab,kw
   #7 #4 OR #5 OR #6
   #8 MeSH descriptor: [Anticoagulants] explode all trees
   #9 ("DOAC" OR "NOAC" OR "apixaban" OR "dabigatran" OR "rivaroxaban" OR "edoxaban"):ti,ab,kw
   #10 #8 OR #9 #11 #3 AND #7 AND #10

Methodological Note on Targeted Sub-study Retrieval:

Because detailed heart failure data (including left ventricular ejection fraction strata) are often unpublished in the primary pivotal trial manuscripts, an adjunctive targeted search was performed. The acronyms of the four pivotal phase III trials (ARISTOTLE, RE-LY, ROCKET-AF, ENGAGE AF-TIMI 48) were queried in conjunction with "heart failure" to capture dedicated secondary publications and

post-hoc sub-analyses that contained the precise event counts and Kaplan-Meier curves necessary for the IPD reconstruction and RMST derivation.

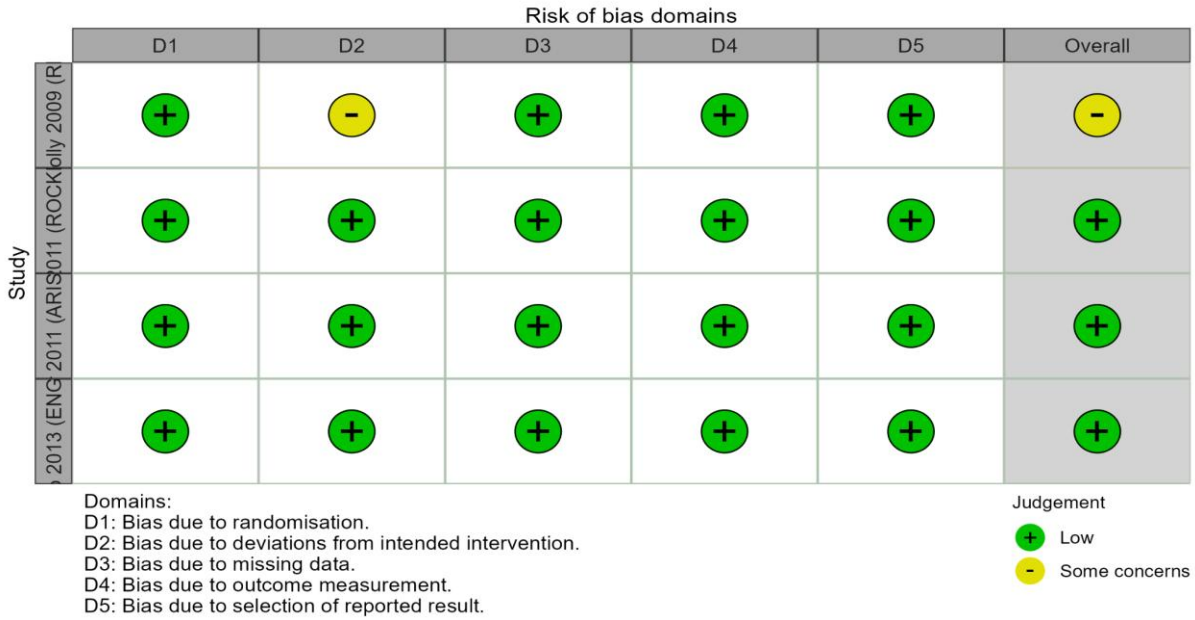

Figure S1A. RoB 2.0 traffic light plot

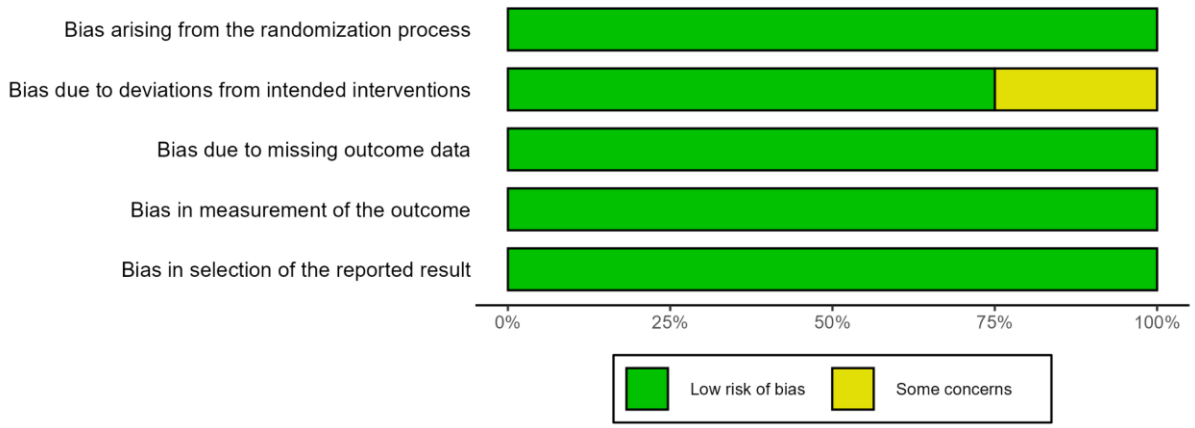

Figure S1B. RoB 2.0 weighted bar chart

DATA CHARACTERISTICS & TRANSITIVITY

Table S1. Detailed Baseline Characteristics of the Included Trials

Study (Year)	Treatments (HF Cohort N)	Mean Age (years)	Male (%)	Mean CHA_2DS_2-VASc	NYHA Class III/IV (%)	Mean TTR in VKA arm (%)	Transitivity Assessment (Clinical Similarity)
ARISTOTLE (2011)	Apixaban 5 mg vs. VKA(N = 6,422)	70.0	70.4	~3.3	31.0	62.0	Comparable baseline risk; highly representative of standard AF-HF population.
RE-LY (2009)	Dabigatran 110/150 mg vs. VKA(N = 4,904)	72.0	64	~3.3	30.0	64.0	Comparable baseline risk; closely matches ARISTOTLE in HF symptom severity.
ROCKET-AF (2011)	Rivaroxaban 20 mg vs. VKA(N = 9,033)	72.0	60	~4.8*	20.0	55.0	Enrolled a slightly higher-risk $CHADS_2$ cohort with a lower VKA TTR; adjusted via random-effects model.
ENGAGE AF-TIMI 48 (2013)	Edoxaban 60 mg vs. VKA(N = 5,403)	71.0	60.5	~3.3	~26.0	~65.0	Comparable baseline risk; TTR and symptom severity align well with the overall network architecture.

Data are presented as means or percentages as reported in the respective heart failure (HF) substudies of the pivotal trials.

Note on Transitivity: In network meta-analysis, the transitivity assumption requires that effect modifiers are reasonably similarly distributed across the available evidence base. Given the large sample sizes of the pivotal DOAC trials, standard statistical tests for heterogeneity (e.g., ANOVA P-values) across baselines are typically overpowered and consistently yield $P < 0.05$ for trivial, non-clinical differences. Therefore, a qualitative clinical assessment of transitivity is provided. Overall, the trials demonstrated sufficient clinical similarity in terms of age, stroke risk, and HF severity to satisfy the transitivity assumption, validating the indirect comparisons centered on the common VKA comparator.

**In ROCKET-AF, stroke risk was reported primarily using the $CHADS_2$ score (mean ~4.8); patients generally had a slightly higher baseline stroke risk profile compared to other trials, which is accounted for by the use of the Bayesian random-effects model.*

Abbreviations: AF: atrial fibrillation; HF: heart failure; VKA: vitamin K antagonist; NYHA: New York Heart Association; TTR: time in therapeutic range.

