Sex-specific efficacy and safety of cryoballoon versus radiofrequency ablation for atrial fibrillation: An individual patient data meta-analysis

Jeanne du Fay de Lavallaz, MD, PhD[1,2]; Patrick Badertscher, MD[1,3]; Atsushi Kobori, MD, PhD[4]; Karl-Heinz Kuck, MD[5]; Josep Brugada, MD[6]; Serge Boveda, MD, PhD[7,8]; Rui Providência, MD, PhD[7,9,10]; Ziad Khoueiry, MD[11]; Armin Luik, MD[12]; Fabien Squara, MD[13]; Ioanna Kosmidou, MD, PhD[14]; Karapet V Davtyan, MD, PhD[15]; Arif Elvan, MD, PhD[16,17]; Nicasio Perez-Castellano, MD, PhD[18]; Ross Hunter, MD[9], Richard Schilling, MD[9]; Sven Knecht, MS[1,19]; Kojodjojo, MBBS, PhD, FHRS[20]; Jeremiah Wasserlauf, MD, MS[21]; Hakan Oral, MD[22]; Mario Matta, MD[23]; Sandeep Jain, MD[24]; Matteo Anselmino, MD[25]; Michael Kühne, MD[1,19]

[1] Cardiovascular Research Institute, University Hospital Basel, Basel, Switzerland
[2] Rush University, Department of Internal Medicine, Chicago, IL, USA
[3] Charleston Hospital, Department of Electrophysiology, SC, USA
[4] Kobe City Medical Center General Hospital, Division of Cardiovascular Medicine, Kobe, Japan
[5] Department of Cardiology, Asklepios Klinik St. Georg, Hamburg, Germany
[6] Hospital Clínic de Barcelona, Servicio de Cardiología, Barcelona, Spain
[7] Clinique Pasteur, Heart Rhythm Department, Toulouse, France
[8] Universitair Ziekenhuis Brussel, Vrije Universiteit Brussel, Brussels, Belgium.
[9] St. Bartholomew's Hospital, Barts Health NHS Trust, London, United Kingdom
[10] Institute of Health Informatics Research, University College of London, London, United Kingdom
[11] Clinique St. Pierre, Department of Cardiology, Perpignan, France
[12] Medizinische Klinik IV, Städtisches Klinikum Karlsruhe, Karlsruhe, Germany
[13] University Hospital of Nice, Pasteur Hospital, Department of Cardiology, Nice, France.
[14] New York Presbyterian Hospital/Columbia University Medical Center, New York, NY, USA
[15] National Medical Research Center for Preventive Medicine of the Ministry of Healthcare of the Russian Federation, Heart Rhythm and Conduction Disorder, Moscow, Russia
[16] Isala Heart Centre, Zwolle, The Netherlands
[17] Diagram, Zwolle, The Netherlands
[18] Cardiovascular Institute, Department of Cardiology, Instituto de Investigación Sanitaria del Hospital Clínico San Carlos (IdISSC), CIBER de Enfermedades Cardiovasculares.
[19] University Hospital Basel, Department of Cardiology, Basel, Switzerland
[20] National University Hospital, Singapore, Singapore
[21] Northwestern University Feinberg School of Medicine, Department of Electrophysiology, Chicago, IL, USA
[22] University of Michigan, Cardiac Arrhythmia Service, Ann Arbor, MI, USA
[23] Sant'Andrea Hospital, Cardiology Division - Electrophysiology, Vercelli, Italy
[24] University of Pittsburgh School of Medicine, UPMC Heart and Vascular Institute, Center for Atrial Fibrillation, Pittsburgh, PA, USA
[25] Città della Salute e della Scienza di Torino Hospital, Department of Medical Sciences and Division of Cardiology, University of Turin, Turin, Italy.

Running Title: Sex-specific individual patient data meta-analysis of atrial fibrillation ablation

Correspondence to:
Professor Michael Kühne, Department of Cardiology, University Hospital Basel, Petersgraben 4, 4051 Basel, Switzerland. E-mail: michael.kuehne@usb.ch

Contributors :
Rod Passman, MD, MSCE[26], Georgiy Yu Simonyan, MD[27]; Donald Siddoway, MD[28]; Miki Yokokawa, MD[29]; Sangeeta Lathkar-Pradhan, MBBS[29]; Thomas Buist, MD[30]; Christiane Pudenz, PhD[31], Fred Kueffer, PhD[32], Bastian Fries, MD[33].

[26] Northwestern University Feinberg School of Medicine, Department of Electrophysiology, Chicago, IL, USA
[27] National Medical Research Center for Preventive Medicine of the Ministry of Healthcare of the Russian Federation, Department of Heart Rhythm and Conduction Disorder, Moscow, Russia
[28] University of Pittsburgh School of Medicine, UPMC Heart and Vascular Institute, Center for Atrial Fibrillation, Pittsburgh, PA, USA
[29] University of Michigan, Cardiac Arrhythmia Service, Ann Arbor, MI, USA
[30] Isala Heart Centre, Zwolle, The Netherlands
[31] Cardiovascular Research Institute, University Hospital Basel, Basel, Switzerland
[32] Medtronic, Mounds View, Minnesota.
[33] Medizinische Klinik IV, Städtisches Klinikum Karlsruhe, Karlsruhe, Germany

Abstract (336/350)

Background

Atrial fibrillation (AF) is a growing health burden and pulmonary vein isolation (PVI) using cryoballoon (CB) or radiofrequency (RF) represents an attractive therapeutic option. Both energy sources have been shown to be equally safe and efficacious. While sex-specific differences in the epidemiology, pathophysiology and clinical presentation of AF are recognized, the comparative sex-specific characteristics of CB versus RF ablation have not been definitively assessed and classical meta-analytic techniques are not feasible due to the lack of published subgroup analyses.

Methods

We performed a structured electronic database search of the literature for randomized controlled trials (RCTs) and observational prospective studies comparing CB and RF ablation efficacy with at least 1 year follow-up and contacted all authors for individual patient-data. After collecting and merging individual patient data from 18 datasets, we investigated the sex-specific efficacy, safety and procedural characteristics of CB versus RF in a uniform and sex-specific manner. Kaplan Meier and multi-level models were used to assess the effect of female sex and of the type of energy source on efficacy (procedure failure defined as recurrence of atrial arrhythmia, re-ablation and start of anti-arrhythmic medication), safety (peri-procedural complications) and procedural characteristics.

Results

Data from 18 studies were gathered, representing 6819 patients (4840 men, 1979 women). While women were at higher risk of recurrence, CB was associated with less failures than RF in models correcting for the most important comorbidities (Risk reduction of 9%, 95% CI 0-17%**,** p=0.028). An analysis stratified by sex again showed

a better efficacy of CB in men (p-value = 0.027) but not in women (p-value = 0.939). CB but not female sex was associated with a slightly higher risk of complication (CB OR= 1.01, 95%-CI 1.00-1.03, p=0.048). The total procedure time was shorter when CB was used (-23min with CB, p<0.001).

Conclusion

Women are at higher risk for ablation failure than men and cryoballoon ablation is associated with less long-term failures in men but not in women. Technological device improvement specifically for female patients could lead to higher success rates in women.

Introduction

Atrial fibrillation (AF) represents a growing health problem and is currently leading to an increasing burden of morbidity, mortality and hospitalizations worldwide.[1,2]

Sex-specific differences in the epidemiology, pathophysiology, and clinical presentation of AF are recognized.[3] While the prevalence of AF is higher in men than women, women live longer and the cumulative lifetime risk of developing AF has been reported to be significantly higher in women than men after 40 years old.[3,4] Furthermore, women with AF show higher mortality rates[2], lower quality of life[4], lower tolerability of anti-arrhythmic drugs[4] and higher stroke incidence than men[5]. Therefore, definitive AF treatment could be particularly beneficial to this patient population.

Cryoballoon (CB) or radiofrequency (RF) are two commonly used energy sources for AF ablation and have been shown to be equally safe and effective in the limited number of available randomized controlled trials (RCTs)[6–10], which randomized a total of 1359 patients.

Female sex has been associated with increased risk of arrhythmia recurrence, with a different complications profile and cardiovascular rehospitalizations after catheter ablation for AF.[11,12] However, little is known about the comparative efficacy and safety of both ablation technologies in male versus female patients.

To investigate this important sex-specific question, we conducted an individual-patient data meta-analysis of RCTs and large observational prospective studies comparing RF and CB ablation of AF in men and women.

Methods

This systematic review was registered on PROSPERO (CRD42019125515) and was approved by the local ethics committee of Basel (Ethikkomission Nordwest und Zentralschweiz Project ID 2018-01529)

Search, study selection, call for data, individual patient data collection and datasets merging

Details regarding the search, study selection and contact of the authors are available in the supplemental appendix. In brief, we searched publication databases for the terms “atrial fibrillation”, “pulmonary vein ablation”, “radiofrequency” and “cryo*” on March 28th 2018 and March 15th 2019. We included studies if they met the following pre-specified criteria: 1) Randomized controlled trials (RCT) or prospective observational studies (POS), 2) with at least 40 patients per group (CB versus RF) for POS, 3) with patients undergoing their first ablation, 4) using a first- or second-generation CB and , non-irrigated, non-contact-force irrigated or contact-force guided RF catheters, 5) investigating an efficacy outcome of time-to-failure (defined as recurrence of atrial arrhythmia, re-ablation and re-start of anti-arrhythmic medication) and/or a safety outcome (defined as the percentage of recorded complications) and 6) following patients for at least 12 months per original publication or by author confirmation. We contacted first and/or corresponding authors of each publication at least 3 times at 2-month intervals. We discussed the availability of required variables and collected the datasets which were qualifying per inclusion criteria in order to conduct a one-step analysis. For one study, regulations did not allow for sharing of individual patient data (the Fire&Ice study). In order to integrate the data to this IPD meta-analysis, an investigator of the current project (JdFdL) programmed the

analysis which was then ran on the Fire&Ice data set by a statistician of the Fire&Ice study team. The estimates were provided for a 2-step analysis.

After collection, we renamed and relevelled all variables with the help of the dataset providers in order to obtain a homogenous dataset structure allowing for pooling of 17 datasets. As previously mentioned, the 18th dataset from the Fire&Ice study was kept at the Medtronic Headquarter (Minneapolis) and was not merged with the other data sets.

Endpoints

We assessed the effect of sex on efficacy, safety and procedural endpoints of PVI using either CB or RF. The efficacy endpoint was the recurrence of arrhythmia, re-ablations or re-start of anti-arrhythmic medications following a 90-day blanking period. Atrial arrhythmias were defined as AF, atrial flutter or atrial tachycardia. This endpoint was assessed at 360, 720 and 1080 days of follow-up, as longer-term follow-up was available only in patients undergoing an RF ablation. As some studies were designed to report only the recurrence of arrhythmia[6,7,13–19], we planned a sensitivity analysis using only arrhythmia recurrence as an endpoint. We thereby tested the hypothesis that re-ablations and re-start of medications would not have a major impact on the results, as they are most likely surrogate events of arrhythmia recurrence.[20]

The safety endpoint was the composite of all recorded peri-procedural complications, which encompassed death, cerebrovascular events, serious treatment-related adverse events and groin complications. Data were separately extracted for the following complications: peri-procedural death, acute myocardial infarction (AMI), stroke, pericardial effusion, tamponade, phrenic nerve palsy, groin complications, pulmonary vein stenosis and esophageal injury.

The procedural endpoints were the duration of the total procedure and the fluoroscopy time.

Assessment of study quality

Study quality was assessed according to two pre-specified tools : the Cochrane Collaboration tool for RCTs[21] and a modified Newcastle-Ottawa Scale (NOS) for non-randomized observational studies (criteria listed in the supplement).

Statistical analysis

The analysis was performed according to the recommendations of the Cochrane Collaboration[21] and the reporting was in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement.[22] (Supplemental table 1).

Continuous variables are presented as mean ± standard deviation (SD) when normally distributed and median with interquartile ranges (IQR) when non-normally distributed. Categorical variables are expressed as numbers and percentages. Mann-Whitney-U test was applied for comparison of continuous variables and Fisher's exact test for comparison of categorical variables. Confidence intervals of percentages were computed according to Agresti and Coull[23].

Missing data were imputed and a one-step analysis was conducted on the merged dataset. The same analysis was conducted separately on the 18th dataset (Fire&Ice study) and either the estimates were merged in a two-step analysis (for the models) or the time-to-event results were integrated in the one-step analysis of the 17 other datasets (for the Kaplan-Meier analyses, details in the supplemental).

For the efficacy endpoint, Kaplan Meier representing time-to-failure of CB versus RF were first constructed using all 18 studies. The analysis was stratified by sex, by the exact catheter subtype (CB 1st versus 2nd generation and contact-force versus non-

contact force RF) and by AF type (paroxysmal versus persistent). Differences between groups were tested using a log rank test at 360, 720 and 1080 days after the proportional hazards assumption was checked using a statistical test based on the scaled Schoenfeld residuals. As all studies planned a 90-day blanking period in their design, all time-to-event analyses were started after the 90th day post-ablation. Furthermore, to account for clustering of the studies and the influence of important comorbidities, multi-level Cox proportional hazard models taking into account the type of catheter intervention (RF vs CB), sex, and the interaction of both parameters were fitted. Correction for covariates highlighted in previous literature as decisive for the recurrence of AF following ablation[11] (See supplemental methods) was applied. The interaction term between the catheter type and sex was removed when non-significant.[24] Similar Cox models stratified by sex were derived. Models were derived once in the large merged dataset containing all 17 studies and once in the Fire&Ice dataset, allowing for a two-stage analysis using classical meta-analytic techniques (Restricted maximum Likelihood (REML) random-effects model).

A sensitivity analysis on the endpoint of arrhythmia recurrence only was conducted in the merged dataset using similar multi-level Cox proportion hazard models.

For the safety endpoint, a multi-level logistic regression model investigating the association between the catheter type, sex, the interaction of both terms and important comorbidities with peri-procedural complications was fitted.

For the procedural endpoints, multi-level linear regression models taking similar parameters into account were fitted to investigate their impact on fluoroscopy and total procedure time.

To investigate heterogeneity between individual studies for all endpoints, a two-stage analysis of individual studies was conducted using simplified models and pooled

using a Restricted maximum Likelihood (REML) random-effects model. A similar random-effect model was used to pool the peri-procedural complications and procedural characteristics depending on the ablation device used.

Heterogeneity was determined using I^2 as measure. Significant heterogeneity was defined as an I^2 statistic of >50%. Heterogeneity was investigated for three pre-specified variables (publication year, mean age of the enrolled patients and study type, namely RCT versus non-RCT) using meta-regressions, as we hypothesized that these variables would be available for all studies and could play a possible role in results divergence.

Evidence for publication bias was assessed graphically using funnel plots and the Egger test.

All statistical analyses were performed using the Statistical Software “R” (R Foundation for Statistical Computing, Vienna, Austria). Detailed explanation of the statistical analysis is available in the supplemental.

Results

Selected studies

A total of 1081 studies were identified and 29 authors were finally contacted (Supp. Fig. 1, Supp table 2). Nine authors did not wish to participate, 1 did not respond and 4 publications were linked to two datasets, leaving 18 datasets (5 RCTs, 13 POSs) available for analysis. The characteristics of the 18 studies are presented in Supp table 3-4.

Nine studies[6,7,13–19] reported exclusively recurrence of arrhythmia without re-ablation or re-start of anti-arrhythmic medications.

The 18 studies accounted for a total of 6819 patients (4840 men, 1979 women, 2501 ablations with CB, 4318 ablation with RF).

As some patients were lost to follow-up during the 90-day blanking period, 6581 patients were available for the efficacy analysis. Due to missing data, 5725 and 6308 patients were available for the analysis of fluoroscopy time and total procedure time, respectively.

The mean duration of follow-up in included studies varied from 8.8 to 51.6 months and monitoring was appropriate in all studies, using either Holter ECGs or Loop recorders (Supplemental table 5). Some studies presented with median a follow-up shorter than 12 months, as some patients were lost to follow-up before this time point.

Four[6,10,25,26] out of five RCTs had an age-limiting inclusion criterion set at 75 or 79 years old.

Baseline Patient Characteristics

Baseline patient characteristics by sex are presented in table 1. Baseline patient characteristics by energy source are presented in Supp table 6.

Women were older, presented more often with a severely dilated left atrium (LA), on average, had a better left ventricular ejection fraction (LVEF). While the number of men and women undergoing an ablation with either CB or RF was balanced in the F&I study, more women than men were ablated using CB in the merged dataset in the 17 studies. Patients undergoing an ablation with RF presented more often with hypertension, CHF, with a previous stroke, or with LVSD.

Efficacy analysis

In the merged dataset, on the 5831 patients available for analysis, 2198 patients experienced a failure (1951 experienced an arrhythmia recurrence during the overall follow-up, 155 underwent a redo ablation and 92 were re-started on arrhythmic medications). In the F&I data set, a recurrence was observed in 281 patients (167 arrhythmia, 14 redo and 100 re-started on medications) (Supp. Table 7). In both the Fire&Ice study as well as the merged dataset, women experienced more failure of the efficacy outcome than men. In the merged dataset, patients undergoing an ablation with CB experienced significantly less recurrences than patients undergoing an ablation with RF but this difference was not observed in the Fire&Ice dataset. Kaplan Meier of the event-free survival in the merged dataset combined with data from the F&I study are represented in Figure 1. While men undergoing an ablation with CB experienced less recurrences at 2 and 3 years follow-up, this was not the case for women. In the overall population, the advantage of an ablation using CB was present starting at two years follow-up. A better performance of CB 2nd generation over CB 1st generation and of RF with contact force over RF without contact force was observed in the overall cohort (Supplemental figure 2).

The cox proportional hazard models correcting for a large number of clinically relevant covariates and fitted both in the merged and F&I dataset are presented in

Table 2. As no interaction was present between female sex and catheter type (p for interaction=0.661 in the merged data set and 0.290 in the Fire&Ice data set), and the interaction term was therefore removed from the final model.[24]

The combined hazard ratio of the catheter type for both data sets showed a modest but significant improvement in efficacy up to 3 years follow-up when CB was used in the overall cohort (Figure 2A, risk reduction of 9%, 95% CI 0-17%**,** p=0.028). An analysis stratified by sex again showed a better efficacy of CB in men (Supplemental Table 8, Figure 2C, p-value = 0.027) but not in women (Supplemental table 9, Figure 2B, p-value = 0.939).

The combined hazard ratio of sex for both data sets showed that women were at higher risk of experiencing a recurrence (Supplemental Figure 3)

Heterogeneity between the merged dataset and the F&I study for all analyses of the efficacy endpoint was low.

Similar results were observed in the merged dataset when the efficacy endpoint of recurrence of arrhythmia only (without redo ablations or re-start of antiarrhythmic medications) was considered (Supplemental Table 10).

Safety analysis

The numbers of periprocedural complications for both the merged dataset and the F&I dataset depending on the type of catheter used or on the patients' sex are presented in Table 3. Women presented with a higher rate of complications in the merged dataset, which was driven by the groin complications, phrenic nerve palsies and tamponades. CB also presented with a higher complication rate, driven by a larger number of phrenic nerve palsies. The logistic regressions accounting for several clinical relevant covariates and modelling the occurrence of periprocedural complications in both the merged and F&I data set are presented in Supplemental

Table 11. The pooled odds ratios from these models showed that CB but not female sex was associated with a slightly higher risk of complication (CB OR= 1.01, 95%-CI 1.00-1.03, p=0.048, and female sex OR=1.05, 95%-CI 0.88-1.24 , p=0.597, Figure 3). The heterogeneity between the merged dataset and the F&I dataset was low for the safety endpoint.

Procedural endpoints analysis

The median fluoroscopy and total procedure time for the merged and F&I datasets are presented in Table 4.

A linear regression model accounting for several clinical relevant covariates was derived to model these two endpoints (Supplemental Table 12 and 13) and the estimates for the pooled estimates for the merged and the F&I datasets are presented in Figure 4.

While a much shorter total procedure time was observed when CB was used (-23min with CB, p<0.001), results were contradictory and very heterogeneous regarding fluoroscopy time (+2.02min with CB, p=0.5), where CB fluoroscopy time was longer in the merged dataset but shorter in the F&I dataset.

There was a non-significant trend for longer procedure and fluoroscopy time in women (and +1.74min total procedure time, p=0.23 and +0.67min fluoroscopy time, p=0.35).

Heterogeneity analyses

The pooled estimates for all data sets using a simplified model accounting only for the type of catheter and sex are presented in Supplemental Figures 4 for efficacy, 5 for safety and 6 and 7 for procedural endpoints.

While efficacy and procedural estimates by catheter type were very heterogenous between studies, safety results were more homogenous (Supplemental figure 4A, 5A, 6A and 7A). The sex-specific estimates of all observed outcomes also presented with little heterogenicity (Supplemental figure 4B, 5B, 6B and 7B). Mean age of the enrolled patients, year of publication and study design (RCT versus OP) were investigated as source of heterogeneity but none of these parameters significantly contributed to the heterogeneity between studies for any of the endpoints (Supplemental table 14)

Study quality and publication bias

The quality of the included dataset was summarized in Supplemental table 15 and 16 and Supplemental Figure 8. A Funnel plot of the efficacy outcome by catheter type appeared symmetrical (Supplemental Figure 9) and an Egger test did not find any publication bias (p-value of Egger test =0.88).

