

DESIGN OF A CLUSTER RANDOMIZATION TRIAL TO EVALUATE A
CLINICAL PATHWAY

MSc. Thesis - Health Research Methodology
Girish M. Nair - McMaster University; Student No: 0426738

A CLUSTER RANDOMIZATION TRIAL TO TEST THE EFFECTIVENESS OF A
CLINICAL PATHWAY IN MANAGING ATRIAL FIBRILLATION IN THE
EMERGENCY DEPARTMENT

By

GIRISH M. NAIR, M.B.B.S., M.D., F.R.C.P. (C).

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Science

McMaster University

© Copyright by Girish M. Nair, July 2012

MASTER OF SCIENCE (2012)

McMaster University

(Health Research Methodology)

Hamilton, Ontario

MSc. Thesis - Health Research Methodology
Girish M. Nair - McMaster University; Student No: 0426738

TITLE: A Cluster Randomization Trial to Test the Effectiveness of a Clinical
Pathway in Managing Atrial Fibrillation in the Emergency
Department

AUTHOR: Girish M. Nair, M.B.B.S., M.D., F.R.C.P (C).

SUPERVISOR: Stuart J. Connolly, M.D., F.R.C.P (C).

EXAMINING COMMITTEE: Andrew Worster M.D., C.C.F.P. (E.M.), F.C.F.P.

Robin Roberts Ph.D., and

Ian Stiell M.D., F.R.C.P (C).

NUMBER OF PAGES: xv, 141

ABSTRACT

BACKGROUND:

Atrial Fibrillation (AF) is the most prevalent arrhythmia seen in clinical practice. The prevalence of AF is roughly 0.1% in those younger than 55 years and increases to about 9% in those older than 80 years. Emergency Department (ED) visit rates for patients with symptomatic, recent onset atrial fibrillation or atrial flutter (RAFF) are on the rise, which in turn has resulted in high hospital admission rates (38-45%) and associated health care costs. Variations in management of patients with RAFF at the institutional and ED physician level make it challenging to develop guidelines for management of RAFF in the ED. Optimal management strategies need to be developed for RAFF patients presenting to the ED with the goal of reducing hospital admission rates and health care costs.

STUDY QUESTION:

The primary objective of this study is to evaluate the effectiveness of a clinical pathway for the management of low-risk (without concomitant medical conditions needing hospital admission) RAFF patients presenting to the ED. The hypothesis is that a flexible and evidence-based clinical pathway will help ED physicians better manage AF patients, thereby reducing hospital admission rates.

STUDY DESIGN:

A prospective, blinded, stratified, two-arm cluster-randomized trial will be conducted, with EDs across Canada serving as the unit of randomization.

Patients presenting to the ED with RAFF will be included in the trial if they satisfy all inclusion criteria. The intervention is a clinical pathway for the management of RAFF. Clusters randomized to the control arm will be managed according to the standard institutional protocol for managing AF in the ED. Once approval from all institutional research ethic boards has been obtained we will commence patient recruitment.

STUDY POPULATION:

Patient inclusion criteria:

- i. Adult patients (age > 18 yrs.) presenting to the ED with RAFF as the primary cause of their symptoms
- ii. The patient is able to provide informed consent

Patient exclusion criteria:

- i. Acute coronary syndrome
- ii. Pneumonia
- iii. Severe CHF or hemodynamic compromise
- iv. Major Trauma and/or haemorrhage
- v. Severe exacerbation of bronchial asthma or COPD

- vi. Acute or chronic renal failure needing dialysis or hospital admission
- vii. Any other comorbidity at presentation that necessitates direct hospital admission

PRIMARY AND SECONDARY OUTCOMES:

The primary outcome of this trial will be the proportion of low-risk, RAAF patients admitted to the hospital from the ED. Secondary outcomes include:

- i. Proportion of patients in normal sinus rhythm at the time of discharge from the ED or admission to hospital
- ii. Proportion of patients with adequate heart rate control while in AF (a resting heart rate ≤ 100 bpm for at least one hour) at time of discharge from the ED or admission to hospital
- iii. Proportion of patients returning to the ED with AF within the first thirty days of their index visit
- iv. Safety outcomes- A composite safety outcome including the following:
 - a) Hypotension needing intervention or admission
 - b) Bradycardia needing intervention or admission
 - c) Torsades de pointe or Ventricular tachycardia/fibrillation

- d) Any adverse reaction to medications used for management of AF resulting in hospital admission

Economic analysis- We will conduct a formal cost benefit analysis to assess the health cost savings (if any) as a result of the clinical pathway

STATISTICAL CONSIDERATIONS:

We will conduct an intention-to-treat analysis at the individual level using proportions and means according to the variable in question with an alpha level of 0.05 and power of 0.80 for the primary outcome. EDs in Ontario with an annual volume of >140 RAAF visits will be selected for the trial. The proposed cRCT (assuming a 30% RRR) will be conducted over a two year time period. A total of 13 clusters will need to be recruited in each intervention arm (a total of 26 clusters for the trial). A total of 3500 ED visits for RAFF will need to be included from these 26 clusters.

SUMMARY:

This thesis explores the methodological issues relevant to the design of a cRCT evaluating a clinical pathway in the management of acute onset, low risk AF patients presenting to the ED.

ACKNOWLEDGEMENTS

I wish to thank Stuart Connolly, Andrew Worster and Robin Roberts, my thesis supervisors, for their encouragement, supervision and thoughtful guidance. Their constructive criticism has helped me to develop an in-depth understanding of the methodological aspects of cluster-randomized trials. I extend my sincere gratitude to them for their steadfast support and patience that made this task painless. I wish to thank Ian Stiell, my external examiner, for his insightful and thorough evaluation of my thesis. His comments and suggestions have significantly improved the quality of my thesis.

Next, I would like to thank Carlos Morillo and Syamkumar Menon for their initial work that helped me come up with the study question for my thesis. Their suggestions, constructive criticism and strong encouragement helped me immensely.

I would also like to thank Jeff Healey and Parameswaran Nair for taking the time to critically evaluate my work.

I am deeply indebted to my wife Vidhya and daughter Arundhati, for their love and undying support that helped me complete my Master's degree.

Finally, I would like to thank Alexia Mars, my administrative assistant who helped me with the manuscript and created time for me, from my busy clinical schedule, to complete my Master's course work.

Table of Contents

Title Page.....	i
Descriptive Note.....	ii
Abstract.....	iii
Acknowledgements.....	vii
Table of Contents.....	viii
CHAPTER 1- INTRODUCTION AND REVIEW OF LITERATURE	1
1.1: Rising burden of hospital admissions for Atrial Fibrillation	1
1.2: Factors responsible for hospital admission of Atrial Fibrillation patients	3
1.3: Consequences of hospital admission for Atrial Fibrillation	5
1.4: Interventions to reduce hospital admissions for Atrial Fibrillation	5
1.5: Emergency department physician survey- Hamilton Health Sciences	8
1.6: Atrial Fibrillation management- Hamilton Health Sciences	9
1.7: Variation in ED management of Atrial Fibrillation	11
1.8: Variation in ED physician management of Atrial Fibrillation	13
1.9: Clinical Pathways- definition and efficacy	15

1.10: Need for clinical trial	18
CHAPTER 2- STUDY DESIGN CONSIDERATIONS	22
2.1: Introduction	22
2.2: Selection of appropriate trial design	23
2.2.1: Observational studies	23
2.2.2: Quasi-experimental designs	23
2.2.3: Randomized trials	27
2.3: Selection of the intervention	29
2.4: Study hypotheses	30
2.5: Study outcomes	31
2.5.1: Primary outcome	31
2.5.2: Secondary outcomes	31
2.6: Cluster definition and eligibility	32
2.6.1: Cluster inclusion criteria	32
2.6.2: Cluster exclusion criteria	32
2.7: Unit of inference	33

2.7.1: Patient inclusion criteria	33
2.7.2: Patient exclusion criteria	33
2.8: Participant recruitment	34
2.9: Randomization procedures	35
2.9.1: Restricted randomization	35
2.9.2: Allocation concealment	37
2.9.3: Validity of restricted randomization	38
CHAPTER 3: MEASUREMENTS AND DATA COLLECTION	39
3.1: Introduction	39
3.2: Definition of Atrial Fibrillation and Atrial Flutter	39
3.3: Primary outcome	41
3.4: Secondary outcomes	42
3.5: Follow-up	43
3.6: Data collection	44

CHAPTER 4- METHODOLOGICAL AND STATISTICAL CHALLENGES

IN THE DESIGN OF A CLUSTER RANDOMIZATION TRIAL	45
4.1: Methodological issues in sample size calculation	45
4.1.1: General issues of sample size calculation	45
4.1.2: Impact of cluster randomization on design and analysis	47
4.1.3: Intracluster (intraclass) correlation and intercluster variation	48
4.1.4: Estimated means and their variances	53
4.1.5: Subsample size and statistical efficacy	56
4.1.6: Issues involving losses to follow-up	59
4.1.7: Strategies for achieving desired power	60
4.2: Methodological issues in data analysis of binary outcomes	63
4.2.1: General considerations - analysing binary (dichotomous) outcome data	63
4.2.2: Selecting the unit of analysis	64
4.2.3: The completely randomized design	69
4.2.3.1: Comparison of proportions	69

4.2.3.2: Standard Pearson chi-square test	70
4.2.3.3: Cluster level analyses	71
1. Two sample t-test comparing average values of the event rates	71
2. Non-Parametric approaches	72
4.2.3.4: Individual Level analyses	74
1. Adjusted chi-square approach	74
2. Ratio-estimator approach	75
3. Parametric modelling	76
4. Generalized estimating equations approach	77
4.2.3.5: Summary of results	80
4.2.4: Adjusting for covariates	81
4.2.5: Analysis strategies for trials with a small number of clusters per group	82
4.2.6: Sample size calculation	83
4.2.6.1: Estimation of the ICC for binary data	83
4.2.6.2: Estimation of the sample size	85

CHAPTER 5- POTENTIAL DESIGN CHALLENGES AND STUDY

LIMITATIONS	90
5.1: Introduction	90
5.2: Threats to internal validity	90
5.3: Threats to external validity	90
5.4: Minimizing commission and omission errors	92
5.5: Pilot study	95
5.6: Issues involving losses to follow-up	97
5.7: Strategies for achieving desired power	98
5.8: Missing data	99
5.9: Sensitivity analyses	100
CHAPTER 6- ETHICAL AND LEGAL ISSUES	101
6.1: Introduction	101
6.2: Source of funding and potential conflicts of interest	101
6.3: Ethical issues in cluster randomized trials	101
6.4: Health Canada approval and trial registration	104
6.5: Trial Monitoring	104

6.6: Patient confidentiality	105
CHAPTER 7- SUMMARY	106
7.1: Conclusions	106
7.2: Limitations of the study	107
7.2.1: Cluster level limitations	107
7.2.2: Patient level limitations	108
7.2.3: Physician level limitations	108
7.2.4: Limitations of the intervention	109
References.....	110
List of Figures.....	125
Figure 1:.....	125
Figure 2:.....	126
Figure 3:.....	128
List of Tables.....	129
Table 1:.....	129
Table 2:.....	130

Table 3:.....	130
Table 4:.....	131
Abbreviations.....	133
Appendix.....	134
Appendix1.....	134
Appendix2.....	138
Appendix3.....	139

Chapter 1 – Introduction and Review of Literature

1.1: Rising burden of hospital admissions for Atrial Fibrillation (AF)

Atrial Fibrillation (AF) is the most prevalent arrhythmia seen in clinical practice. The different clinical types of AF include-

First detected episode of AF: This can be symptomatic or asymptomatic and may be self-limited. There may be uncertainty regarding duration of the episode and about previous undetected episodes.

Paroxysmal AF: Is defined as recurrent AF (≥ 2 episodes) that terminates spontaneously within seven days.

Persistent AF: Is defined as AF, which is sustained beyond seven days, or lasting less than seven days but requiring pharmacologic or electrical cardioversion. Included within the category of persistent AF is “longstanding persistent AF”, which is defined as continuous AF or greater than one-year duration.

Permanent AF: In cases where all attempts at restoration of sinus rhythm, including pharmacologic agents, DC cardioversion and catheter or surgical ablation, has failed. This also is used to refer to the situation where both the patient and the treating physicians have concluded that it is futile to attempt restoration of sinus rhythm.

Recent onset AF: Is defined as AF with onset less than 48 hours prior to presentation. The patient can state the time of onset of AF, if he or she is confident of when AF started. In patients undergoing continuous cardiac monitoring, EKG strips can be used to confirm the time of onset of AF and its duration.

Atrial Flutter: Is a supraventricular arrhythmia that commonly occurs in association with AF. The typical form of AFI is characterized by a saw-tooth pattern of regular atrial activation called flutter (f) waves on the ECG, particularly visible in leads II, III, aVF and V1.

The prevalence of AF is roughly 0.1% in those younger than 55 years and increases to about 9% in those older than 80 years. Population studies have shown that over 80-90% of patients with AF have paroxysmal or persistent AF. The incidence of permanent AF in these studies is less than 10% (1). AF is associated with significant morbidity in the form of strokes, heart failure and adverse quality of life (resulting from frequent visits to the ED and/or hospital admissions). In addition AF is independently associated with increased mortality (2,3). Emergency Department visit rates for recent onset AF or atrial flutter (RAFF) increase with age, ranging from 0.2 per 1,000 US population for those ≤ 50 years to 6.7 per 1,000 US population for those aged ≥ 80 years and older (4). According to the US National Hospital Discharge Survey between 1985 and 1999 hospitalizations for AF increased from 154,086 to 376,487 for a first-listed diagnosis and from 787,750 to 2,283,673 for any diagnosis (5).

Atzema et al. conducted a retrospective cohort analysis of all patients with a primary ED diagnosis of AF or atrial flutter (AFI) seen in 158 Ontario EDs with at least ten or more annual ED visits for AF, between April 2002 and March 2008. The vast majority (80-90%) of these patients had paroxysmal or persistent AF. A total of 103,618 ED AF visits were made by 82,723 patients, increasing from 15,818 in fiscal year (FY) 2002 to 18,598 in FY 2007 and seems to reflect the trend seen in the US (6).

Patients admitted with a primary diagnosis of recent onset AF or AFI (RAAF; onset usually within the past 48-72 hours) had a mean length of hospital stay of 4 days and incurred hospital charges of approximately \$7,000 per admission (7,8,9). Understanding ED admission and management practices is essential for planning interventions to reduce hospital admissions for AF and in turn related health care costs.

