Clinical trials with new oral anticoagulants

Additive value of indirect comparisons also named network meta-analyses

J. Harenberg¹; C. Weiss²

¹Department of Clinical Pharmacology, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany;
²Department of Biometry and Statistics, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany

Keywords
Oral anticoagulants, apixaban, dabigatran, rivaroxaban, atrial fibrillation, total knee replacement, total hip replacement, prophylaxis of thromboembolism

Summary
To compare the efficacy and safety of the new oral anticoagulants (NOAC), ideally head-to-head clinical trials should be performed. Given the expense of such an undertaking, it is highly unlikely that such a comparison would be performed. Therefore, there is a need for an unbiased comparative assessment of the benefits and risks of the NOACs, based on the available trial data. Indirect or mixed treatment comparisons may be an useful tool to overcome these limitations also known as network meta-analysis (NMA).

The aim of this paper is to give an overview on published NMAs for dabigatran, rivaroxaban and apixaban, each assessed against warfarin in patients with atrial fibrillation, and against enoxaparin in patients undergoing total knee and total hip replacement surgery, in order to obtain insights into the comparability of the adopted methodological techniques.

Patients with non-valvular atrial fibrillation (NVAF) are exposed to an increased risk for ischaemic stroke and systemic embolism. Mortality and morbidity following stroke associated with NVAF are higher than in patients who suffer ischaemic strokes in the absence of NVAF (1). In these patients with NVAF (2) anticoagulation with vitamin K antagonists (VKA) reduces the incidence of
• ischaemic stroke,
• systemic embolism, and
• mortality.

However, severe bleeding complications including intracranial haemorrhage occur (3). VKA usage requires frequent dose adjustments to optimize the time in the therapeutic range (TTR) of international normalized ratio (INR) values between 2 and 3 (4). Due to these limitations, patients at risk for bleeding complications may be undertreated with VKA or receive less effective medications such as aspirin or even no prophylaxis (5).

Venous thromboembolism is one of the major complications following primary elective total hip (THR) and knee replacement (THR) surgery resulting in considerable preventable morbidity and mortality (6). Prophylaxis of thromboembolic events is currently performed by subcutaneous injection of low molecular weight heparins (LMWHs) or fondaparinux for 10 days in TKR or 35 days in THR (7). Limitations of the regimes include (8)
• heparin-induced thrombocytopenia type I and type II,
• local haematoma and allergy,
• transient increase of liver enzymes, and
• the systemic administration.

Correspondence to:
Job Harenberg, MD, Professor of Medicine
Department of Clinical Pharmacology
Medical Faculty Mannheim, University of Heidelberg
Maybachstr. 14, 68169 Mannheim, Germany
Tel. +49(0)621/383 96 23, Fax +49(0)621/383 96 22
E-mail: job.harenberg@medma.uni-heidelberg.de

Klinische Studien mit neuen oralen Antikoagulanzien – Der Zusatznutzen indirekter Vergleiche, auch Netzwerk-Metaanalysen genannt
Hämostaseologie 2013;33: 62–70
received: November 30, 2012
accepted in revised form: January 7, 2013
prepublished online: January 24, 2013
DOI:10.5482/HAMO-12-11-0021
As new oral anticoagulants (NOACs) direct factor Xa and thrombin inhibitors have been developed to overcome the limitations of conventional anticoagulants (9).

- The oral direct factor Xa inhibitors rivaroxaban and apixaban specifically target factor Xa.
- The oral direct thrombin inhibitor dabigatran specifically inhibits the coagulation enzyme thrombin.

All new oral direct coagulation inhibitors are given at fixed doses without laboratory guided dose adjustment.

The three NOACs tested in clinical trials are dabigatran etexilate, which has antithrombin activity, and rivaroxaban and apixaban, which have anti-factor Xa activity (9). In NVAF, results for all three agents are now available from large prospective, randomized trials against standard VKA (warfarin) therapy for the prevention of ischaemic stroke and systemic embolism. The reported trials are the

- RE-LY trial using 110 mg bid and 150 mg bid dabigatran etexilate (10, 11),
- ROCKET-AF trial testing 20 mg od rivaroxaban (12),
- ARISTOTLE trial with 5 mg bid apixaban (13).