Table S2. Detailed Inclusion and Exclusion Criteria of the Heart Failure Subpopulations in Pivotal DOAC Trials

Study (HF Substudy)	General Trial Inclusion (AF Cohort)	Specific HF Definition (Subpopulation Included in NMA)	Phenotype Stratification (HFrEF vs. HFpEF)	Key Exclusion Criteria
ARISTOTLE (McMurray et al., 2014)	Nonvalvular AF or flutter with ≥1 additional stroke risk factor.	Investigator-reported history of HF or documented left ventricular ejection fraction (LVEF) ≤ 40%.	Stratified using LVEF ≤ 40% (HFrEF) and LVEF > 40% (HFpEF).	Moderate/severe mitral stenosis; prosthetic heart valve; severe renal impairment (CrCl < 25 mL/min); active bleeding.
RE-LY (Ferreira et al., 2013)	Nonvalvular AF with ≥1 additional stroke risk factor.	Documented symptomatic HF (NYHA class II–IV) within 6 months prior to screening, or LVEF < 40%.	Stratified using LVEF < 40% (HFrEF) and LVEF ≥ 40% (HFpEF).	Severe heart valve disorder; recent stroke (within 14 days); severe renal impairment (CrCl < 30 mL/min); active liver disease.
ROCKET-AF (van Diepen et al., 2013)	Nonvalvular AF with a history of stroke/TIA or a $CHADS_2$ score ≥ 2.	Investigator-reported clinical history of HF or documented LVEF ≤ 35% (used as a qualifying risk factor).	Stratified primarily by the presence of clinical HF. LVEF data were reported but with varying cutoffs (predominantly ≤ 40% for reduced EF comparisons).	Planned cardioversion; severe renal impairment (CrCl < 30 mL/min); active internal bleeding; non-cardiovascular life expectancy < 2 years.
ENGAGE AF-TIMI 48 (Magnani et al., 2016)	Nonvalvular AF with a $CHADS_2$ score ≥ 2.	Clinical history of HF, NYHA class II–IV, or objectively reduced LVEF at baseline.	Stratified into LVEF ≤ 40% (HFrEF) and LVEF > 40% (HFpEF/unmeasured, evaluated via clinical symptoms).	Severe renal impairment (CrCl < 30 mL/min); acute coronary syndrome within 30 days; high risk of bleeding; dual antiplatelet therapy requirement.

Note on HF Definitions: The definitions of heart failure across the pivotal trials were clinically consistent, predominantly relying on investigator-reported clinical history, NYHA functional class (II–IV), or documented objective metrics such as left ventricular ejection fraction (LVEF). For the phenotype-specific network meta-analysis, an LVEF threshold of 40% (or equivalent clinical cutoffs as reported by the trial's dedicated HF substudy) was utilized to distinguish between Heart Failure with reduced Ejection Fraction (HFrEF) and Heart Failure with preserved Ejection Fraction (HFpEF). Minor variations in LVEF thresholds (e.g., <40% vs. ≤40%) were considered clinically negligible for the purpose of this network framework.

Abbreviations: AF: atrial fibrillation; HF: heart failure; DOAC: direct oral anticoagulant; LVEF:

left ventricular ejection fraction; NYHA: New York Heart Association; CrCl: creatinine clearance; TIA: transient ischemic attack.