Discussion

Gathering sufficient data on women is challenging in cardiovascular trials, as they tend to be under-represented.[27–29] Particularly in devices or interventional trials, such as trials investigating ICDs implantations or coronary angiographies, the proportion of enrolled women generally account for about a third of the cohort.[30–33] Moreover, despite the recognition of the growing importance of sex-based differences in medicine,[34] the reporting of sex-specific analyses was found to be low in a large number of publications[35,36] and the ones comparing CB to RF are no exception. The lack of sufficient published sex-specific subgroup analyses therefore hinders any classical meta-analytic conclusion[37]. We therefore conducted this large individual patient data meta-analysis to investigate the efficacy, safety and procedural outcomes of CB versus RF ablation in patients undergoing a first ablation and focused more specifically on sex-specific outcomes. We report seven main findings. **First**, only 29.02% percent of the enrolled patients were women and this proportion ranged from 21.5 to 46.1% in the individual studies. **Second**, CB appeared to be more efficacious in the overall population. **Third**, CB 2nd generation performed better than CB 1st generation and RF contact-force better than irrigated-tip RF without contact-force. **Fourth**, women were at higher risk for efficacy failure than men, independently of the type of catheter used. **Fifth**, in unadjusted analyses (represented in the Kaplan Meier curves), CB appeared to be less effective in women than in men. This was confirmed in large multivariable models correcting for a substantial number of comorbidities where, despite a negative interaction term between the catheter type and sex, stratified sex-specific analyses showed a better efficacy of CB over RF in men but not in women. **Sixth**, CB and female sex were both associated with a higher risk of complications. **Seventh**, CB was associated

with a much shorter overall procedure time. However, fluoroscopy duration was too heterogenous between the studies to allow for any conclusion.

The low number of women enrolled in the studies integrated in this IPD meta-analysis is consistent with the observed lower enrollment of women in cardiovascular and more specifically trials involving an intervention[29,38]. Several factors, either directly related to the female sex or more based on physician bias[4,39], have been suggested to be responsible for the discrepancy in recruitment. Women seem to develop AF later[3,4] and therefore account for a smaller portion of the young AF population. Acknowledging the age-limiting criterion of 4 out of 5 RCTs included in this meta-analysis, this later development and symptom presentation may restrict the overall number of women available for inclusion. Moreover, a possible bias of physician towards a lower referral of women for invasive procedures could further contribute to the low enrollment rate[4,40,41]. The low number of women enrolled in each individual trial therefore consequently limited the conclusions drawn from each of them and the current pooled analysis provides important insights into a larger female population.

As observed in some previous studies and confirmed by the current analysis, women appear to be at higher risk of arrhythmia recurrence[4,12,39,42]. This increased risk of recurrence in women has often be attributed to their comorbidity profile, but our extensive models correcting for a large number of covariables suggest that other sex-specific factors (such as hormones, the enhanced arrhythmogenicity of atrial cells in women or sex-specific variations in LA size[5,43]) could play a role in this difference.

In the overall population, CB performed better than RF starting at 2-year follow-up and this superiority was already observed at 1-year follow-up in males. Interestingly,

women did not benefit more from an ablation with CB compared to RF at any time point of the follow-up. The better performance of CB later in the follow-up raise further questions on the cellular damages and their durability induced either by “freezing” or “burning” the cells. For instance, late or peripheral apoptotic mechanisms as well as deep lesions have been associated with CB[44–47] and could possibly be responsible for delayed efficacy. Several hypotheses could contribute to the absence of superiority of CB in women as compared with men. First, CB technologies might have been more developed for the larger “male cardiac anatomy” and tested in males more than females, therefore limiting the generalizability of the technology to women. While a 23-mm CB is available for smaller-sized pulmonary veins[48], such as supposedly the ones of women’s heart, we showed that women more often presented with severely dilated LA, suggesting that larger or other devices may be required in these patients for adequate occlusion and energy transmission. Second, similar factors as the ones proposed for higher arrhythmia recurrence in female patients could interact with CB more than with RF. For instance, electrical (such as more non-pulmonary veins foci[4,5,43]), endocrine (such as hormone replacement therapy in older women or menopause age[3]) and structural factors (such as more atrial fibrosis or higher inflammation[5,43]) are important sex-specific differences in pathophysiological mechanisms of atrial fibrillation which may also interact with the type of ablation energy selected.

Other cofactors known to play an important role in the recurrence of AF following an ablation (such as LA dilation, AF duration of paroxysmal versus persistent AF[49–51]) were also significant predictors of AF recurrence in the current analysis. However, as these comorbidities have all been integrated in our predictive model, they are likely not factors explaining the higher recurrence risk observed in women or with RF.

CB second generation performed better than CB first generation and similarly to contact-force radiofrequency catheters. This finding highlights the large difference observed depending on the type of technology used and the possible improvements based on hardware refinement[52,53]. Acknowledging an always growing number of patients eligible for AF ablation, further improvement in technologies could rapidly lead to significantly improved outcomes.

In several previous studies, female sex has been associated with an increased rate of complications[4,39,42,54,55]. In the present cohort, we confirmed these observations and found that women presented with a higher rate of complications (driven by groin complications, tamponades and phrenic nerve palsies). However, we could not observe any higher risk in multivariable models. The increased complication rate is therefore most likely due to the older age of presenting women more than to female sex itself.

While CB was clearly associated with a shorter procedure time, data on fluoroscopy were very heterogenous between studies, reflecting the differences in interventions techniques between centers. All results presented in this meta-analysis are largely conditional on operator experience, a factor we could not correct for and which is hardly quantifiable during trial conduction or later studies comparison. Center more at ease with different technologies can therefore produce largely different results,[56,57] contributing to the relevant heterogeneity we observed on the individual study-level regarding efficacy and procedural characteristics.

This individual-patient data meta-analysis presents at least two important strengths: first, we were able to gather data from all randomized controlled trials as well as a large number of observational studies comparing CB versus RF. Second, this is the first analysis on such a large patient pool where extensive corrections for

comorbidities was possible. Acknowledging the important differences in demographics and comorbidities between males and females at presentation, these corrections were essential to address bias in the results.

Several limitations of this individual-patient meta-analysis are however to consider. First, the studies collected were conducted on several years (2010 to 2018), a time span during which the indication for ablation of atrial fibrillation and therefore patient's selection and recruitment changed. Older patients more thoroughly selected depending on their LA size and earlier after their initial AF diagnoses were selected, therefore possibly modifying the enrolled population over time. However, when introduced in meta-regressions investigating the heterogeneity of the results for each study, the publication year of the study did not significantly impacted on the results. Second, some large observational studies[58,59] could not be integrated in the current results. However these studies showed high recurrence rate in women across all catheters used[12] or a similar tendency for a better efficacy of CB but not RF in males but not in females[20]. Integrating these results would most likely have confirmed our current observations. Third, while we could combine all data from the higher-quality RCTs currently available for the comparison of CB versus RF, all of these randomized studies were limited to 1-year follow-up and therefore did not contribute to the differences between CB and RF observed mainly in the longer follow-up. Our results therefore still need to be validated in RCTs following patients over at least 3-year follow-up. Fourth, the exact ablation strategies could not be precisely investigated in all studies. Therefore, some procedural differences such as variation in ablation strategies or operator experience have likely brought heterogeneity to our results. These between-studies variation however likely reflect the current situation in the different institution and bolster the generalizability of our observations. Fifth, we

could not account for every single important sex-specific covariate possibly influencing ablations outcomes such as hormone replacement therapy in older women or non-pulmonary vein foci. Finally, only the first ablation was considered, we therefore cannot comment on any further procedure.

In conclusion, women are at higher risk for ablation failure than men and cryoballoon ablation is associated with less long-term failures in men but not in women. As refinements in ablation technologies, such as the introduction of the second generation CB or RF contact-force catheters were associated with a significant decrease in recurrence rate in this analysis, adaptation of devices specifically for female patients could lead to higher success rate of AF ablation in women.

Funding

This individual patient-data meta-analysis has been funded by the University Hospital of Basel.

The individual studies provided the following funding sources: The datasets provided by S. Boveda, R. Providencia and Z. Khoueiry were founded by a grant from the ART (Association de Rythmologie Toulousaine), Clinique Pasteur, Toulouse, France.

Disclosures

Dr. Anselmino report an educational grant from Abbott and consultant for Biosense Webster. Dr. KV Davtyan serves as a proctor for Medtronic and Abbott. Dr. Kühne has received grants from the Swiss National Science Foundation, grants from the Swiss Heart Foundation, grants from Bayer, and grants from Pfizer-BMS. Dr. Passman report research support, consulting fees, and speaker fees from Medtronic, royalties from UpToDate, and research support from Pfizer and AliveCor

Dr. Kühne has received lecture/consulting fees from Bayer, Boehringer Ingelheim, Pfizer-BMS, Daiichi-Sankyo, Sanofi, Medtronic, Abbott, Biotronik, Boston Scientific, Biosense Webster, Zoll, AstraZeneca, and Novartis.

All other authors have nothing to disclose.

References

1. Lau DH, Nattel S, Kalman JM, Sanders P. Modifiable Risk Factors and Atrial Fibrillation. *Circulation*. 2017;136(6):583-596. doi:10.1161/CIRCULATIONAHA.116.023163
2. Chugh SS, Havmoeller R, Narayanan K, et al. Worldwide epidemiology of atrial fibrillation: A global burden of disease 2010 study. *Circulation*. 2014;129(8):837-847. doi:10.1161/CIRCULATIONAHA.113.005119
3. Gillis AM. Atrial Fibrillation and Ventricular Arrhythmias: Sex Differences in Electrophysiology, Epidemiology, Clinical Presentation, and Clinical Outcomes. *Circulation*. 2017;135(6):593-608. doi:10.1161/CIRCULATIONAHA.116.025312
4. Beck H, Curtis AB. Sex differences in outcomes of ablation of atrial fibrillation. *J Atr Fibrillation*. 2014;6(6):1024. http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L373194719%5Cnhttp://sfx.library.uu.nl/utrecht?sid=EMBASE&issn=19416911&id=doi:&atitle=Sex+differences+in+outcomes+of+ablation+of+atrial+fibrillation&stitle=J.+Atrial+Fibrillation&ti.
5. Ko D, Rahman F, Martins MAP, et al. Atrial fibrillation in women: Treatment. *Nat Rev Cardiol*. 2017;14(2):113-124. doi:10.1038/nrcardio.2016.171
6. Pérez-Castellano N, Fernández-Cavazos R, Moreno J, et al. The COR trial: A randomized study with continuous rhythm monitoring to compare the efficacy of cryoenergy and radiofrequency for pulmonary vein isolation. *Hear Rhythm*. 2014;11(1):8-13. doi:10.1016/j.hrthm.2013.10.014
7. Hunter RJ, Baker V, Finlay MC, et al. Point-by-Point Radiofrequency Ablation Versus the Cryoballoon or a Novel Combined Approach: A Randomized Trial Comparing 3 Methods of Pulmonary Vein Isolation for Paroxysmal Atrial

Fibrillation (The Cryo Versus RF Trial). *J Cardiovasc Electrophysiol*. 2015;26(12):1307-1314. doi:10.1111/jce.12846

8. Luik A, Radzewitz A, Kieser M, et al. Cryoballoon Versus Open Irrigated Radiofrequency Ablation in Patients With Paroxysmal Atrial FibrillationCLINICAL PERSPECTIVE. *Circulation*. 2015;132(14):1311-1319. doi:10.1161/CIRCULATIONAHA.115.016871
9. Schmidt B, Gunawardene M, Krieg D, et al. A prospective randomized single-center study on the risk of asymptomatic cerebral lesions comparing irrigated radiofrequency current ablation with the cryoballoon and the laser balloon. *J Cardiovasc Electrophysiol*. 2013;24(8):869-874. doi:10.1111/jce.12151
10. Kuck K-HH, Brugada J, Furnkranz A, et al. Cryoballoon or Radiofrequency Ablation for Paroxysmal Atrial Fibrillation. *N Engl J Med*. 2016;374(23):2235-2245. doi:10.1056/NEJMoa1602014
11. Kuck K-H, Brugada J, Fürnkranz A, et al. Impact of Female Sex on Clinical Outcomes in the FIRE AND ICE Trial of Catheter Ablation for Atrial Fibrillation. *Circ Arrhythm Electrophysiol*. 2018;11(5):e006204. doi:10.1161/CIRCEP.118.006204
12. Zylla MM, Brachmann J, Lewalter T, et al. Sex-related outcome of atrial fibrillation ablation: Insights from the German Ablation Registry. *Hear Rhythm*. 2016;13(9):1837-1844. doi:10.1016/j.hrthm.2016.06.005
13. Khoueiry Z, Albenque JP, Providencia R, et al. Outcomes after cryoablation vs. radiofrequency in patients with paroxysmal atrial fibrillation: impact of pulmonary veins anatomy. *Europace*. 2016;18(9):1343-1351. doi:10.1093/europace/euv419
14. Siddoway D, Friehling M, Voigt A, Saba S, Jain S. Improved resource

utilization with similar efficacy during early adoption of cryoballoon pulmonary vein isolation as compared to radiofrequency ablation for paroxysmal atrial fibrillation. *J Atr Fibrillation*. 2015;7(5):15-19. http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L603772779.

15. Knecht S, Von Felten S, Conen D, et al. Propensity-score-matched comparison of cryoballoon versus radiofrequency catheter ablation of paroxysmal atrial fibrillation. *Eur Heart J*. 2013;34:753. http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L71260313.
16. Matta M, Anselmino M, Ferraris F, Scaglione M, Gaita F. Cryoballoon vs. radiofrequency contact force ablation for paroxysmal atrial fibrillation: a propensity score analysis. *J Cardiovasc Med*. 2018;19(4):141-147. doi:10.2459/jcm.0000000000000633
17. Providencia R, Defaye P, Lambiase PD, et al. Results from a multicentre comparison of cryoballoon vs. radiofrequency ablation for paroxysmal atrial fibrillation: is cryoablation more reproducible? *Europace*. 2017;19(1):48-57. doi:10.1093/europace/euw080
18. Squara F, Zhao A, Marijon E, et al. Comparison between radiofrequency with contact force-sensing and second-generation cryoballoon for paroxysmal atrial fibrillation catheter ablation: a multicentre European evaluation. *Europace*. 2015;17(5):718-724. doi:10.1093/europace/euv060
19. Wasserlauf J, Passman R, Giedrimas E, et al. Cryoballoon versus radiofrequency catheter ablation for atrial fibrillation. *J Cardiovasc Electrophysiol*. 2014;25(5):568-569. doi:10.1111/jce.2014.25

20. Mörtsell D, Arbelo E, Dagres N, et al. Cryoballoon vs. radiofrequency ablation for atrial fibrillation: A study of outcome and safety based on the ESC-EHRA atrial fibrillation ablation long-term registry and the Swedish catheter ablation registry. *Europace*. 2019;21(4):581-589. doi:10.1093/europace/euy239
21. Higgins JPT, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343(7829):1-9. doi:10.1136/bmj.d5928
22. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement (Chinese edition). *J Chinese Integr Med*. 2009;7(9):889-896. doi:10.3736/jcim20090918
23. Agresti A, Caffo B. Simple and Effective Confidence Intervals for Proportions and Differences of Proportions Result from Adding Two Successes and Two Failures. *Am Stat*. 2000;54(4):280-288.
24. Beck CW, Bliwise NG. Interactions are critical. *CBE Life Sci Educ*. 2014;13(3):371-372. doi:10.1187/cbe.14-05-0086
25. Davtyan K, Shatakhtsyan V, Poghosyan H, et al. Radiofrequency versus Cryoballoon Ablation of Atrial Fibrillation: An Evaluation Using ECG, Holter Monitoring, and Implantable Loop Recorders to Monitor Absolute and Clinical Effectiveness. *Biomed Res Int*. 2018;2018:1-7. doi:10.1155/2018/3629384
26. Luik A, Radzewitz A, Kieser M, et al. Cryoballoon Versus Open Irrigated Radiofrequency Ablation in Patients With Paroxysmal Atrial Fibrillation: The Prospective, Randomized, Controlled, Noninferiority FreezeAF Study. *Circulation*. 2015;132(14):1311-1319. doi:10.1161/circulationaha.115.016871
27. Tsang W, Alter DA, Wijeysundera HC, Zhang T, Ko DT. The impact of cardiovascular disease prevalence on women's enrollment in landmark

randomized cardiovascular trials: a systematic review. *J Gen Intern Med*. 2012;27(1):93-98. doi:10.1007/s11606-011-1768-8

28. Harris DJ, Douglas PS. Enrollment of women in cardiovascular clinical trials funded by the National Heart, Lung, and Blood Institute. *N Engl J Med*. 2000;343(7):475-480. doi:10.1056/NEJM200008173430706
29. Clayton JA, Arnegard ME. Taking cardiology clinical trials to the next level: A call to action. *Clin Cardiol*. 2018;41(2):179-184. doi:10.1002/clc.22907
30. Kragholm K, Halim SA, Yang Q, et al. Sex-Stratified Trends in Enrollment, Patient Characteristics, Treatment, and Outcomes Among Non–ST-Segment Elevation Acute Coronary Syndrome Patients. *Circ Cardiovasc Qual Outcomes*. 2015;8(4):357-367. doi:10.1161/CIRCOUTCOMES.114.001615
31. Nolan MR, Nguyen T-L. Analysis and Reporting of Sex Differences in Phase III Medical Device Clinical Trials—How Are We Doing? *J Women's Heal*. 2013;22(5):399-401. doi:10.1089/jwh.2013.4400
32. Dhruva SS, Redberg RF. Clinical trial enrollment and progress in women's health. *JAMA*. 2011;305(12):1197; author reply 1197-8. doi:10.1001/jama.2011.347
33. Dhruva SS, Redberg RF. The Need for Sex-Specific Data Prior to Food and Drug Adminstration Approval. *J Am Coll Cardiol*. 2010;55(3):261. doi:10.1016/j.jacc.2009.08.053
34. Regitz-Zagrosek V, Seeland U. Sex and gender differences in clinical medicine. *Handb Exp Pharmacol*. 2012;(214):3-22. doi:10.1007/978-3-642-30726-3_1
35. Blauwet LA, Hayes SN, Mcmanus D, Redberg RF, Walsh MN. Low rate of sex-specific result reporting in cardiovascular trials. *Mayo Clin Proc*.

2007;82(2):166-170. doi:10.4065/82.2.166

36. Dhruva SS, Redberg RF. Variations Between Clinical Trial Participants and Medicare Beneficiaries in Evidence Used for Medicare National Coverage Decisions. *Arch Intern Med*. 2008;168(2):136. doi:10.1001/archinternmed.2007.56

37. du Fay de Lavallaz J, Clerc O, Pudenz C, Illigens B, Kühne M. Sex-specific efficacy and safety of cryoballoon versus radiofrequency ablation for atrial fibrillation: A systematic review and meta-analysis. *J Cardiovasc Electrophysiol*. 2019.