All three studies used dose-adjusted warfarin as standard of care control, using a prospective open label blinded endpoint evaluation (RE-LY) or double blind, double-dummy study design (ROCKET AF, ARISTOTLE). These trials showed non-inferiority or superiority for the efficacy outcome of combined ischaemic stroke and systemic embolism. Further, the rates of major bleeding complications were similar or even reduced in comparison with warfarin for all three NOACs.

To compare the efficacy and safety of the NOACs, ideally head-to-head clinical trials should be performed. Given the expense of such an undertaking, it is highly unlikely that such a comparison would be performed. Therefore, there is a need for an unbiased comparative assessment of the benefits and risks of the NOACs, based on the available trial data. Indirect or mixed treatment comparisons may be an useful tool to overcome these limitations also known as network meta-analysis (NMA) (14, 15). This approach allows for indirect comparisons of the three compounds for efficacy and safety outcomes, using information obtained from the connected network of published NVAF trial data (16).

Direct comparisons between two or more treatment regimens are performed in regular clinical trials. Indirect comparisons (also named NMAs) can be performed between the published direct comparisons if the direct comparisons include the same comparator such as enoxaparin or warfarin (Fig. 1). They may also be performed if some of the studies include other comparators such as placebo or aspirin. Such comparisons require various statistical adjustments and the number of trials need to be high.

- The advantages of NMAs are to gain information of treatment regimes, which are highly unlikely to become available due to organizational or economic reasons in near future. They may also generate information about the comparison of efficacy and safety of treatment options which otherwise would be unlikely to be obtained. This may help the medical society to improve patient care.
- The limitations of NMAs include differences in
  - biographic patient’s data,
  - inclusion and exclusion criteria,
  - definition and documentation of efficacy or safety endpoints.

Heterogeneity between studies may substantially affect those analyses as well as inclusion of only subgroups. As a historical note the first NMA was published in 2003, five NMAs were published in 2007, and more than 50 NMAs in the first eleven months of 2012.

The aim of this paper is to give an overview on published NMAs for dabigatran, rivaroxaban and apixaban, each assessed against warfarin in patients with NVAF, and against enoxaparin in patients undergoing TKR and THR surgery, in order to obtain insights into the comparability of the adopted techniques. The high number of NMAs of NOACs in the different indications prompted us to analyse the methodologies using published guidelines (19, 20).

Methodsological considerations

The methodological techniques of NMAs are not generally applicable and may be subject to biases and other limitations. Assumptions of similarity and consistency are particularly sensitive to error thereby rendering results questionable (17, 18). Ongoing research and periodic revisions of the recommendations should further support validity and critical interpretation of these methods.

The international society for pharmacoeconomics and outcome research (ISPOR) provided guidelines for the interpretation of indirect treatment comparisons and network metaanalyses for health-care decision making (19, 20). When these techniques are applied adherence to the ISPOR checklist, diligence and transparency should be employed. The standardization of methods as facilitated by this checklist should enhance the overall credibility, transparency and applicability of ITC (indirect treatment comparisons) and NMA (19).

The major points are summarized as follows:

- The studies included into our indirect comparison of new oral anticoagulants clearly stated the same
  - rationale to prevent embolic diseases in patients with NVAF or VTE in TKR and THR surgery,
  - eligibility criteria based on the definition of NVAF and additional risk factors based on the CHADS2 score (C cardiac failure, H hypertension, A age ≥75 years, D diabetes mellitus, S stroke, 2 two points for stroke) as well as for VTE prophylaxis in orthopaedic surgery.

- The studies were prospective, randomized, multicentric, adequately powered and controlled to adjusted warfarin therapy.