NETWORK GEOMETRY & MODEL FIT

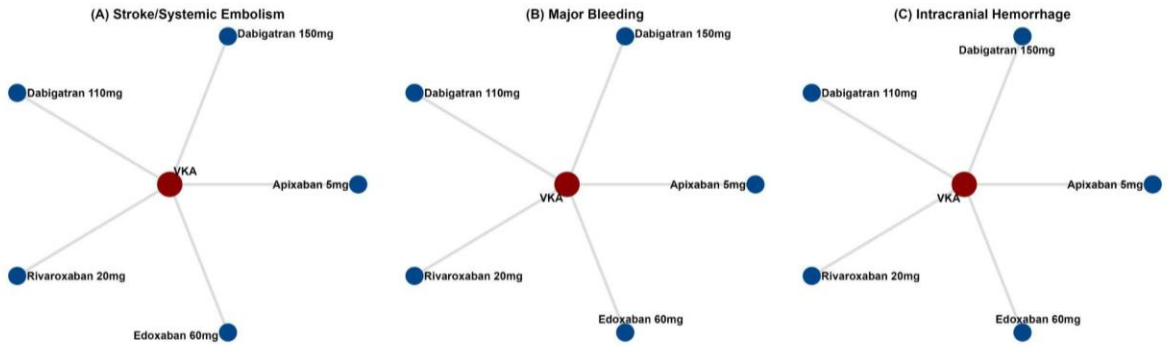

Figure S2. Network geometry for primary and secondary outcomes. (A) Stroke and systemic embolism; (B) Major bleeding; (C) Intracranial hemorrhage. Nodes represent individual treatment arms, while edges signify direct head-to-head randomized evidence. The star-shaped configuration highlights the central role of Vitamin K Antagonists (VKAs) as the common reference for indirect comparisons of direct oral anticoagulants (DOACs).

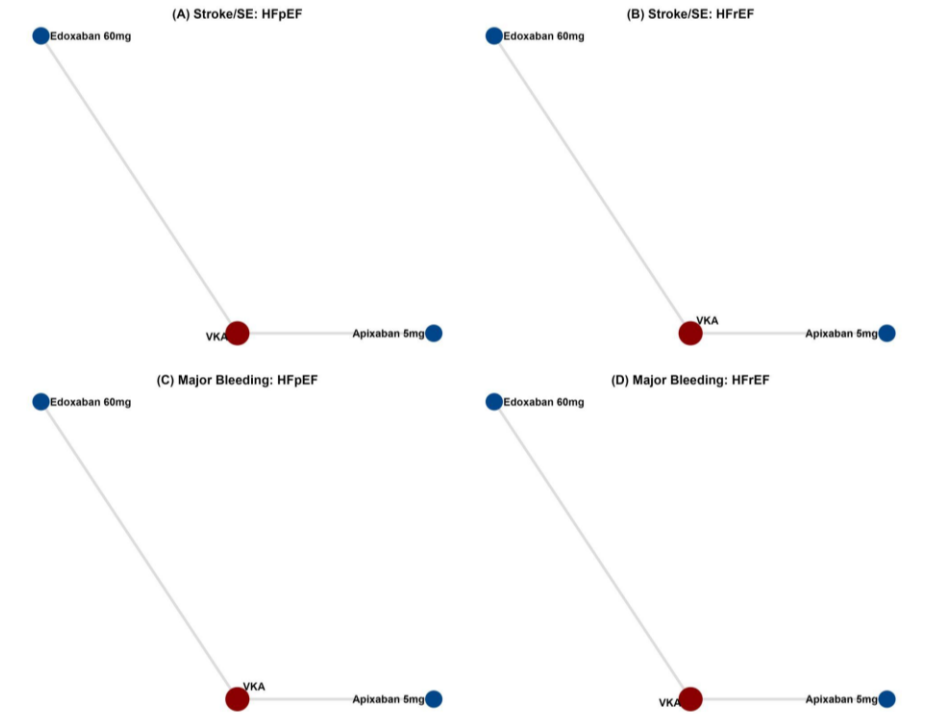

Figure S3. Network geometry across heart failure ejection fraction phenotypes. (A) Stroke/systemic embolism in HFpEF; (B) Stroke/systemic embolism in HFrEF; (C) Major bleeding in HFpEF; (D) Major bleeding in HFrEF. The network architecture remains consistent across phenotypes to ensure robust and comparable estimation of treatment effects in both preserved and reduced ejection fraction populations.

Table S3. Comparison of DIC between Fixed- and Random-Effects Models

Clinical Outcome	Model Type	Residual Deviance	DIC	ΔDIC (FE - RE)	Model Selection
Stroke/Systemic Embolism	Fixed-Effects (FE)	4.96	9.93	-0.24	RE Model
	Random-Effects (RE)	5.09	10.17		
Major Bleeding	Fixed-Effects (FE)	5.00	10.00	0.01	RE Model
	Random-Effects (RE)	5.01	10.01		
Intracranial Hemorrhage	Fixed-Effects (FE)	4.99	9.99	0.02	RE Model
	Random-Effects (RE)	4.98	9.97		

Notes: The Deviance Information Criterion (DIC) values were obtained from Bayesian MCMC simulations. A DIC difference (ΔDIC) of < 3 indicates statistically equivalent model fit. The Random-Effects (RE) model was consistently selected for all outcomes to account for clinical heterogeneity across studies (such as variations in baseline stroke risk and TTR values in the VKA control group) and to yield more conservative estimates to support the generalization of results to a heterogeneous AF-HF population.

MAIN NETWORK META-ANALYSIS RESULTS

Table S4. Integrated Network League Table for Efficacy and Safety Outcomes of Direct Oral Anticoagulants versus Vitamin K Antagonists in Heart Failure Patients.