38. Dhruva SS, Bero LA, Redberg RF. Gender bias in studies for Food and Drug Administration premarket approval of cardiovascular devices. *Circ Cardiovasc Qual Outcomes*. 2011;4(2):165-171. doi:10.1161/CIRCOUTCOMES.110.958215

39. Patel D, Mohanty P, Di Biase L, et al. Outcomes and complications of catheter ablation for atrial fibrillation in females. *Hear Rhythm*. 2010;7(2):167-172. doi:10.1016/j.hrthm.2009.10.025

40. Shehab A, Al-Dabbagh B, AlHabib KF, et al. Gender Disparities in the Presentation, Management and Outcomes of Acute Coronary Syndrome Patients: Data from the 2nd Gulf Registry of Acute Coronary Events (Gulf RACE-2). *PLoS One*. 2013;8(2). doi:10.1371/journal.pone.0055508

41. Hernandez AF, Fonarow GC, Liang L, et al. Sex and racial differences in the use of implantable cardioverter-defibrillators among patients hospitalized with heart failure. *JAMA*. 2007;298(13):1525-1532. doi:10.1001/jama.298.13.1525

42. ZHANG X-DD, Tan H-WW, Gu JJ-NN, et al. Efficacy and safety of catheter ablation for long-standing persistent atrial fibrillation in women. *PACE - Pacing*

Clin Electrophysiol. 2013;36(10):1236-1244. doi:10.1111/pace.12212

43. Ko D, Rahman F, Schnabel RB, Yin X, Benjamin EJ, Christophersen IE. Atrial fibrillation in women: Epidemiology, pathophysiology, presentation, and prognosis. *Nat Rev Cardiol*. 2016;13(6):321-332. doi:10.1038/nrcardio.2016.45

44. Erinjeri JP, Clark TWI. Cryoablation: Mechanism of action and devices. *J Vasc Interv Radiol*. 2010;21(SUPPL. 8):S187-S191. doi:10.1016/j.jvir.2009.12.403

45. Baust JG, Gage AA. The molecular basis of cryosurgery. *BJU Int*. 2005;95(9):1187-1191. doi:10.1111/j.1464-410x.2005.05502.x

46. Ghavidel AA, Javadpour H, Shafiee M, Tabatabaie MB, Raiesi K, Hosseini S. Cryoablation for surgical treatment of chronic atrial fibrillation combined with mitral valve surgery: a clinical observation. *Eur J Cardio-thoracic Surg*. 2008;33(6):1043-1048. doi:10.1016/j.ejcts.2008.03.019

47. Hirao T, Nitta J, Adachi A, Takahashi Y, Goya M, Hirao K. First confirmation of histologic changes in the human heart after cryoballoon ablation. *Hear Case Reports*. 2019;5(2):93-96. doi:10.1016/j.hrcr.2018.10.012

48. Hartl S, Dorwarth U, Bunz B, et al. Lessons from individualized cryoballoon sizing. Is there a role for the small balloon? *J Cardiol*. 2017;70(4):374-381. doi:10.1016/j.jjcc.2016.12.016

49. Sultan A, Lüker J, Andresen D, et al. Predictors of Atrial Fibrillation Recurrence after Catheter Ablation: Data from the German Ablation Registry. *Sci Rep*. 2017;7(1):1-7. doi:10.1038/s41598-017-16938-6

50. Shin SH, Park MY, Oh WJ, et al. Left Atrial Volume Is a Predictor of Atrial Fibrillation Recurrence After Catheter Ablation. *J Am Soc Echocardiogr*. 2008;21(6):697-702. doi:10.1016/j.echo.2007.10.022

51. Lee SH, Tai CT, Hsieh MH, et al. Predictors of early and late recurrence of

atrial fibrillation after catheter ablation of paroxysmal atrial fibrillation. *J Interv Card Electrophysiol*. 2004;10(3):221-226. doi:10.1023/B:JICE.0000026915.02503.92

52. Ullah W, Schilling RJ, Wong T. Contact force and atrial fibrillation ablation. *J Atr Fibrillation*. 2016;8(5):74-80.

53. Giovanni G Di, Wauters K, Chierchia GB, et al. One-year follow-up after single procedure cryoballoon ablation: A comparison between the first and second generation balloon. *J Cardiovasc Electrophysiol*. 2014;25(8):834-839. doi:10.1111/jce.12409

54. Hoyt H, Bhonsale A, Chilukuri K, et al. Complications arising from catheter ablation of atrial fibrillation: Temporal trends and predictors. *Hear Rhythm*. 2011;8(12):1869-1874. doi:10.1016/j.hrthm.2011.07.025

55. Spragg DD, Dalal D, Cheema A, et al. Complications of catheter ablation for atrial fibrillation: Incidence and predictors. *J Cardiovasc Electrophysiol*. 2008;19(6):627-631. doi:10.1111/j.1540-8167.2008.01181.x

56. Dahme T. Cryoballoon ablation in high versus low volume centers – Does experience make a difference? *Int J Cardiol*. 2018;272:227-228. doi:10.1016/j.ijcard.2018.08.042

57. Su W, Aryana A, Passman R, et al. Cryoballoon Best Practices II: Practical guide to procedural monitoring and dosing during atrial fibrillation ablation from the perspective of experienced users. *Hear Rhythm*. 2018;15(9):1348-1355. doi:10.1016/j.hrthm.2018.04.021

58. Mortsell D, Arbelo E, Dagres N, et al. Cryoballoon vs. radiofrequency ablation for atrial fibrillation: a study of outcome and safety based on the ESC-EHRA atrial fibrillation ablation long-term registry and the Swedish catheter ablation

registry. *Eur Eur pacing, arrhythmias, Card Electrophysiol J Work groups Card pacing, arrhythmias, Card Cell Electrophysiol Eur Soc Cardiol*. October 2018. doi:10.1093/europace/euy239

59. Schmidt M, Dorwarth U, Andresen D, et al. RF versus cryoballoon in atrial fibrillation ablation: Outcome data from the German ablation registry I. *Eur Heart J*. 2013;34:650-651. http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L71259931.

Figures&Tables - Sex-specific efficacy and safety of cryoballoon versus radiofrequency ablation for atrial fibrillation: An individual patient-data meta-analysis

Table 1 :

	Merged dataset				Fire&Ice			
	All patients	**Male**	**Female**	**P value**	**All patients**	**Male**	**Female**	**P value**
Number of patients	6069	4383	1686		750	457	293	
Age-years (median [IQR])	62.0 [54.7, 68.0]	61.0 [53.5, 67.0]	65.0 [58.3, 70.2]	<0.001	61.0 [54.0, 68.0]	58.0 [51.0, 64.0]	65.0 [59.0, 70.0]	<0.001
Sex (females)	1686 (28)	0 (0)	1686 (100)	<0.001	293 (39)	0 (0)	293 (100)	<0.001
Patients characteristics								
AF type				<0.001				-
paroxysmal	4423 (73)	3086 (71)	1337 (79)		750 (100)	457 (100)	293 (100)	
persistent	1213 (20)	957 (22)	256 (15)		0 (0)	0 (0)	0 (0)	
longstanding persistent	253 (4)	202 (5)	51 (3)		0 (0)	0 (0)	0 (0)	
other[1]	168 (3)	129 (3)	39 (2)		0 (0)	0 (0)	0 (0)	
Duration of AF	4.7 (5.0)	4.7 (5.0)	4.6 (4.8)	0.480	4.6 (5.2)	4.6 (5.2)	4.6 (5.2)	0.987
BMI	26.6 (4.6)	26.8 (4.3)	26.2 (5.5)	<0.001	27.9 (4.6)	27.7 (4.0)	28.2 (5.5)	0.157
hypertension	2459 (47)	1707 (45)	752 (51)	0.001	436 (58)	253 (55)	183 (62)	0.065
DM	526 (10)	390 (10)	136 (9)	0.199	60 (8)	39 (9)	21 (7)	0.602
CHF	413 (8)	304 (9)	109 (8)	0.466	209 (28)	110 (24)	99 (34)	0.005
stroke	360 (8)	233 (7)	127 (9)	0.004	28 (4)	16 (4)	12 (4)	0.825
vascular disease	425 (10)	354 (12)	71 (6)	<0.001	48 (6)	33 (7)	15 (5)	0.320
Anti-arrhythmic drug Type I	2397 (47)	1732 (47)	665 (48)	0.308	-	-	-	-
Anti-arrhythmic drug Type III	1885 (37)	1360 (37)	525 (38)	0.380	-	-	-	-
Anti-arrhythmic drug Type II	2081 (41)	1477 (40)	604 (42)	0.131	-	-	-	-
any anti-arrhythmic drug	4153 (73)	2964 (72)	1189 (75)	0.054	289 (39)	175 (38)	114 (39)	0.927

[1] left atrial tachycardia or flutter

	Merged dataset				Fire&Ice			
	All patients	**Male**	**Female**	**P value**	**All patients**	**Male**	**Female**	**P value**
Measure of LA				0.004				<0.001
normal	2088 (40)	1538 (41)	550 (38)		204 (35)	137 (39)	67 (30)	
mildly abnormal	973 (19)	684 (18)	289 (20)		83 (14)	49 (14)	34 (15)	
moderately abnormal	1073 (21)	796 (21)	277 (19)		251 (43)	160 (45)	91 (40)	
severely abnormal	1096 (21)	751 (20)	345 (24)		40 (7)	7 (2)	33 (15)	
LVEF	60.2 (7.9)	59.7 (8.1)	61.6 (7.1)	<0.001	62.6 (6.6)	62.0 (6.8)	63.7 (6.3)	0.004
LVSD	268 (5)	219 (5)	49 (3)	0.001	13 (3)	11 (4)	2 (1)	0.142
Catheter data								
Catheter type: RF	3937 (65)	2885 (66)	1052 (62)	0.013	381 (51)	241 (53)	140 (48)	0.212
Catheter details				0.213				0.161
Cryoballoon 1st generation	1012 (17)	708 (17)	304 (19)		90 (12)	52 (11)	38 (13)	
Cryoballoon 2nd generation	962 (16)	688 (16)	274 (17)		279 (37)	164 (36)	115 (39)	
RF contact force	1282 (22)	938 (22)	344 (21)		94 (13)	67 (15)	27 (9)	
RF irrigated no contact force	2155 (36)	1572 (37)	583 (36)		287 (38)	174 (38)	113 (39)	
RF not irrigated	500 (8)	375 (9)	125 (8)		0 (0)	0 (0)	0 (0)	

Table 1 – patients characteristics in the merged data set and in the Fire&Ice study. AF = Atrial Fibrillation, BMI=Body mass index, CB=Cryoballoon, CHF = Congestive Heart Failure, DM = Diabetes Mellitus, LA=Left atrium, LVEF = Left ventricular ejection fraction, LVSD =Left Ventricular Systolic Dysfunction, RF= radiofrequency.

Table 2

	Merged data set				Fire&Ice dataset			
	estimate	2.5 %	97.5 %	P value	estimate	2.5 %	97.5 %	P value
Sex: Female	1.120	1.013	1.238	0.026	1.302	1.003	1.689	0.047
Catheter: CB	0.899	0.812	0.995	0.041	0.918	0.724	1.165	0.484
age	1.002	0.997	1.007	0.362	0.999	0.985	1.014	0.890
LA measure: Mild abnormal	1.098	0.947	1.273	0.212	1.074	0.727	1.589	0.719
LA measure: Moderately abnormal	1.038	0.904	1.192	0.594	0.788	0.583	1.066	0.121
LA measure: Severely abnormal	1.217	1.029	1.441	0.023	1.140	0.704	1.845	0.593
LVSD	0.932	0.732	1.185	0.564	0.512	0.125	2.088	0.350
Vascular disease	0.990	0.842	1.164	0.901	0.732	0.424	1.263	0.262
Hypertension	1.060	0.957	1.174	0.261	1.079	0.829	1.406	0.572
DM	1.016	0.862	1.199	0.848	1.213	0.785	1.872	0.384
AF duration	1.026	1.016	1.036	<0.001	1.021	1.000	1.042	0.048
CHF	1.012	0.838	1.223	0.898	1.214	0.941	1.565	0.135
BMI	1.009	0.990	1.028	0.337	0.980	0.952	1.009	0.173
Stroke	1.029	0.828	1.279	0.790	0.842	0.443	1.599	0.599
AF type: Paroxysmal	0.517	0.462	0.579	<0.001	-	-	-	-

Table 2 – Cox proportional hazard model to predict the recurrence of the efficacy endpoint in the overall population. As all patients from the F&I study had paroxysmal AF, this variable was not used in their model. AF = atrial Fibrillation, BMI = body mass index CB = Cryoballoon, CHF = Congestive heart failure, DM = Diabetes Mellitus, LA = left atrium, LVSD = Left ventricular systolic dysfunction.

Table 3

	Merged data set				Fire&Ice			
	All patients	**Male**	**Female**	**P value**	**All patients**	**Male**	**Female**	**P value**
Number of patients	6069	4383	1686		750	457	293	
At least one complication	367 (6)	236 (5)	131 (8)	0.001	63 (8)	31 (7)	32 (11)	0.058
Groin complication	123 (2)	74 (2)	49 (3)	0.004	23 (3)	9 (2)	14 (5)	0.048
Oesophageal fistula	3 (0)	2 (0)	1 (0)	1.000	0 (0)	0 (0)	0 (0)	-
pericardial effusion	60 (1)	44 (1)	16 (1)	1.000	0 (0)	0 (0)	0 (0)	-
phrenic nerve palsy	73 (1)	43 (1)	30 (2)	0.017	10 (1)	5 (1)	5 (2)	0.524
pulmonary vein stenosis	5 (0)	5 (0)	0 (0)	0.331	0 (0)	0 (0)	0 (0)	-
stroke or TIA	22 (0)	19 (0)	3 (0)	0.159	4 (1)	2 (0)	2 (1)	0.646
tamponade	22 (0)	11 (0)	11 (1)	0.029	6 (1)	2 (0)	4 (1)	0.216
	All patients	**CB**	**RF**	**P value**	**All patients**	**CB**	**RF**	**P value**
Number of patients	6069	2132	3937		750	369	381	
At least one complication	367 (6)	149 (7)	218 (6)	0.028	63 (8)	28 (8)	35 (9)	0.511
Groin complication	123 (2)	46 (2)	77 (2)	0.633	23 (3)	7 (2)	16 (4)	0.089
Oesophageal fistula	3 (0)	0 (0)	3 (0)	0.556	0 (0)	0 (0)	0 (0)	-
pericardial effusion	60 (1)	8 (0)	52 (1)	<0.001	0 (0)	0 (0)	0 (0)	-
phrenic nerve palsy	73 (1)	59 (3)	14 (0)	<0.001	10 (1)	10 (3)	0 (0)	0.001
pulmonary vein stenosis	5 (0)	3 (0)	2 (0)	0.352	0 (0)	0 (0)	0 (0)	-
stroke or TIA	22 (0)	6 (0)	16 (0)	0.509	63 (8)	28 (8)	35 (9)	0.511
tamponade	22 (0)	10 (0)	12 (0)	0.371	23 (3)	7 (2)	16 (4)	0.089
Groin complication	123 (2)	46 (2)	77 (2)	0.633	23 (3)	7 (2)	16 (4)	0.089

Table 3 – Occurrence of complications by sex and catheter type in the merged and Fire&Ice data sets

Table 4:

	Merged data set				Fire&Ice			
	All patients	**Male**	**Female**	**P value**	**All patients**	**Male**	**Female**	**P value**
Number of patients	5558	4013	1545		750	457	293	
Procedure time (median [IQR])	155.0 [120.0, 204.0]	154.0 [120.0, 204.0]	160.0 [120.0, 204.0]	0.450	120.0 [100.0, 156.5]	120.0 [100.0, 157.0]	120.0 [100.0, 155.0]	0.772
Number of patients	4975	3559	1416		750	457	293	
Fluoroscopy time (median [IQR])	27.0 [18.0, 38.8]	26.5 [18.0, 39.0]	27.0 [19.0, 38.0]	0.226	16.0 [10.0, 24.0]	16.0 [10.0, 24.0]	16.0 [10.0, 24.0]	0.333
	All patients	**CB**	**RF**	**P value**	**All patients**	**CB**	**RF**	**P value**
Number of patients	5558	1837	3721		750	369	381	
Procedure time (median [IQR])	155.0 [120.0, 204.0]	140.0 [110.0, 176.0]	170.0 [120.0, 220.0]	<0.001	120.0 [100.0, 156.5]	120.0 [97.0, 146.0]	130.0 [100.0, 175.0]	<0.001
Number of patients	4975	1737	3238		750	369	381	
Fluoroscopy time (median [IQR])	27.0 [18.0, 38.8]	25.5 [18.0, 36.1]	27.0 [18.0, 40.0]	0.033	16.0 [10.0, 24.0]	19.0 [14.0, 26.0]	13.0 [8.0, 21.0]	<0.001

Table 4 - Description of the procedural outcomes by sex and catheter type in the merged and Fire&Ice data sets. IQR = Interquartile range.

Figure 1

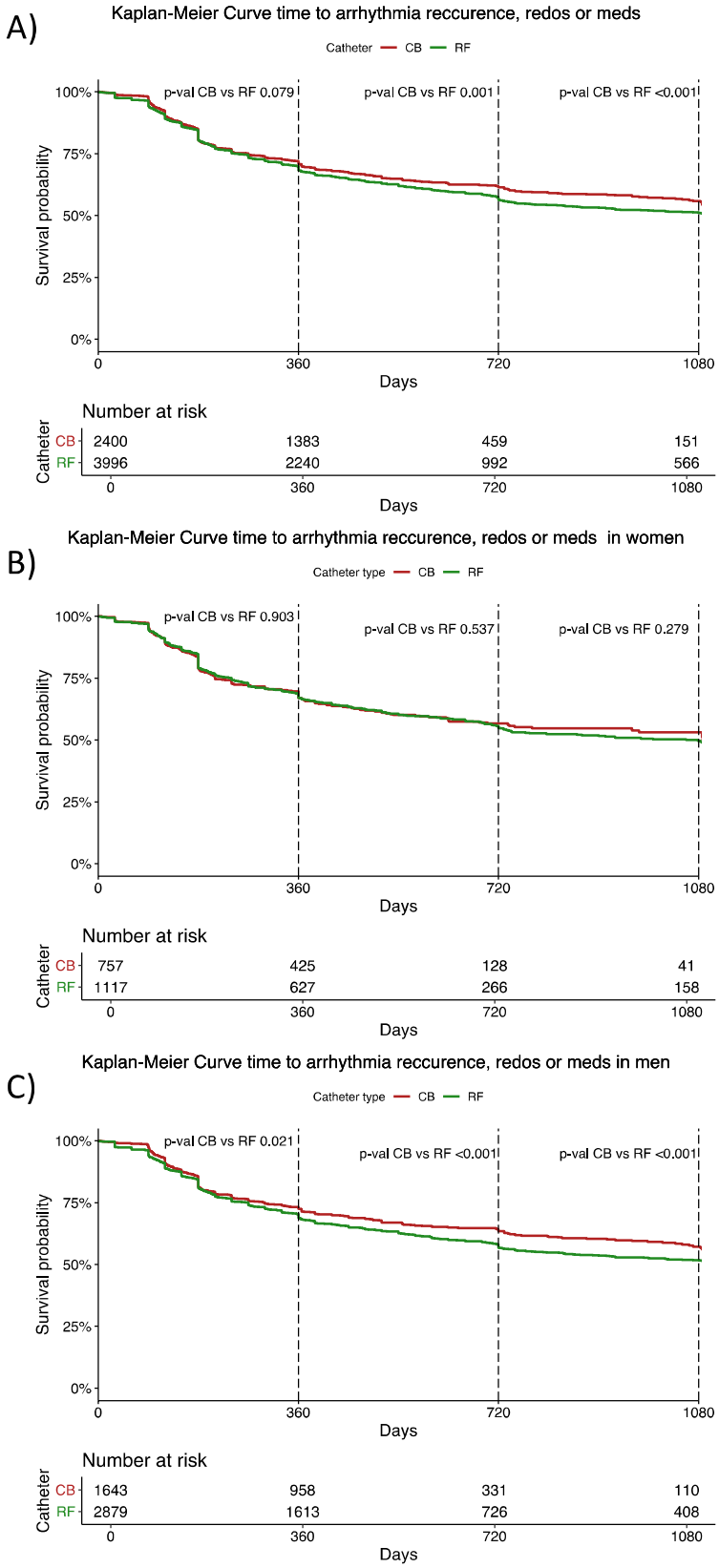


Figure 1 – Kaplan Meier representing the event-free survival in the merged data and F&I dataset for A) the overall cohort, B) women and C) men separately. CB = Cryoballoon, RF = Radiofrequency catheters.

Figure 2

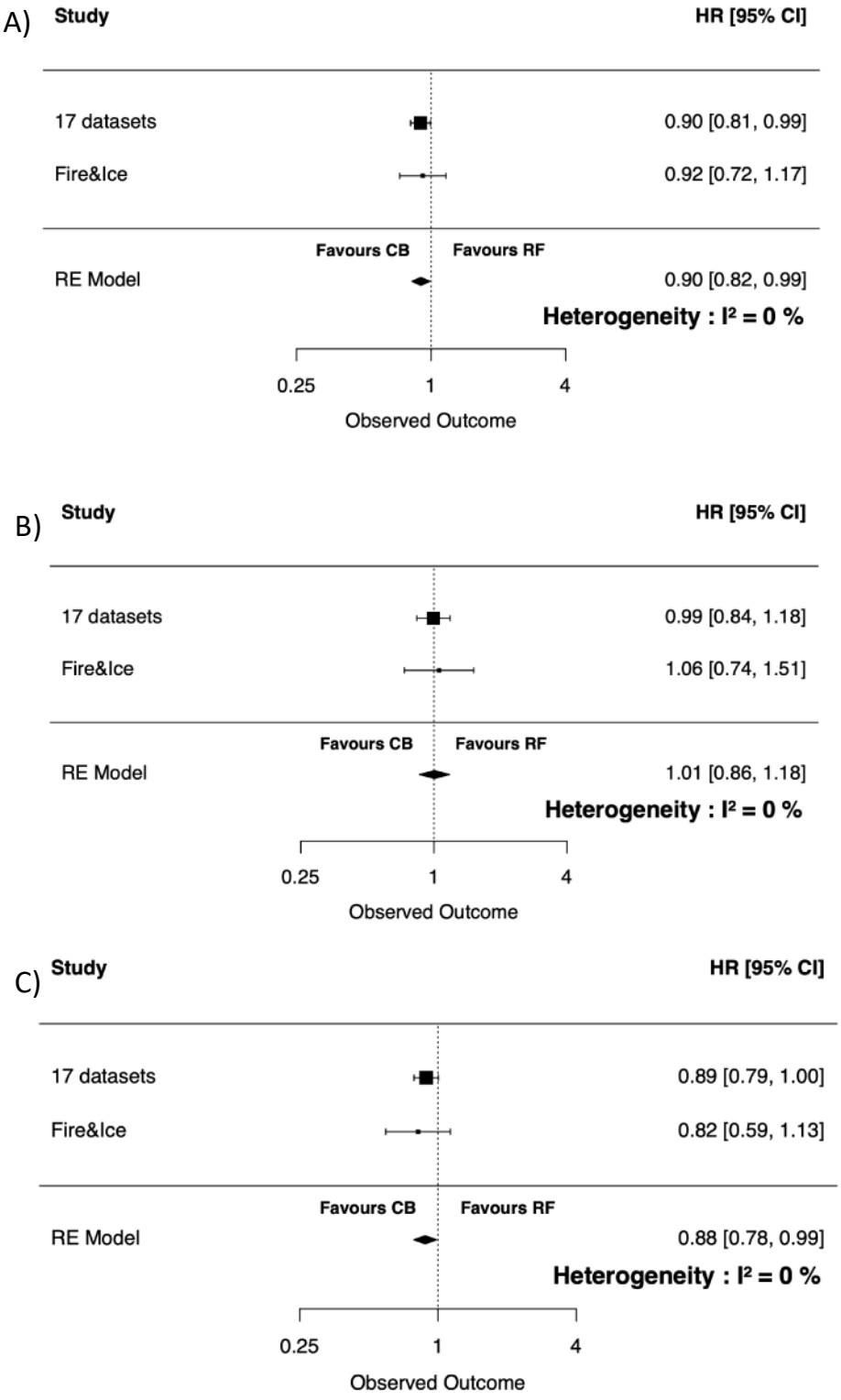


Figure 2 – Pooled estimates of the mixed-effect cox proportional hazard models by catheter type for arrhythmia recurrence, redos and re-start of medications up to three years follow-up in the merged and Fire&Ice datasets for A) the overall cohort, B) women and C) men.