1.2: Factors responsible for hospital admission of AF patients

The absolute number of ED visits for AF in the US increased 88% from 300,000 between 1993 and 1994 to 564,000 between 2003 and 2004. Of the patients presenting to the ED with a primary diagnosis of AF, 64% were admitted to the hospital and this admission rate remained constant during the 12 year-study period. Admission rates did not vary significantly with age. There were major regional differences in admission rates across the US varying from a low of 48% to a high of 76%. Patients who were admitted to the hospital had similar

characteristics to those who were not admitted. There were no significant differences in age, sex, metropolitan status or insurance type. Multivariate analysis identified the presence of CHF as the only factor associated with a higher risk of admission. However, only 14% of all admitted patients had CHF. This study failed to find any difference in the clinical status of patients that were admitted to the hospital compared to those discharged from the ED, suggesting lack of uniform criteria for admission of such patients. (4). Another study identified other factors responsible for increase in hospital admission of patients with AAF (acute AF- onset within past 48-72 hours; as defined by the authors) presenting to the ED. The recognition of the fact that anticoagulation can reduce the thromboembolic risk in patients with AF has resulted in increased admission rates in AAF patients for initiation of anticoagulation. The lack of access to follow up care is another potential factor responsible for lower threshold of hospital admission. (8,9,10). These findings suggest that AF admission is not routinely based on clinical characteristics of patients and that practice patterns, which vary by region, are arbitrary. In short many patients presenting to the ED with new onset AF may not need admission.

A study comparing ED utilization in the US and Ontario showed similar rates of ED utilization (11). Data from Ontario reflects the general trend of rising admission rates for AF patients from the ED as seen in the US and Europe. However, admission rates for patients presenting to the ED, with a primary diagnosis of AF, in the province of Ontario are generally lower than those in the

US and have held steady during the period 2002-2007 (40-45% vs. 60-76%) (4,6).

1.3: Consequences of hospital admission for AF

Hospital admission is associated with significant health care costs and it is estimated that patients admitted for AAF had a mean length of hospital stay of four days and hospital charges of approximately US \$7,000 (7). In a study from Hamilton Health Sciences the mean length-of-stay for patients admitted with a primary diagnosis of AF was 6.93 days, even though most of the patients in the study were at very low risk for vascular events (12). Inappropriate hospitalization also results in societal costs of low productivity, increased risk of complications and unnecessary procedures. Last but not least, admission to hospital can be a highly unpleasant experience for a patient (13,14).

1.4: Interventions to reduce hospital admissions for AF

Not all patients presenting with AAF will need hospital admission. Mulcahy et al estimated that approximately one third of patients presenting to the ED with AAF were stable enough to be safely treated as outpatients (4,5,15). In stable patients with AAF hospitalization can be avoided by using a combination of initial rate control and /or therapies to encourage conversion to sinus rhythm. In a retrospective analysis of 655 consecutive patients presenting to the ED of a teaching hospital with a diagnosis of AF, 280 (44%) patients were found to have AAF. Of these patients, 97% were managed in the ED with rate or rhythm

(pharmacological or electrical cardioversion) strategies by ED physicians and discharged directly from the ED (14). Other studies have suggested that over 90% of patients presenting to the ED with AAF (low risk for thromboembolism and without concomitant medical conditions requiring admission) can be safely managed and discharged directly from the ED (16,17,18). To manage the problem of increasing hospitalization of low risk AAF patients, there has been growing interest in an ED observation unit (EDOU) as an alternative to inpatient admission in the United States. In this setting, patients with other conditions are aggressively managed with accelerated treatment or diagnostic protocols, resulting in high discharge rates, lower costs, and improved patient satisfaction. These units are expensive to set up and although they reduce length of hospital stay they are not likely necessary for the typical low risk AAF patient. A few pilot trials have shown that implementation of a clinical pathway and provision of rapid access to outpatient subspecialty consultation for patients with AAF have resulted in reduction of ED visits and hospital admissions (19,20,21,22).

Relative merits and demerits of AF management strategies in the ED
(20,21,22,23,24,26,27,28):

Intervention	Merits	Demerits
Rhythm Control (Electrical/pharmacological cardioversion)	<ol style="list-style-type: none"> 1. High probability (95%) of restoring sinus rhythm and sending patient home (eg. Agg Ottawa Protocol, American studies see ref) 2. 50-60% of EDs in Canada use DCCV for eligible patients with AF 	<ol style="list-style-type: none"> 1. Not all patients will be eligible (unknown AF duration/ dur >48 hrs, as TEE cannot be done on an urgent basis in most Canadian centers) 2. Not all ED physicians centers use DCCV in hemodynamically stable patients
	<ol style="list-style-type: none"> 3. If effective patient reverts to sinus rhythm and DCCV is not required 	<ol style="list-style-type: none"> 4. Acute conversion rate only about 50% 5. Patients with LVH and repolarization abnormalities and QT prolongation will have to be excluded
Rate Control	<ol style="list-style-type: none"> 1. Can be used in a wider spectrum of patients (AF greater than 48 hrs, and in centers without the facilities for DCCV) 	<ol style="list-style-type: none"> 1. Most patients will still be in AF and therefore may make it difficult for the ED physician to discharge them 2. Choice between calcium channel blockers and beta blockers (Data shows that 60% of ED physicians use Metoprolol and 30% Diltiazem. 3. Review of literature does not show a significant advantage to using calcium channel blockers over beta blockers for rate control. Calcium channel blockers can

		be used in those patients with bronchial asthma.
Clinical Pathway	1. Flexibility to use any of the different strategies/medications	1. Cumbersome to use in clinical practice and may not be adhered to by ED physicians (especially if it is very complex).

1.5: Emergency department physician survey- Hamilton Health Sciences

A review of the literature has identified the fact that the incidence of ED visits and hospital admissions for AAF is on the rise. Published literature suggests that there is scope for improvement in management of these patients. At the same time there are many practical considerations preventing frontline emergency physicians, internists and cardiologists from adhering to guidelines and optimally managing AAF patients. In order to understand these factors we conducted a survey of ED physicians working in Hamilton Health Sciences (HHS) (Appendix 1). We received responses from five emergency physicians. The important issues identified from the survey are as follows-

1. There was no consistent protocol for management of low risk AAF patients.
2. Lack of rapid and easy access to an expert AF clinic was identified as an important factor influencing hospital admission.

3. Lack of resources including easy access to transesophageal echocardiography (TEE; to rule out left atrial clot) was an important factor preventing emergency physicians from adopting rhythm control strategies.
4. Selection of antiarrhythmic medications for pharmacological cardioversion was not always according to current guidelines.

1.6: Atrial Fibrillation Management- Hamilton Health Sciences

A retrospective analysis performed at the Hamilton Health Sciences (HHS) between June 2004 and September 2006 (26 months) provides insights into the management of patients presenting with a primary diagnosis of AF to the ED. The results of this study were presented at the annual (2008) scientific sessions of the Canadian Cardiovascular Society (CCS) (12).

A total of 546 patients were seen in the ED with a diagnosis of AF. The mean age of the patients was 68.4 ± 13.6 years. 70% of the patients had paroxysmal AF lasting less than 48-72 hours and without structural heart disease. The main symptoms at presentation were palpitations (51%), heart failure symptoms (48%), chest pain (36%) and dizziness (28%). 229 (42.4%) patients were discharged directly from the ED and 311(57.6%) patients were admitted for further management. The hospital admission rate at HHS was higher than the Ontario admission rate for patients presenting to the ED with a primary diagnosis of AF during the same time period (Range 40-43%) (6). The commonest indication for admission was inadequate heart rate control (86%). The mean time

taken for rate/rhythm control was 3.1 hours in the discharged group and 18.1 hours in the admitted group ($P < 0.001$). Of the admitted patients, 48% had low risk for thromboembolism with a CHADS₂ score < 2 . The CHADS₂ score is a clinical scoring system to estimate the thromboembolic risk in patients with AF. It takes into account the presence of congestive heart failure (CHF), hypertension, diabetes mellitus, age of the patient and history of prior stroke to risk stratify AF patients (Detailed explanation of CHADS₂ score – Appendix 1). In addition about half of the patients with RAAF were found to have no associated comorbidities (eg. acute coronary syndrome, pneumonia, pulmonary thromboembolism or CHF) that necessitated admission as per current guidelines (22,23,24). The main reason responsible for unsuccessful rate control was the use of subtherapeutic doses of rate limiting agents. For instance, in our study we found that the mean doses of calcium channel blockers, beta blockers and digoxin used for rate control of AAF in the ED were about 50% of the recommended dose in the current guidelines (11,22,23,24).

This study highlights the following points:

1. A majority (70%) of patients presenting to the ED at Hamilton health sciences have AAF without significant comorbidities
2. At least half of these patients probably could have been discharged directly from the ED after adequate rate of rhythm control

3. Inadequate heart rate control was the commonest factor responsible for hospital admission, due largely to inadequate dosing of drugs
4. Electrical and pharmacological cardioversion was tried in a minority of patients (<10%)
5. More information is needed to understand the reasons for admission and discharge of RAAF patients in the ED setting
6. There is an opportunity to improve care of AF patients, to reduce unnecessary hospitalization and possibly to improve outcomes

1.7: Variation in ED management of Atrial Fibrillation

Acute onset AF is commonly managed in the ED. However, there is insufficient evidence to help ED physicians choose between the two competing treatment strategies, rate vs. rhythm control. Stiell and colleagues conducted a cross-sectional survey of adult AAF cases in eight academic EDs during a 12-month period. The main aim of the survey was to understand the variation in practice among sites for managing RAFF in the ED (25).

Among the 1,068 patients included in the study 59.4% (interhospital range 42% to 85%) were managed with rhythm control. There was significant variation in management strategies used in each of these hospitals with regards to use of rate control drugs, choice of antiarrhythmic agent used for pharmacological cardioversion, use of electrical cardioversion and referral to the internal medicine

or cardiology service (who are responsible for making the decision to admit or discharge patients). 83 percent of patients presenting to the ED with AAF were discharged home in a stable condition (interhospital range- 73-90%).

Data from HHS, over a two-year period, showed that in a similar cohort of patients, 57% had to be admitted to the hospital as a result of inadequate rate or rhythm control (11). Atzema and colleagues conducted a retrospective cohort analysis of patients with a primary ED diagnosis of AF seen at the 158 EDs in Ontario, Canada, with more than 10 annual AF visits, during the period April 2006 to March 2008 (6). Logistic regression analysis was performed to compare mortality outcomes in patients discharged from the ED with two cohorts of admitted AF patients, defined a priori: (1) excluding patients with serious hospital discharge diagnoses (eg. acute coronary syndromes, pneumonia, congestive heart failure, etc) and (2) including only patients with hypertension, diabetes, hyperlipidemia, and COPD. This survey found that 12,320 (44.6%) patients with a primary diagnosis of AF and without significant comorbidity were admitted from the ED. One of the limitations of this large database study is that it does not identify the proportion of patients with RAFF, the kind of patient relevant to the proposed trial.

The data presented above highlights the variation in management of RAAF presenting to the ED. In addition it is clear that about half of such patients are admitted to the hospital. As described previously hospital admissions are expensive and may be avoided in some cases. At the same time the study from

the Ottawa region hospitals indicates a lower admission rate (3-16%) indicating that there is significant variation in admission rates within Ontario. Also the study from the Ottawa Hospital shows that with an aggressive management protocol about 97% of RAFF patients can be discharged home from the ED. This highlights the need for interventions to optimize management of AAF patients in the ED and reduce hospital admissions (25,26,27).

1.8: Variation in ED physician management of Atrial Fibrillation

The conventional approach to management of low risk RAFF by ED physicians is to initiate rate control and basic work-up in the ED followed by referral for definitive management. (25). The role of electrical or pharmacological cardioversion of low risk RAFF has not been clearly addressed in current guidelines, even though this has been proven to be safe and effective (26,27,28,29,30,31,32). Restoration of sinus rhythm is the best possible disposition for patients with RAFF presenting to the ED, both in terms of patient satisfaction and reduction in length of ED stay and associated health care costs.

Borgundvaag and Ovens conducted a survey of all members of the Canadian Association of Emergency Physicians (CAEP) to understand management preferences (rhythm vs. rate control) for low risk AAF patients presenting to the ED. They also explored associations between management practices and the characteristics of the treating physicians. A total of 663 responses were obtained from the 1255 physicians surveyed (52.8%). 622

(95%), 514 (78%) and 242 (38%) respondents reported routine performance of rate control, chemical cardioversion and electrical cardioversion respectively.

Physicians working in high-volume EDs (>50 000 visits/yr) were significantly more likely to self-manage rate control and chemical/electrical cardioversion than those working in lower volume EDs. Residency training was associated with higher performance of electrical (44% v. 31%, $p < 0.01$) but not chemical cardioversion or rate control, although, amongst residency trained physicians, those with FRCP-level training were significantly more likely to perform both chemical (86% v. 76%, $p < 0.05$) and electrical (57% v. 37%, $p < 0.01$) cardioversion (33).

This survey demonstrated that a majority of Canadian ED physicians actively manage low risk AAF independently. However, significant variations were noted especially related to the use of rate vs. rhythm control and in the use of pharmacological and electrical cardioversion. Some of the factors responsible for these variations were attributed to different practice environments, levels of training and lack of evidence to guide best practice.

This variation in individual physician practices in the management of AAF in the ED has important implications. The final disposition of a patient with AAF presenting to the ED will be influenced by the particular treatment strategy adopted by the treating ED physician ie. rate vs. rhythm control. This will also play a role in the calculation of the Intra-cluster Correlation Coefficient (ICC) and other statistical analyses of clinical trials assessing interventions in the management of AF in the ED. Variation in physician practice patterns will

influence the design of interventions to hospital admissions in patients with AF presenting to the ED. Interventions mandating rhythm or rate control strategies (including pharmacological /electrical cardioversion) exclusively will not be easy to implement given the fact that ED physicians require the flexibility of using rate and/or rhythm control strategy in patients with AF. Therefore, the ideal intervention in this clinical situation will be a clinical pathway that incorporates rate and rhythm control strategies that can be used concurrently or sequentially.