	Apixaban 5mg	**Dabigatran 110mg**	**Dabigatran 150mg**	**Edoxaban 60mg**	**Rivaroxaban 20mg**	**VKA**
Apixaban 5mg	Apix_5	1.30 (0.69-2.47)	0.98 (0.51-1.85)	1.15 (0.65-2.03)	1.06 (0.60-1.88)	1.31 (0.87-1.98)
Dabigatran 110mg	0.92 (0.51-1.64)	Dabi_110	0.75 (0.37-1.54)	0.87 (0.47-1.65)	0.81 (0.43-1.55)	1.01 (0.61-1.66)
Dabigatran 150mg	0.96 (0.54-1.73)	1.05 (0.57-1.92)	Dabi_150	1.16 (0.62-2.23)	1.08 (0.57-2.07)	1.33 (0.79-2.21)
Edoxaban 60mg	0.93 (0.54-1.60)	1.02 (0.57-1.79)	0.97 (0.54-1.69)	Edox_60	0.92 (0.53-1.64)	1.15 (0.77-1.69)
Rivaroxaban 20mg	0.73 (0.42-1.26)	0.80 (0.45-1.40)	0.76 (0.43-1.33)	0.78 (0.47-1.33)	Riva_20	1.24 (0.82-1.87)
VKA	0.76 (0.52-1.13)	0.83 (0.54-1.27)	0.79 (0.52-1.20)	0.82 (0.57-1.19)	1.04 (0.72-1.50)	VKA

Notes: Estimates are Hazard Ratios (95% Credible Intervals) for Column vs Row. The upper-right triangle displays the comparative efficacy (Stroke/Systemic Embolism), whereas the lower-left triangle presents the comparative safety (Major Bleeding). An HR < 1.00 favors the column-defining treatment.

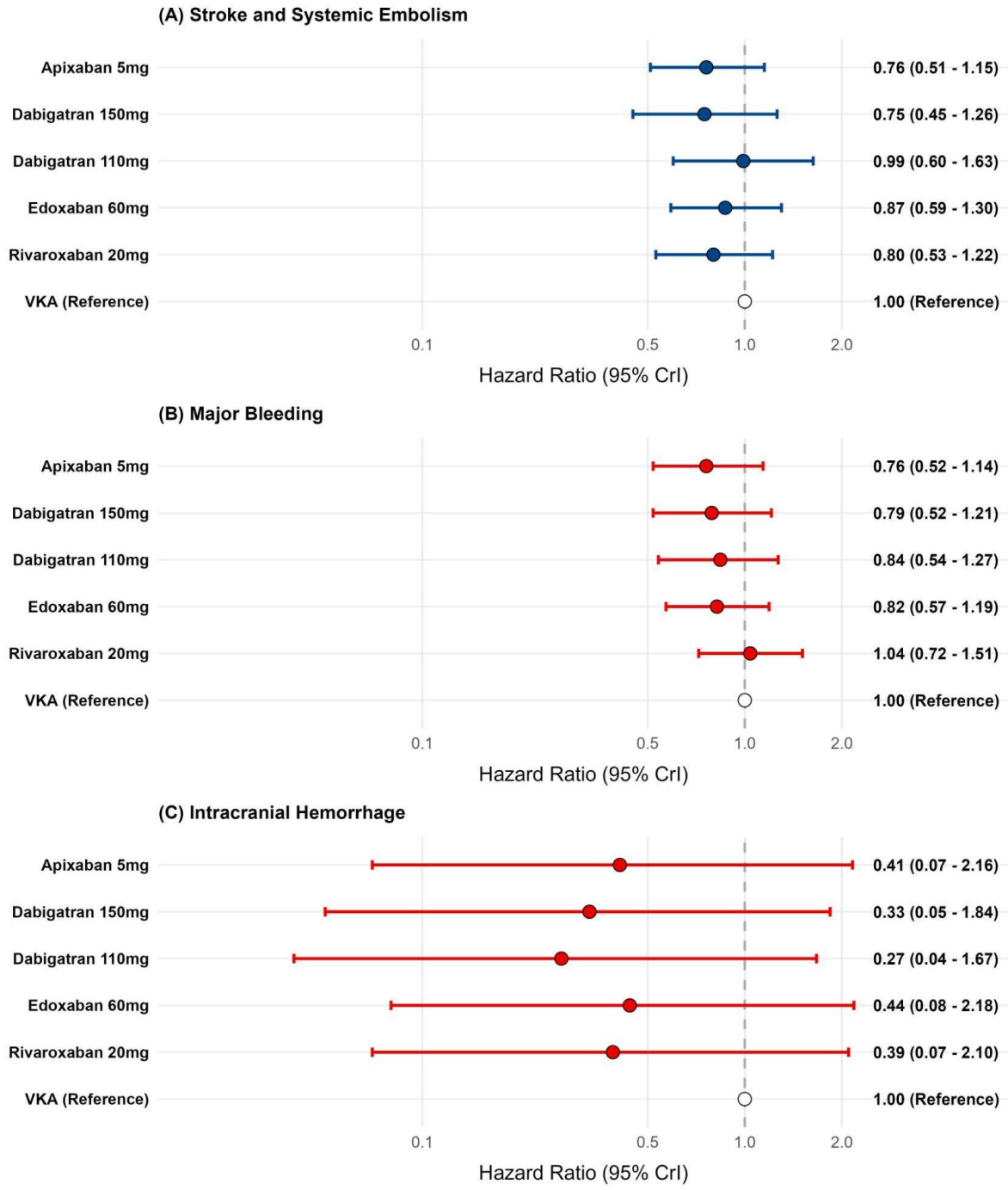

Figure S4. Network meta-analysis results for efficacy and safety outcomes of individual DOACs versus Vitamin K Antagonists. Forest plots depict the relative treatment effects for (A) stroke and systemic embolism, (B) major bleeding, and (C) intracranial hemorrhage. Data are presented as hazard ratios (HR) with corresponding 95% credible intervals (CrI), plotted on a logarithmic scale. An HR < 1.0 indicates a relative risk reduction in favor of the specified direct oral anticoagulant (DOAC) compared to the reference vitamin K antagonist (VKA) arm. The VKA reference is explicitly anchored at an HR of 1.00.