Figure 3

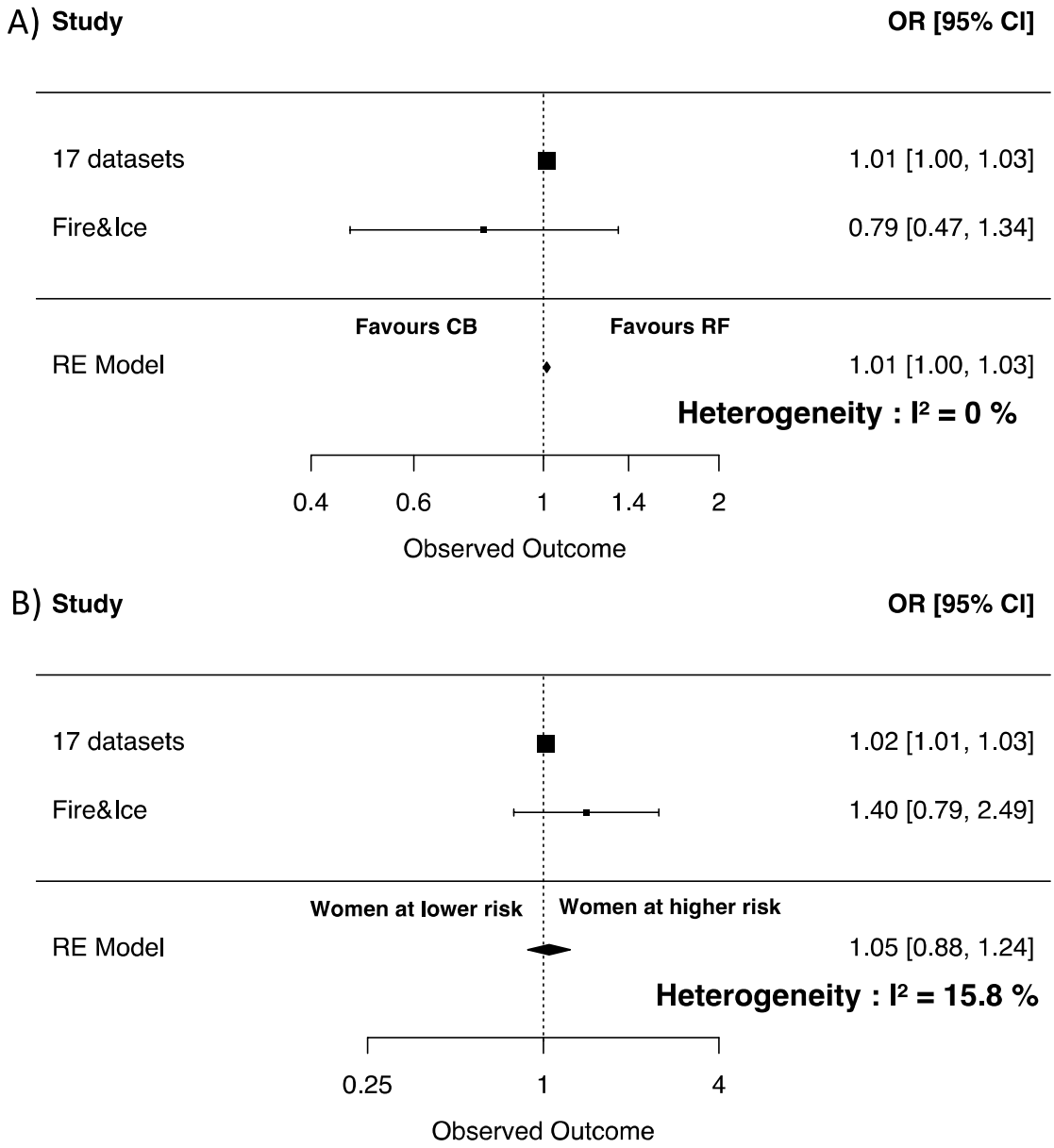


Figure 3 – Pooled estimates of the mixed-effect logistic model for periprocedural complications in the merged and Fire&Ice datasets by A) the catheter type and B) the patient's sex.

Figure 4

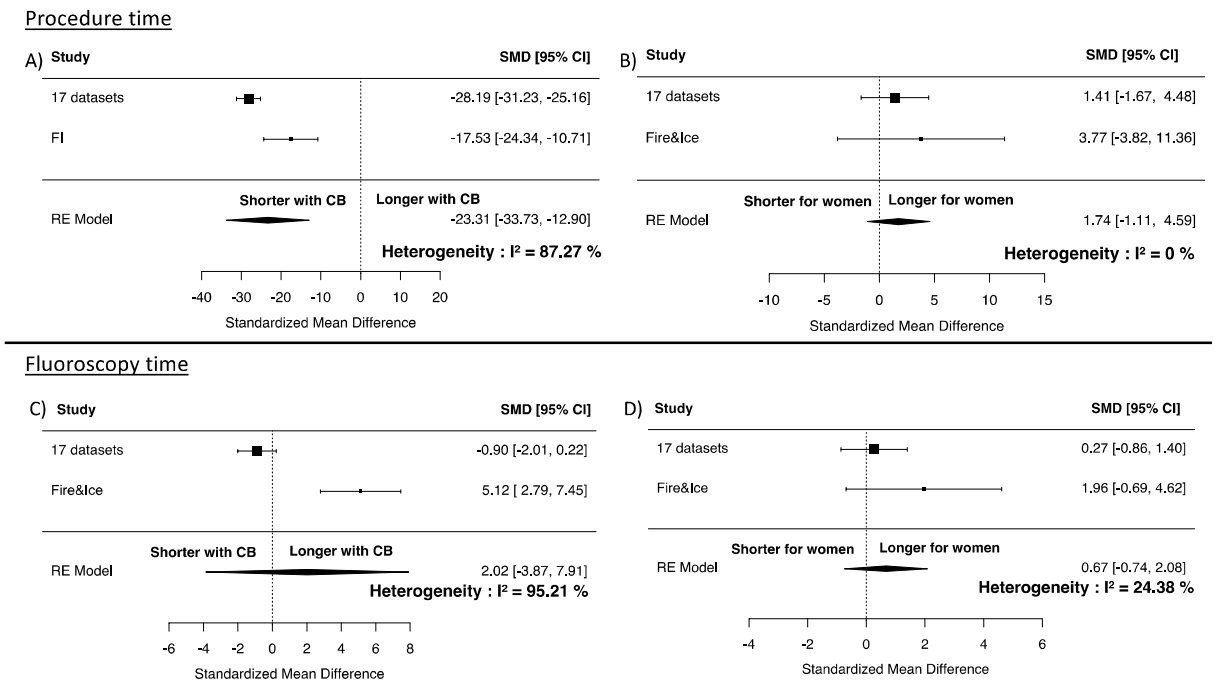


Figure 4 – Pooled estimates from multivariable linear models for the procedural endpoints by catheter types and by sex. CB = Cryoballoon.

Supplemental material - Sex-specific efficacy and safety of cryoballoon versus radiofrequency ablation for atrial fibrillation: : An individual patient-data meta-analysis

Table of content

Supplemental methods:

Search criteria

A comprehensive search was conducted in PubMed, MEDLINE, the Cochrane database and Embase. The search design was conducted with the assistance of a research librarian. The terms “atrial fibrillation”, “pulmonary vein ablation”, “radiofrequency” and “cryo*” were used for the search. The study registry Clinicaltrial.gov was manually searched using the same terms. No limitations of dates or languages were applied. The search was conducted on the 28th of March 2018 and repeated on the 15th of March 2019 to account for the past year.

Pubmed

(((atrial fibrillation[MeSH Terms] OR atrial fibrillation[Title/Abstract] OR pulmonary vein isolation[Title/Abstract]))) AND ((cryosurgery[MeSH Terms] OR cryoballoon[Title/Abstract] OR cryoablation[Title/Abstract] OR cryothermal[Title/Abstract]) AND (catheter ablation, radiofrequency[MeSH Terms] OR radiofrequenc*[Title/Abstract])) NOT (animals [mh] NOT humans [mh])

Embase and Medline

('atrial fibrillation'/exp/mj OR 'atrial fibrillation':ab,ti OR 'pulmonary vein isolation':ab,ti) AND ('cryoablation'/exp/mj OR 'cryoablation':ab,ti OR 'cryosurgery'/exp/mj OR 'cryosurgery':ab,ti OR 'cryoballoon':ab,ti OR 'cryothermal':ab,ti) AND ('catheter ablation'/exp/mj OR 'radiofrequency ablation':ab,ti) AND ([embase]/lim NOT ([embase]/lim AND [medline]/lim) OR ([medline]/lim NOT ([embase]/lim AND [medline]/lim) NOT ([embase classic]/lim AND [medline]/lim)))

Cochrane

- #1 MeSH descriptor: [Cryosurgery] explode all trees
- #2 MeSH descriptor: [Atrial Fibrillation] explode all trees
- #3 MeSH descriptor: [Catheter Ablation] explode all trees
- #4 atrial fibrillation:ti,ab,kw (Word variations have been searched)
- #5 "pulmonary vein isolation":ti,ab,kw (Word variations have been searched)
- #6 cryoballoon:ti,ab,kw (Word variations have been searched)
- #7 cryoablation:ti,ab,kw (Word variations have been searched)
- #8 croythermal:ti,ab,kw (Word variations have been searched)
- #9 radiofrequency:ti,ab,kw (Word variations have been searched)
- #10 (#2 or #4 or #5) and (#1 or #6 or #7 or #8) and (#3 or #9)

Study selection

Studies that met the following pre-specified criteria were included : 1) RCTs or prospective observational studies, 2) with at least 40 patients per group (CB versus RF), 3) with patients undergoing their first ablation, 4) using a first- or second-generation CB and irrigated or contact-force guided RF catheters, 5) investigating an efficacy outcome of time-to-failure (defined as recurrence of atrial arrhythmia, re-ablation and re-start of anti-arrhythmic medication) and a safety outcome (defined as the percentage of recorded complications, not further specified) and 6) following patients for at least 12 months. When it was unclear if the datasets behind the publications retrieved by the search would met the requirements, authors were contacted. As recommended by the literature for systematic reviews of rapidly evolving technologies, we did not focus exclusively on RCTs but also included observational studies.[1]

Selection of studies was performed by two independent readers (JdFdL, University Hospital of Basel, CP, University Hospital of Basel). Disagreements were solved by discussion.

Details of the individual studies were extracted in a dedicated RedCap database (Vanderbilt University Medical Center) for quality assessment.

Call for data

For each identified study, either the first, last or corresponding author was contacted per e-mail depending on the contact information available. A first call-for-data was sent 1st May 2018 and two more times thereafter in two-month intervals. Co-authorships but no monetary incentives were offered for participation. Supplemental table 2 summarizes the contacted authors whose datasets were deemed suitable for the current project after verification of the inclusion criteria, the datasets sought and the answers received. Difficulties linked to original patient consents not explicitly allowing for data sharing were found for one of the study, the Fire&Ice study. An agreement was found with the industrial partner managing the dataset (Medtronic) that the analysis would be programmed by the investigators of the current project (JdFdL) using simulated data displaying the exact same structure than the real dataset, that the final analysis would be run using this code at the Headquarter of Medtronic in Minneapolis by the statistician responsible of the dataset (FK) and that all estimates would then be shared with the investigators of the current project. The code was written by JdFdL, independently of any influence coming from Medtronic, in

order to store diverse baseline tables, outcome tables, time-event tables for the Kaplan Meier and model estimates with relevant confidence intervals in separate files, which respected the ethical limitations and could not be linked with individual patient data but allowed the dataset to be integrated in the current project. As several corrections were undertaken during the analysis, the patients of the Fire&Ice dataset were analyzed as per-protocol.

Individual patient data collection and datasets merging

All authors who accepted to participate were provided the following list of required variables.

- Baseline characteristics: Age, sex, AF classification (paroxysmal, persistent, permanent), years since first AF diagnosis, BMI, left atrial diameter and/or volume, left ventricular ejection fraction, left ventricular hypertrophy, hypertension, diabetes type 2 , coronary artery disease, CHA_2DS_2-VASc Score, previous direct current cardioversion, previous stroke, previous TIA, previous myocardial infarction, use of antiarrhythmic drug, patients without any cardiac medications
- Procedural characteristics: Type of radiofrequency catheter (contact force versus non-contact force) or cryoballoon (1^{st} or 2^{nd} generation), total duration of the procedure, total fluoroscopy time
- Center-specific data: monitoring system to detect recurrence (transtelephonic EGC, Holter)
- Data collected during the blanking period: repeated intervention, direct current cardioversion, use of antiarrhythmic drugs.
- Follow-up data: time to last follow-up contact
- Safety data: death, cause of death, rehospitalisations for cardiovascular causes, stroke or transient ischemic attack, cardiac arrhythmias (apart from the recurrence of AF) related to the intervention, cardiac tamponade or pericardial effusion, groin-site complications (Vascular pseudoaneurysm, arteriovenous fistula, device-related infection, hematoma, puncture-site hemorrhage, groin pain), phrenic-nerve injury (transient or persistent >3 months), atrio-esophageal fistulae, pulmonary-vein stenosis, procedure-related deaths
- Efficacy data: Time-to-recurrence of recurrent atrial arrhythmia, of antiarrhythmic drug treatment or of repeated ablation

While not all variables were available in all data sets, none of the studies was lacking essential variables hindering their participation.
All datasets were sent to the University Hospital of Basel where they were stored on a protected server. Their variables were renamed and releveled in order to allow for merging to one single homogenous dataset. All data wrangling was conducted using the Statistical Software “R” (R Foundation for Statistical Computing, Vienna, Austria). Correct merging was controlled using logical cleaning rules and estimates obtained in the individual datasets were compared with the original publications to verify plausibility.

Variable definitions:

- Left ventricular systolic dysfunction (LVSD) : defined as left ventricular ejection fraction <45%
- Left atrial dilation : As not all studies provided the same unit for left atrial dilation (diameter, volume or body-size indexed volume) a table by Lang et al[2] (see page 9) was used to divided the LA dilation of each patient in four categories : Normal, mild dilation, moderate dilation and severe dilation.
- AF type: AF was classified as paroxysmal, persistent, long-standing persistent or others.

Criteria to evaluate study quality

General criteria

- Explicit definition of :
    - Objectives
    - Inclusion criteria
    - Exclusion crtieria

Randomized controlled trials (Cochrane Collaboration tool for RCTs) :

- Random sequence generation
- Allocation concealment
- Single blinding (participants)
- Double blinding (operators and participants)
- Blinding of outcome assessment
- Complete outcome data
- No selective reporting

Observational studies (modified New Castle Ottawa Scale):

- Selection:
    - Representativeness of the exposed cohorts : +1
    - Case definition adequate : +1
    - Consecutive series of cases +1
    - Demonstration that the outcome of interest was not present at the start of the study (confirmation of proper ablation): +1
- Outcome:
    - Assessment of outcome (independent blind) : +1
    - Adequate monitoring (Holter, loop recorder) : +1
    - Adequacy and completeness of follow-up: +1

Statistical analysis

Missing data were imputed in the merged data set using the “MICE” package (Stef van Buuren, Karin Groothuis-Oudshoorn, R statistical software[3]). An algorithm accounting for multi-level data was used and recognized a study-specific identification number across the 17 datasets. Variables containing >30% of missing were not imputed and dropped from the datasets. Convergence of the algorithm and distribution of the imputed data points were verified in diagnostic plots. Outcome data were used for the imputation of comorbidities but were not themselves imputed. Missing covariates were imputed in the same way in the Fire&Ice data set.

Efficacy analysis.

Kaplan Meier analyses: Thanks to the time-event tables from the merged data set and from the F&I data set, a single time-event table could be reconstructed for the overall cohort and all subgroup of interest and allowed for the construction of Kaplan Meier curves accounting for all data (17 merged datasets as well as the F&I dataset). Analyses for the overall cohort including the F&I data, as well as stratified by sex and catheter types were conducted. Differences between groups were tested using a logrank test at 360, 720 and 1080 days after the proportional hazards assumption was checked using a statistical test based on the scaled Schoenfeld residuals. As all studies planned a 90-day blanking period in their design, all time-to-event analyses were started after the 90th day post-ablation.

Multivariable mixed-effect Cox proportional analysis: In order to account for a maximal number of important covariates highlighted in previous literature as decisive for the recurrence of AF following ablation[4], multivariable mixed-effect cox proportional analysis models were fitted to perform time-dependent efficacy analyses. The study ID was the random effect and was used as strata in the cox model. Fixed effect covariates were : sex, catheter type, age, measure of LA dilation, left ventricular systolic dysfunction, vascular disease, hypertension, diabetes, AF duration since first diagnosis in years, congestive heart failure, stroke and AF type (paroxysmal versus non-paroxysmal). To test for the interaction between catheter type and sex, an interaction term was added but was not kept in the analysis if not significant. An analysis stratified by sex was also performed. Similar Cox models stratified by sex were derived. Models were derived once in the large fused dataset containing all 17 studies and once in the Fire&Ice dataset, allowing for a two-stage

analysis using classical meta-analytic techniques (Restricted maximum Likelihood (REML) random-effects model).
A sensitivity analysis on the endpoint of arrhythmia recurrence only was conducted in the merged dataset using similar multi-level Cox proportion hazard models.

Safety analysis:

Occurrence of complications by sex and catheter types are described in number and percentages and differences were tested using a Fisher exact test.
Multivariable mixed-effect logistic models were fitted to predict the risk for complication in the overall cohort and separately in men and women. Again, the models corrected for the same covariables as for the efficacy endpoint and the interaction term between sex and catheter type was removed from the model when non-significant. Estimates of the merged data set and the F&I study were merged in a 2-step meta-analysis using REML random-effects model

Procedural endpoints

Total procedure time and fluoroscopy time were modelled using mixed-effect linear models again accounting for the same covariates. Again, the interaction term between sex and catheter type was tested but removed when non significant and estimates of the merged data set and the F&I study were merged in a 2-step meta-analysis using REML random-effects model.
For the efficacy and the safety models, at least 10 events needed to be present in order to correct for one covariate.

Heterogeneity assessment

A two-stage analysis was used with all provided datasets separately to assess heterogeneity between studies. Simplified models accounting only for sex and the type of catheter used were fitted in all studies for the efficacy and safety endpoint and estimates were pooled using REML random-effects models. An I^2 of >50% was considered representative of a significant heterogeneity. A forest plot of the aggregated studies was established for the efficacy endpoint and an Egger test performed to test for a potential publication bias.