1.9: Clinical Pathways- definition and efficacy

Clinical pathways (also known as care pathways, critical pathways or care maps) are tools used to manage quality in healthcare concerning the standardization of health care processes. It is a tool that guides care management for a well-defined group of patients (particular case type eg: AF, subset eg: acute onset AF) for a well-defined period of time. The pathway states the goals and key elements of care based on evidence and best practice and organizes, sequences and times the major interventions of a multidisciplinary health care team (doctors, nurses, pharmacists etc.) in a structured fashion. The main goals of a clinical pathway are to detail essential steps in the care of patients, with specific clinical problems, in the hope of reducing physician and institutional level variance in care and medical errors. They also facilitate translation of national guidelines into local protocols, minimize delays and maximize resource utilization. In addition clinical pathways aim to improve quality

of care and communication with patients by providing a clearly written summary of care (34,35).

In 2003 a review of hospitals in the USA reported that more than 80% of hospitals had implemented one or more clinical pathways for management of health related conditions. There is debate about the effectiveness of clinical pathways as results of studies evaluating the application of clinical pathways are inconsistent and suffer from biases (34,35,36). A systematic review examining the effect of clinical pathways on length-of-stay, hospital costs and patient outcomes was undertaken to provide a framework for health care organization considering the effectiveness of clinical pathways as a management strategy. The review included randomized and non-randomized controlled trials evaluating clinical pathways in the management of adults and children in hospitals. Seventeen clinical trials representing 4,070 patients were included in the systematic review. In general, the majority of studies reporting economic data i.e. length-of-stay (LOS) and hospital costs showed a positive impact. Out of 16 reporting effects on LOS, 12 found significant shortening. Furthermore, in a subgroup-analysis, clinical pathways for invasive procedures showed a stronger LOS reduction (weighted mean difference (WMD) of 2.5 days versus 0.8 days). There was no evidence of differences in readmission to hospitals or in-hospital complications. The overall Odds Ratio (OR) for re-admission was 1.1 (95% CI: 0.57 to 2.08) and for in-hospital complications, the overall OR was 0.7 (95% CI: 0.49 to 1.0). Six studies examined costs, and four showed significantly lower

costs for the pathway group. However, heterogeneity between studies reporting on LOS and cost effects was substantial. As a result of the relatively small number of studies meeting inclusion criteria, this evidence is not conclusive enough to provide a replicable framework for all pathway strategies. Considering the clinical areas for implementation, clinical pathways seem to be effective especially for reducing hospital length of stay (36).

Data from the US, Europe and Canada indicate that about 50% of all AF patients presenting to the ED are admitted. Clinical pathways have been implemented in the ED to test their efficacy in reducing hospital admissions. Zimetbaum et al. implemented a clinical pathway for management of AF in the ED and compared hospital admission rates before and after implementation of the pathway. After implementation of the clinical pathway the admission rate for AF patients presenting to the ED dropped from 74% to 38% ($p < 0.0001$). The clinical pathway used in this trial was based on clinical guidelines and allowed the treating ED physician to use either a rhythm or rate control strategy for management of AF (20).

Stiell and colleagues instituted a clinical pathway for the management of low risk AF patients presenting to the ED at the Ottawa Hospital ED. The authors were able to discharge 97% of patients from the ED after a median interval of 5 hours after presentation to the ED. The group used protocol advocating pharmacological and/or electrical cardioversion as first line management and were successful in restoring sinus rhythm in 92% patients at the time of

discharge. Only 9% patients returned to the ED with AF within 7 days of discharge (27).

1.10: Need for clinical trial

Management of RAFF in the ED is variable at both the institutional and individual ED physician level (15,33). Many institutions and ED physicians adopt a conservative rate control management strategy and 40-50% of low risk RAFF patients are admitted from the ED every year in Ontario. At the same time the study from the Ottawa region hospitals indicates a lower admission rate (3-16%) indicating that there is significant variation in admission rates within Ontario. Also the study from the Ottawa Hospital shows that with an aggressive management protocol about 97% of RAFF patients can be discharged home from the ED. (5,6,8,9,12,17,25,27). Current guidelines for AF management offer very little guidance for management of RAFF, do not offer a clear mandate for either rhythm or rate control strategy (22,23,24).

The ability to restore sinus rhythm in the patient with RAFF presenting to the ED is ideal in low-risk RAFF patients who present to the ED. From the patient's perspective this is an ideal disposition, and in the absence of any complications related to the cardioversion, they can be discharged. However, many institutions and ED physicians do not routinely employ rate control strategies eg. pharmacological or electrical cardioversion due to a variety of reasons (15,33).

Relatively few studies have evaluated optimal management strategies for management of RAFF in the ED. Different investigators have evaluated rate and rhythm control strategies for managing RAFF patients in the ED. Investigators evaluating rate control strategy for management of RAFF patients in the ED have compared beta-blockers with calcium channel blockers or digoxin (18, 29, 30,31,32). A retrospective chart-review study performed by Stiell and colleagues has shown that a systematic approach to rhythm control using intravenous Procainamide followed by electrical cardioversion (in non-responders) was successful in restoring sinus rhythm in 92% of patients and 97% of patients were discharged from the ED (27). Small case series and randomized trials using pharmacological or electrical cardioversion in the ED have demonstrated that a rhythm control strategy is safe and effective for managing a majority of RAFF patients presenting to the ED (14, 16, 17, 18, 19).

One of the challenges of managing patients with RAFF in the ED is the fact that ED physicians and institutions vary in their preference for rhythm or rate control strategy (33, 12). In addition the absence of clear protocols for management of RAFF in the ED results in sub-optimal management as is evident from the study from HHS, where sub-therapeutic doses of rate controlling agents was used in over 50% of cases and 57% patients with RAFF were admitted from the ED (12). To overcome this issue clinical pathways have been developed for managing a variety of clinical conditions such as- acute coronary syndromes, pulmonary thromboembolism etc. (19,27,34,36). Clinical pathways are tools

developed for the management of a well-defined group of patients (particular case type eg: AF, subset eg: acute onset AF) for a well-defined period of time.

The pathway states the goals and key elements of care based on evidence and best practice and organizes, sequences and times the major interventions of a multidisciplinary health care team (doctors, nurses, pharmacists etc.) in a structured fashion. The main goals of a clinical pathway are to detail essential steps in the care of patients, with specific clinical problems, in the hope of reducing physician and institutional level variance in care and medical errors.

They also facilitate translation of national guidelines into local protocols, minimize delays and maximize resource utilization. In addition clinical pathways aim to improve quality of care and communication with patients by providing a clearly written summary of care (33, 34). A systematic review evaluating the effectiveness of clinical pathways in various medical conditions has demonstrated a positive impact on the length of hospital stay and health care costs (36).

Zimetbaum et al. implemented a clinical pathway for management of AF in the ED and compared hospital admission rates before and after implementation of the pathway. After implementation of the clinical pathway the admission rate for AF patients presenting to the ED dropped from 74% to 38% ($p < 0.0001$). The clinical pathway used in this trial was based on clinical guidelines and allowed the treating ED physician to use either a rhythm or rate control strategy for management of AF (20).

It is clear that management of RAFF is sub-optimal in many institutions and is in turn responsible for increase in hospital admissions and rising health care costs (6,7,8,9). Well-designed clinical trials are required to prove that a clinical pathway designed to provide evidence-based management of RAFF in the ED can reduce hospital admission rates (24,25). The choice of experimental study design ('before and after' study, randomized etc.) and the type of intervention (rhythm control strategy, flexible clinical pathway etc.) has to be selected in consultation with thought leaders in the field of emergency medicine and ED physicians. The issues involved in selecting the appropriate trial design and intervention will be discussed in detail in subsequent sections.

Chapter 2 – Study Design Considerations

2.1: Introduction:

As discussed in the previous section the choice of experimental study design ('before and after' study, randomized etc.) and the type of intervention (rhythm control strategy, flexible clinical pathway etc.) has to be selected in consultation with thought leaders in the field of emergency medicine and ED physicians. As described in this section the specific intervention to reduce AF admissions from the ED is debatable. We acknowledge that pharmacological and /or electrical cardioversion for rhythm control and rate control measures can be used as interventions in the ED. However, there is no consensus among ED physicians as to specific interventions (eg. pharmacological vs. DC cardioversion, rate vs. rhythm control). Therefore, we have decided to use an evidence based clinical pathway, offering ED physicians the flexibility to use rhythm (pharmacological and or electrical) and or a rate control, as the intervention to reduce hospital admission rates for RAFF patients. The trial design selected for the study should provide the strongest evidence for effectiveness of the intervention. At the same it should be feasible to implement the trial with available resources. In addition the intervention (rate vs. rhythm control, clinical pathway) has to be accepted and adopted by a majority of the stakeholders in the trial ie. ED physicians. These key issues have to be addressed by investigators planning to conduct a trial evaluating optimal management of RAFF in the ED.

2.2 Selection of appropriate trial design:

A number of potential study-designs are available to help investigators evaluate different interventions (in different settings for different targeted clinicians and behaviours) with respect to the degree to which they allow observed effects to be attributed to the intervention with confidence, resources needed to deliver the intervention and cost-effectiveness (37).

2.2.1: Observational Studies- Observational (descriptive) studies of single groups provide understanding of the management of clinical conditions and are useful for hypotheses generation. However, they are not useful for evaluation of interventions in the management of clinical conditions as they are unable to account for factors other than the intervention that may affect the outcomes being measured. It is never possible to completely rule out unrecognized bias with confidence even when recognizable differences have been adjusted for (37).

2.2.2: Quasi-experimental designs- These studies are conducted when there are practical and ethical barriers to conducting randomized trials (37,38). Some of the commonly used designs in guideline implementation studies include-

- i. Before and after studies (uncontrolled and controlled)- Uncontrolled before and after studies measure outcomes before and after the introduction of an intervention (eg. clinical pathway) and any

observed differences in outcome are assumed to be as a result of the intervention. These are relatively simple to conduct and are superior to observational studies. However, the single most important flaw of this design is that secular trends or sudden changes in incidence of outcomes make it difficult to attribute the observed changes to the intervention. In addition the effect of the intervention is confounded by the Hawthorne effect (the non-specific beneficial effect on performance of taking part in research), which could lead to overestimation of the effectiveness of an intervention. A meta-analysis evaluating studies with uncontrolled before and after design has confirmed the tendency of this study design to overestimate the benefits from interventions. In general this study design should not be used to evaluate the effects of guideline implementation strategies and results of such studies have to be interpreted with caution (37,39).

In controlled before and after studies a control population is identified which is expected to have secular trends or sudden changes in outcomes similar to the study population. Data is collected from both groups during the same time periods before and after the intervention is implemented. Any observed differences are assumed to be due to the effect of the intervention. The major challenges with this design include- a) difficulty in finding

comparable control group b) differences in outcome at baseline in apparently well matched control and study groups c) before and after outcomes are assessed separately for the control and study group and differences may be spurious. These limitations limit the confidence of the effectiveness attributed to interventions in these studies. Also, in many cases a randomized controlled trial may be just as easily undertaken as is a controlled before and after trial (37).

- ii. Time series designs- The purpose of time series designs is to detect whether an intervention has had a significantly greater effect on the outcome than the underlying trend for the particular outcome. This design is principally reserved for testing the effects of interventions when it is difficult to randomize or identify an appropriate control group. Data are collected at multiple time points before and after the intervention and this allows estimation of the underlying trend of the study outcome. Time series designs increase the confidence with which the change in point estimate can be attributed to the intervention. However, this cannot account for other clinical variables that may affect the outcome, thereby introducing bias. Often these studies collect insufficient time data points resulting in inappropriate statistical evaluation and overestimation of the effect of the intervention (37, 38)

- iii. Stepped wedge design (wait-list design)- In the stepped wedge trial design an intervention is sequentially rolled-out to the trial participants (either as individuals or clusters) over a number of time periods. There is random allocation of the intervention to the study subjects and at the end of the trial all study subjects or clusters would have received the intervention. Data is collected at each point where a new group (step) receives the intervention. There are two situations in which a stepped wedge design is considered advantageous to a traditional parallel randomized trial- a) if the investigators have a strong belief that the intervention is likely to do more good than harm, rather than a prior belief of equipoise, and it may be unethical to withhold therapy from a proportion of subjects (eg. AIDS trials), and b) situations in which logistical, practical or financial constraints mean that the intervention can only be implemented in stages and random allocation of the intervention is likely to be morally and politically acceptable. However, this trial design is likely to lead to a longer trial duration than a traditional parallel arm randomised trial. There is also the potential for contamination in situations in which a clinical pathway is being evaluated by exposing the care providers (ED physicians) in the proposed trial. Blinding of outcome assessors is very essential in this study design as often care-givers (ED physicians) cannot be

blinded. Finally, this study design has been applied in a limited number of situations and methodological descriptions including methods of randomisation, sample size calculations and statistical analyses have not yet been completely evaluated. In view of the above listed shortcomings it will be difficult to use this trial design for evaluating the proposed study question (38,40)

2.2.3: Randomized trials- Randomized trials are considered to be the most robust method of assessing health care innovations. They estimate the effect of an intervention through direct comparison with a randomly allocated control group that receives either no intervention or an alternative intervention. The randomization process usually ensures that both known and unknown variables capable of introducing bias are evenly distributed between the two comparison groups. When evaluating clinical pathway implementation single patient randomized trials may be less than ideal. There is the danger that the treatment offered to control patients will be contaminated by the ED physician's prior experiences of applying the experimental management strategy (the intervention) to patients in the intervention group. This may result in underestimation of the true effects of the clinical pathway (the intervention)(37,41,42,43).

To overcome these problems it is necessary to implement cluster randomization trials. A cluster randomization trial (c-RCT) is one in which intact units of individuals (cluster), rather than the individuals themselves,

are randomized to different intervention groups. Trials randomizing clusters are being widely used to evaluate non-therapeutic interventions, including- lifestyle modifications, educational programs and clinical pathways. The units of randomization in such trials are varied including households or families, neighbourhoods or communities and hospitals. The increasing popularity of this design among health researchers has produced an extensive body of methodology and clinical trials across multiple disciplines. Even though cRCTs overcome the problem of contamination to a large extent it has implications for the design and analysis of results as discussed in subsequent sections (37,41,42).

The factors listed below makes it necessary to conduct a c-RCT to evaluate clinical pathways in the management of RAAF in the ED:

1. Influence of the intervention on ED physician management practices: In the proposed trial to evaluate a clinical pathway for management of RAAF in the ED, the clinical pathway is the intervention. Individual patient randomization to treatment guided by the clinical pathway or standard institutional practice will introduce significant bias. An ED physician using the clinical pathway for managing some of his patients might alter his practice while managing patients randomized to the standard institutional protocol. This mandates that each ED will have to serve as the unit for cluster for randomization. In this fashion ED physicians will not be required to use two different protocols while

managing AF patients in the ED. This will prevent ED physicians from making changes in their practice to reflect the clinical pathway thereby, reducing the chance of bias.