- The search strategy was based on Medline and Embase search strategies to identify the relevant publications according the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) strategy (21).

- Only studies reporting the incidence of predefined primary outcome-events; the secondary outcome measures included...
Fig. 1 Direct (solid line) and indirect (dotted) comparisons of trials; adapted from (26, 43)

a) NOACs or warfarin including patients with NVAF
b) NOAC or enoxaparin in patients undergoing primary elective TKR and THR and on prophylaxis of VTE (indirect comparison also called network meta-analysis)

![Diagram of drug comparisons](image-url)

- major bleeding complications,
- intracranial haemorrhage,
- myocardial infarction and mortality and using dose adjusted warfarin as comparator or major bleeding and mortality in TKR and THR surgery.
- Figures with results for the paired comparisons were included as point estimates and 95% confidence intervals.
- Possible biases and inconsistencies between the studies caused by gender could be excluded.

Age, creatinine, CHADS2 score and INR ranges were reported differently and therefore could not be analysed statistically for the AF-trials.

**Heterogeneity** is another relevant aspect when several studies were performed for one indication. For each clinical endpoint (such as thromboembolism, major bleeding or mortality) homogeneity has to be demonstrated between studies before inclusion into the metaanalysis. This may not always be respected (22); at least at the level of studies on the these NOACs this did not apply here as only one study was available for each NOAC in NVAF.

**Indirect comparisons**

**NOACs in NVAF**

Within a time frame of only three months four NMAs were published. Four NMAs were published on this topic (22–25). All NMAs employed near identical statistical methods and conclude that some differences exist regarding the occurrence of embolic events, intracranial haemorrhage, major bleeding, myocardial infarction, but not on mortality between the NOACs. These studies were compared recently and some minor differences were identified (26) (Fig. 1).

The NMAs reporting on the efficacy and safety of NOACs in NVAF differ in the statistical approach in that three authors used the odds ratio (OR) for endpoint calculations (22), (24), (25) and one used the hazard ratio (HR) (23) (Tab. 1). The advantage of the OR is its independence from the time point at which an event occurs in the study. In contrast, the exact date of an event has to be known for all trials to calculate the HR. The detailed event dates are not available from the NOAC publications, and thus, the majority of authors elected to use the OR (26).

Another indirect comparison of the NOACs was published for a subgroup of patients at increased risk for embolism (CHADS2 score ≥3) ‘to address a lack of comparability between trial populations caused by the restriction of ROCKET-AF to high-risk patients’ (27) (Tab. 1). In this context, some confusion may arise by this publication because patients at risk for embolic events (and hence, eligible for anticoagulation) are defined at a CHADS2 score of ≥3 in their study, in contrast to those with ≥1 stroke risk factors, as recommended in guidelines (28, 29). Also, the studies may not be adequately powered for sub-analysis (only a third of the RE-LY and ARISTOTLE populations were CHADS2 score ≥3, in contrast to ROCKET-AF) (30). The RE-LY reported major bleed and ICH was reported separately. In addition, ROCKET-AF had 101 patients with CHADS2 score not reported for the rivaroxaban group versus only 4 in the warfarin group. The NOACs did not include dabigatran 110 mg bid. This dose was suggested by the EMA for those patients included. Accordingly, analysis of homogeneity between the selected subgroups of patients would have been relevant. These considerations limit the interpretation of this NMA (30).