Supplemental tables

Supplemental table 1 – PRISMA IPD Checklist

Complete before submission

PRISMA-IPD Section/topic	Item No	Checklist item	Reported on page
Title			
Title	1	Identify the report as a systematic review and meta-analysis of individual participant data.	
Abstract			
Structured summary	2	Provide a structured summary including as applicable:	
		Background: state research question and main objectives, with information on participants, interventions, comparators and outcomes.	
		Methods: report eligibility criteria; data sources including dates of last bibliographic search or elicitation, noting that IPD were sought; methods of assessing risk of bias.	
		Results: provide number and type of studies and participants identified and number (%) obtained; summary effect estimates for main outcomes (benefits and harms) with confidence intervals and measures of statistical heterogeneity. Describe the direction and size of summary effects in terms meaningful to those who would put findings into practice.	
		Discussion: state main strengths and limitations of the evidence, general interpretation of the results and any important implications.	
		Other: report primary funding source, registration number and registry name for the systematic review and IPD meta-analysis.	
Introduction			
Rationale	3	Describe the rationale for the review in the context of what is already known.	
Objectives	4	Provide an explicit statement of the questions being addressed with reference, as applicable, to participants, interventions, comparisons, outcomes and study design (PICOS). Include any hypotheses that relate to particular types of participant-level subgroups.	
Methods			
Protocol and registration	5	Indicate if a protocol exists and where it can be accessed. If available, provide registration information including registration number and registry name. Provide publication details, if applicable.	
Eligibility criteria	6	Specify inclusion and exclusion criteria including those relating to participants, interventions, comparisons, outcomes, study design and characteristics (e.g. years when conducted, required minimum follow-up). Note whether these were applied at the study or individual level i.e. whether eligible participants were included (and ineligible participants	

		excluded) from a study that included a wider population than specified by the review inclusion criteria. The rationale for criteria should be stated.	
Identifying studies - information sources	7	Describe all methods of identifying published and unpublished studies including, as applicable: which bibliographic databases were searched with dates of coverage; details of any hand searching including of conference proceedings; use of study registers and agency or company databases; contact with the original research team and experts in the field; open adverts and surveys. Give the date of last search or elicitation.	
Identifying studies - search	8	Present the full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	
Study selection processes	9	State the process for determining which studies were eligible for inclusion.	
Data collection processes	10	Describe how IPD were requested, collected and managed, including any processes for querying and confirming data with investigators. If IPD were not sought from any eligible study, the reason for this should be stated (for each such study).	
		If applicable, describe how any studies for which IPD were not available were dealt with. This should include whether, how and what aggregate data were sought or extracted from study reports and publications (such as extracting data independently in duplicate) and any processes for obtaining and confirming these data with investigators.	
Data items	11	Describe how the information and variables to be collected were chosen. List and define all study level and participant level data that were sought, including baseline and follow-up information. If applicable, describe methods of standardising or translating variables within the IPD datasets to ensure common scales or measurements across studies.	
IPD integrity	A1	Describe what aspects of IPD were subject to data checking (such as sequence generation, data consistency and completeness, baseline imbalance) and how this was done.	
Risk of bias assessment in individual studies.	12	Describe methods used to assess risk of bias in the individual studies and whether this was applied separately for each outcome. If applicable, describe how findings of IPD checking were used to inform the assessment. Report if and how risk of bias assessment was used in any data synthesis.	
Specification of outcomes and effect measures	13	State all treatment comparisons of interests. State all outcomes addressed and define them in detail. State whether they were pre-specified for the review and, if applicable, whether they were primary/main or secondary/additional outcomes. Give the principal measures of effect (such as risk ratio, hazard ratio, difference in means) used for each outcome.	

Synthesis methods	14	Describe the meta-analysis methods used to synthesise IPD. Specify any statistical methods and models used. Issues should include (but are not restricted to): • Use of a one-stage or two-stage approach. • How effect estimates were generated separately within each study and combined across studies (where applicable). • Specification of one-stage models (where applicable) including how clustering of patients within studies was accounted for. • Use of fixed or random effects models and any other model assumptions, such as proportional hazards. • How (summary) survival curves were generated (where applicable). • Methods for quantifying statistical heterogeneity (such as I^2 and τ^2). • How studies providing IPD and not providing IPD were analysed together (where applicable). • How missing data within the IPD were dealt with (where applicable).	
Exploration of variation in effects	A2	If applicable, describe any methods used to explore variation in effects by study or participant level characteristics (such as estimation of interactions between effect and covariates). State all participant-level characteristics that were analysed as potential effect modifiers, and whether these were pre-specified.	
Risk of bias across studies	15	Specify any assessment of risk of bias relating to the accumulated body of evidence, including any pertaining to not obtaining IPD for particular studies, outcomes or other variables.	
Additional analyses	16	Describe methods of any additional analyses, including sensitivity analyses. State which of these were pre-specified.	
Results			
Study selection and IPD obtained	17	Give numbers of studies screened, assessed for eligibility, and included in the systematic review with reasons for exclusions at each stage. Indicate the number of studies and participants for which IPD were sought and for which IPD were obtained. For those studies where IPD were not available, give the numbers of studies and participants for which aggregate data were available. Report reasons for non-availability of IPD. Include a flow diagram.	
Study characteristics	18	For each study, present information on key study and participant characteristics (such as description of interventions, numbers of participants, demographic data, unavailability of outcomes, funding source, and if applicable duration of follow-up). Provide (main) citations for each study. Where applicable, also report similar study characteristics for any studies not providing IPD.	
IPD integrity	A3	Report any important issues identified in checking IPD or state that there were none.	
Risk of bias within studies	19	Present data on risk of bias assessments. If applicable, describe whether data checking led to the up-weighting or down-weighting of these assessments. Consider how any potential bias impacts on the robustness of meta-analysis conclusions.	

Results of individual studies	20	For each comparison and for each main outcome (benefit or harm), for each individual study report the number of eligible participants for which data were obtained and show simple summary data for each intervention group (including, where applicable, the number of events), effect estimates and confidence intervals. These may be tabulated or included on a forest plot.	
Results of syntheses	21	Present summary effects for each meta-analysis undertaken, including confidence intervals and measures of statistical heterogeneity. State whether the analysis was pre-specified, and report the numbers of studies and participants and, where applicable, the number of events on which it is based.	
		When exploring variation in effects due to patient or study characteristics, present summary interaction estimates for each characteristic examined, including confidence intervals and measures of statistical heterogeneity. State whether the analysis was pre-specified. State whether any interaction is consistent across trials.	
		Provide a description of the direction and size of effect in terms meaningful to those who would put findings into practice.	
Risk of bias across studies	22	Present results of any assessment of risk of bias relating to the accumulated body of evidence, including any pertaining to the availability and representativeness of available studies, outcomes or other variables.	
Additional analyses	23	Give results of any additional analyses (e.g. sensitivity analyses). If applicable, this should also include any analyses that incorporate aggregate data for studies that do not have IPD. If applicable, summarise the main meta-analysis results following the inclusion or exclusion of studies for which IPD were not available.	
Discussion			
Summary of evidence	24	Summarise the main findings, including the strength of evidence for each main outcome.	
Strengths and limitations	25	Discuss any important strengths and limitations of the evidence including the benefits of access to IPD and any limitations arising from IPD that were not available.	
Conclusions	26	Provide a general interpretation of the findings in the context of other evidence.	
Implications	A4	Consider relevance to key groups (such as policy makers, service providers and service users). Consider implications for future research.	
Funding			
Funding	27	Describe sources of funding and other support (such as supply of IPD), and the role in the systematic review of those providing such support.	

Supplemental table 2 – Contacts authors and answers

Study provider	Publication retrieved by the search	Answer to call-for-data
Fire&Ice publication committee	Kuck et al., Cryoballoon or Radiofrequency Ablation for Paroxysmal Atrial Fibrillation, New England Journal of Medicine, 2016	Yes
Dr. Hunter, Dr. Shilling	Hunter et al., Point-by-Point Radiofrequency Ablation Versus the Cryoballoon or a Novel Combined Approach: A Randomized Trial Comparing 3 Methods of Pulmonary Vein Isolation for Paroxysmal Atrial Fibrillation (The Cryo Versus RF Trial), Journal of Cardiovasc Electrophysiology, 2015	Yes
Dr. Luik	Luik et al., Cryoballoon Versus Open Irrigated Radiofrequency Ablation in Patients With Paroxysmal Atrial Fibrillation The Prospective, Randomized, Controlled, Noninferiority FreezeAF Study, Circulation, 2015	Yes
Dr. Pérez-Castellano	Perez-Castellano et al., The COR trial: A randomized study with continuous rhythm monitoring to compare the efficacy of cryoenergy and radiofrequency for pulmonary vein isolation, Heart Rhythm, 2014	Yes
Dr. Jourda	Jourda et al., Contact-force guided radiofrequency vs. second-generation balloon cryotherapy for pulmonary vein isolation in patients with paroxysmal atrial fibrillation-a prospective evaluation, Europace, 2015	Yes[1]
Dr. Boveda	Boveda et al., Outcomes after cryoballoon or radiofrequency ablation for persistent atrial fibrillation: a multicentric propensity-score matched study, Journal of interventional Cardiac Electrophysiology, 2016	Yes*
Dr. Khoueiry	Khoueiry et al, Outcomes after cryoablation vs. radiofrequency in patients with paroxysmal atrial fibrillation: impact of pulmonary veins anatomy, Europace, 2016	Yes*
Dr. Providencia	Providencia et al., Results from a multicentre comparison of cryoballoon vs. radiofrequency ablation for paroxysmal atrial fibrillation: is cryoablation more reproducible?, Europace, 2017	Yes*
Dr. Knecht	Knecht et al., Long-term comparison of cryoballoon and radiofrequency ablation of paroxysmal atrial fibrillation: A propensity score matched analysis , International Journal of Cardiology, 2014	Yes
Dr. Kojodjoo	Kojodjoo et al., Pulmonary venous isolation by antral ablation with a large cryoballoon for treatment of paroxysmal and persistent atrial fibrillation: medium-term outcomes and non-randomised comparison with pulmonary venous isolation by radiofrequency ablation, Heart, 2010	Yes
Dr. Squara	Comparison between radiofrequency with contact force-sensing and second-generation cryoballoon for paroxysmal atrial fibrillation catheter ablation: a multicentre European evaluation, Europace, 2015	Yes
Dr. Oral	Yokokawa et al., Ablation of paroxysmal atrial fibrillation using a second-generation cryoballoon catheter or contact-force sensing radiofrequency ablation catheter: A comparison of costs and long-term clinical outcomes, Journal of Cardiovascular Electrophysiology	Yes

[1] Related to the 2 datasets provided by Dr. Khoueiry and Dr. Providência

Dr. Matta	Matta et al., Cryoballoon vs. radiofrequency contact force ablation for paroxysmal atrial fibrillation: a propensity score analysis, Journal of Cardiovascular Medicine, 2018	Yes
Dr. Anselmino	Incidence of silent cerebral thromboembolic lesions after atrial fibrillation ablation may change according to technology used: comparison of irrigated radiofrequency, multipolar nonirrigated catheter and cryoballoon, Journal of Cardiovascular Electrophysiology, 2011[2]	Yes
Dr. Kosmidou	Kosmidou et al., Comparing Safety and Efficacy of Irrigated Radiofrequency Catheter Ablation Versus Combined Cryoballoon and Catheter Ablation for Persistent Atrial Fibrillation, Journal of Atrial Fibrillation, 2013	Yes
Dr. Davtyan	Davtyan et al., Radiofrequency versus Cryoballoon Ablation of Atrial Fibrillation: An Evaluation Using ECG, Holter Monitoring, and Implantable Loop Recorders to Monitor Absolute and Clinical Effectiveness, BioMed Research International, 2018	Yes
Dr. Elvan	Buist et al., Arrhythmia-free survival and pulmonary vein reconnection patterns after second-generation cryoballoon and contact-force radiofrequency pulmonary vein isolation, Clinical Research in Cardiology, 2018	Yes
Dr. Wasserlauf	Wasserlauf et al., Cryoballoon versus Radiofrequency Catheter Ablation for Paroxysmal Atrial Fibrillation, PACE, 2014	Yes
Dr. Kobori	Kobori et al., Abstract: 3295, Durability of pulmonary vein isolation by various kind of ablation catheter, European Heart Journal, 2016	Yes
Dr. Jain	Siddoway et al., Improved Resource Utilization With Similar Efficacy During Early Adoption of Cryoballoon Pulmonary Vein Isolation as Compared to Radiofrequency Ablation for Paroxysmal Atrial Fibrillation, Journal of Atrial Fibrillation, 2015	Yes
Dr. Blumstrom-Lundqvist	Moertsell et al., Cryoballoon vs. radiofrequency ablation for atrial fibrillation: a study of outcome and safety based on the ESC-EHRA atrial fibrillation ablation long-term registry and the Swedish catheter ablation registry, Europace, 2018	? (EHRA)
Dr. Dagres	Chen et al., Catheter ablation for atrial fibrillation: results from the first European Snapshot Survey on Procedural Routines for Atrial Fibrillation Ablation (ESS-PRAFA) Part II, Europace, 2015	? (EHRA)
Dr. Tanaka	Tanaka et al., Abstract 18922: Is Treatment for Paroxysmal Atrial Fibrillation With Second-generation Cryoballoon Ablation More Reproducible Than That With Contact Force Guided Radiofrequency Ablation?, Circulation, 2017	No (No answer)
Dr. Chun	Complications in Catheter Ablation of Atrial Fibrillation in 3,000 Consecutive Procedures?, JACC Clinical Electrophysiology, 2017	No
Dr. Bordignon	Schmidt et al., A prospective randomized single-center study on the risk of asymptomatic cerebral lesions comparing irrigated radiofrequency current ablation with the cryoballoon and the laser balloon, Journal of Cardiovascular Electrophysiology, 2013	No
Dr. Tondo	Tondo et al., Pulmonary vein isolation cryoablation for patients with persistent and long-standing persistent atrial fibrillation: Clinical outcomes from the real-world multicenter observational project, Heart Rhythm, 2018	No

[2]Dr. Anselmino was contacted for the listed study but as these data were not suitable for the IPD-meta analysis, he shared data from his center's registry.

Dr. Chierchia	Cherchia et al., Pericardial effusion in atrial fibrillation ablation: a comparison between cryoballoon and radiofrequency pulmonary vein isolation, Europace, 2010	No
Dr. de Asmundis	Ciconte et al., Circumferential pulmonary vein isolation as index procedure for persistent atrial fibrillation: a comparison between radiofrequency catheter ablation and second-generation cryoballoon ablation, Europace, 2015	No
Dr. Schmidt	Schmidt et al., German ablation registry: Cryoballoon vs. radiofrequency ablation in paroxysmal atrial fibrillation--One-year outcome data, Heart Rhythm, 2016	No
Dr. Straube	First-line catheter ablation of paroxysmal atrial fibrillation: outcome of radiofrequency vs. cryoballoon pulmonary vein isolation, Europace, 2016	No

Supplemental table 3 – Details of the provided studies

Author	Type	Number of patients	Country	Type of center	Publication date	Percentage of women in the cohort
Fire&Ice	RCT open label	750	International	multicentric	2016-06-09	39.06
Hunter	RCT open label	155	England	monocentric	2015-12-01	33.55
Luik	RCT open label	315	Germany	multicentric	2015-12-06	38.34
Perez Castellano	RCT open label	50	England	monocentric	2014-01-01	22.00
Khoueiry	observational prospective	1460	France	multicentric	2016-05-18	26.64
Providencia	observational prospective	994	France	multicentric	2016-06-05	23.99
Knecht	observational prospective	201	Switzerland	monocentric	2014-07-02	27.86
Kojodjojo	observational prospective	177	England	monocentric	2010-05-04	23.16
Wasserlauf	observational prospective	148	USA	monocentric	2015-04-01	31.76
Anselmino	observational prospective	97	Italy	monocentric	2011-09-01	25.77
Squara	observational prospective	376	international	multicentric	2015-04-03	25.27
Elvan	observational prospective	269	The Netherlands	monocentric	2018-02-06	29.00
Kosmidou	observational prospective	296	USA	monocentric	2013-10-31	28.23
Matta	observational prospective	128	Italy	multicentric	2018-04-19	21.49
Jain	observational prospective	100	USA	monocentric	2015-02-01	27.00
Kobori	observational prospective	1117	Japan	monocentric	2016-08-31	28.14
Davtyan	RCT single blind	89	Russia	monocentric	2018-03-12	46.07
Oral	observational prospective	146	USA	monocentric	2017-10-23	34.93

Supplemental table 4 – Details of the inclusion/exclusion and endpoints for each included study

Dataset provider	Publication year	Inclusion criteria	Exclusion criteria	Primary endpoints	Secondary endpoints
Fire&Ice	2016	1) Symptomatic PAF with at least 2 episodes and at least one episode documented (30sec, documented with ECG within last 12 months). 2) Documented treatment failure for effectiveness of at least one anti-arrhythmic drug (AAD type I or III, including Beta-blocker and AAD intolerance). 3) >=18 and <=75y.o. 4) Mentally and linguistically able to understand the trial 5) Consent	1) Pregnant women or women of childbearing potential not on adequate birth control. 2) Previous LA ablation or surgery 3) Cardiac surgery or percutaneous coronary intervention within 3 months prior to enrolment 4) Unstable angina pectoris 5) Myocardial infarction within 3 months prior to enrolment 6) Stroke or transient ischemic attack within six months prior to enrolment. 7) New York Heart Association (NYHA) class III or IV congestive heart failure. 8) EF < 35 % 9) Anteroposterior LA diameter > 55 mm (by trans-thoracic echocardiography (TTE or TEE) within three months to prior enrolment). Please refer to the original publication[5] for a complete list	Primary efficacy endpoint: First documented clinical failure occurring more than 90 days after the index ablation procedure. Failure was defined as the recurrence of AF (>30sec), atrial flutter or atrial tachycardia, prescription of antiarrhythmic drugs (I or III) or repeat ablation. Primary safety endpoint: Composite of death from any cause, stroke/TIA from any cause and serious adverse events (cardiac arrhythmia other than AF and other serious adverse events causally related to the treatment).	1) Death from any cause 2) Death from arrhythmia 3) Total duration of the procedure 4) Total fluoroscopy time 5) Hospitalization for cardiovascular causes

Dataset provider	Publication year	Inclusion criteria	Exclusion criteria	Primary endpoints	Secondary endpoints
Hunter	2015	Patients with symptomatic paroxysmal AF refractory to >=1 antiarrhythmic drug	1) Persistent AF 2) Potentially reversible cause of AF 3) Any contra-indication to ablation 4) Severe valvular disease 5) Prior left atrial ablation	The success rate at 1 years (Freedom from documented AF/AT Lasting >=30sec (whether symptomatic or not) following a 3-month blanking period) following a single procedure without antiarrhythmic drugs.	1) Complication rates 2) Procedure times 3) Fluoroscopy times
Luik	2015	At least 1 episode of AF confirmed by ECG and documentation of at least 1 ineffective AAD treatment (including beta-blockers)	1) Previously underwent LA ablation or surgery 2) LA >55mm 3) LA thrombus 4) Unstable angina, MI within the previous 3 months, transluminal coronary angioplasty within the preivous 3 months 5) EF <40% 6) HF grade III/IV 7) Stroke or TIA within the last 6 months 8) Pregnancy 9) Life-expectancy <1y	Absence of atrial arrhythmias in combination with absence of persistent complications during the 6- and 12-month follow-up periods. Persistent complications: PV stenosis, PNP, cerebrovascular accidents, bleedings, vascular complications occurring during or within after 48h post-procedure. Failure : Atrial arrythmia >30sec post blanking period.	1) Procedural data 2) total radiation exposure 3) total procedure duration 4) occurrence of adverse events including PNP, pericardial effusion and vascular complications.