2. Personal interactions among ED physicians – Physicians working in a particular ED might interact with each other and share their experiences with each other. A physician who has used the clinical pathway could influence the views of another physician working in the same ED introducing bias into the study. Cluster randomization will eliminate this scenario, as all patients in a particular ED will be managed using either the clinical pathway (intervention) or using the standard institutional protocol (control).

2.3: Selection of the intervention:

The optimal management strategy for RAFF patients in the ED is open to debate. The two major management strategies include pharmacological or electrical cardioversion (rhythm control strategy) and rate control of RAFF using medications (beta-blockers/calcium channel blockers). The relative merits and demerits of these strategies have been discussed previously.

However, EDs and ED physicians have demonstrated significant variation in their preferences (rate vs. rhythm control) for managing RAAF. This makes it very challenging (albeit not impossible) to compare rate and rhythm control strategies in parallel-arm randomized trials. Moreover, guidelines on the

management of AF in the ED have been ambiguous about the optimal strategy for management of AF in the ED and have left it to the discretion of the ED physician to select a strategy (rate vs. rhythm). A reasonable compromise would be to use a clinical pathway that provides the ED physician flexibility in terms of rhythm or rate control strategy while managing patients with RAFF. This is also useful in situations where the physician and patient are unsure of the duration of onset of RAFF, TEE facilities are not available to rule out the presence of LA or LA appendage clot and/or the ED physician is unable to or considers the patient at high risk for cardioversion under sedation.

As discussed in the previous section clinical pathways have shown to be effective in reducing hospital admission rates, length of hospital stay and health care related costs in a variety of clinical situations. Therefore, we have decided to use an evidence based clinical pathway as the intervention for the proposed cRCT. (8,9,14,16,17,20,22,23,24,25)

2.4: Study hypotheses

The primary objective of this study is to evaluate the effectiveness of a clinical pathway for the management of low risk AF patients presenting to the ED. The hypothesis is that a flexible and evidence-based clinical pathway will help ED physicians better manage AF patients, thereby reducing hospital admission rates (Figures 1 & 2 and Tables 1,2,3 & 4).

2.5: Study outcomes

2.5.1: Primary outcome - The primary outcome of this trial will be the proportion of low-risk, acute-onset AF patients admitted to the hospital from the ED.

2.5.2: Secondary outcomes –

- i. Proportion of patients in normal sinus rhythm at the time of discharge from the ED or admission to hospital
- ii. Proportion of patients with adequate heart rate control while in AF (a resting heart rate \leq 100 bpm for at least one hour) at time of discharge from the ED or admission to hospital
- iii. Proportion of patients returning to the ED with AF within the first thirty days of their index visit
- iv. Safety outcomes- A composite safety outcome including the following-
 - e) Hypotension needing intervention or admission
 - f) Bradycardia needing intervention or admission
 - g) Torsades de pointe or Ventricular flutter/fibrillation
 - h) Any adverse reaction to medications used for management of AF resulting in hospital admission

Economic analysis- a formal cost benefit analysis will be performed to assess the health cost savings (if any) as a result of the clinical pathway

2.6: Cluster definition and eligibility

2.6.1: Cluster inclusion criteria

- i. EDs located in teaching and non-teaching hospitals across Canada
- ii. EDs with an annual volume of > 140 low risk AF patients. The investigators will survey EDs in Canada to identify EDs with an annual low risk AF volume in excess of 140 patients.
- iii. Individual EDs or groups of EDs that are staffed by the same group of ED physicians will be the unit for cluster randomization. For example- the same group of ED physicians staff the EDs at the Hamilton General Hospital and the Juravinski Hospital. Therefore, these two EDs will be grouped as a single cluster for purposes of randomization.
- iv. Hospital ED physician group approval of the protocol
- v. Hospital administration approval of the protocol
- vi. Hospital research ethics board approval of the protocol

2.6.2: Cluster exclusion criteria

- i. Pediatric EDs

2.7: Unit of inference

2.7.1: Patient inclusion criteria

- i. Adult patients (age > 18 yrs) presenting to the ED with symptomatic AF or Atrial Flutter (AFI) as the primary problem

2.7.2: Patient exclusion criteria (Please refer to Appendix 2 for detailed explanations)

- i. Acute coronary syndrome
- ii. Pneumonia
- iii. Severe CHF or hemodynamic compromise
- iv. Major bleeding
- v. Severe exacerbation of bronchial asthma or COPD
- vi. Acute or chronic renal failure needing dialysis or hospital admission
- vii. Any other comorbidity at presentation that necessitates direct hospital admission

2.8: Participant recruitment

Study participants will be recruited from eligible EDs across Canada, when the ED physician has confirmed an EKG diagnosis of AF or AFI and the patient meets the required inclusion/exclusion criteria for the trial. The patient's name and basic demographic data will be entered into a pre-printed form. The ED physician will manage the patient as per standard institutional protocol or the AF clinical pathway depending on the cluster randomization. Patient information will be recorded in the provided data collection sheet and will be faxed to PHRI using the iDataFax system. To improve compliance with data acquisition in the emergency department a research nurse will help with collection of data during regular working hours. The nurse will also check data collection sheets of patients recruited between 5 pm and 8 am (outside of regular hours) to identify deficiencies and complete missing data fields. Research assistants will follow patient outcomes and will report to the investigators, adjudication committee members and the cluster representation mechanism group. The principal investigators will not be involved in recruiting study participants. When required they will provide advice and assistance to the most responsible physician in direct charge of participant care when there are doubts regarding EKG diagnosis of the presenting arrhythmia, interpretation of inclusion/exclusion criteria, withdrawal from the study or ethical issues.

2.9 Randomization procedures

The principal reasons for using randomization techniques in clinical trials is to ensure-

- i. Impartiality of allocation
- ii. Transparency for replication
- iii. To balance confounding factors
- iv. Blinding
- v. Formal justification for statistical inference

These factors are equally applicable to cRCTs. However, practical constraints may force investigators to recruit relatively small number of clusters and this may result in unbalanced comparison groups. This shortcoming of cRCT may necessitate the use of restricted randomization as described in the next section (39)

2.9.1: Restricted randomization

Restricted randomization involves selecting randomly from a set of allocations fulfilling certain restrictions. Stratified and matched-pair designs are examples of restricted randomization. Stratified randomization designs impose fewer constraints as allocations are selected from a larger set of the total number of clusters possible in unrestricted randomization. Stratified randomization may yield unbalanced comparison groups, especially when there are several variables under consideration. To overcome this limitation homogenous stratification criteria are first applied to the selected clusters and then additional pre-specified restrictions are used to ensure balanced comparison groups. Restricted stratification

allows comparisons to be made within homogenous strata that have been evenly balanced with respect to pre-specified clinical variables (39).

In the proposed study restricted randomization will be used to stratify clusters in the two comparison groups (EDs) based on a) the admission rates for RAFF patients from the ED, and b) whether the ED is located in a teaching/tertiary care hospital or in a non-teaching/community hospital. Teaching/tertiary care hospitals usually employ a multi-disciplinary approach between ED physician, specialists and sub-specialists. These centers also have medical students and residents involved in providing care to patients. In addition they are more likely to have institutional protocols for managing disease conditions. These centers typically have higher patient volumes and patients with a greater level of complexity with regards to disease state and management variables. On the other hand non-teaching/community hospitals have lower patient volumes and complexity. In addition in such hospitals the management strategies are more likely to revolve around individual physician preferences (40).

The stratified design is more likely to minimize bias by balancing the two treatment groups. In addition other factors likely to influence statistical analysis should be identified and accounted for during the design of the trial. Three general types of criteria are usually chosen for stratification and include- clinical covariates, statistical factors (eg.

sample size) and logistical/political factors. A computer program designed by the Population Health Research Institute (PHRI) computer sciences department will be used to stratify allocation based on the different variables identified. However, if such a technique makes site selection and random allocation very restrictive or impractical, some of the stratification criteria may have to be relaxed or abandoned altogether for the sake of feasibility (39, 40)

2.9.2: Allocation concealment

In cRCTs it is very important to ensure that subject allocation to treatment arms is concealed till the patient has been enrolled in the trial. Pre-emptive knowledge of the treatment arm to which a particular cluster may be randomized, may bias investigators and influence the decision to include the cluster in the trial. To prevent this it should be made clear during consent for recruitment of clusters that random allocation will be enforced, and that the assigned treatment protocol will be revealed to the cluster only after the first study subject has been recruited (38).

In the proposed trial after the stratified sequence generation is applied, EDs agreeing to participate will be assigned to institutional management of low-risk, acute-onset AF patients or to management using the proposed clinical pathway. Thus treatment allocation of each cluster is predetermined. This potentially introduces selection bias, both at the

cluster level and individual patient level. In order to ascertain the extent of and to minimize selection bias investigators audit case report forms and patient care documents from all clusters. The purpose of this audit is to identify whether all eligible patients were included in the study and if they were treated as per the assigned treatment protocol (clinical pathway or standard institutional protocol) (39,40).

2.9.3 Validity of restricted randomization

A completely valid cluster randomization protocol ensures that each pair of clusters has an equal probability of being assigned to each of the treatment groups. If this condition is not satisfied correlations may result between clusters in each intervention arm introducing bias. An invalid design will induce incorrect Type I error and incorrect coverage of confidence intervals (39). Investigators should check the validity of the randomization scheme by assessing the number of acceptable allocations. If the number of unrestricted (unacceptable) allocations is larger and the number of restricted allocations small (<100) significant bias can be expected. In such a situation it might be necessary to relax the stratification criteria to obtain acceptable number of clusters in each strata. (39,40).

Chapter 3 – Measurements and data collection

3.1: introduction

This chapter will address the selection of tools for measuring outcomes and issues related to data collection in the proposed clinical trial.

3.2: Definition of Atrial Fibrillation and Atrial Flutter

AF is a common supraventricular arrhythmia that is characterized by chaotic and uncoordinated contraction of the atrium. The common electrocardiographic (ECG) manifestations of AF include the presence of irregular fibrillatory waves and, in patients with intact atrioventricular (AV) conduction, the presence of an irregular ventricular response. The currently accepted classification of atrial fibrillation is as follows-

First detected episode of AF: This can be symptomatic or asymptomatic and may be self-limited. There may be uncertainty regarding duration of the episode and about previous undetected episodes.

Paroxysmal AF: Is defined as recurrent AF (≥ 2 episodes) that terminates spontaneously within seven days.

Persistent AF: Is defined as AF, which is sustained beyond seven days, or lasting less than seven days but requiring pharmacologic or electrical cardioversion. Included within the category of persistent AF is “longstanding

persistent AF”, which is defined as continuous AF or greater than one-year duration.

Permanent AF: In cases where all attempts at restoration of sinus rhythm, including pharmacologic agents, DC cardioversion and catheter or surgical ablation, has failed. This also is used to refer to the situation where both the patient and the treating physicians have concluded that it is futile to attempt restoration of sinus rhythm.

Recent onset AF: Is defined as AF with onset less than 48 hours prior to presentation. The patient can state the time of onset of AF, if he or she is confident of when AF started. In patients undergoing continuous cardiac monitoring, EKG strips can be used to confirm the time of onset of AF and its duration.

Atrial Flutter: Is a supraventricular arrhythmia that commonly occurs in association with AF. The typical form of AFI is characterized by a saw-tooth pattern of regular atrial activation called flutter (f) waves on the ECG, particularly visible in leads II, III, aVF and V1. If untreated the atrial rate typically ranges from 240 to 320 beats per minute with f waves inverted in ECG leads II, III and aVF and upright in lead V1. The direction of activation of the right atrium (RA) may be reversed, resulting in upright f waves in the above mentioned ECG leads.

AFI may degenerate into AF, and AF may convert to AFI. In addition occasionally the two arrhythmias may be confused with each other especially if

fibrillatory atrial activity is prominent in more than one ECG lead. In the ED the management steps for AF and AFI are identical and hence for the purpose of the study any patient with symptomatic or asymptomatic AF or AFI presenting to the ED will be included in the study, provided he/she meets the inclusion and exclusion criteria. (22, 23, 27)

3.3: Primary outcome

In cRCTs a single pre-defined primary endpoint that is statistically significant carries more weight than differences in composite endpoints compared to RCTs. For this reason the proposed cRCT will use the proportion of low-risk, acute-onset AF patients admitted to the hospital from the ED as its primary outcome variable in the context of the specific intervention (management as per the clinical pathway versus management as per institutional practice). The trial will allow the treating ED physician five-hours to manage the patient in the ED. The exact duration of ED management after which the patient will be admitted to the hospital or discharged may vary among institutions. The investigators will survey different EDs to develop consensus on the duration of ED management after which the patient will be admitted to the hospital if discharge criteria are not met. We are using a 5 hour duration has been tentatively used in this protocol based on a survey from HHS ED physicians (Appendix 1). If the pre-specified clinical endpoints for discharge from the ED are not met within this time period or if the patient's clinical status deteriorates he/she will be admitted to the hospital. All other patients will be discharged from the ED.

An independent adjudication committee will confirm primary and secondary outcomes after a detailed review of the data collection sheets and patient care records. The committee will be blinded with regards to the cluster intervention allocation. The committee will consist of three members, with at least one cardiologist and one ED physician. None of the principal investigators will be part of the adjudication committee.

3.4: Secondary outcomes

The secondary outcomes of interest include-

- i. The ability of the clinical pathway to improve heart rate control and/or restore sinus rhythm in patients with RAFF.
- ii. The rate of recurrent ED visits within thirty days of the index visit. The expectation is that the clinical pathway will provide AF patients with rapid access to a specialist AF clinic that can provide appropriate and ongoing management of AF resulting in fewer repeat ED visits.
- iii. The following safety endpoints (section 2.3.1.iv) will be collected. For the purpose of the study the following definitions will be used (18,19, 20, 23, 26)-
 - a. **Death:** Death from any cause. Confirmation of death within 30 days will be done by review of hospital records and records from the coroner's office.