A recently published NMA performed an indirect comparison of these anticoagulants in those patients by pooling the results of the control groups of the three studies treated with warfarin (31, 32). In addition, they also pooled the results of the NOACs and compared them to the results they calculated for the pooled data of the warfarin groups of the three studies. The authors concluded that an "overall superiority of NOAC over warfarin is largely influenced by the reduction of haemorrhagic stroke". They further conclude "dabigatran 150 mg bid seems to have the best risk/benefit profile" (Tab. 1). In this context, some confusion may arise by the publication because the authors have added another limitation to the problematic analysis of indirect comparisons of the NOACs. The variance of their results is increased if the out-
### Tab. 1 Overview of results of indirect comparisons – also named network meta-analyses – using new oral anticoagulants (NOACs) versus warfarin in non valvular atrial fibrillation (NVAF) for prophylaxis of stroke and systemic embolism

<table>
<thead>
<tr>
<th>citation</th>
<th>Wells</th>
<th>Harenberg</th>
<th>Lip</th>
<th>Mantha</th>
<th>Schneeweis</th>
<th>Testa</th>
<th>Miller</th>
<th>Dentali</th>
<th>Rasmussen</th>
<th>Alonso</th>
</tr>
</thead>
<tbody>
<tr>
<td>stroke or systemic embolism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D 150 vs R</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 110 vs R</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 150 vs A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 110 vs A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>R vs A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 110 vs 150</td>
<td>+</td>
<td>+</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 150 vs HM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>major bleeding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D 150 vs R</td>
<td>-</td>
<td>-</td>
<td>--</td>
<td>+</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 110 vs R</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 150 vs A</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 110 vs A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>R vs A</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 110 vs 150</td>
<td>+</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 150 vs HM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>intracranial haemorrhage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D 150 vs R</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 110 vs R</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 150 vs A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 110 vs A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>R vs A</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 110 vs 150</td>
<td>+</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 150 vs HM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>mortality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D 150 vs R</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 110 vs R</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 150 vs A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 110 vs A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>R vs A</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 110 vs 150</td>
<td>+</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 150 vs HM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>myocardial infarction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D 150 vs R</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 110 vs R</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 150 vs A</td>
<td>+</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>+</td>
<td>na</td>
</tr>
<tr>
<td>D 110 vs A</td>
<td>+</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>R vs A</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
<tr>
<td>D 110 vs 150</td>
<td>-</td>
<td>-</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>-</td>
<td>na</td>
</tr>
</tbody>
</table>
come data of the control groups of the RE-LY, ROCKET AF, and ARISTOTLE trials were pooled. This limits the interpretation of these data.

One NMA of NOACs in NVAF included also phase II studies of smaller sizes and of shorter treatment periods and not reporting all secondary endpoints (33) (▶Tab. 1). The authors also pooled the results of efficacy and safety endpoints of the NOACs despite the known pharmacological variations between them. They also included results of unpublished trials making the interpretation of the data even more difficult. Asymmetrical funnel plots for systemic embolism and myocardial infarction may indicate a bias of the studies included into their analysis (33).

Another NMA analysed several subgroups of bleeding complications without finding differences between the NOACs in NVAF. However, ISTH (international society of thrombosis and haemostasis) definition of bleeding was significantly less frequent in patients receiving apixaban compared to rivaroxaban. In addition, this analysis published data on a combination of the results of all NOACs on the primary efficacy and secondary safety endpoints (34). Such analyses remain difficult to interpret because the NOACs were initially developed separately as individual drugs due to their distinct pharmacological properties. Despite the pharmacological differences the results of the clinical trial are nonetheless pooled and compared to INR-adjusted warfarin, data that were also pooled from the studies included into the analysis.

Self-monitoring and self-management improves the efficacy and safety of INR-adjusted VKA therapy compared to usual management. An indirect comparison with the data of dabigatran’s RE-LY trial suggested that the treatments have similar impact on thrombosis, bleeding and death (35) (▶Tab. 1). The authors stated as limitation that the incidence of endpoint in the self-management group is too low to draw a definite conclusion (34). It is well known that younger patients perform self-management of warfarin therapy. Patients suffer mainly from artificial heart valve replacement and have few concomitant diseases and medication. In contrast patients of the RE-LY study displayed an age of 75 years or older, were multimorbid and most probably took more concomitant medications. These differences add substantial limitations for the comparability of the patient groups included into this NMA. A comparison of the NMAs performed with NOACs in NVAF to date is given (▶Tab. 1).