Dataset provider	Publication year	Inclusion criteria	Exclusion criteria	Primary endpoints	Secondary endpoints
Perez Castellano	2014	Symptomatic recurrent paroxysmal AF (>2 episodes in the last 6 months(refractory to 1 or more antiarrhythmic drugs (class I or III) and an anatomical pattern consisting of 4 single PVs.	1) Aged <18 or >75 prior to AF ablation 2) Prior cardiac surgery 3) moderate to severe valvular heart disease 4) Anteroposterior diameter of the left atrium >50mm 5) hyperthyroidism 6) intracardiac thrombus 7) Contraindications for anticoagulant therapy 8) Concomitant acute illness 9) Pregnancy 10) Unavailability for 1y FU at the center	Proportion of patients who remained free from AF recurrences without taking antiarrhythmic drugs 12 months after ablation. AF recurrence was defined as a clinical or subclinical AF episode >=2min recorded on ECG, Holter monitor or ICM after a 3-month blanking period.	1) AF-free survival after a 3-month blanking period 2) Cumulative AF burden 12 months after ablation (number of AF episodes and percentage of time in AF) 3) Proportion of PVs that remained isolated at the end of the procedure 4) Ablation time 5) Procedure time.
Davtyan	2018	1) ≥1 documented ECG occurrence of nonvalvular symptomatic paroxysmal AF lasting > 30 seconds within 90 days of enrollment that was refractory or intolerant to ≥1 antiarrhythmic drug (including beta blockers) 2) age ≥ 18 and ≤79 years 3) left atrial diameter < 50mm (anteroposterior) by parasternal long axis view 4) left ventricular ejection fraction ≥50% during sinus rhythm (estimated by Simpson's method)	1) a patient history of myocardial infarction or cardiac surgery within 90 days of enrollment 2) a patient history of stroke or transient ischemic attack within 1 year of enrollment 3) any uncontrolled thyroid dysfunction 4) a patient who was contraindicated or had an inability to maintain anticoagulation via oral pharmaceutical drug.	12-month assessment of the freedom from atrial arrhythmia occurrence when comparing the effectiveness of RF vs CB catheter ablation. Freedom from arrhythmia was denoted by the lack of detection of AF (>=30sec in duration) atrial flutter or atrial tachycardia episodes.	Evaluation of the postablation 90-day blanking period to determine the duration period needed by each catheter type to achieve long-term stable normal sinus rhythm after the healing of injuries associated with the cardiac ablation lesion formation as determined by ILR examination

Dataset provider	Publication year	Inclusion criteria	Exclusion criteria	Primary endpoints	Secondary endpoints
Jourda	2014	1) Patients with paroxysmal AF 2) AF refractory to at least one anti-arrhythmic drug from Class I or III 3) Patients undergoing a first PVI procedure	1) Need of additional lines (roof, mitral isthmus) or ablation of complex-fractionated electrograms during the procedure.	Recurrence up to 12 months after a 3 months blanking period, with a recurrence defined as any symptomatic or asymptomatic atrial arrhythmia lasting >30s. Patients with recurrence of AF during the blanking period with no response to cardioversion (either pharmacological or direct-current) were classified as having a relapse.	1) Procedural efficacy endpoint defined as the portion of effective PVI with PV isolation confirmed by entry and exit block after a waiting time of 20 minutes. 2) Periprocedural safety: All-cause periprocedural death, thromboembolisms (stroke, TIA and systemic or pulmonary embolism) and major bleeding (cardiac tamponade, bleeding necessitating intervention such as thrombin injection, surgery or transfusion, massive haemoptysis, haemothorax, retroperitoneal bleeding or any other life-threatening bleeding).
Khoueiry	2016	1) Patients with paroxysmal AF refractory to at least one class I or class III antiarrhythmic drug.	1) Patients with persistent AF 2) Previous left atrial ablation procedure 3) Left atrial thrombus	Freedom from atrial arrhythmias after a blanking period of 3 months and a single atrial ablation procedure, where recurrence was defined as any symptomatic or asymptomatic atrial arrhythmia lasting >30s following the 3 months blanking period after catheter ablation.	1) Procedural endpoint: Complete PVI in each group 2) Periprocedural safety: All-cause periprocedural death, thromboembolisms (stroke, TIA and systemic or pulmonary embolism) and major bleeding (cardiac tamponade, bleeding necessitating intervention such as thrombin injection, surgery or transfusion, massive haemoptysis, haemothorax, retroperitoneal bleeding or any other life-threatening bleeding).

Dataset provider	Publication year	Inclusion criteria	Exclusion criteria	Primary endpoints	Secondary endpoints
Boveda	2016	1) Patients older than 18 undergoing a first catheter ablation for persistent AF refractory to at least one anti-arrhythmic drug agent	1) Previous ablation 2) RF ablation using a one-shot technique 3) RF using a contact-force sensing catheter	AF/atrial tachycardia recurrence, defined as any symptomatic or asymptomatic atrial arrhythmia lasting >30s after a 3-month blanking period.	1) Safety endpoints: vascular complications (if requiring intervention or prolongation of admission), thromboembolism (TIA, stroke and/or systemic embolism), phrenic nerve palsy persisting after the procedure, pericardial effusion (if causing hemodynamic instability and/or requiring pericardiocentesis or prolonged monitoring) and procedure-related death. 2) Procedural endpoints
Providencia	2016	1) Patients older than 18 undergoing catheter ablation of paroxysmal AF refractory to at least one anti-arrhythmic drug agent	1) Patients with a previous AF ablation procedure 2) Patients with persistent AF 3) Patients with atrial flutter 4) Patients undergoing a procedure with any other single-shot technique	AF/atrial tachycardia recurrence, defined as any symptomatic or asymptomatic atrial arrhythmia lasting >30s after the blanking period	1) Safety endpoints: vascular complications (if requiring intervention or prolongation of admission), thromboembolism (TIA, stroke and/or systemic embolism), phrenic nerve palsy persisting after the procedure, pericardial effusion (if causing hemodynamic instability and/or requiring pericardiocentesis or prolonged monitoring) and procedure-related death. 2) Procedural endpoints

Dataset provider	Publication year	Inclusion criteria	Exclusion criteria	Primary endpoints	Secondary endpoints
Knecht	2014	Patients with paroxysmal AF undergoing PVI either using RF or balloon-based cryoablation	1) Persistent or permanent AF 2) A history of any previous left atrial procedure (surgical or percutaneous) 3) The use of a magnetic Navigation system	1) Overall single-procedure efficacy of CB-PVI and RF-PVI (CB-PVI versus RF-PVI) 2) Single-procedure efficacy of the subset of CB-PVI using the CB alone for PVI compared to RF-PVI (CB-only versus RF-PVI). Single-procedure efficacy was defined as the time to the first recurrence of AF or any atrial tachycardia	1) Procedure-related complications 2) Procedure duration (defined as vascular access to sheath removal) 3) fluoroscopy duration
Kojodjojo	2010	1) Patients with paroxysmal or early persistent AF undergoing their first AF ablation procedure	1) In the RF Group: Those undergoing repeat procedures for recurrent AF 2) In the RF Group: Those undergoing pulmonary venous Isolation using other investigational devices 3) In the RF Group: Those who received additional left atrial ablation lesions such as ablation of complex fractionated atrial electrograms or linear lesions	Recurrence was defined as any documented episode of AF (both symptomatic and asymptomatic) or atrial tachycardia lasting for more than 30s	1) Complications 2) Procedural and fluoroscopy time.

Dataset provider	Publication year	Inclusion criteria	Exclusion criteria	Primary endpoints	Secondary endpoints
Squara	2015	Patients presenting with paroxysmal AF refractory to antiarrhythmic therapy, and were undergoing a first PVI procedure using either (i) a CF-sensing RF catheter or (ii) a second-generation CB	1) follow-up of <6 months at the time of the study 2) additional linear ablations except for cavo-tricuspid isthmus (CTI) line 3) complex-fractionated electrograms (CFAEs) ablation 4) repeat AF ablation 5) he lack of the use of a three-dimensional (3D) electroanatomic mapping system when RF was the selected energy.	Freedom from any atrial arrhythmia lasting more than 30s for up to 18 months after a single procedure	1) Procedural endpoint : PVI with confirmed entrance and exit blocks 2) Safety endpoint : Comparison of significant adverse events between groups
Wasserlauf	2015	1) All patients had documented pAF per ECG 2) No patients had undergone prior ablation for AF.	1) Patients among the first three cases per provider 2) Lack of required rhythm follow-up	Freedom from documented AF, atrial flutter and atrial tachycardia of any duration on surface ECG or <30sec on rhythm monitoring without requirement for antiarrhythmic medication.	1) Procedural endpoints
Kosmidou	2013	Patients with a history of anti arrhythmic drug-refractory persistent AF, who underwent initial AF ablation including pulmonary vein isolation and atrial substrate modification.		Assess peri-procedural and clinical outcomes in patients undergoing cryoballoon and catheter ablation when compared to patients undergoing standard radio frequency catheter ablation for PVI and left atrial substrate modification in patients with persistent AF	

Dataset provider	Publication year	Inclusion criteria	Exclusion criteria	Primary endpoints	Secondary endpoints
Elvan	2018	Patients who were suitable for primo PVI for drug-refractory paroxysmal or early persistent atrial fibrillation.	1) procedures with additional ablation strategies (complex fractionated atrial electrograms, linear ablations or right atrium ablation) were excluded.	The primary effectiveness endpoint was single procedure arrhythmia-free survival, defined as patients without AF/atrial flutter/atrial tachycardia recurrence after the blanking period of 90 days after PVI.	
Matta	2018	Consecutive patients suffering from paroxysmal atrial fibrillation referred for a first procedure of transcatheter ablation.		1-year freedom from arrhythmic recurrences.	1) Acute success in pulmonary vein isolation, defined as pulmonary vein isolation 20min after the end of ablation with the predefined catheter. 2) Safety, measured as the incidence of major complications (life-threatening complications or those requiring interventions or prolonging hospital stay) 3) fluoroscopy time and procedural duration
Oral	2017	Consecutive patients with PAF who underwent catheter ablation of AF.	1) Patients with a prior ablation	Compare the costs and long-term outcomes of atrial PVI of CBA using the second-generation cryoballoon catheter and a contact force-sensing irrigated-tip ablation catheter in patients with PAF.	
Kobori	2016	Patients undergoing catheter ablation for AF		Durability of PVI depending on kind of catheters	

Dataset provider	Publication year	Inclusion criteria	Exclusion criteria	Primary endpoints	Secondary endpoints
Jain	2015	1) Undergoing CB or RF ablation 2) Symptomatic PAF, which was defined as self-terminating atrial fibrillation lasting no longer than 7 days 3) not having previously required cardioversion	1) persistent or permanent atrial fibrillation 2) previous pulmonary vein isolation 3) insufficient follow-up	Recurrence was defined as an episode of an atrial arrhythmia (atrial fibrillation, atrial tachycardia, atrial flutter) lasting > 30 seconds on event monitoring or a recurrence of typical symptoms, if not captured on monitoring.	1) Complications 2) Procedure time 3) Fluoroscopy time 4) Hospital length of stay.
Anselmino	-	Patients with a history of paroxysmal AF and refractory to antiarrhythmic drugs undergoing PVI.	1) Age >80 2) Severe valvular heart disease 3) Acute coronary syndrome in the last 3 months 4) Previous catheter ablation 5) Previous pacemaker or ICD implantation	-	

Supplemental table 5 – Details regarding the follow-up and monitoring for each included study

					Monitoring								
Study provider	**Publication year**	**Mean FU**	**Patients lost to FU reported**	**Blanking period of 3 months**	**ECG**	**ECG at which months**	**Holter 12h**	**Holter 12h at which months**	**Holter 24h**	**Holter 24h at which months**	**Holter 7d**	**Holter 7d at which months**	**loop recorder for part of the cohort**
Fire&Ice	2016	18	Yes	Yes	Yes	3, 6, 12	No		Yes	3, 6, 12	No		No
Hunter	2015	12	Yes	Yes	No		No		No		Yes	3, 6, 12	No
Luik	2015	12	Yes	Yes	No		No		Yes	3, 9,	Yes	6, 12	No
Perez Castellano	2014	12	Yes	Yes	No		No		No		No		Yes
Davtyan	2018	12	Yes	Yes	Yes	1, 3, 6, 12	No		Yes	1, 3, 6, 12	No		Yes
Jourda	2014	12	No	Yes	Yes	1, 3, 6, 9, 12	No		Yes	1, 3, 6, 9, 12	No		No
Khoueiry	2016	14	No	Yes	Yes	1, 3, 6, 9, 12	No		Yes	1, 3, 6, 9, 12	No		No
Boveda	2016	15.6	Yes	Yes	Yes	1, 3, 6, 9, 12	No		Yes	1, 3, 6, 9, 12	No		No
Providencia	2016	14	Yes	Yes	Yes	1, 3, 6, 9, 12	No		Yes	1, 3, 6, 9, 12	No		No
Knecht	2014	28	No	Yes	Yes	3, 6, 12	No		Yes	3, 6,	Yes	12	No
Kojodjojo	2010	13.3	No	Yes	No		No		Yes	1, 3, 6, 12	No		Yes
Squara	2015	12	No	No	Yes	1, 3, 6, 12	No		Yes	1, 3, 6, 12	No		No
Wasserlauf	2015	8.76	Yes	Yes	Yes	3, 6, 12	No		Yes	6, 12	Yes	3,	Yes
Kosmidou	2013	13.2	No	Yes	Yes	1, 3, 6, 9, 12	Yes	1, 3, 6, 9, 12	No		No		Yes

					Monitoring								
Study provider	**Publication year**	**Mean FU**	**Patients lost to FU reported**	**Blanking period of 3 months**	**ECG**	**ECG at which months**	**Holter 12h**	**Holter 12h at which months**	**Holter 24h**	**Holter 24h at which months**	**Holter 7d**	**Holter 7d at which months**	**loop recorder for part of the cohort**
Elvan	2018	12.9	No	Yes	Yes	3, 6, 12	No		Yes	6, 12	No		No
Matta	2018	12	No	Yes	Yes	3, 12	No		Yes	3, 12	No		No
Oral	2017	25	No	Yes	Yes	3, 6, 12	No		No		No		Yes
Kobori	2016	51.6	No	No	No		No		No		No		No
Jain	2015	-	No	Yes	Yes	1, 3, 6,	No		No		No		No
Anselmino	2016	12	No	Yes	Yes	1, 3, 12	Yes	1, 3, 12	No		No		No

Supplemental table 6 – Baseline characteristics of all included patients by catheter type

	Merged data set				Fire&Ice			
	All patients	CB	RF	P value	All patients	CB	RF	P value
Number of patients	6069	2132	3937		750	369	381	
Age-years (median [IQR])	62.0 [54.7, 68.0]	61.5 [54.0, 67.5]	62.5 [55.0, 68.4]	<0.001	61.0 [54.0, 68.0]	60.0 [54.0, 68.0]	61.0 [53.5, 68.0]	0.834
Sex (females)	1686 (28)	634 (30)	1052 (27)	0.013	293 (39)	153 (41)	140 (37)	0.212
Patients characteristics								
AF type				<0.001				-
paroxysmal	4423 (73)	1923 (90)	2500 (64)		750 (100)	369 (100)	381 (100)	
persistent	1213 (20)	196 (9)	1017 (26)		0 (0)	0 (0)	0 (0)	
longstanding persistent	253 (4)	5 (0)	248 (6)		0 (0)	0 (0)	0 (0)	
Other[3]	168 (3)	3 (0)	165 (4)		0 (0)	0 (0)	0 (0)	
Duration of AF	4.7 (5.0)	4.7 (5.1)	4.7 (4.9)	0.701	4.6 (5.2)	4.6 (5.1)	4.7 (5.3)	0.955
BMI	26.6 (4.6)	27.1 (4.8)	26.4 (4.6)	<0.001	27.9 (4.6)	27.9 (4.7)	27.8 (4.5)	0.782
hypertension	2459 (47)	884 (44)	1575 (48)	0.006	436 (58)	211 (57)	225 (59)	0.656
DM	526 (10)	189 (9)	337 (10)	0.298	60 (8)	35 (10)	25 (7)	0.176
CHF	413 (8)	110 (6)	303 (10)	<0.001	209 (28)	107 (29)	102 (27)	0.565
stroke	360 (8)	110 (6)	250 (8)	0.004	28 (4)	15 (4)	13 (3)	0.780
vascular disease	425 (10)	128 (9)	297 (11)	0.098	48 (6)	26 (7)	22 (6)	0.574
Anti-arrhythmic drug Type I	2397 (47)	774 (48)	1623 (47)	0.579	-	-	-	-
Anti-arrhythmic drug Type III	1885 (37)	605 (37)	1280 (37)	0.817	-	-	-	-
Anti-arrhythmic drug Type II	2081 (41)	701 (42)	1380 (40)	0.281	-	-	-	-
any anti-arrhythmic drug	4153 (73)	1435 (74)	2718 (73)	0.442	289 (39)	135 (37)	154 (40)	0.316
Measure of LA				<0.001				0.022
normal	2088 (40)	765 (44)	1323 (38)		204 (35)	82 (29)	122 (41)	
mildly abnormal	973 (19)	380 (22)	593 (17)		83 (14)	45 (16)	38 (13)	

[3] left atrial tachycardia or flutter

	Merged data set				Fire&Ice			
	All patients	CB	RF	P value	All patients	CB	RF	P value
moderately abnormal	1073 (21)	297 (17)	776 (22)		251 (43)	129 (46)	122 (41)	
severely abnormal	1096 (21)	285 (17)	811 (23)		40 (7)	24 (9)	16 (5)	
LVEF	60.2 (7.9)	60.6 (7.2)	60.0 (8.2)	0.004	62.6 (6.6)	63.1 (6.7)	62.2 (6.6)	0.149
LVSD	268 (5)	68 (3)	200 (5)	0.002	13 (3)	6 (2)	7 (3)	1.000
Catheter data								
Catheter type: RF	3937 (65)	0 (0)	3937 (100)	<0.001	381 (51)	0 (0)	381 (100)	<0.001
Catheter details				<0.001				<0.001
Cryoballoon 1st generation	1012 (17)	1012 (51)	0 (0)		90 (12)	90 (24)	0 (0)	
Cryoballoon 2nd generation	962 (16)	962 (49)	0 (0)		279 (37)	279 (76)	0 (0)	
RF contact force	1282 (22)	0 (0)	1282 (33)		94 (13)	0 (0)	94 (25)	
RF irrigated no contact force	2155 (36)	0 (0)	2155 (55)		287 (38)	0 (0)	287 (75)	
RF not irrigated	500 (8)	0 (0)	500 (13)		0 (0)	0 (0)	0 (0)	

AF = Atrial Fibrillation, BMI=Body mass index, CB=Cryoballoon, CHF = Congestive Heart Failure, DM = Diabetes Mellitus, LA=Left atrium, LVEF = Left ventricular ejection fraction, LVSD =Left Ventricular Systolic Dysfunction, RF= radiofrequency.

Supplemental table 7 – Efficacy outcomes by sex and by sex and catheter type

	Merged data set				Fire&Ice			
	All patients	**Male**	**Female**	**P value**	**All patients**	**Male**	**Female**	**P value**
Number of patients	5831	4205	1626		750	457	293	
Failure : recurrence of AF/AFlutt/LAT or redo or medication	2198 (38)	1552 (37)	646 (40)	0.047	281 (37)	156 (34)	125 (43)	0.020
Failure : type of failure				<0.001				0.079
None	3633 (62)	2653 (63)	980 (60)		469 (63)	301 (66)	168 (57)	
AF	709 (12)	454 (11)	255 (16)		146 (19)	76 (17)	70 (24)	
Aflut	89 (2)	62 (1)	27 (2)		21 (3)	11 (2)	10 (3)	
LAT	18 (0)	13 (0)	5 (0)		-	-	-	
Atrial arrhythmia, not further specified[4]	1135 (19)	840 (20)	295 (18)		-	-	-	
Meds	92 (2)	69 (2)	23 (1)		100 (13)	59 (13)	41 (14)	
Redo	155 (3)	114 (3)	41 (3)		14 (2)	10 (2)	4 (1)	
	All patients	**CB**	**RF**	**P value**	**All patients**	**CB**	**RF**	**P value**
Number of patients	5831	2083	3748		750	457	293	
Failure : recurrence of AF/AFlutt/LAT or redo or medication	2198 (38)	675 (32)	1523 (41)	<0.001	281 (37)	134 (36)	147 (39)	0.546
Failure : type of failure				<0.001				0.786
None	3633 (62)	1408 (68)	2225 (59)		469 (63)	235 (64)	234 (61)	
AF	709 (12)	239 (11)	470 (13)		146 (19)	68 (18)	78 (20)	
Aflut	89 (2)	16 (1)	73 (2)		21 (3)	8 (2)	13 (3)	
LAT	18 (0)	4 (0)	14 (0)		-	-	-	

[4] Some studies recorded the recurrence of arrhythmia without further specification of the subtype (atrial fibrillation, left atrial tachycardia or atrial flutter)

	Merged data set				Fire&Ice			
	All patients	**Male**	**Female**	**P value**	**All patients**	**Male**	**Female**	**P value**
Atrial arrhythmia, not further specified[4]	1135 (19)	350 (17)	785 (21)		-	-	-	
Meds	92 (2)	39 (2)	53 (1)		100 (13)	51 (14)	49 (13)	
Redo	155 (3)	27 (1)	128 (3)		14 (2)	7 (2)	7 (2)	

AF = Atrial fibrillation, Aflut = Atrial flutter, CB= Cryoballoon, RF = Radiofrequency, LAT = Left atrial tachycardia, Meds = Medication.

Supplemental table 8 – Mixed-effect multivariable Cox Model for the efficacy endpoint of arrhythmia recurrence, redos or re-start of medication in men

	Merged data set				Fire&Ice			
	estimate	2.5 %	97.5 %	P value	estimate	2.5 %	97.5 %	P value
Catheter: CB	0.884	0.781	0.999	0.048	0.819	0.592	1.132	0.226
age	1.003	0.997	1.008	0.358	0.998	0.980	1.016	0.831
LA measure: Mild abnormal	1.196	1.006	1.423	0.043	1.100	0.645	1.877	0.725
LA measure: Moderately abnormal	1.060	0.896	1.254	0.492	0.822	0.563	1.200	0.309
LA measure: Severely abnormal	1.191	0.973	1.459	0.089	0.970	0.247	3.801	0.964
LVSD	1.014	0.776	1.324	0.919	0.356	0.048	2.629	0.311
Vascular disease	0.984	0.820	1.182	0.866	0.723	0.362	1.443	0.358
Hypertension	1.031	0.915	1.162	0.614	1.058	0.736	1.522	0.759
DM	0.991	0.817	1.203	0.930	1.443	0.829	2.510	0.194
AF duration	1.026	1.014	1.038	<0.001	1.019	0.992	1.048	0.174
CHF	0.991	0.800	1.228	0.936	1.184	0.829	1.692	0.353
BMI	1.014	0.992	1.036	0.201	0.969	0.924	1.017	0.201
Stroke	1.022	0.792	1.320	0.865	0.549	0.199	1.514	0.247
AF type: Paroxysmal	0.542	0.475	0.619	<0.001	-	-	-	-

AF = atrial Fibrillation, BMI = body mass index CB = Cryoballoon, CHF = Congestive heart failure, DM = Diabetes Mellitus, LA = left atrium, LVSD = Left ventricular systolic dysfunction.