Abbreviations: CrI: credible interval; DOAC: direct oral anticoagulant; HR: hazard ratio; VKA: vitamin K antagonist.

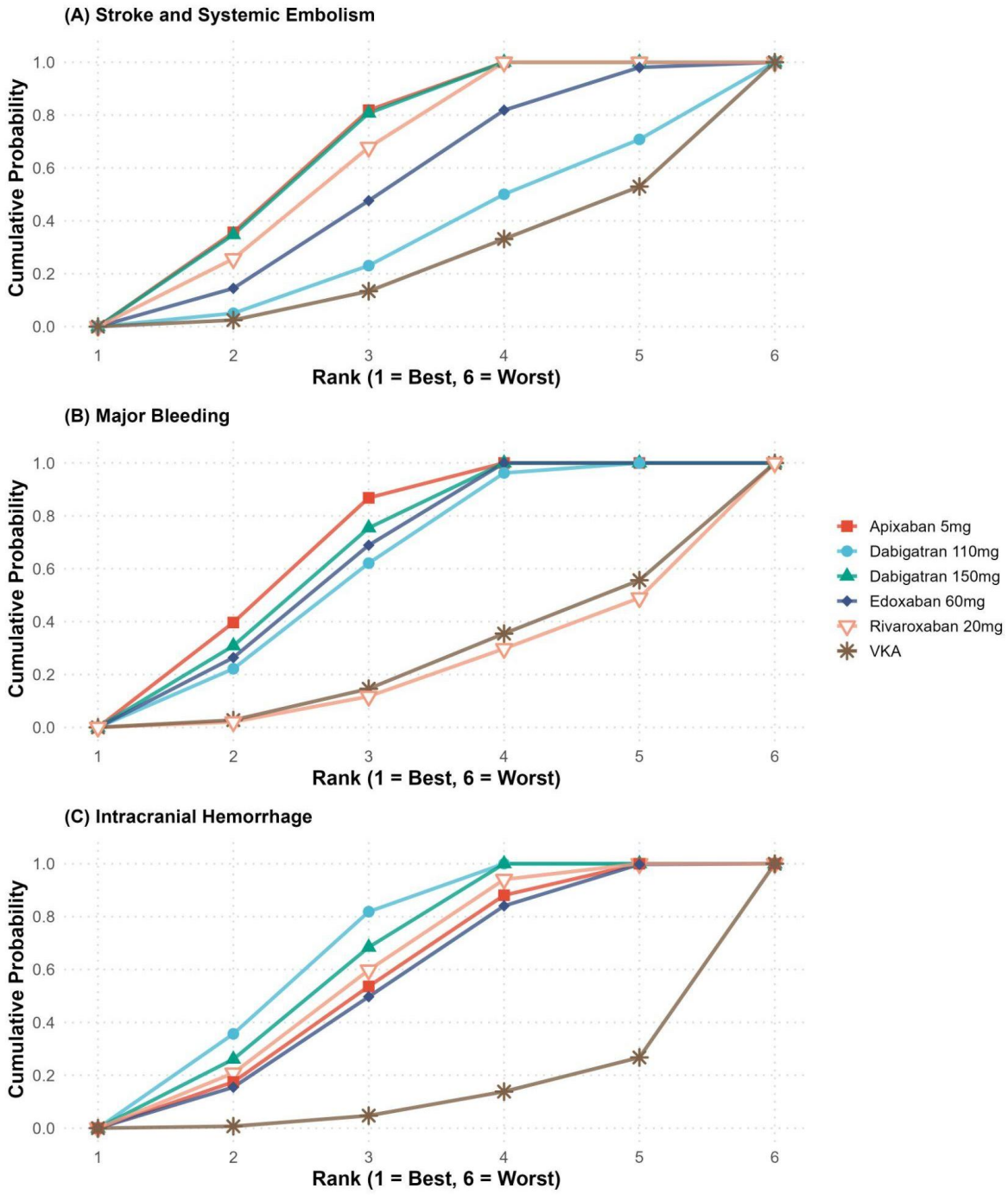

Figure S5. Cumulative ranking probability (SUCRA) curves for each antithrombotic regimen. The figure displays the surface under the cumulative ranking curve (SUCRA) probability distributions across three key clinical outcomes: (A) stroke and systemic embolism, (B) major bleeding, and (C) intracranial hemorrhage. The x-axis denotes the hierarchical ranking (from 1 = optimal to 6 = least favorable), while the y-axis indicates the cumulative probability of a given treatment securing that rank or higher. Trajectories that rapidly ascend towards the top-left quadrant denote a higher statistical likelihood of being the superior therapeutic option. Apixaban 5 mg consistently demonstrates a leading rank profile across both efficacy and safety domains, whereas vitamin K antagonists (VKA) persistently localize to the lower-right quadrant, reflecting their subordinate ranking profile in the network.