**NOACs in elective TKR and THR surgery**

Several network meta-analyses were carried out including studies of NOACs versus LMWHs in elective TKR and THR surgery. The direct comparison of an NOAC used a LMWH as comparator. Homogeneity within the study programs for dabigatran (two doses), rivaroxaban, and apixaban were achieved only if enoxaparin 40 mg od was used as comparator (43). Therefore, the indirect comparison between the NOACs could be performed only with this set of homogenous data of the study programs (▶Fig. 1b).

Initially, some NMAs were performed using only some of the four treatment regimens of the NOACs in this indication. Trkulja et al. did not find differences in the efficacy and safety between rivaroxaban and dabigatran in this indication (36). Le-Reun et al. described a superior efficacy of rivaroxaban over dabigatran 220 mg for prevention of VTE and VTE related death as well as major VTE in THR and TKR surgery. There were no differences for major bleeding complications (37). Loke et al. confirmed these data and added that rivaroxaban was superior also to dabigatran 150 mg daily and that the results were independent of the dose of enoxaparin used as comparator (38).

Cohen et al. included apixaban as well as several LMWHs and fondaparinux for prophylaxis of VTE after TKR and THR. They pooled the results of the control groups of all studies for the comparison to the other anticoagulants including the NOACs (39). Studies of control groups cannot be pooled across study programs.
because of differences in definitions of VTE and major bleeding (40). Maratea et al. reported on the efficacy of all NOACs following total TKR and THR orthopaedic surgery (Tab. 2). They did not include safety parameters into their NMA. This adds a substantial limitation to their study since major bleeding complication and mortality are relevant aspects for patients undergoing VTE prophylaxis postoperatively (41). Another limitation of NMA becomes evident if the end points are pooled to a net clinical benefit from studies for prevention of VTE in patients undergoing elective TKR and THR surgery. Accordingly, the main conclusion of this paper of a higher efficacy of new anticoagulants were generally associated with a higher bleeding tendency and that the NOACs did not differ significantly for efficacy and safety reflect the contradictory statements based on such pooling of data (42).

One NMA compared all four treatment regimens against each other without pooling of data, but by using the data of the individual control groups for every study (43). A cluster analysis preceded the meta-analysis and the following NMA to identify homogenous groups within a study pro-

| Tab. 2 Overview of results of indirect comparisons – also named network meta-analyses – using new oral anticoagulants (NOACs) versus enoxaparin in total knee and hip replacement (TKR and THR) surgery for postoperative prophylaxis of venous thromboembolism |
|---|---|---|---|---|---|---|
| citation | Trkulja | LeReun | Loke | Cohen | Marateka | Harenberg |
| 10 ± 5 days | 34 ± 5 days |
| venous thromboembolism | D 150 vs D 220 | na | na | na | na | na | - | - | na |
| D 150 vs R | + | + | na | na | + | + | + | - | major VTE |
| D 220 vs R | + | + | + | na | na | + | + | + | |
| D 150 vs A | na | na | na | na | + | + | + | na | + |
| R vs A | na | na | na | na | - | - | na | - | - | major VTE |
| major bleeding | D 150 vs D 220 | na | na | na | na | na | - | - | na |
| D 150 vs R | - | - | na | na | na | na | - | - | - |
| D 220 vs R | - | - | - | na | na | na | - | - | |
| D 150 vs A | na | na | na | na | - | - | na | - | - | + A, CRMB |
| D 220 vs A | na | na | na | na | - | - | na | - | - | |
| R vs A | na | na | na | na | - | - | na | - | - | |
| mortality | D 150 vs D 220 | na | na | na | na | na | - | - | na |
| D 150 vs R | na | na | na | na | na | na | - | - | na |
| D 220 vs R | na | na | na | na | na | na | - | - | na |
| D 150 vs A | na | na | na | na | na | na | - | - | na |
| D 220 vs A | na | na | na | na | na | na | - | - | na |
| R vs A | na | na | na | na | na | na | - | - | na |
| p-values | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| PRISMA criteria | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| ISPOR criteria | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| cluster analysis | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| odds ratio | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| NOAC pooling | √ | √ | 