Supplemental table 9 – Mixed-effect multivariable Cox Model for the efficacy endpoint of arrhythmia recurrence, redos or re-start of medication in women

	Merged data set				Fire&Ice			
	estimate	2.5 %	97.5 %	P value	estimate	2.5 %	97.5 %	P value
Catheter: CB	0.958	0.796	1.154	0.653	1.055	0.735	1.514	0.771
age	1.001	0.991	1.010	0.895	1.001	0.976	1.027	0.929
LA measure: Mild abnormal	0.909	0.697	1.185	0.480	1.065	0.561	2.023	0.846
LA measure: Moderately abnormal	0.943	0.733	1.214	0.649	0.733	0.439	1.224	0.233
LA measure: Severely abnormal	1.224	0.926	1.618	0.155	1.172	0.666	2.061	0.583
LVSD	0.663	0.370	1.189	0.168	1.010	0.133	7.693	0.992
Vascular disease	1.021	0.721	1.446	0.907	0.790	0.311	2.004	0.619
Hypertension	1.127	0.929	1.366	0.225	1.106	0.736	1.664	0.628
DM	1.047	0.767	1.430	0.770	0.969	0.470	1.997	0.931
AF duration	1.027	1.009	1.046	0.004	1.026	0.995	1.058	0.104
CHF	1.067	0.738	1.544	0.727	1.291	0.889	1.874	0.180
BMI	1.000	0.975	1.025	0.999	0.988	0.951	1.027	0.542
Stroke	1.040	0.757	1.428	0.809	1.257	0.535	2.953	0.600
AF type: Paroxysmal	0.453	0.363	0.567	<0.001	-	-	-	-

AF = atrial Fibrillation, BMI = body mass index CB = Cryoballoon, CHF = Congestive heart failure, DM = Diabetes Mellitus, LA = left atrium, LVSD = Left ventricular systolic dysfunction.

Supplemental table 10 – Mixed-effect multivariable Cox Model for the efficacy endpoint of arrhythmia recurrence only

Overall pop	estimate	2.5 %	97.5 %	p.value
Sex: Female	1.119	1.008	1.241	0.034
Catheter: CB	0.872	0.784	0.968	0.011
age	1.004	0.999	1.009	0.147
LA measure: Mild abnormal	1.065	0.913	1.242	0.422
LA measure: Moderately abnormal	1.022	0.886	1.177	0.768
LA measure: Severely abnormal	1.155	0.962	1.386	0.120
LVSD	0.891	0.685	1.159	0.390
Vascular disease	1.015	0.855	1.206	0.861
Hypertension	1.081	0.973	1.202	0.148
DM	0.958	0.803	1.144	0.636
AF duration	1.024	1.013	1.034	<0.001
CHF	1.039	0.850	1.271	0.706
BMI	1.010	0.991	1.030	0.295
Stroke	0.956	0.755	1.211	0.706
AF Type: Paroxysmal	0.517	0.459	0.582	<0.001

AF = atrial Fibrillation, BMI = body mass index CB = Cryoballoon, CHF = Congestive heart failure, DM = Diabetes Mellitus, LA = left atrium, LVSD = Left ventricular systolic dysfunction.

Supplemental table 11 – Mixed-effect multivariable logistic regression model for the safety endpoint of periprocedural complications

	Merged data set				Fire&Ice data set			
Overall pop	estimate	2.5 %	97.5 %	P value	estimate	2.5 %	97.5 %	P value
Intercept	1.051	0.976	1.133	0.185	0.006	<0.001	0.111	0.001
Sex: Female	1.020	1.006	1.034	0.004	1.403	0.791	2.491	0.246
Catheter: CB	1.014	1.001	1.028	0.035	0.792	0.466	1.345	0.388
age	1.001	1.000	1.001	0.017	1.034	0.999	1.071	0.060
LA measure: Mild abnormal	0.998	0.979	1.017	0.830	0.994	0.434	2.278	0.989
LA measure: Moderately abnormal	1.001	0.983	1.020	0.875	0.835	0.415	1.679	0.610
LA measure: Severely abnormal	1.012	0.991	1.033	0.269	0.759	0.239	2.406	0.638
LVSD	0.981	0.948	1.015	0.265	1.744	0.194	15.665	0.619
Vascular disease	1.004	0.982	1.026	0.750	1.212	0.448	3.274	0.705
Hypertension	0.985	0.971	0.998	0.027	1.064	0.585	1.935	0.840
DM	1.005	0.983	1.027	0.677	1.107	0.434	2.822	0.831
AF duration	1.000	0.998	1.001	0.640	0.982	0.930	1.037	0.509
CHF	1.016	0.990	1.044	0.229	1.153	0.655	2.031	0.621
BMI	0.999	0.997	1.001	0.292	1.024	0.965	1.087	0.434
Stroke	1.014	0.990	1.039	0.257	0.728	0.164	3.240	0.677
AF type: Paroxysmal	0.989	0.974	1.005	0.166	-	-	-	-

AF = atrial Fibrillation, BMI = body mass index CB = Cryoballoon, CHF = Congestive heart failure, DM = Diabetes Mellitus, LA = left atrium, LVSD = Left ventricular systolic dysfunction.

Supplemental table 12 – Mixed-effect multivariable linear regression model for the total procedural time

	Merged data set				Fire&Ice data set			
	estimate	2.5 %	97.5 %	P value	estimate	2.5 %	97.5 %	P value
Intercept	186.154	154.776	217.531	<0.001	154.51	120.44	188.57	<0.001
Sex: Female	1.399	-1.691	4.489	0.375	3.76	-3.83	11.37	0.331
Catheter: CB	-28.255	-31.295	-25.215	<0.001	-17.52	-24.35	-10.70	<0.001
age	-0.380	-0.522	-0.239	<0.001	-0.51	-0.91	-0.11	0.012
LA measure: Mild abnormal	6.537	2.374	10.700	0.002	1.25	-9.97	12.48	0.826
LA measure: Moderately abnormal	11.323	7.240	15.406	<0.001	-4.26	-12.63	4.11	0.318
LA measure: Severely abnormal	16.865	12.258	21.472	<0.001	-16.91	-31.38	-2.45	0.022
LVSD	7.021	-0.380	14.422	0.063	-9.96	-45.48	25.54	0.582
Vascular disease	2.264	-3.487	8.015	0.434	7.47	-6.59	21.54	0.297
Hypertension	-0.867	-3.964	2.230	0.583	-6.27	-13.87	1.31	0.105
DM	2.465	-2.715	7.646	0.349	-5.25	-18.22	7.72	0.427
AF duration	0.603	0.200	1.006	0.005	0.19	-0.47	0.85	0.568
CHF	8.697	3.069	14.324	0.003	0.97	-6.69	8.65	0.803
BMI	0.783	0.342	1.223	0.001	0.80	-0.01	1.61	0.052
Stroke	0.684	-5.315	6.683	0.821	-17.37	-35.46	0.72	0.060
AF type: Paroxysmal	-21.895	-25.430	-18.360	<0.001	-	-	-	-

AF = atrial Fibrillation, BMI = body mass index CB = Cryoballoon, CHF = Congestive heart failure, DM = Diabetes Mellitus, LA = left atrium, LVSD = Left ventricular systolic dysfunction.

Supplemental table 13 – Mixed-effect multivariable linear regression model for the fluoroscopy time

	Merged data set				Fire&Ice data set			
	estimate	2.5 %	97.5 %	P value	estimate	2.5 %	97.5 %	P value
Intercept	38.152	28.938	47.365	<0.001	23.04	11.41	34.67	<0.001
Sex: Female	0.286	-0.843	1.416	0.619	1.96	-0.69	4.62	0.148
Catheter: CB	-0.873	-1.982	0.236	0.123	5.12	2.78	7.45	<0.001
age	-0.096	-0.149	-0.042	<0.001	-0.04	-0.17	0.09	0.567
LA measure: Mild abnormal	0.212	-1.296	1.719	0.783	-0.98	-5.22	3.26	0.648
LA measure: Moderately abnormal	1.462	-0.086	3.009	0.064	-1.94	-5.30	1.41	0.253
LA measure: Severely abnormal	1.451	-0.272	3.173	0.099	-2.71	-9.69	4.26	0.434
LVSD	0.481	-2.161	3.123	0.721	-2.50	-14.58	9.56	0.684
Vascular disease	1.550	-0.527	3.626	0.141	-1.89	-6.67	2.88	0.438
Hypertension	-0.109	-1.254	1.036	0.852	-1.70	-4.30	0.89	0.198
DM	0.275	-1.453	2.002	0.755	-0.68	-5.12	3.76	0.764
AF duration	0.170	0.042	0.299	0.010	0.01	-0.21	0.24	0.909
CHF	2.960	0.801	5.119	0.007	0.06	-2.54	2.68	0.959
BMI	0.039	-0.116	0.194	0.621	-0.08	-0.36	0.19	0.545
Stroke	0.985	-1.183	3.152	0.370	-3.41	-9.57	2.74	0.277
AF type : Paroxysmal	-5.605	-6.940	-4.269	<0.001	-	-	-	-

AF = atrial Fibrillation, BMI = body mass index CB = Cryoballoon, CHF = Congestive heart failure, DM = Diabetes Mellitus, LA = left atrium, LVSD = Left ventricular systolic dysfunction.

Supplemental table 14 – Meta-regression investigating the influence of three predefined covariables on the results heterogeneity among the 18 data sets

	Heterogeneity in sex-specific results			Heterogeneity in catheter-specific results		
		estimate	P value		estimate	P value
Recurrence of arrhythmia, redo or re-start of medications	study year	1.004	0.898	study year	0.926	0.117
	median age	0.999	0.976	median age	0.993	0.883
	study type: RCT	1.288	0.015	study type: RCT	1.215	0.391
Periprocedural complications	study year	1.028	0.741	study year	1.131	0.2
	median age	1.034	0.554	median age	1.016	0.841
	study type: RCT	0.988	0.961	study type: RCT	1.608	0.259
Total procedure time	study year	1.371	0.046	study year	-2.424	0.67
	median age	-0.084	0.901	median age	-3.444	0.397
	study type: RCT	3.086	0.322	study type: RCT	28.933	0.15
Fluoroscopy time	study year	-0.015	0.943	study year	0.16	0.894
	median age	-0.2	0.402	median age	-0.11	0.899
	study type: RCT	0.989	0.315	study type: RCT	0.913	0.838

RCT = randomized controlled trial.

Supplemental table 15 – Quality assessment of the randomized controlled trials provided

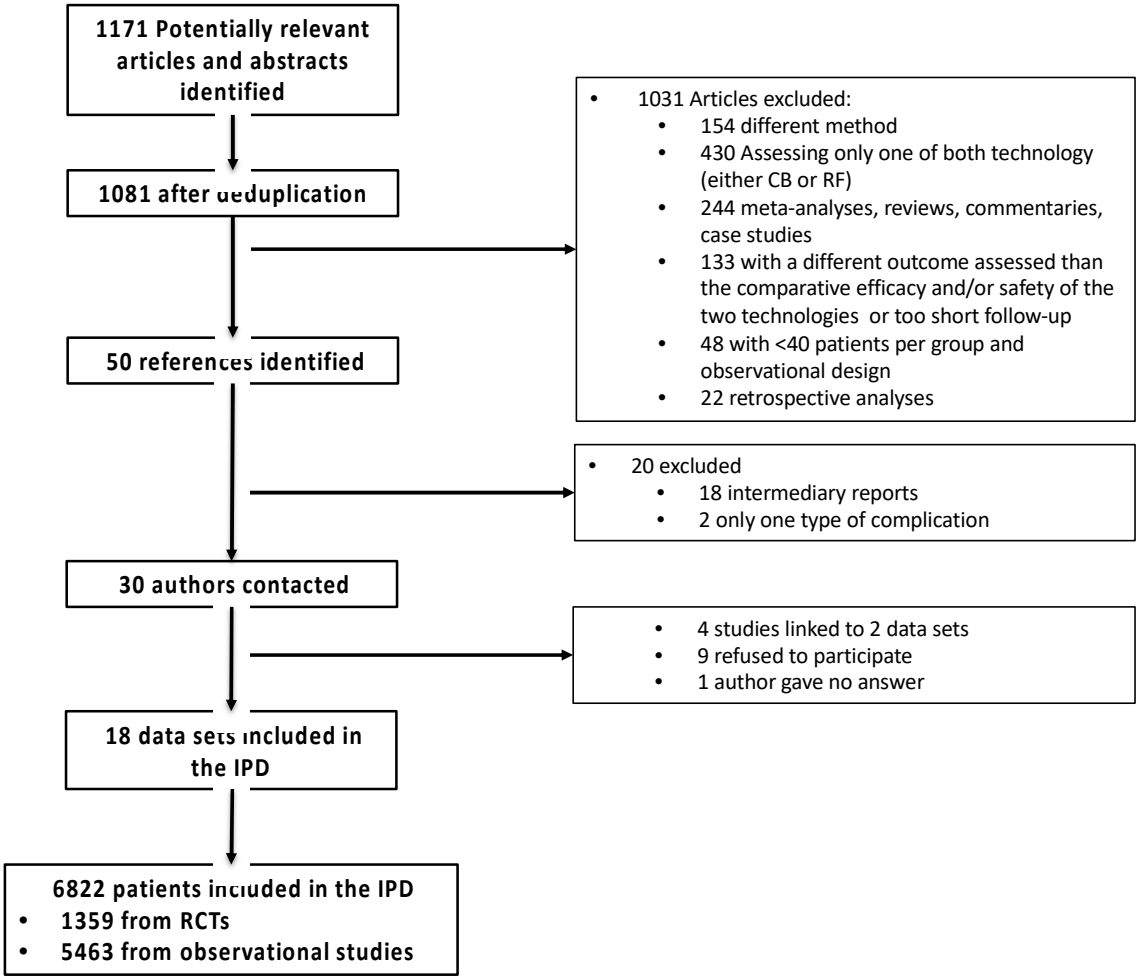
Study provider	Publication year	Random sequence generation	Allocation concealment	Single blinding	Double blinding	Blinding of outcome assessment	Complete outcome data	No selective reporting
Fire&Ice	2016	yes	yes	no	no	yes	no	yes
Hunter	2015	yes	yes	no	no	yes	yes	yes
Luik	2015	yes	yes	unclear	no	yes	no	yes
Perez Castellano	2014	yes	unclear	unclear	no	yes	yes	yes
Davtyan	2018	unclear	unclear	no	no	unclear	yes	unclear

Supplemental table 16 – Quality assessment of the prospective observational studies provided

Study provider	Publication year	Representativeness of the cohort	Case definition	Consecutive cases	Ablation effective	Blind outcome assessment	Adequate monitoring	Adequate and complete FU
Jourda	2014	*	*	no star	*	unclear	*	unclear
Khoueiry	2016	*	*	*	*	unclear	*	unclear
Boveda	2016	*	*	*	*	unclear	*	no star
Providencia	2016	*	*	*	*	unclear	*	no star
Knecht	2014	*	*	*	*	unclear	*	unclear
Kojodjojo	2010	*	*	*	*	unclear	*	unclear
Ciconte	2015	*	*	*	*	unclear	*	*
Squara	2015	*	*	*	*	unclear	*	unclear
Wasserlauf	2015	*	*	no star	*	unclear	*	no star
Kosmidou	2013	*	*	*	*	unclear	*	unclear
Elvan	2018	*	*	*	*	unclear	*	unclear
Matta	2018	*	*	*	*	unclear	*	unclear
Oral	2017	*	*	*	*	unclear	*	unclear
Kobori	2016	*	unclear	*	*	unclear	unclear	unclear
Jain	2015	*	*	*	*	no star	*	*
Anselmino	2016	*	*	*	*	no star	no star	no star

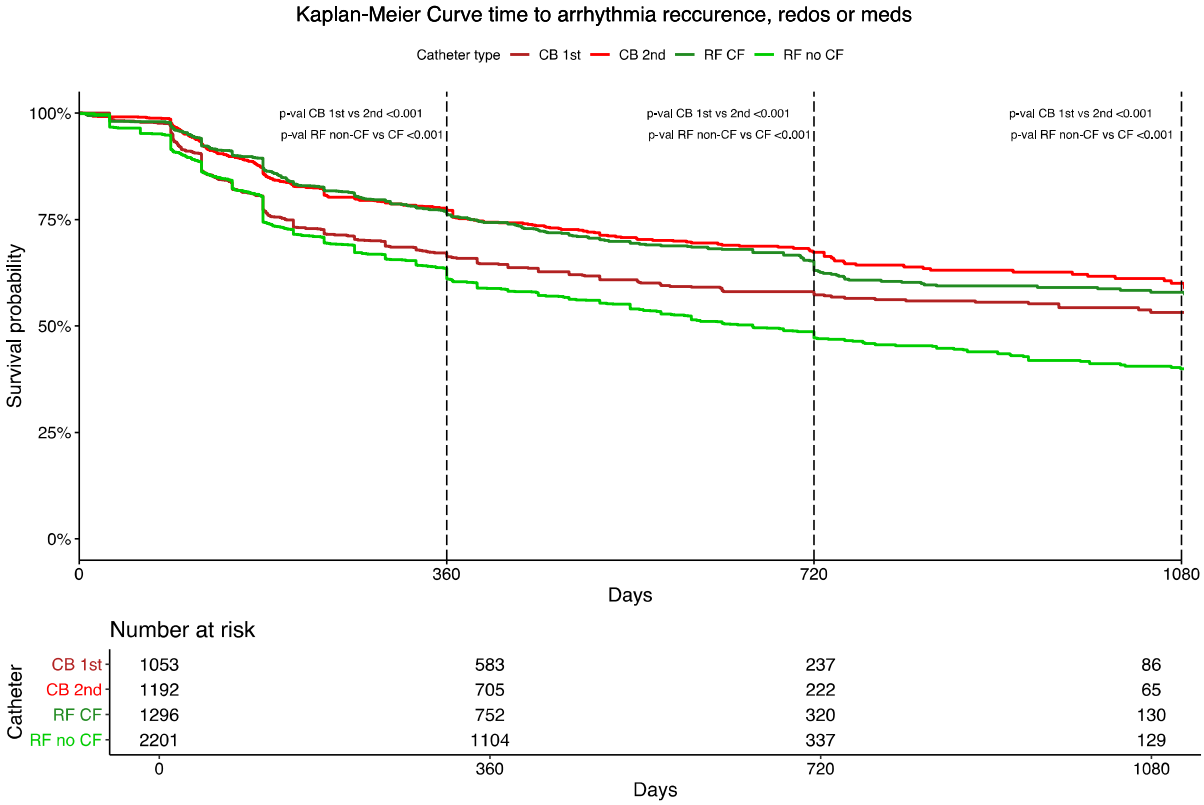
Supplemental figures

Supplemental figure 1 – Study and patient chartflow

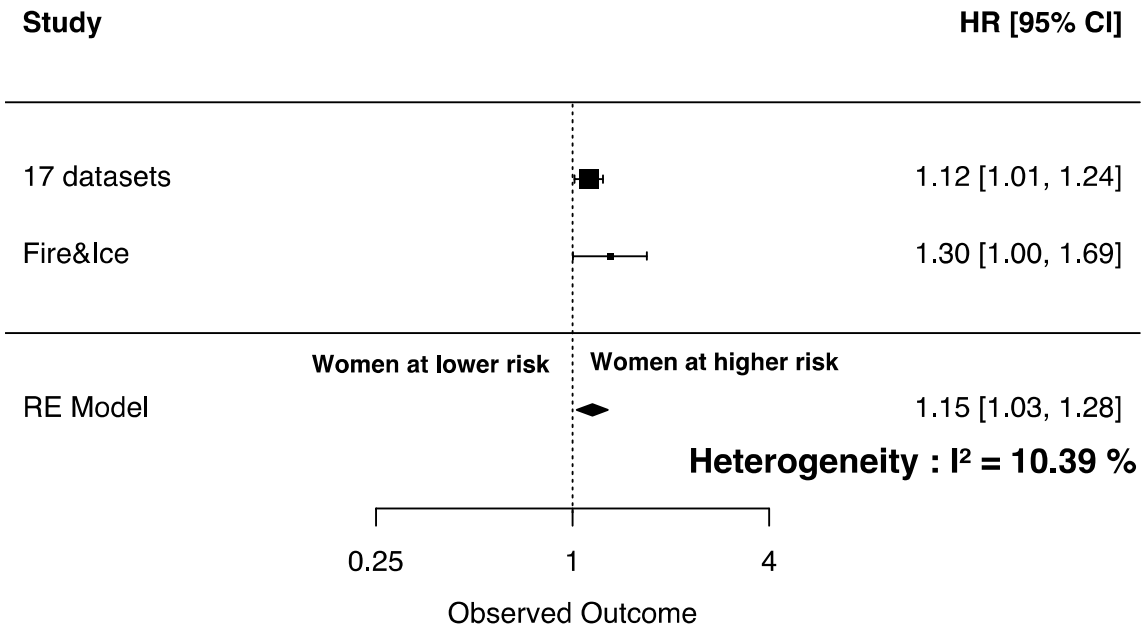


Among the nine studies who refused to participates, 4 at least were based on two datasets. Moreover some authors could not be contacted and some publications originally contained only safety data. Therefore, the exact number of patients who would have been possible suitable for this meta-analysis is difficult to estimate.