Table S5. Network Meta-Analysis Results: Hazard Ratios for DOACs versus VKAs

Outcome	Comparison	Hazard Ratio (95% CrI)	Interpretation
Stroke/Systemic Embolism	Apixaban 5 mg vs. VKA	0.76 (0.50–1.15)	The risk is comparable to VKA (trend toward benefit)
	Dabigatran 150 mg vs. VKA	0.75 (0.45–1.26)	The risk is comparable to VKA (trend toward benefit)
	Edoxaban 60 mg vs. VKA	0.87 (0.59–1.30)	The risk is comparable to VKA.
	Rivaroxaban 20 mg vs. VKA	0.81 (0.53–1.22)	The risk is comparable to VKA.
Major Bleeding	Apixaban 5 mg vs. VKA	0.76 (0.52–1.13)	The risk is comparable to VKA (trend toward safety).
	Dabigatran 150 mg vs. VKA	0.79 (0.52–1.20)	The risk is comparable to VKA.
	Edoxaban 60 mg vs. VKA	0.82 (0.57–1.19)	The risk is comparable to VKA.
	Rivaroxaban 20 mg vs. VKA	1.04 (0.72–1.50)	The risk is comparable to VKA.
Intracranial Hemorrhage	Apixaban 5 mg vs. VKA	0.42 (0.08–2.17)	The risk is comparable to VKA (trend toward reduction).
	Edoxaban 60 mg vs. VKA	0.44 (0.08–2.19)	The risk is comparable to VKA.
	Rivaroxaban vs. VKA	0.39 (0.07–2.09)	The risk is comparable to VKA.

Note: Hazard ratios (HR) represent the relative treatment effects derived from the Bayesian consistency model. An HR value < 1.00 indicates a trend favoring the DOAC over VKA. While all 95% credible intervals currently cross 1.00 (indicating comparable relative risk), absolute event-free survival benefits are captured in the complementary RMST analysis.

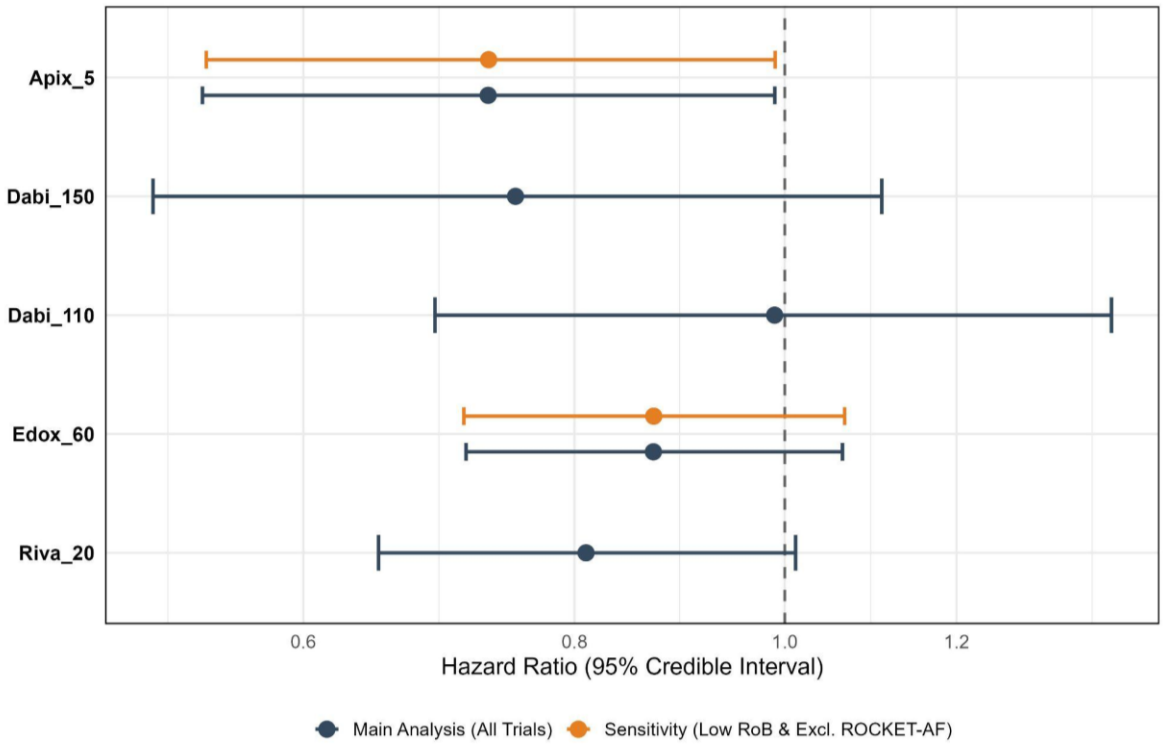

Figure S6. Sensitivity analysis of the network meta-analysis evaluating the robustness of primary estimates. Forest plot comparing the hazard ratios (HR) and 95% credible intervals (CrI) of the main network meta-analysis (dark blue, including all eligible trials) versus the prespecified sensitivity analysis (orange). The sensitivity analysis was performed to test the stability of the pooled estimates by restricting the network to high-quality trials, specifically excluding studies with a 'Moderate' or 'High' risk of bias (e.g., RE-LY) and excluding the ROCKET-AF trial. All estimates are generated from a Bayesian Markov Chain Monte Carlo (MCMC) consistency model. The vertical dashed line at HR = 1.00 represents the line of no effect, with Vitamin K Antagonist (VKA) serving as the common reference baseline.

Abbreviations: CrI: Credible Interval; HR: Hazard Ratio; NMA: Network Meta-Analysis; RoB: Risk of Bias; VKA: Vitamin K Antagonist; Apix_5: Apixaban 5 mg; Dabi_110: Dabigatran 110 mg; Dabi_150: Dabigatran 150 mg; Edox_60: Edoxaban 60 mg; Riva_20: Rivaroxaban 20 mg.