NOAC: new oral anticoagulant; TK(H)R: total knee (hip) replacement; VTE: venous thromboembolism; CRMB: clinical relevant major bleeding; D: dabigatran; A: apixaban; R: rivaroxaban; PRISMA and ISPOR: see main document; na: not analysed; + significant; – not significant; √ applied.
Program. Clustering algorithms are developed for datasets that are too large and complex for manual analysis. Briefly, clustering partitions objects into clusters, such that objects with similar characteristics are clustered together and dissimilar objects are in different clusters. It is anticipated that no prior knowledge exists on any object classifications (44–46). Only the separation for treatment durations of 10 ± 5 days and 34 ± 5 days led to homogenous groups, whereas other variables such as type and dose of LMWH, gender, age and other variables did not lead to homogeneous groups. The relevant additional findings were that

1. rivaroxaban performed better than both doses of dabigatran for 10 ± 5 days and 34 ± 5 days of prophylaxis on the incidence of the composite of VTE events,
2. apixaban and rivaroxaban did not differ for this outcome for both treatment periods at 35 days.
3. There were no differences in major bleeding events between all NOAC treatment regimens.

Whether the trend for a lower incidence of major bleeding with apixaban compared to rivaroxaban during 34 ± 5 days of prophylaxis is relevant remains to be confirmed in additional trials (43) (Tab. 2). NMAs without proving for heterogeneity between trials, without using OR and 95% CI of each trial program, and without comparison for statistical significant differences between the NOACs for their efficacy and safety (47) may be easier to perform but do not follow the ISPOR guidelines described.

An indirect comparison between the NOAC trial-programs was published to overcome problems of direct comparisons between these new compounds. However reporting of the defined endpoints may differ markedly between the study programs and in particular, for the bleeding events (40). Pooling data with divergent endpoint criteria will introduce a bias that may influence the result of the meta-analysis (43). The original approach for an indirect comparison intended to minimize bias by using the OR and 95% CI of endpoints of homogeneous trials included into network meta-analysis (14,15). Therefore, we decided to use this approach to indirectly compare the NOACs from the respective trial programs. However, due to the lack of standardized definitions for VTE and bleeding, we adopted cluster analysis to identify groups of those studies within the trial programs leading to homogeneity for these endpoints as well as mortality before performing meta-analysis for each NOAC followed by NMA. A comparison of the results of the NMAs using NOACs compared to enoxaparin following TKR and THR surgery is summarized (Tab. 2).

**Treatment with NOACs**

NOACs are currently investigated in clinical trials for treatment of acute deep vein thrombosis and stable pulmonary embolism. Rivaroxaban was approved by the FDA and EMA for both indications. Dabigatran and apixaban are at different stages of the clinical development. A NMA was published including the data of abstracts (48). This is not in agreement with the PRISMA and ISPOR guidelines for performing metaanalysis or NMAs. Further NMAs will be available in this field as soon as the corresponding studies are published in peer reviewed journals.

**Discussion**

In the absence of a direct comparison of treatments of interest, an indirect comparison can provide useful information about differences in efficacy and safety aspects between competing interventions. NOACs represent an important example as they have never been compared in head-to-head trials. In trials on NOACs, low molecular weight heparins or warfarin have been used as common comparators. Therefore, an indirect treatment comparison (ITC) can utilize the relative effects of the treatments versus the common comparator. NMA is an emerging tool for ITC. The international society for pharmacoconomics and outcome research (ISPOR) provided a checklist of good research practices for reporting network meta-analyses (19,20).