- b. **Hypotension:** Is defined as a drop in systolic or diastolic blood pressure that necessitates intravenous fluid resuscitation, discontinuation of antiarrhythmic or rate limiting cardiac medications, use of positive inotropes and/or necessitates immediate admission to the hospital.
- c. **Bradycardia:** Defined as a heart rate < 50 bpm that necessitates interventions such as temporary transcutaneous or trans venous pacing and/or hospital admission.
- d. **Torsades de pointe or Ventricular flutter/fibrillation:** These are rapid, sustained ventricular arrhythmias resulting in hypotension or cardiac arrest.

3.5: Follow-up

Patients will be contacted telephonically at 30 days to find out if they have been to the ED for a repeat visit. In addition research staff will review ED records, data collection forms and patient hospital records from all participating clusters to identify secondary outcomes. In the event of a safety outcome or death, all hospital records and the coroner's report will be reviewed. In case a study patient is readmitted to a non-participant cluster relevant hospital records detailing the management of the patient will be obtained and the patient will be contacted telephonically to confirm out of protocol visits. ED databases from participating centers will be checked to identify patients lost to follow up and to determine if they had subsequent visits to EDs not enrolled in the trial. Data will be collected

for both the occurrence and place of occurrence of safety outcomes (i.e. ED, hospital or out of hospital).

3.6: Data collection

The ED physician treating the study participant will be responsible for entering relevant clinical information in the data collection form. This information will be transmitted to PHRI using the iDataFax system and will be securely stored in a password-protected server. To improve compliance with data acquisition in the emergency department a research nurse will help with collection of data during regular working hours. The nurse will also check data collection sheets of patients recruited between 5 pm and 8 am (outside of regular hours) to identify deficiencies and complete missing data fields. A research assistant and statistician will periodically review the database to solve queries and enter missing data. Off line secure intranet data acquisition software will be available in every cluster for back up in case telephone and internet connections are not accessible at the time of patient recruitment. Finally, a hard copy data collection form will be provided in case a computer is not available. All participating physicians will be trained in data entry during workshops prior to the start of the study.

Chapter 4 – Methodological and statistical challenges in the design of a cluster randomization trial

4.1: Methodological issues in sample size calculation

4.1.1: General issues of sample size calculation

A quantitatively justified sample size calculation is a fundamental pre-requisite for a well designed randomized controlled trial and the cluster number should provide sufficient statistical power to detect differences between groups considered to be of clinical interest (41). The major issues in sample size calculation in cRCTs are (42,43):

1. Appropriate formulas for calculating sample size are relatively complex and inaccessible
2. The proper use of sample size calculation formulas requires prior assessment of the intracluster correlation coefficient either directly or through comparable information on the between cluster component of variation. Both of these parameters may not be very familiar to investigators, complicating the task of obtaining relevant past data that may be used for sample size calculation
3. Studies enrolling hundreds or even thousands of patients may give the misleading impression of extensive statistical power, when in fact the effective sample size, after taking into account the clustering effect, is actually quite small.

4. The power of a cluster randomization trial depends more on the number of units randomized than on their size. However, the number of units that can be realistically studied may be small due to reasons of logistics and cost.
5. Limited resources may also lead to a diluted intensity of effect in the experimental group, especially if these resources have to be spread over a wide geographic area thereby limiting the number of clusters included in trials. Trials enrolling clusters across large countries like Canada face this feasibility issue.
6. Issues with blinding, where members of the control group become aware of the intervention, result in dilution of the effect of the intervention. This will result in bias as a result of an inadequate sample size. However, this is not likely to be an issue in the proposed cRCT as patients will not be aware whether they are being managed according to standard institutional practice or the proposed clinical pathway.
7. Loss to follow up due to the unique challenges related to the cRCT design can result in an inadequate sample size, especially if this factor is not accounted for in sample size calculations. This is not an issue in the case of the proposed cRCT as the primary outcome is determined at the end of the index ED room encounter.
8. Many cRCT trials are prevention trials testing the effect of the intervention in a relatively healthy population. The event rate of the

primary outcome may be low and patient compliance with the intervention may be poor. Both of these factors may result in over-estimation of the benefits of the intervention. This in turn will result in an under-estimation of the sample size required for an adequately powered experiment.

Therefore, the major concern in cRCTs at the community level is that in large-scale public health interventions the size of the effect may be meagre in relation to the effort expended. Often the available resources may be inadequate to detect even medium effects. When a small sized effect is all that is meaningfully expected from an intervention, it is very important to make sample size calculations taking into account all the above-mentioned factors. This will increase the probability of detecting an intervention expected to produce only a small sized effect (44,45)

4.1.2: Impact of cluster randomization on design and analysis

Many of the challenges encountered in the design of cRCTs arise because inferences are frequently intended to apply at the individual level, while randomization is at the cluster level. As is often the case inferences regarding the efficacy of an intervention have to be made at the individual level. In this situation the lack of statistical independence among members in a cluster will invalidate standard approaches that ignore the cluster effect (those used in patient level randomized controlled trials) to both

estimation of sample size and analysis of the trial data. Application of standard sample size formulas will lead to under estimation of the required sample size. In addition applying standard statistical methods to cRCT with the assumption that there is no between cluster variation will tend to bias the observed p-values downward (i.e. exaggerate the p-value), thus risking a spurious claim of statistical significance. Another commonly committed error in cRCTs is randomization by cluster accompanied by a sample size assessment appropriate to randomization at the individual level. This leads to a type I error substantially above that which is planned for and the latter may lead to a substantially elevated type II error. (46,47). In addition to the dramatic effect of cluster randomization on sample size considerations and on the approach to statistical analysis there are several other issues related tot the conduct and interpretation of cRCTs. These include issues related to the role of stratification, blinding, informed consent and loss to follow up.

4.1.3 Intracluster (intraclass) correlation and intercluster variation

The degree of similarity among responses within a cluster of randomization is typically measured by a parameter known as the intracluster (intraclass) correlation coefficient (ρ or ICC). This parameter may be interpreted as the standard Pearson correlation between any two responses in the same cluster. A positive ICC is equivalent to assuming

that the variation between observations in different clusters exceeds the variation within clusters. Under these conditions the design is characterized by between cluster variations, which means that the clusters cannot be assumed to be interchangeable with regard to the experimental endpoint (37,38, 43,47, 48).

The lack of interchangeability of variation between clusters will differ from trial to trial may result from the following factors (49,50,51,52,53):

1. **Subject selection:** For example, in a trial randomizing medical practices, the characteristics of patients belonging to a practice could be related to age or sex differences among physicians. In addition the outcomes of two or more patients treated by the same physician or group of physicians (eg. ED physician group) could share the influence of that physician group's treatment style or preferences.
2. **Influence of covariates at the cluster level:** It is well known that postoperative outcomes including success and complication rates vary systematically among surgeons. The effect of clustering is particularly strong in health care trials in which the intervention is applied to the provider of care rather than the patient directly. The between cluster variation will reflect the variation among surgeons and the differences in individual patients comprising the cluster.

3. Tendency of conditions to be more prevalent within the cluster:

Environmental, genetic and socio-economic factors may vary between clusters in a trial affecting the degree of between-cluster variation.

4. Effect of personal interactions between individuals in a cluster

randomized to a particular intervention: Educational strategies or therapies provided in a group setting could lead to a sharing of information or practices that might result in a clustering effect.

The ICC is dependent on the relationship of the between to within-cluster variance and is defined as the proportion of the total variation in the outcome that can be attributed to the difference between clusters.

Suppose we measure some kind of patient outcome y from multiple patients (indexed by the subscript j) and from multiple clusters (indexed by the subscript i). Then we can think of the j^{th} observation in the i^{th} cluster as:

$$y_{ij} = \mu_i + e_{ij}$$

where, μ_i ($i = 1$ to k , the number of clusters) is the true mean for the i^{th} cluster and e_{ij} ($j = 1$ to m_i , the number of patients in cluster i) is part of the measurement, which causes one patient's observation to be different from another's in the same cluster (ie. within-cluster variation or σ_w^2) and is simply the variance of e_{ij} .

In addition, perhaps due to the patient mix (potentially explainable through patient covariates such as age and sex) or, more likely, heterogeneity in clinical practice/expertise, one must anticipate extra variation from cluster to cluster over and above that contributed by patient-to-patient variation within the cluster. In other words, the y_{ij} cluster round a slightly different mean μ_i for each cluster. Unless, in truth, all $\mu_i = \mu$ (ie. all equal; an unlikely scenario), the μ_i themselves contribute an extra “between cluster” (inter-cluster) source of variation; denoted by σ_b^2 . Therefore,

$$\text{ICC}(\rho) = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_w^2} = \frac{\sigma_b^2}{\sigma^2}$$

where σ_b^2 is the between-cluster variance component and σ_w^2 is the within-cluster variance component (figure 3).

The relative size of the between and within-cluster variance components is best quantified in terms of the ICC, which in words is the proportion of total variance in an observation that is being contributed by between-cluster sources. Rearranging the formula above we can see that,

$$\sigma_w^2 = (1 - \rho) \sigma^2$$

and,

$$\sigma_b^2 = \rho \sigma^2$$

A number of papers have detailed ICC estimates in community, school-based and worksite studies. However, ICC estimates from epidemiological studies and surveys are not readily transferable to the implementation research setting. The primary reason for this is that information at the appropriate cluster level (for example, the ED) required for implementation research studies cannot be calculated from population based surveys and also because the outcomes associated with implementation research focuses on measures of behaviour change (54,55,56,57). This is best measured using process variable, such as adherence to guidelines or compliance with best practice recommendations. In addition the setting in which the study is conducted affects ICC estimation. In primary care settings, the ICCs for process variables appear to be of an order of magnitude higher compared with those for outcome variables. This can be explained by a greater than expected variation in patient compliance and behavior as compared with physician behavior. In addition, ICCs for process variables from secondary care settings are generally of an order higher than those obtained from primary care settings. This is likely related to more consistent management practices within hospitals as compared with general practices, where general practitioners act fairly autonomously (58). In light of the above factors a range of possible values of ICC need to be considered when estimating sample size. The reasons being:

1. The estimates of ICC are dependent on the study design. For instance the estimate of ICC from a study employing stratified randomization may be smaller than that obtained from a trial using a completely randomized design.
2. Estimates of ICC from a small number of clusters may be erroneous.
3. Currently, very little information is available concerning the generalizability of estimates of ICC (59).

4.1.4: Estimated means and their variances

It is reasonable to assume that the within-cluster measurement component e_{ij} is statistically independent from the between-cluster variation in μ_i so that the total variance of y_{ij} (which we will separately designate as σ^2) is,

$$\begin{aligned}\text{Var}(y_{ij}) &= \text{Var}(\mu_i) + \text{Var}(e_{ij}) \\ &= \sigma_b^2 + \sigma_w^2 = \sigma^2\end{aligned}$$

The i^{th} cluster mean would be estimated in the conventional way, summing the observations within the cluster and dividing by the number of observations. The observed cluster mean $\bar{y}_{i\cdot}$ would estimate the true cluster mean μ_i ,

$$\bar{y}_{i\cdot} = \frac{\sum_{j=1}^{j=m_i} y_{ij}}{m_i} = \hat{\mu}_i$$

and,

$$\text{Var} (\bar{y}_{i_j}) = \frac{\sigma_w^2}{m_i}$$

The overall mean \bar{y} would estimate the true mean (μ) of the cluster means

(μ_i):

$$\bar{y} = \frac{\sum_{i=1}^{i=k} \sum_{j=1}^{j=m_i} y_{ij}}{\sum_{i=1}^{i=k} m_i}$$

and,

$$\text{Var} (\bar{y}) = \frac{\sigma_b^2}{k} + \frac{\sigma_w^2}{\sum_{i=1}^{i=k} m_i}$$

If all $m_i = m$ then,

$$\text{Var} (\bar{y}) = \frac{\sigma_b^2}{k} + \frac{\sigma_w^2}{k \times m} \quad \text{formula (1)}$$

Clearly, increasing the cluster size m will “dampen down” the effect of the within-cluster variance to produce a more reliable estimate of μ_i . However, only increasing the number of clusters (k) can reduce the influence of the between cluster variance. The optimal choice of k and m depends on the relative sizes of the within-cluster (σ_w^2) and between-cluster (σ_b^2) variance.

The relative size of the between and within-cluster variance components is best quantified in terms of the ICC, which in words is the

proportion of total variance in an observation that is being contributed by between-cluster sources (see section 3.1.3 for details)-

$$\sigma_w^2 = (1 - \rho) \sigma^2$$

and,

$$\sigma_b^2 = \rho \sigma^2$$

Substituting this in the formula (1) described earlier in the and rearranging leads us to,

$$Var(\bar{y}) = \frac{\sigma^2}{km} [(m - 1) \rho + 1]$$

Note that the quantity $(m - 1) \rho + 1$ is the proportional increase in variance created by the clustering effect (sometimes called the design effect).

When conducting a cRCT with two treatment arms each in k clusters with m patients per cluster, the variance of the difference in means would be the sum of the two individual mean variances. Assuming that each group has similar sized variances this would be given by the formula (58,61):

$$Var(\bar{y}_1 - \bar{y}_2) = 2 \frac{\sigma^2}{km} [(m - 1) \rho + 1]$$

4.1.5: Subsample size and statistical efficacy

Some cRCTs are designed to enrol only a sample of eligible cluster members, i.e. adopt a subsampling strategy. Subsampling comes into play once the number of units (schools, hospitals) for randomization has been selected and a method has been identified for sampling subjects (inclusion criteria). Increasing the subsample size affects the precision with which the ICC is estimated. However, this does not enhance its underlying value. To understand this the relative effect on power of increasing the number of clusters versus increasing the subsample size needs to be studied. For sample size calculation, statistical efficiency can be assessed for different fixed values of ICC while varying the number of clusters per intervention group, k and the subsample size, m . The variance of the mean difference $\bar{Y}_1 - \bar{Y}_2$ as computed over all study participants is given by the equation:

$$Var(\bar{Y}_1 - \bar{Y}_2) = \frac{2\sigma^2}{km} [(m - 1)\rho + 1]$$

From this equation it is clear that as k (the number of clusters) becomes large, the variance $[Var(\bar{Y}_1 - \bar{Y}_2)]$ approaches zero, implying that the power of the study may be increased indefinitely by recruiting more clusters. However, as m (the number of subsamples within a cluster) becomes large with the number of clusters (k) fixed, the variance $[Var(\bar{Y}_1 - \bar{Y}_2)]$ approaches a lower bound limited by the value of σ^2 (the

variance in the response of an individual to the intervention and is the sum of the variance within [σ_w^2] and between clusters [σ_b^2]). This means that the power for testing the effect of an intervention will then become largely insensitive to the value of m (the number of subsamples within a cluster) (60). Therefore, it is clear that up to a point, the same level of precision for estimating the effect of intervention may be achieved by choosing different values of m and k . In some trials the value of m will be fixed due to practical considerations. At times there may be an opportunity to increase k at the expense of m , or vice versa. In making this choice caution should be exercised in reducing the number of clusters (k) and increasing the subsample size (m) simply on statistical grounds. This is because reducing the number of clusters randomized has the effect of limiting the generalizability of the results. In addition there is generally much more uncertainty in the likely value of the between cluster variance [σ_b^2] compared to the within cluster variance [σ_w^2]. Thus, a strategy that places too much faith on a low ICC (small between cluster variance [σ_b^2] compared to the within cluster variance [σ_w^2]) is not desirable. From the statistical perspective increasing the number of clusters randomized while reducing the average subsample size may be the preferred strategy (49, 61).