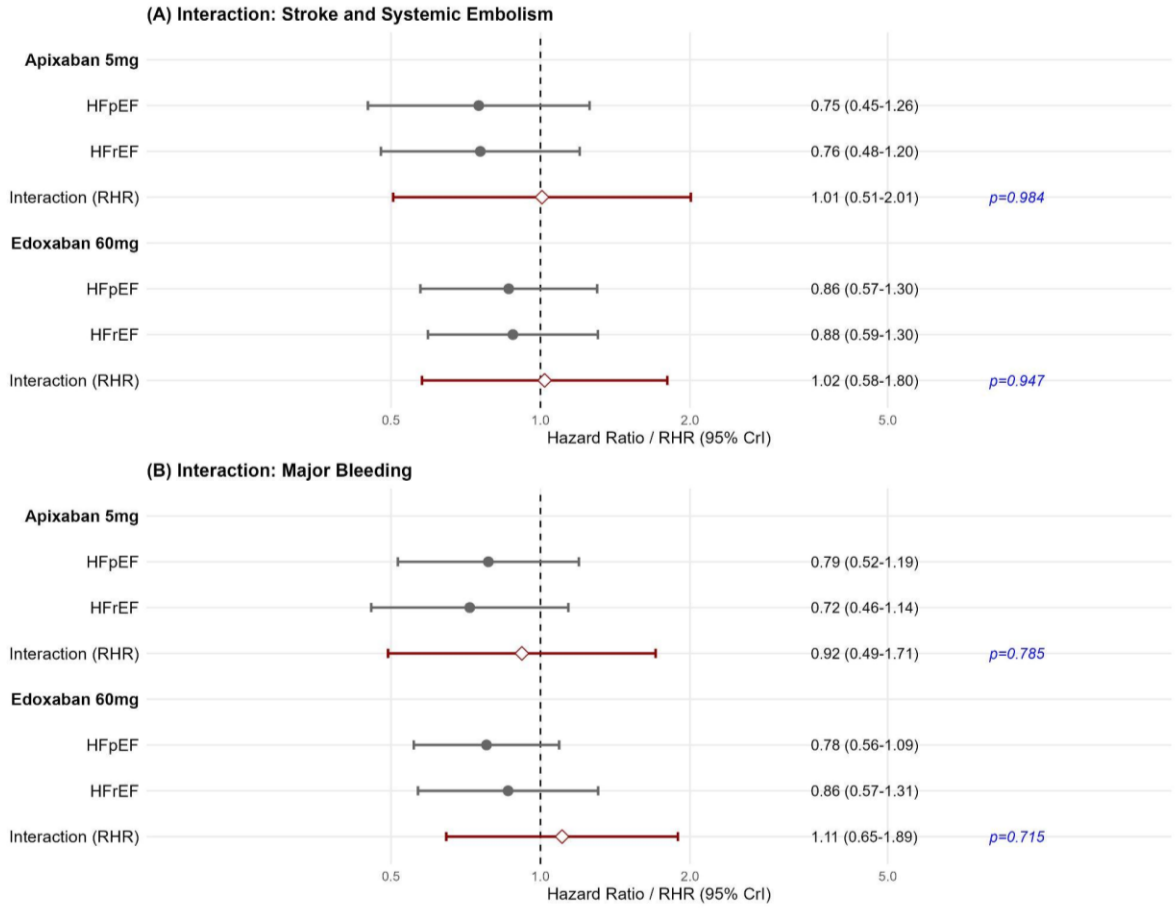

Figure S7. Interaction analysis of treatment effects by heart failure phenotype (HFrEF vs. HFpEF). The forest plots illustrate the subgroup-specific hazard ratios (vs. Vitamin K Antagonists) for (A) stroke and systemic embolism and (B) major bleeding. For each DOAC, the interaction is formally tested using the Ratio of Hazard Ratios (RHR), represented by the diamond symbol. An RHR of 1.00 indicates perfect consistency of the treatment effect across both ejection fraction phenotypes. The p-values for interaction (*p-int*) were derived from the posterior distributions of the subgroup differences. In all analyses, the interaction p-values were non-significant ($p > 0.05$), suggesting that the relative efficacy and safety of Apixaban and Edoxaban compared to VKA are preserved regardless of the heart failure phenotype.

Abbreviations: CrI: credible interval; HFpEF: heart failure with preserved ejection fraction; HFrEF: heart failure with reduced ejection fraction; RHR: ratio of hazard ratios.