Supplemental figure 2 – Kaplan Meier for the efficacy of CB first and second generation and RF with or without contact force for the efficacy endpoint of arrhythmia recurrence, redo or re-start of medication.

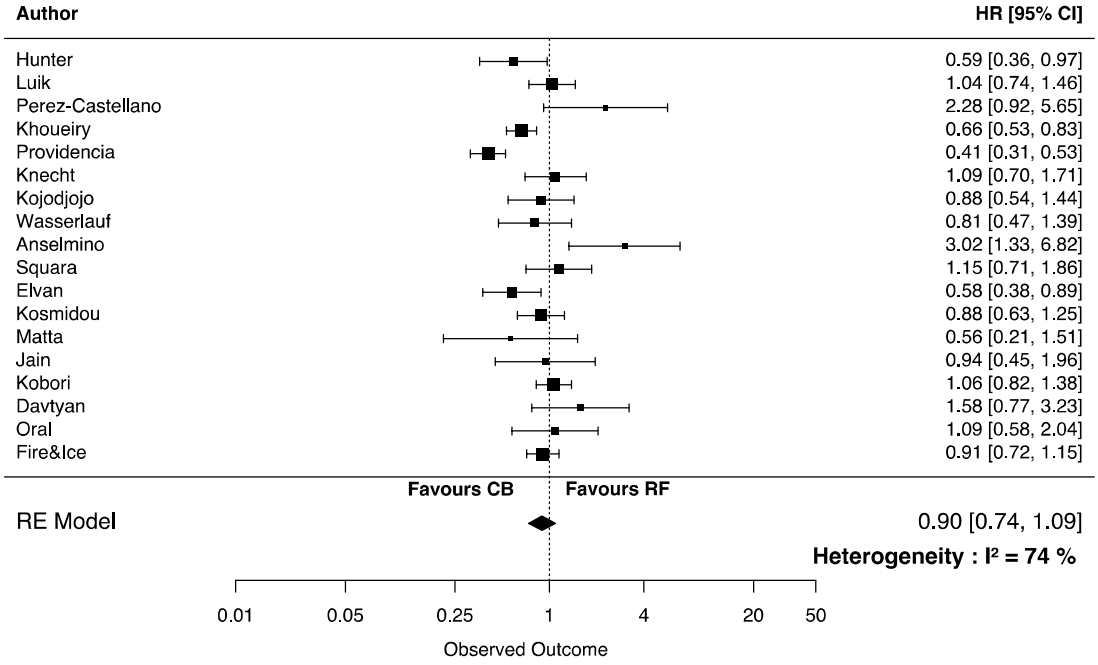


Supplemental figure 3 – Pooled estimates from multivariable cox models for the risk of women to experience an arrhythmia recurrence, undergo a redo or be re-started on medication.

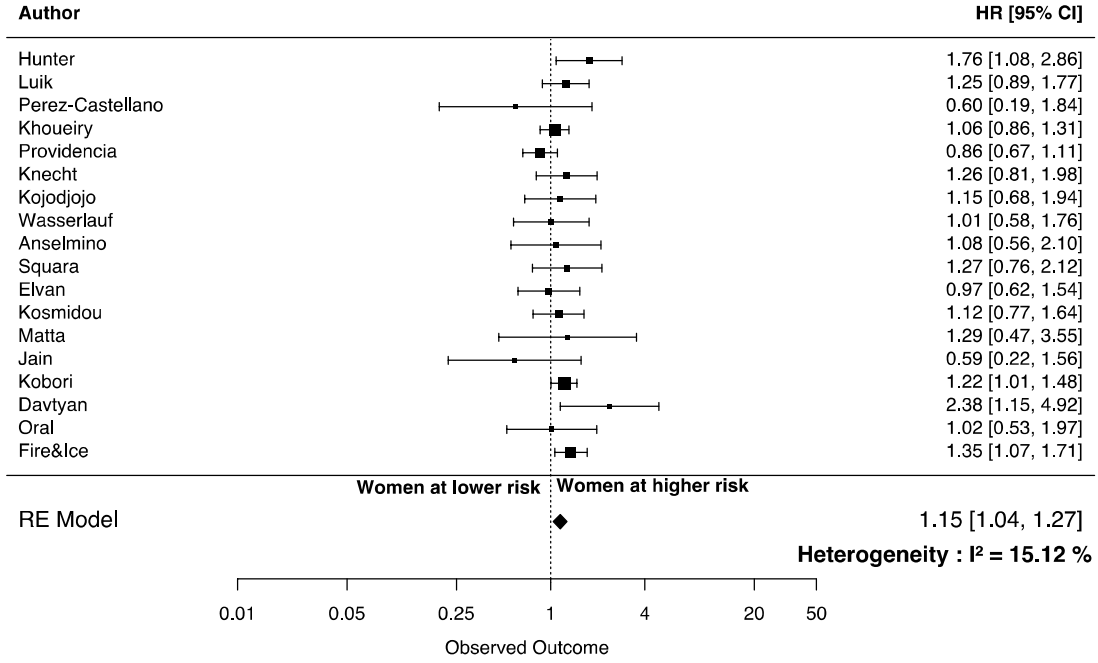


Supplemental figure 4 – Heterogeneity between studies for the recurrence of arrhythmia, redo or re-start of medications

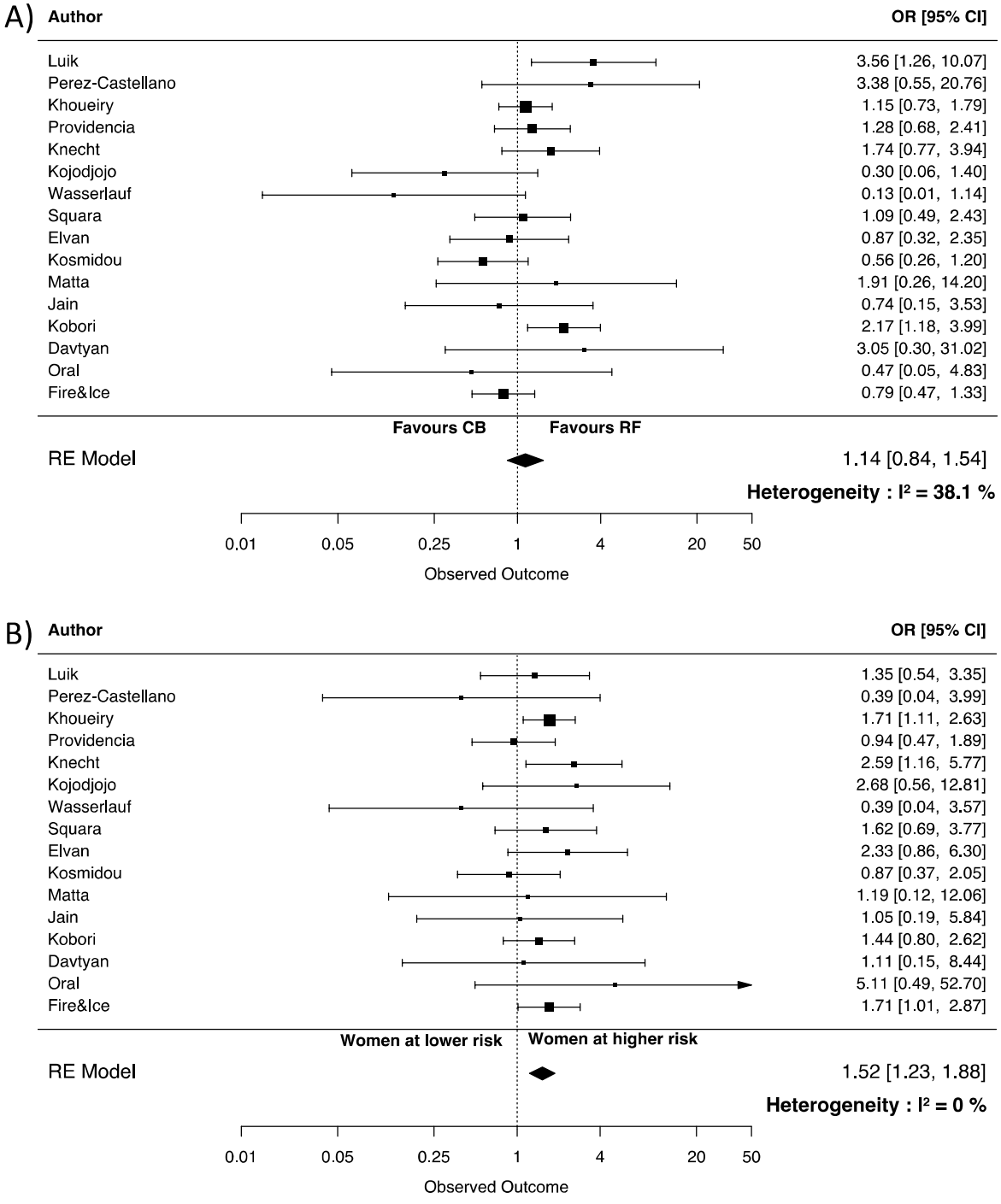
A)



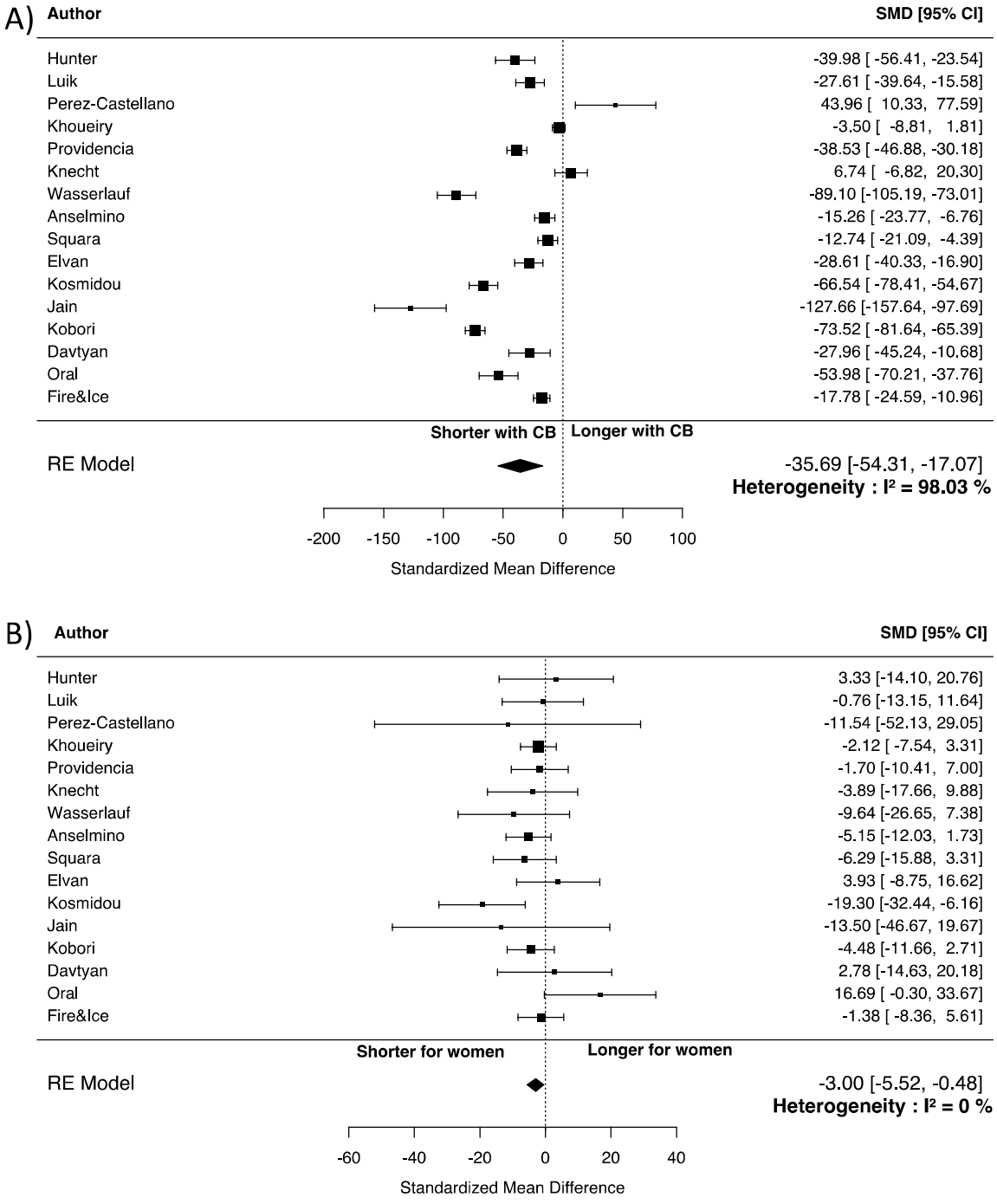
B)



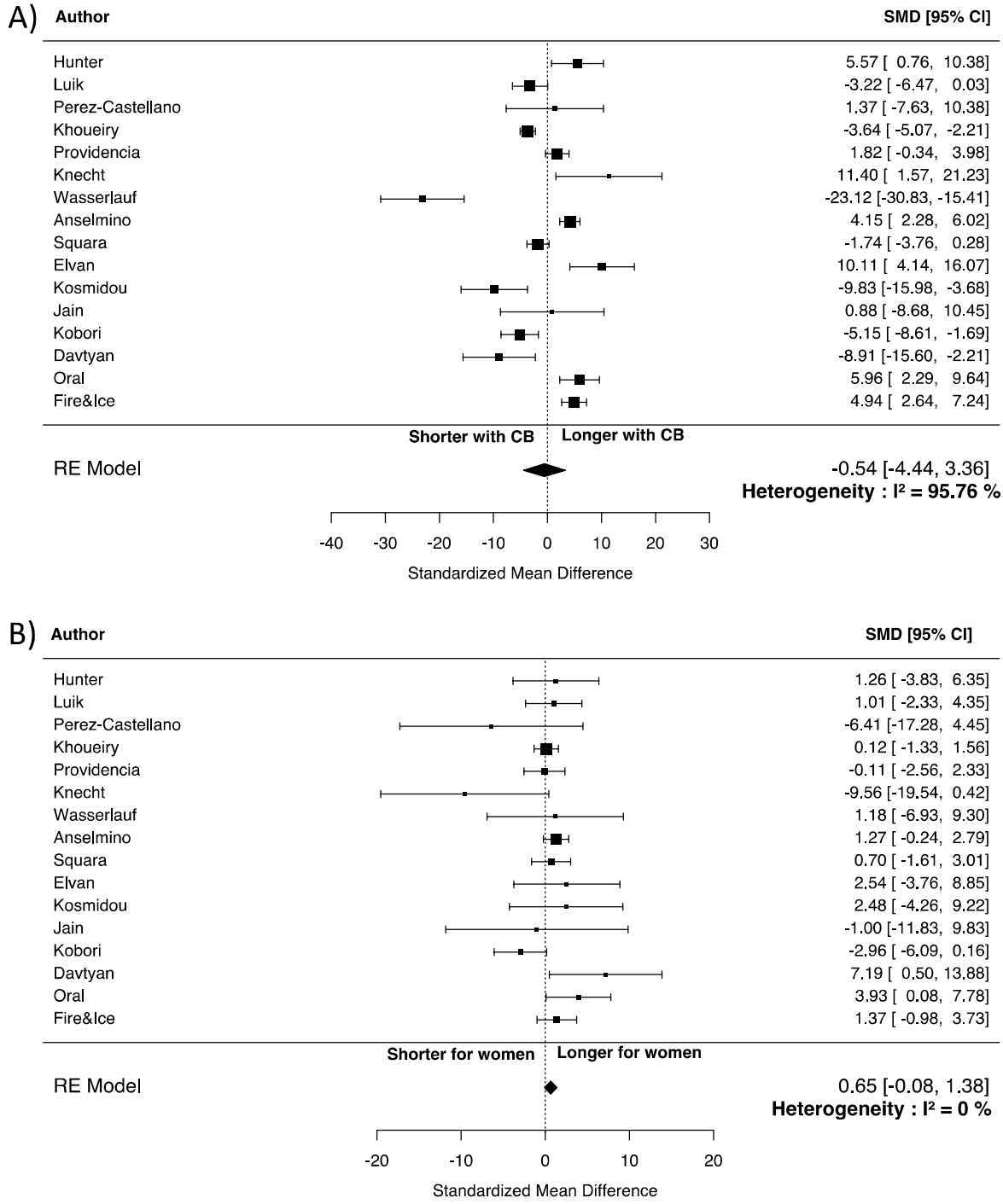
Supplemental figure 5 – Heterogeneity between studies for the occurrence of peri-procedural complications



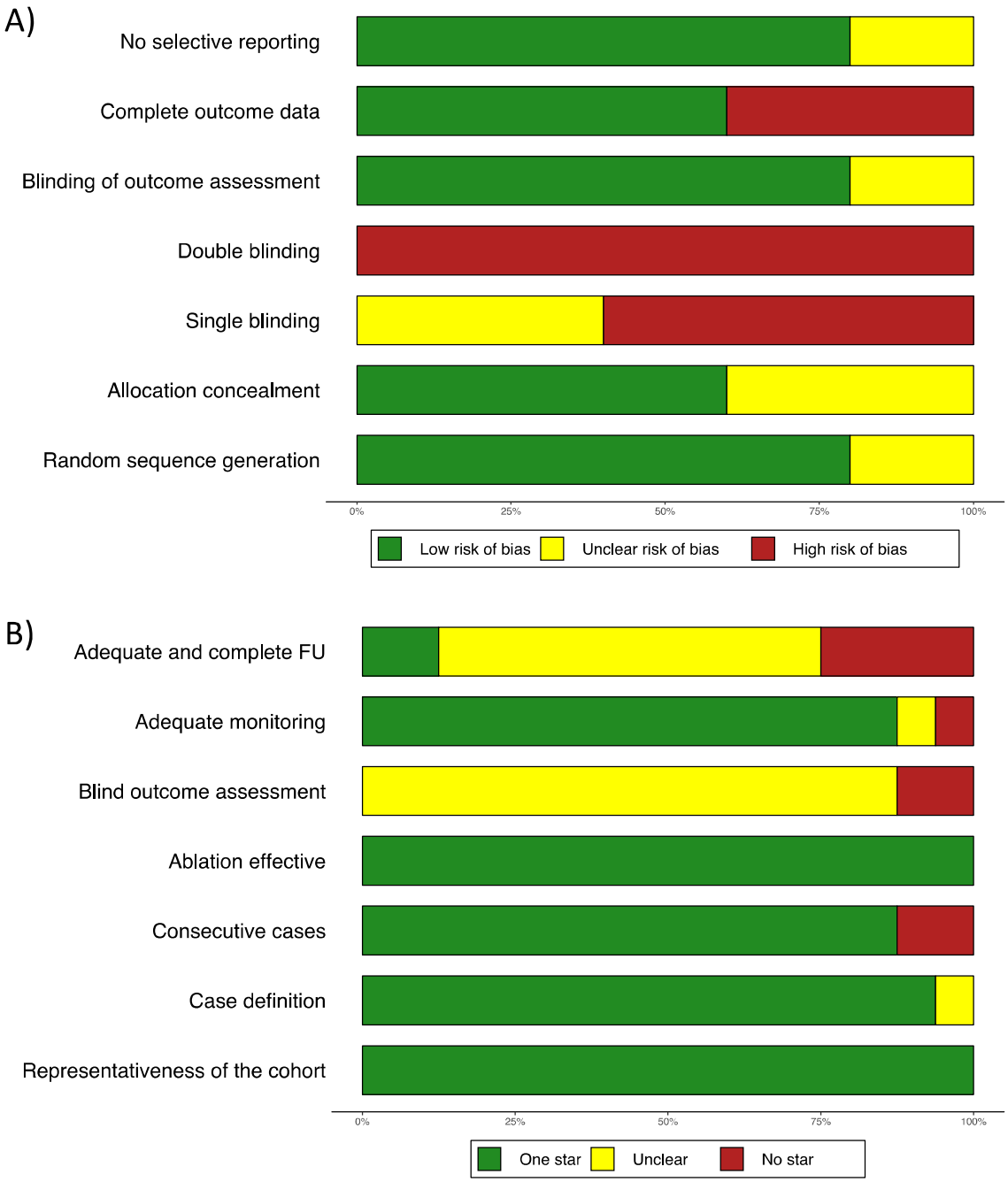
Supplemental figure 6 – Heterogeneity between studies for the total procedure time



Supplemental figure 7 – Heterogeneity between studies for the fluoroscopy time



Supplemental figure 8 – Summary of A) randomized controlled trials and B) prospective observational studies quality



Supplemental figure 9 – Funnel plot of the efficacy endpoint by catheter type