Limitations of NMA of NOAC in patients with NAVF include that warfarin treatment open in the RE-LY study versus double dummy in the two other studies; differences in time in therapeutic range of the international nationalized ratio between the studies; differences in biographic data of the patients across the studies such as gender, CHADS2 score, creatinine clearance, and concomitant use of aspirin (26). These limitations are neglected by several authors pooling the data of the control groups for the direct comparison as well as of treatment groups for a indirect comparison of the NOACs.

Although the methodological issues regarding ITC and NMA are recognized, these methods are increasingly applied to trials on NOACs, as it is assumed that compiling and analysing information from all available sources as an evolving process may improve health-care decision making. For that reason, the goal of the Task Force is to inform and educate policymakers and health care professionals about the studies and identify areas of future research such as NMAs and to sort out methodological weaknesses (19,20).

NMAs typically encounter a number of statistical problems. Heterogeneity is one relevant aspect when several studies were performed for one indication. For each clinical endpoint (such as thromboembolism, major bleeding or mortality) homogeneity has to be demonstrated between studies before performing the metaanalysis. This may not always be respected.

If only one but large study (more than 6000 patients per group) is available for each NOAC in NVAF this did not apply. It has also to be taken into consideration, that the pharmacology of NOACs differs substantially and clinical data on NOACs may thus not necessarily be pooled versus the comparator. Therefore, pooling of data of the NOACs or of the control groups may be problematic.

Several NMAs have been conducted after the publication of the RE-LY and ROCKET AF, and ARISTOTLE trials for prevention of embolism in atrial fibrillation (22–25, 27, 30–34) and of many
studies for prevention of thromboembolism following total hip and knee replacement surgery (35–41). Sub-analyses mainly reflect the interest of experts and health communities to optimize treatment for patients. This motivation and the opportunities inherent to NMAs may not always be fully appreciated when related publications are criticised. Even small progresses in science may improve patient care, and NMAs appear as a useful tool in the decision making process of doctors and patients to change or not to change from conventional anticoagulation to a NOAC. However, the complex methodology of ITC and NMA render the interpretation of these specific re-analyses difficult.

The three NOACs dabigatran, rivaroxaban, and apixaban differ (9) in:
• mode of action (factor IIa and factor Xa inhibition),
• pharmacology (prodrug, association and dissociation constants),
• pharmacokinetic and pharmacodynamic parameters (absorption and elimination half lives, metabolism),
• drug interactions and
• side effects.

For this reason the NOACs are regarded as different drugs and therefore their data obtained from clinical studies cannot be pooled. The analyses which pooled the data of the NOACs lead to a substantial confusion in the new area of anticoagulation where substantial sets of data are now available leading to an improvement patient care suffering from non-valvular atrial fibrillation after over 50 years of using vitamin-K antagonists as the only available blood thinner.

Conclusion

The additive value of indirect treatment comparisons becomes difficult to substantiate due to the manifold methodological problems. Comparing the literature of NMAs on NOACs in NVAF and in TKR/THR surgery the most frequent methodological issues refer to:
• use of hazard ratio instead of odds ratio,
• analysing not all treatment regimens performed,
• not analysing safety endpoints,
• analysis of subgroup of patients,
• pooling of data of the NOACs as well as of the comparator, and
• not respecting heterogeneity within a trial program.

Standardization of methods is required including search of approaches to identify homogenous groups such as hierarchical clustering. The large number of NMAs makes it more difficult for practitioners and patients to identify those following published guidelines. Despite all these limitations all NMAs identify a non-inferiority versus INR-adjusted warfarin in atrial fibrillation and NMH-prophylaxis of VTE following elective TKR/THR surgery regarding efficacy and frequently also of safety. In addition to the results of the NMAs the cost-effectiveness evaluations (49–53) may facilitate patients and health care professionals to improve patient care by NOAC in NVAF and elective TKR/THR surgery in favour of NOACs.

Conflict of interest

The authors declare that they do not have to declare any conflicts of interest.

References

23. Lip GY, Larsen TB, Skjøth F et al. Indirect comparisons of new oral anticoagulant drugs for efficacy and safety when used for stroke prevention in...