CONVERGENCE & CERTAINTY OF EVIDENCE

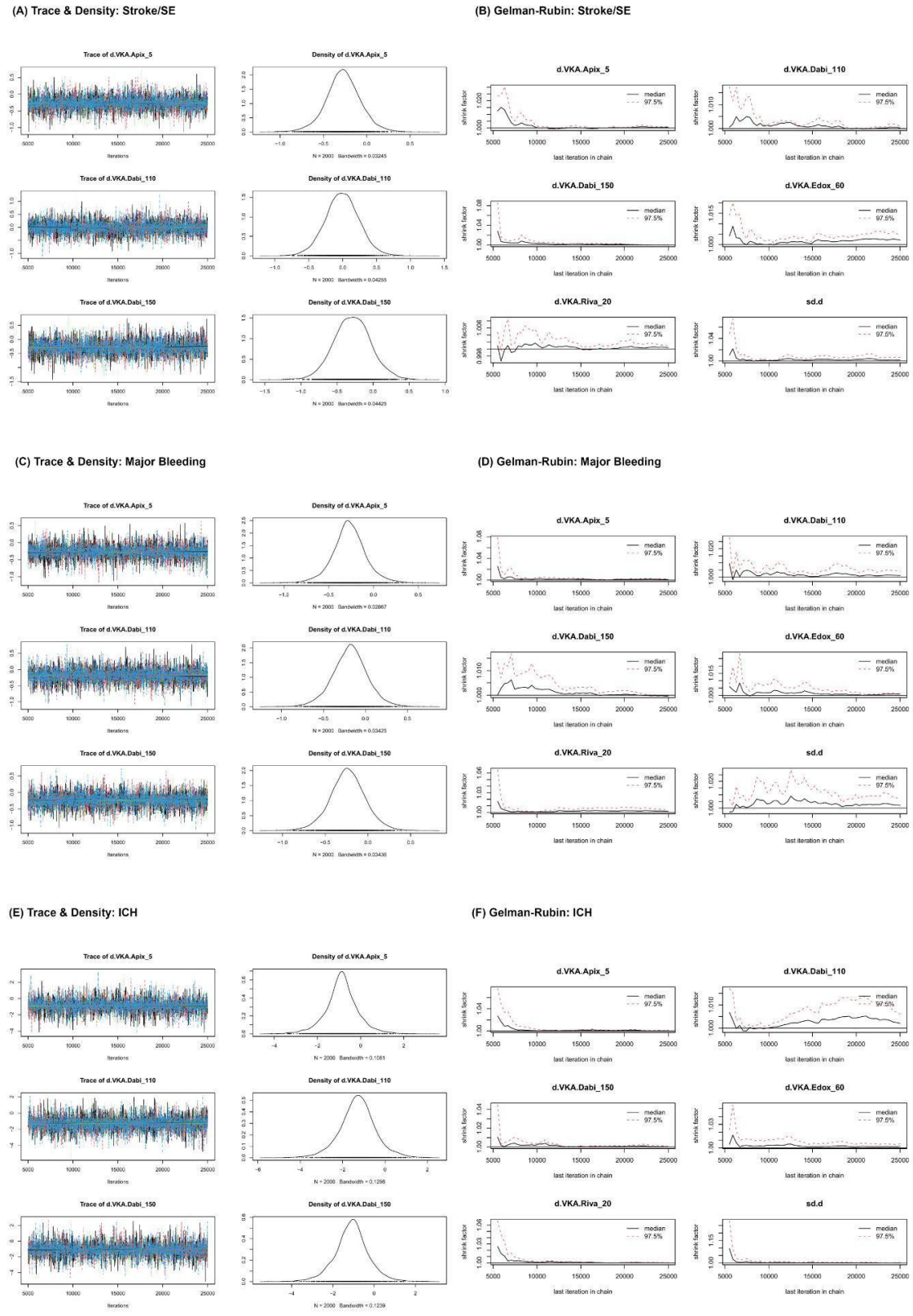

Figure S8. Markov Chain Monte Carlo (MCMC) convergence diagnostics for primary outcomes. The panels confirm the stability and convergence of the Bayesian models. (A, C) Trace plots and posterior

density plots for stroke and systemic embolism and intracranial hemorrhage, respectively. The trace plots demonstrate visual evidence of adequate mixing across the four independent chains with no discernible directional trends (often described as a "hairy caterpillar" appearance). The density plots show smooth, unimodal posterior distributions, confirming sufficient sampling from the target distributions. (B, D) Gelman-Rubin diagnostic plots (shrink factors) for the corresponding outcomes. The Potential Scale Reduction Factor (PSRF) rapidly converges to 1.0 across all iterations, indicating that the variance between the multiple MCMC chains is virtually indistinguishable from the variance within the individual chains, thus firmly establishing model convergence.

Table S6. Certainty of Evidence Assessment Using the GRADE Approach for Network Meta-Analysis

Study & Outcome	No. of studies (Total N)	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Certainty assessment	Certainty
Stroke/SE (Apixaban vs VKA)	4 (25,762)	⊕⊕⊕⊕ – no downgrade	No downgrade (Consistent benefit across heart failure subtypes)	No downgrade (Direct evidence from ARISTOTLE HF sub-study)	Downgrade -1 (CrI [0.50-1.15] crosses 1.00)	No serious concerns	⊕⊕⊕◯	Moderate
Major Bleeding (Apixaban vs VKA)	4 (25,762)	⊕⊕⊕⊕ – no downgrade	No downgrade (Consistent safety profile vs VKA)	No downgrade (Direct evidence from ARISTOTLE HF sub-study)	Downgrade -1 (CrI [0.52-1.13] crosses 1.00)	No serious concerns	⊕⊕⊕◯	Moderate
Stroke/SE (Other DOACs vs VKA)	4 (25,762)	⊕⊕⊕⊕ – no downgrade	No downgrade (Point estimates favor DOAC)	No downgrade (Direct evidence from ROCKET-AF & ENGAGE-AF HF sub-studies)	Downgrade -1 (All DOACs CrIs cross 1.00)	No serious concerns	⊕⊕⊕◯	Moderate
Major Bleeding (Edoxaban vs VKA)	4 (25,762)	⊕⊕⊕⊕ – no downgrade	No downgrade (Consistent reduction in bleeding)	No downgrade (Direct evidence from ENGAGE-AF HF sub-study)	Downgrade -1 (CrI [0.57-1.19] crosses 1.00)	No serious concerns	⊕⊕⊕◯	Moderate

Major Bleeding (Rivaroxaban/ Dabigatran vs VKA)	4 (25,762)	⊕⊕⊕⊕ – no downgrade	No downgrade	No downgrade	Downgrade −1 (CrIs for Rivaroxaban [0.72-1.50] and Dabigatran 110 mg [0.54-1.27] cross 1.00)	No serious concerns	⊕⊕⊕◯	Moderate
Intracranial Hemorrhage (Overall DOAC vs VKA)	4 (25,762)	⊕⊕⊕⊕ – no downgrade	No downgrade (Highly consistent effect across all DOAC trials)	No downgrade	Downgrade -1 (CrIs are wide and cross 1.00, e.g., Apixaban [0.08-2.17])	No serious concerns	⊕⊕⊕◯	Moderate