Another scenario often encountered while calculating sample size for cRCTs is when the investigators cannot exceed a specified maximum

value for clusters that can be randomized due to cost and practical considerations. However, some control can be retained over the subsample size (m). In such an event the sample size requirements are specified in terms of the number of subjects required per cluster. If k_{max} denotes the maximum number of cluster that can be assigned per intervention group, then the number of subjects required per cluster (m) is given by the equation:

$$m = \frac{(1 - \rho)}{\left(\frac{k_{max}}{k_{(m=1)}} - \rho\right)}$$

where,

$$k_{(m=1)} = \frac{(t_{\alpha} + t_{\beta})^2 (2\sigma^2)}{(\mu_1 + \mu_2)^2}$$

denotes the number of clusters per intervention group which would be required if $m=1$. The value t has $2(k_{max} - 1)$ degrees of freedom and

$$\frac{k_{max}}{k_{(m=1)}} = \frac{[(m-1)\rho + 1]}{m}. \text{ The restriction imposed by these algebraic}$$

relationships is that the subsample size must be less than $\frac{k_{max}}{k_{(m=1)}} (58)$.

The use of a student's t-distribution rather than normal distribution is attributed to the fact that the number of clusters and subsamples are usually small and the distribution often cannot be assumed to be normal.

4.1.6: Issues involving losses to follow-up

The possibility of lost to follow-up is a potentially serious source of bias in all randomized trials and can be particularly important in cRCTs having a relatively long follow-up period. For instance, a 5%/year loss to follow-up in a cohort of subjects measured at baseline reduces the size of the cohort by nearly 25% over 5 years. Trials, in which the interventions are applied at the cluster level, with little or no attention given to individual study participants, may result in a much higher overall attrition rate. This has led to investigators adopting a strategy of oversampling along with aggressive follow-up to account for this.

Cluster RCTs must also deal with the possibility of entire clusters dropping out. Provided that the reasons for loss to follow-up are not related to intervention, oversampling at baseline will help to compensate for both types of losses. In addition oversampling also helps to account for the inevitable uncertainty associated with prior assessments of parameters needed for the application of sample size formulas, which are in any case approximations. Another source of bias is non-adherence to the randomization scheme in trials where the effectiveness of the intervention depends directly on the subject's active participation. In addition refusal to participate in the trial post-randomization is an extreme form of non-adherence resulting in bias. Both these forms of bias can be accounted for

by using an “intent-to-treat” analysis. Oversampling of cohorts is needed to account for these two situations (62).

4.1.7: Strategies for achieving desired power

One of the greatest challenges facing investigators during the design of cRCTs is adequately powering the trial to test the primary hypothesis. A number of strategies have been identified to reduce the chance of under powering trials. These are summarized below (61,63,64,65):

1. **Strategies to reduce between cluster variability:** Trial subjects and clusters are rarely representative of the general population limiting the ‘external validity’ of trials. On the other hand, investigator-driven selection criteria can help increase the power of trials. This is accomplished by placing geographic or other relevant restrictions on the clusters to be randomized; thereby reducing between cluster variability. However, this may substantially reduce the generalizability of the clinical trial, especially if the inclusion criteria are very restrictive and result in selection of a study population that is not representative of the general population.
2. **Justified data collection:** There is natural tendency on the part of investigators to request large amounts of data from study subjects. This can reduce subject compliance and retention rates of clusters and

subjects within clusters. Therefore, all data points that need to be collected in the trial should be carefully justified.

3. **Realistic estimates of minimally important effect of the intervention:** Overestimation of the beneficial effect of an intervention and the availability of limited number of available clusters for recruitment can reduce the power of the trial. Therefore, realistic estimates of the beneficial effect of an intervention, adjustments for loss to follow-up and treatment 'cross-over' should be used for sample size calculation.
4. **Increased cluster recruitment in the control group:** Greater gains in power can be obtained by increasing the number of clusters enrolled rather than by increasing the number of subjects recruited in each cluster. In cases of administrative or economical difficulties limiting cluster recruitment the potential gain in power achieved by increasing the number of subjects recruited to each cluster and the number of clusters recruited to the control arm of the trial can be explored.
5. **Using surrogate and combined endpoints:** To increase the power of trials, where infrequent binary outcomes (such as death) are the primary endpoint, using composite endpoints (major adverse cardiac events in cardiovascular trials) and surrogate markers for the primary outcome (low infant birth weight for perinatal mortality) should be used. Particular care should be used when dichotomizing continuous

endpoints since this can result in profound loss of power. An alternative approach would be to adjust an analysis based on cluster means for cluster-aggregated covariates such as age or socio-economic level.

6. **Stratified or matched-pair designs:** Matched pair or stratified designs are useful only when baseline variables strongly related to the primary outcome (cluster covariates) is used for the analysis. However, this strategy comes at the expense of limitations and increased complexity of the matched pair analysis.
7. **Use of repeated measurements:** A minimal gain in power may also be obtained by taking repeated assessments over time from the same subjects or from different samples of subjects. In such cases the gain in efficiency obtained through the use of repeated patient measurement in cRCTs will be via suppressing within-cluster variation only and must, always be balanced against the added cost and administrative complexity that will be involved.
8. **Increasing consistency of the intervention:** In implementation trials evaluating clinical pathways or care plans meticulous training of care givers (physicians/nurses) will result in consistent use of the intervention. This will reduce the variation in treatment effect between clusters and improve precision of the estimate of benefit.

4.2: Methodological issues in data analysis of binary outcomes

4.2.1: General considerations - analysing binary (dichotomous) outcome data

Methods for analysing binary (dichotomous) outcome data in cRCTs are not as well established as those for continuous outcome data. Consequently, binary outcome data have been treated as continuous in statistical analyses or clustering has been completely ignored. Analytic issues are complicated by the absence of a unique multivariate extension of the binomial distribution analogous to the multivariate normal distribution. In addition, there is no single statistical approach that is uniformly superior in this situation. Thus multiple analytical approaches need to be considered in the analyses of even the simplest cRCTs where clusters of fixed size are randomized without baseline covariate adjustment at either the cluster or individual level. Moreover, most of the available methods need a large number of clusters per intervention group in order to assure validity. As a general principle, the statistical methods based on individual within-cluster observations must adequately account for the (potentially different) stochastic nature of the individual and cluster-level sources of variation.

The simplest approach to analysing binary outcome data from a cRCT is to obtain a single summary score for each cluster and to conduct

the analysis at the same level as random allocation using standard statistical methods. However, sound scientific practice should adopt an approach that corresponds to the underlying design. Methods specifically suited to the completely randomized trial design, the matched-pair design and the stratified design have been described. However, only analytic issues related to binary outcomes in the completely randomized trial design will be discussed here as the proposed cRCT, evaluating the clinical pathway for management of patients with low risk AF in the ED, will not involve a matched-pair or stratified design.

Most RCTs are designed with the expectation that the randomized cohorts will be of approximately equal size. However, as a result of disproportionate loss to follow-up of study subjects, and, more importantly, the variation in cluster size, the study groups have a greater probability of being unequal in size. Even in cRCTs involving units like medical practices, hospitals etc., some variation in cluster size and number can be expected.

4.2.2: Selecting the unit of analysis

The unit of inference in cRCTs can be analyzed at the individual or cluster level. In many trials an interest in cluster level inferences leads to collection of data only at the cluster level. The challenge of selecting a unit of analysis is particularly acute when data are available from individual

study subjects. This may result in uncertainty regarding what is the most appropriate level of inference.

When individual – level data is available; the lack of statistical independence among individuals within a cluster has to be accounted for. One way of handling this issue is by using meaningful summary measure, such as the cluster average value, to serve as the unit of analysis. Standard statistical methods can then be applied to the collapsed data. This largely removes the problem of non-independence since the subsequent significance tests and confidence intervals would be based on the variation among cluster summary values rather than on the variation among individuals. However, variation in the cluster size will create heteroscedasticity in the cluster means and the complication of having a mixture of two estimated variance components with different degrees of freedom clearly threaten the validity of such comparisons.

Cluster level analyses can also be used for complex summary scores (e.g. trend over time) and, more generally, for any study outcome. Cluster level analyses are most appropriate when the primary outcomes of interest focus more on the randomized unit as whole than on the individual subjects and, in particular, when the between-cluster variance is likely to dominate (ie. a large ICC). In the proposed cRCT evaluating a clinical pathway for the management of AF in the ED the intervention (clinical pathway) is administered at the cluster level. This means each ED

(cluster) will be either randomized to the clinical pathway (the intervention) or standard institutional protocol (control). The primary outcome is the proportion of patients admitted to the hospital from the ED in each cluster. Therefore, it seems quite appropriate and natural to analyze at the cluster level in the proposed trial.

Another important scenario is the analysis of quantitative outcome variables when a cluster has a fixed number of subjects. In this situation exact statistical inferences concerning the effect of intervention can be constructed using standard analysis of variance. The resulting test statistic is also algebraically identical to the test statistic obtained using a cluster-level analysis. Therefore, for a quantitative outcome variable a cluster level analysis is fully efficient when all clusters are of the same size. However, this equivalence for cluster level and individual level analyses seen with quantitative variables holds only approximately for binary outcome variables (50).

The number of subjects per cluster will exhibit considerable variability either by subject attrition or by design in practice. This requires weighted cluster-level analyses to account for such variability. The precision of these analyses is maximized when the cluster level summary score is weighted inversely proportional to the reciprocal of its estimated variance. This means that the optimal weights are usually a function of the cluster sizes and the degree of intracluster correlation. If there are only a

small number of clusters per intervention group the estimation of weights will be imprecise leading to loss of power relative to an unweighted analysis. Thus, in practice large gains in power due to weighting are not assured. Moreover, when there are a large number of clusters, appropriately weighted cluster-level analyses are equivalent to individual-level analyses (66, 67, 68).

The cluster-level approach to analysis may not be statistically the most efficient approach in situations where clusters vary widely in size. The cluster size variation will result in heteroscedasticity and impact the validity of results due to differing variances in the clusters. Adjusting for imbalances in baseline predictors of study outcome may result in potential gain in power. Individual level regression analyses using regression models is preferable in this situation as they can be used to examine the effects of both cluster level and individual level predictors. Estimates of the ICC are also more accurately obtained during individual-level analysis. The most commonly employed regression models include those based on random effects models, such as: Poisson regression for event rates, logistic regression for proportions and linear mixed models for quantitative outcomes. An important advantage of individual-level regression methods is that the effects of modeled covariates are estimated and presented in conjunction with the intervention effects. However, in general, individual level analyses tend to be more complicated and hard for the reader to

understand. They also tend to rely a lot on strong assumption that may or may not be appropriate. If these assumptions are not true the gain in precision may not be accurate (69, 70).

Challenges in identifying and distinguishing the unit of inference, unit of analysis and the impact of clustering are not unique to cluster randomization trials. Data arising from cRCTs may be analyzed at the patient level, family level, health care provider level or practice level. The resulting analyses can be more complicated than is the case of simple RCTs. Therefore; decisions regarding the unit of analysis and the impact of clustering depend much more on subject matter expertise, provided clustering is not a consequence of study design (70).

Advantages of cluster level analyses:

1. Easy to conduct and explain
2. Applicable to any outcome variable
3. Permits construction of exact statistical inferences
4. Allows adjustment for baseline imbalance and/or variability in cluster size
5. Provides power comparable to individual level analyses when properly weighted

Advantages of individual level analyses-

1. Standard statistical procedures can be used in the absence of clustering effect (ie. relatively modest ICC) and, when there are relatively few clusters.
2. Efficiently handles the interplay of within and between cluster variations.
3. Allows direct assessment of the joint effects of cluster-level and individual- level predictors of outcome.
4. Can be extended to include analyses of multilevel data.
5. Yields more precise estimates of ICC
6. Capable of providing more efficient estimates of the effect of intervention than unweighted analyses, when there are many clusters per group with highly variable cluster sizes.

4.2.3: The completely randomized design

4.2.3.1: Comparison of proportions

The primary objective of cRCT design is to compare the proportion of individuals (patients with AF in the proposed trial) in different intervention groups (clinical pathway versus standard institutional protocol) that have a specified characteristic (hospital admission) by a fixed point in time. We will use a two-group intervention model in our discussions (71).

4.2.3.2: Standard Pearson chi-square test

Investigators often use this test to assess the effect of intervention in cRCTs. The fundamental assumption while applying this test is that the sample observations are statistically independent. This assumption generally holds true for simple RCTs. However, in cRCTs this assumption is almost always violated, as it is more likely that responses taken on subjects within a cluster are more similar than response taken on subjects in different clusters i.e. ICC is positive. This implies that analyses performed using standard chi-square tests is likely to push p values down, thereby giving the false impression that results are more statistically significant than they really are. The magnitude of the bias with this test increases with the value of ICC and the average cluster size. Similar concerns also exist with other standard procedures for comparing proportions, such as Fisher's exact test. The basic problem here is that both these tests essentially assume that each individual in a particular treatment group has an identical probability (π) of exhibiting a successful outcome. In the context of a cRCT this is unlikely to be true with each cluster having a distinct outcome rate (π_i). This extra source of, presumably random, between-cluster variation is simply not allowed for by these tests. In most situations this leads to an exaggeration of the p value (ie. overestimates the treatment effect). Inappropriate application of such statistical analyses to clustered data is sometimes referred to as the 'unit

of analysis error'. The above-mentioned analytical methods are not going to be used for data analysis in the proposed RCT (71,72).

4.2.3.3: Cluster level analyses

- 1. Two sample t-test comparing average values of the event rates:** A variety of alternative approaches have been adopted to avoid the pitfalls involved in applying the Pearson chi-square test to clustered data. The assumption used in this statistical approach is that cluster-specific event rates are normally distributed across clusters and share a common variance within each arm of the trial. The test is used to determine if the difference between the average values of the event rates in the 2 groups of the trial is statistically significant using the standard two-sample (unpaired) t-test. A theoretical objection that can be raised against this approach is that the assumptions required for the validity of the t-test are not strictly satisfied in this setting. This is certainly the case if there is considerable variation in cluster size, since the cluster proportion will have variance: $\pi(1-\pi)/m_i$ even if there is no additional cluster effect (ie. ICC=0). Nevertheless, extensive simulation research has shown that the t-test is remarkably robust to violations of the underlying assumptions, both in terms of the distribution/homogeneity and varying strength of the ICC. This is particularly the case when there are equal numbers of clusters assigned to each intervention group, which would be expected in a cRCT.

Another criticism that may be raised is that the test completely ignores any variation in cluster size, giving each of the observed event rates the same weight. One way of overcoming this bias is to use weights proportional to the reciprocal of the variance of the mean in each group. This can be used provided there are enough clusters and that the ICC is accurately estimated. A corollary to this is that when all clusters are of large size, the variation in weighted mean depends less on the number of units within a cluster (m), and unweighted analyses will tend to be as efficient as weighted analyses (71,73,74).

2. **Non-Parametric approaches:** A non-parametric approach that makes no assumptions about the distribution of the event rates may be preferable for testing the effect of an intervention. If the null hypothesis of “no treatment effect” is true one can argue that an observed cluster outcome rate would be the same irrespective of which treatment was allocated to the cluster. Under this assumption, whatever difference is observed in mean outcome rate between treatments can only be due to the play-of-chance inherent in randomization. For the class of non-parametric tests, the sampling distribution of the mean treatment difference can be derived empirically by re-randomizing the observed cluster outcome rates (a randomization test) or mathematically based on the ranks of the cluster outcome rates (eg. A rank-based test such as the Wilcoxon Rank Sum Test). The Wilcoxon rank sum test also known as the Mann-Whitney U-

test, requires that the two intervention groups be pooled and the event rates ranked by size. The underlying rationale is that if the intervention has no effect, the sum of the ranks for each group should be the same. There is no question as to the validity of applying non-parametric methods to the ranks of the original observations. However, the validity is gained at the expense of precision (i.e. a reduced power) compared to the t-test. The t-test's improved power is largely derived from the willingness to accept the assumption of normality (which may be incorrect), but also the loss of information in using ranks rather than the cluster outcomes rates themselves. This loss of information is reflected by the fact that unless the number of clusters per group is at least four, it is impossible to achieve two-sided statistical significance at $p < 0.05$, regardless of the magnitude of the intervention effect. Another non-parametric approach that does take into account the values of the cluster-specific event rates is Fisher's two-sample permutation test. This procedure is based on considering the number of different ways in which the cluster-specific event rate could be permuted (randomized) between intervention groups while maintaining the same number of clusters per group. No other distributional or model-based assumptions are made. An appropriately selected test statistic (e.g. the difference in mean event rates) is then calculated for each permutation. The two-sided statistical significance of the test is equal to the proportion of test statistics found using the permuted data that are at

least as large, in absolute value, as the test statistic found using the observed data. Akin to the previously mentioned non-parametric tests, at least four clusters per group are needed in a two-group comparison to obtain two-sided statistical significance at $p < 0.05$ (75,76).

4.2.3.4: Individual Level analyses

There are several alternative approaches to cluster level analyses that may be conducted at the individual subject level. The relative merits and demerits of these methods are described below.

1. **Adjusted chi-square approach:** An adjustment to the chi-square test was proposed by Donner and Donald and depends on computing clustering correction factors for each group. An assumption behind this adjustment approach is that the correction factors for both the groups being compared estimate the same population design effect i.e. they are not significantly different. This assumption holds good for randomized experiments, but may not hold true for non-randomized comparisons in which clusters are systematically assigned to receive different interventions. The validity of the adjusted chi-square test statistic does not depend on the assumption that the pairwise correlation between any two observations in the same cluster is constant. In the case of a heterogeneous correlation structure within a group, the ICC measures the average degree of pairwise correlation and the adjusted chi-square

statistic remains valid. If the true within-cluster correlation structure were known, adjustments to the chi-square statistic could be developed making this test more efficient. However, in real life such information is rarely available to put this into practice. (77,78,79).

- 2. Ratio-estimator approach:** This approach is also based on a relatively simple adjustment of the standard Pearson chi-square statistic. The concept of this test is based on regarding the event rate as a ratio rather than as a proportion. The ratio estimator approach simply involves dividing the observed counts in the Pearson chi-square statistic by an estimate of the design effect, thereby reducing the effective sample size. The other feature of this analysis is that this procedure does not explicitly involve the notion of an ICC at all. Therefore, unlike the adjusted chi-square statistic the ratio estimator chi-square statistic does not approach the standard Pearson chi-square statistic if the ICC approaches zero. The advantage of the ratio estimator approach is that it does not require the assumption that the population design effects in the two comparison groups are equal. Therefore, this approach is well suited to non-randomized comparison, particularly those involving systematic differences in mean cluster size from group to group. The number of clusters per group required to ensure the validity of the ratio estimator approach may be fairly large (at least 20 per group). From a statistical perspective this approach works best under conditions that ensure the

optimal performance of the ratio estimator as recommended for its use in sample survey research: a large number of clusters and a relatively small degree of variation in cluster size. However, there are various modifications of this method that may improve its performance in small to moderate sized samples (80, 81, 82)

3. **Parametric modelling:** In this approach the subject responses are assumed to follow a specified probability distribution. One proposed distribution is a beta-binomial distribution, which results when:

- i. The number of responses within a cluster follows a binomial distribution conditional on the cluster specific success rate, and
- ii. The cluster-specific success rates are assumed to vary across clusters in accordance with a beta distribution.

This implies greater variability than a standard binomial distribution, but will reduce to the latter when the ICC is zero. In this sense the beta-binomial distribution is 'overdispersed' relative to the binomial distribution.

This test assumes a common value of the ICC in each group, but not necessarily a common design effect. This is also a computationally intensive technique, requiring an iterative solution using numerical maximization techniques. Likelihood ratio procedures based on other parametric distributions may also be developed for testing the equality of event rates in cluster randomization trials. These include the logistic-normal (logit; inverse of the sigmoidal logistic function; converts data to a

normal distribution) likelihood ratio test and the probit-normal (inverse cumulative distribution function; converts data to a near linear distribution) likelihood ratio test. These differ from the beta-binomial likelihood ratio test with respect to the specific assumptions made regarding the nature of the between-cluster variability. The first of these procedures assumes that the logit transform of the cluster-specific success rates follows a normal distribution across clusters and the latter imposes a similar assumption on the probit transform of the cluster specific success rates. One of the advantages of these methods is that they can be used to adjust for model dependence on individual level covariates. A disadvantage is that the results of these tests are very difficult to interpret. The major disadvantages of likelihood ratio tests are that they are computationally intensive and are very tied to parametric assumptions. Also, these tests may be unreliable in smaller samples (<40 clusters per group) and don't perform well in case of deviations from the assumed model. However, their optimal statistical properties in large samples compensate for these disadvantages. (81, 83, 84)

4. **Generalized estimating equations (GEE) approach:** The GEE approach, developed by Liang and Zeger, can be used to construct an extension of standard logistic regression, which adjusts for the effect of clustering. In the absence of clustering this model reduces to that of standard logistic regression. A principal advantage of the GEE approach

as compared to the parametric methods is that specification of an underlying distribution for the sample observations is not necessary.

There are two distinct approaches to adjust for the effect of clustering using this approach:

- i. The first is a model-based approach, as it requires the specification of a working correlation matrix that describes the pattern of correlation between responses of cluster members. For cRCTs the simplest assumption to make is that responses of cluster members are equally correlated, i.e. exchangeable. If all covariates are measured at the cluster level, such an exchangeable correlation matrix is equivalent to assuming that the average correlation among cluster members does not vary among clusters. The validity of statistical inferences constructed using this strategy is assured provided the average ICC can be assumed to be constant across clusters, similar to the assumption for the adjusted chi-square test described earlier.
- ii. The second strategy that is used with GEE to adjust for the effect of clustering employs 'robust variance estimators' that are constructed using between cluster information. These estimators estimate the true variance even if the working correlation matrix is misspecified. If there is a large number of clusters, inferences obtained using robust variance estimators will become equivalent to those obtained

using the model-based strategy when the working correlation matrix is correctly specified. An advantage of this method is that simple algebraic expressions are available to construct the resulting statistical inferences when comparing responses of subjects from two or more independent groups. This method needs a fairly large number of clusters per intervention group (at least 20).

In the absence of covariate adjustment, procedures simpler than GEE are likely to give similar results, and are preferred in practice. The principal advantage of GEE approach is that it can readily be extended to adjust for a combination of cluster-level and individual-level covariates, yielding consistent and normally distributed estimators of regression coefficients.

(85,86)

4.2.3.5: Summary of tests (adapted from 85):

Type of Test	Individual Test Type	Salient Features
Standard statistical tests to test binary outcomes without accounting for clustering	Standard Pearson chi-square test for comparing proportions	p-value biased downward in the presence of clustering; tends to overestimate benefit of an intervention
Cluster level analyses (parametric and non-parametric)	Two sample t-test	Required assumptions are not satisfied; however, statistically very robust
	Wilcoxon rank sum test	Non-parametric based on ranks; validity is achieved at the expense of statistical precision
	Fisher's two-sample permutation test	Non-parametric based on observations; no distributional or model based assumptions; applicable for small cluster size (as low as 2 per intervention arm)
Individual level analyses taking into account the effect of clustering	Direct adjustment of Pearson chi-square test	Reduces to standard Pearson chi-square test when ICC= 0
	Ratio estimate of chi-square test	Requires large numbers of clusters; suited to non-randomized comparisons

	Likelihood ratio test based on parametric modeling	Heavily model-dependent; computationally intensive; needs large number of clusters per intervention group
	Generalized estimating equations	Requires large number of clusters; useful for including cluster and individual-level covariates in the analysis; computationally intensive

4.2.4: Adjusting for covariates: Analyses that adjust for baseline covariates (both at the cluster and individual level) are usually aimed at reducing bias in the estimated treatment effect caused by randomization imbalance but may also yield improved precision. Random assignment both by subjects and clusters generally ensures that, on an average, baseline covariates will be equally distributed or balanced across intervention groups. Nonetheless, it is possible that some residual, yet substantively important, imbalance may occur by chance. For a given total number of individuals, the probability of an important imbalance in covariates will be high in cRCTs, owing to the lower effective sample size. A complicating feature of such analyses in cRCTs is that there may be different associations at the individual and cluster levels. The GEE extension of logistic regression described in previously is capable of accounting for both individual and cluster level covariates. Adjustment for

individual level covariates using this technique results in very modest reduction in ICC. However, adjustment for covariates accounting for between cluster variations is likely to substantially reduce between cluster variation and thus the estimated ICC. The investigation of separate associations within and between clusters is only relevant for baseline covariates that vary within each cluster. Such an analysis is not possible for covariates such as cluster size or intervention group. In addition individual and cluster level adjustments are possible only in trials with large numbers of clusters and cannot be carried out in trials with small number of clusters (<15) (76, 85, 86, 88)

4.2.5: Analysis strategies for trials with a small number of clusters

per group- A special class of studies are community intervention trials, in which fairly small numbers (often <10) of relatively large units, such as medical practices, factories or entire cities are allocated to different intervention groups. In such situations methodological issues are associated with a continuous outcome variable. When the response is dichotomous (binary) and the number of cluster is small (<10), most of the methods discussed in the previous section such as adjusted chi-square approach, the ratio estimator approach and the GEE method become more questionable, since the large sample approximations underlying these procedures become less valid. Evidence suggests that in such situations the two-sample t-test as applied to the cluster-specific event

rates may be safely applied to data arising from trials with as few as three clusters per group, at least when there is little or no variation in cluster size. This remarkable robustness of the t-test persists even in the presence of deviations from underlying assumptions of normality and homogeneity of variance. However, having only as many observations as there are clusters results in a between cluster variance estimate with very few degrees of freedom and thus low power. An alternate approach to this test would be to obtain exact tests of significance using non-parametric procedures such as the Wilcoxon rank sum test or Fisher's permutation procedure. The limitation is that at least four clusters per group are required to perform these tests. Also, using these tests to analyze studies with small cluster sizes (eg. four clusters) will result in p-values that are not considered conventionally significant. These recommendations should not be taken to support studies with small numbers of clusters since the limited power of these studies remains a problem. The attention to the problem of controlling type I error should not distract investigators from losing sight of a possible type II error. These methods should be applied only when administrative and financial constraints restrict the number of clusters that can be recruited to a study (50, 74, 89,).

4.2.6: Sample size calculation

4.2.6.1: Estimation of the ICC for binary data- Several methods have been proposed to estimate the ICC for binary data. These methods

include analysis of variance (ANOVA) estimator, moment estimators, estimators with a direct probabilistic interpretation, estimators based on direct calculation of correlation within each group, and extended quasi-likelihood and pseudo-likelihood estimators. Simulated experiments have shown that the ANOVA estimators, a few of the moment estimators, and an estimator with a direct probabilistic interpretation performed well with low bias and small standard deviation. Of all these estimators the ANOVA method is extensively used and is valid for both dichotomous and continuous variables.

If we denote X_{IJ} ($i = 1, \dots, k, j = 1, \dots, m_i$) as the binary response for k clusters with the i th cluster containing m_i individuals. Then the total number of successes in the i th cluster is $Y_i = \sum_j^{m_i} X_{IJ}$ and the total number of subjects in the sample is $N = \sum_{i=1}^k m_i$. The ANOVA estimator of intraclass correlation is defined as follows:

$$\hat{\rho}^{AOV} = \frac{MSB(\sigma_b^2) - MSW(\sigma_w^2)}{MSB(\sigma_b^2) + (m - 1)MSW(\sigma_w^2)}$$

where MSB is the mean square between clusters, MSW is the mean square within clusters, and m is the cluster size. When the cluster size varies m_0 is substituted for m in the above equation. Where, m_0 is the average cluster size. For binary data MSB, MSW and m_0 can be defined as follows:

$$MSB (\sigma_b^2) = \frac{1}{k-1} \left[\sum_{i=1}^k \frac{Y_i^2}{m_i} - \frac{1}{N} \left(\sum_{i=1}^k Y_i \right)^2 \right]$$

$$MSW (\sigma_w^2) = \frac{1}{N-k} \left[\sum_{i=1}^k Y_i - \sum_{i=1}^k \frac{Y_i^2}{m_i} \right]$$

and

$$m_0 = \frac{1}{k-1} \left[N - \frac{1}{N} \sum_{i=1}^k m_i^2 \right]$$

In most human studies, a negative ICC value does not make sense; therefore, in such cases, if $MSB < MSW$ then ρ (ICC) is set at 0.

For the proposed study we used data from Stiell's retrospective health records review of practice variation in academic EDs and the SAS® statistical program (Cary, NC, USA) to calculate the $\hat{\rho}AOV$ (ICC; ANOVA technique). This trial had twelve clusters (ED) from the province of Ontario providing a total of 1068 patients with low risk AF presenting to the ED. The average cluster size in this trial was 134 patients. In total 178 patients were admitted to the hospital from the ED, giving an overall hospital admission rate of 0.1667. The calculated $\hat{\rho}AOV$ using this data was 0.0257 (25, 90, 91, 92).

4.2.6.2: Estimation of the sample size- Comparison of proportions: Once an ICC estimate is obtained, we can calculate the sample size needed for

the study. In the proposed cRCT we plan to compare the hospital admission rates of low risk AF patients presenting to the ED in two treatment arms- one group of EDs (clusters) randomized to management using the clinical pathway (experimental group) and the second group of EDs randomized to management using institutional protocol (control group). Assume that k clusters of m individuals each are randomly assigned to each group i ($i = 1, 2$), where $i=1$ denotes the experimental group and $i=2$ denotes the control group. The null hypothesis $H_0: P_1=P_2$ at the two-sided 100α per cent level of significance with power $1-\beta$, where P_1 and P_2 are the admission rates in the experimental (AF pathway) and control (institutional management protocol) groups, respectively.

The required number of subjects (n) per intervention group is given approximately by the formula-

$$n = \frac{(Z_{\frac{\alpha}{2}} + Z_{\beta})^2 [P_1(1 - P_1) + (P_2(1 - P_2))][1 + (m - 1)\rho]}{(P_1 - P_2)^2}$$

Where, $Z_{\frac{\alpha}{2}}$ and Z_{β} denote the critical values of the standard normal distribution corresponding to the error rates α and β , respectively and ρ is the ICC (see previous section for ICC calculation).

The required number of clusters per group (k)=

$$k = n/m$$

For the purposes of our study the average hospitalization rate for low risk AF patients presenting to the ED in Ontario (P2) is 0.1667 (16.67%). The ICC (ρ) calculated from the study by Stiell and colleagues is 0.0257. If we consider a clinically meaningful relative reduction in ED admission rate of 25%, then $P2 = 0.125$ (12.5%). Substituting these values in the equation above gives us a sample size adjusted for clustering of 4939 subjects per intervention group (power of 80%). The number of clusters required per intervention group, based on the study duration, is then calculated as follows-

Study duration of 1 year = $4939/134 = 37$ clusters

Study duration of 2 years = $4939/2 \times 134 = 18.5$ clusters

Study duration of 3 years = $4939/3 \times 134 = 12.3$ clusters.

A range of sample and cluster sizes can be calculated based on the relative risk reduction (RRR) values, duration of follow up etc. These projections are summarized in the table below-

(87, 90, 91, 92, 93, 94)

Duration of Study	RRR = 20% Clusters needed per intervention group	RRR = 25% Clusters needed per intervention group	RRR = 30% Clusters needed per intervention group
One year	60	38	26
Two years	30	19	13
Three years	20	13	9
Total Number of Patients needed for the study	7975	4939	3483

The above sample size calculation is based on data from Stiell's retrospective chart analysis of practice variations in academic EDs in Canada. This review indicated a hospital admission rate of 16.7% for RAFF patients from the ED. This value is much lower than the hospital admission rate for similar patients from the HHS ED (57%) and for all EDs in the province of Ontario (40-45%). Under ideal circumstances the sample size calculations should have used these values. However, the absence of the complete data set providing variance and other information needed to perform sample size calculations prevented us from using this data.

We will conduct an intention-to-treat analysis at the individual level using proportions and means according to the variable in question with an alpha level of 0.05 and power of 0.80 for the primary outcome. EDs in Ontario with an annual volume of >140 RAAF visits will be selected for the trial. The proposed cRCT

(assuming a 30% RRR) will be conducted over a two year time period. A total of 13 clusters will need to be recruited in each intervention arm (a total of 26 clusters for the trial). A total of 3500 ED visits for RAFF will need to be included from these 26 clusters.

We have used Stiell's data, even though it may not be an accurate reflection of the data from all EDs across Ontario, to provide an example of the steps and assumptions involved in sample size calculation for the proposed cRCT. Given the higher hospital admission rates for RAFF in the province of Ontario we expect the sample sizes to be significantly lower than shown above. However, this can only be assumed till formal sample size calculations are performed using accurate data. This is an important issue that will have to be addressed to ascertain if it is feasible to perform a cRCT to assess the clinical pathway.

Chapter 5- Potential design challenges and study limitations

5.1: introduction

The purpose of this chapter is to discuss the sources of bias and the potential shortcomings of the proposed trial. The following topics will be discussed in this section-

5.2: Threats to internal validity

Internal validity refers to the extent to which, differences identified between treatment groups in a randomized trial, is a result of the intervention being evaluated. A well-designed and conducted trial with appropriate statistical analysis will ensure minimal bias. With respect to cRCTs, a great deal of emphasis has been placed on sample size calculation and accounting for clustering of patients (as discussed in section 3). A robust sample size and accurate calculation of the ICC will improve the internal validity of the trial.

The inability to blind caregivers to outcome assessment is another drawback of this cRCT. In our study the adjudication committee members will be blinded to the cluster intervention allocation in order to minimize bias (96, 97,98).

5.3: Threats to external validity

External validity refers to the generalizability of the study findings to similar or related individuals or care settings in the context of medical trials. The external validity of interventions in cRCTs can be evaluated using quality

frameworks which include- Reach, Efficiency, Adoption, Implementation and Maintenance (The RE-AIM Framework developed by Glasgow and colleagues).
(40, 99).

Dimension	Definition
Reach	An adequate description of the extent to which patients included in the trial are representative of the population of interest
Efficacy	Success rate of the intervention in preventing the primary outcome if implemented appropriately
Adoption	An adequate description of the extent to which the study settings are representative of the wider population of settings
Implementation	The extent to which the intervention is implemented as intended in the real world. This comprises of – <ul style="list-style-type: none">i. Acceptability- adherence to intervention components

	<p>targeted directly at health care professionals in clusters</p> <p>ii. Feasibility- Extent to which the health care professional adheres to the intervention (clinical pathway in this case)</p>
Maintenance	This refers to the sustainability of the program over time

The reader will be provided with the following information to help understand the extent to which the trial intervention may be applicable to related patient populations and care settings-

- a) Inclusion criteria for clusters and patients
- b) Clusters recruitment rate
- c) Description of the clusters and individual patient clinical variables
- d) Factors listed in the CONSORT statement for reporting cRCTs

5.4: Minimizing commission and omission errors

An error in the context of clinical medicine is defined as a commission or omission with potentially negative consequences that would have been judged wrong at the time it occurred by expert peers, independent of any negative outcomes. This excludes the natural history of the disease that does not respond

to treatment or foreseeable complications of a properly performed intervention, as well as cases in which there is substantial disagreement as to whether an error occurred. Errors of omission are passive occurrences such as- missed diagnoses or failure to read a laboratory report. On the other hand errors of commission involve incorrect action that is usually not indicated, such as- infusing an inappropriate medication or ordering incorrect diagnostic tests (100,101).

Minimizing omission and commission errors (defined as failure to admit a patient whose AF has not been appropriately controlled and admitting a patient with adequate rate or rhythm control of AF and without any significant comorbidity) is important to ensure an unbiased result. The following steps will be applied to each cluster randomized to the clinical pathway arm to minimize errors-

- i. All ED physicians will be trained during workshops so that they can get accustomed to the steps of the clinical pathway and clarify any doubts regarding clinical definitions and bedside implementation of the pathway
- ii. They will be evaluated with clinical simulations to test their adherence to the clinical pathway. This will involve a period where physicians use the proposed clinical pathway to manage patients in the ED and collect data. An audit of the patient records will be performed to assess compliance with the clinical pathway. It will also help in obtaining feedback from front-line ED physicians regarding the clinical pathway. This will assist in making any changes to the pathway. This exercise should be ideally

carried out before finalizing the clinical pathway and during the pilot phase of the clinical trial.

- iii. A failure mode effect analysis (FMEA) will be applied to identify, anticipate and remedy steps in the clinical pathway that are likely to lead to errors (102). The steps will include:
 - a. Analysis of the clinical pathway to identify potential sources of ambiguity and errors
 - b. Assemble an expert multi-disciplinary team of investigators to administer and manage the trial
 - c. Identify potential failures in implementation of the clinical pathway (patient eligibility, errors in management steps at the patient or physician level etc.)
 - d. Create easy to use flow-charts outlining the study flow and management steps
 - e. Modify the clinical pathway if required, and
 - f. Identify tools to measure and track performance (clinical trial steering committee).
- iv. An electronic tool will be designed to aid ED physicians and support staff to systematically evaluate patient eligibility and criteria for discharge from the ED.

5.5: Pilot study

Pilot studies provide preliminary evidence on the clinical efficacy of an intervention and are routinely performed in many clinical settings. These are also commonly known as “feasibility” studies and are designed to assess critical issues that may influence the clinical trial design. The factors necessitating a pilot trial are listed below –

- i. Process: To assess the feasibility of implementing the steps included in the main trial
- ii. Resources: This step assesses the budgetary and time constraints that may influence the main trial. During the pilot trial data is collected on the time and money spent on implementing the intervention.
- iii. Management: Assesses issues that may be encountered in data collection and personnel management at participating clusters.
- iv. Scientific: Assesses adherence to the clinical pathway, effect of the intervention being evaluated and statistical variance within and among clusters.

It is important that the sample for the pilot trial (patients and clusters) be representative of the target population. It should have the same eligibility criteria as the main trial and should be large enough to provide meaningful data regarding the factors that are being evaluated in the design of the main trial. The disadvantage is that samples (clusters) cannot be used in the main

trial if the randomization and blinding are not ensured in the pilot trial. If any pooling of pilot and main data is considered this should be planned beforehand, described clearly in the protocol with respect to the statistical consequences of such a strategy. The goal is to avoid or minimize bias that may occur due to multiple testing issues and other opportunistic actions by investigators.

In addition, it is important to state the criteria for success of a pilot study. The criteria should be always based on the primary feasibility objectives. This provides the basis for study result interpretation and determining the feasibility of the main trial. Lastly, investigators need to disclose the purpose of pilot trials to participants including the feasibility objectives and criteria for success (103).

Pilot studies have been performed evaluating clinical pathways for management of AF patients in the ED and seem to attest to the feasibility of a cRCT using the clinical pathway for management of AF. In a survey of AF in EDs across Ontario, Dorian and colleagues were able to identify 158 EDs with at least 10 or more visits for AF every year. The annual volume of patients with a primary diagnosis of AF in all of Ontario's EDs ranged from 12,613 to 14,778 over a period of 6 years (2002-2007). HHS data indicates that about 40% of all AF patients attending the ED satisfy inclusion criteria for the proposed cRCT. Therefore, the expected volume of RAAF patients in Ontario EDs is approximately 4-5000/year. These data suggest that it will be feasible

to recruit 5000 patients, over a 2-year period from 38 EDs (sample size required for demonstrating a 25% RRR over a 2 years study period); especially if high volume EDs in Canada are enrolled in the study. Also, as stated previously the current sample size has been calculated using a conservative admission rate of 16.67% based on Stiell's data. We expect the sample size to be lower than the currently presented numbers in view of a higher admission (event) rate from all EDs in Ontario. However, caution must be exercised as this is only an assumption and has to be proven using the complete data set once it is available to the investigators for formal analysis (6,12, 25).

5.6: Issues involving losses to follow-up

The potential for lost-to-follow-up can be a severe problem in cRCTs, especially those with a long follow up period. In trials with interventions designed at the cluster level, with little importance given to individual study patients, the loss to follow-up rate may be high. This factor is compensated for during trial design by a strategy of oversampling and aggressive follow up. In addition investigators will also have to deal with the prospect of entire clusters dropping out of the trial due to logistical or budgetary constraints. Adjustments will be made in cluster number calculations to account for this and the required sample size will be inflated by 10%. Steps need to be taken to ensure that the study participants strictly adhere to the trial protocol and follow-up procedures. (74,87).

5.7: Strategies for achieving desired power

Even carefully designed cRCTs may be underpowered to test the study hypothesis. A number of strategies are available to ensure that the proposed clinical trial will have the desired power to answer the principal hypothesis. The relevant strategies are listed below-

- i. **Selecting appropriate restrictive selection criteria:** Geographic or other selection criteria (teaching vs. non-teaching hospitals) can be used to perform restrictive randomization, thereby reducing between cluster variability. However, it must be remembered that employing this strategy to enhance power comes at the cost of loss of generalizability (61,103).
- ii. **Careful justification of data collection:** Investigators have a tendency to request large amounts of data from study subjects. This comes at the cost of subject compliance and loss to follow-up. Therefore, only data essential to answer the study hypothesis will be collected in our trial (74).
- iii. **Realistic estimation of the minimally important clinical effect and loss to follow-up:** Exaggerated estimates of the efficacy of a clinical intervention and patient compliance can lead to significant loss of power. Therefore, a series of projections for the clinical effect of the intervention and loss to follow-up rates will be used to calculate the sample size. In addition a reserve of clusters that can be randomized will be identified, in the event that this became necessary to preserve the pre-specified level of power (63).

- iv. **Over-sampling clusters** upfront is another strategy that can be used to account for errors in ICC calculation
- v. **Re-assessing ICC during the study** can be carried out at pre-specified interim analyses of the study data. In case the ICC is larger than expected from the preliminary analyses or pilot trial provisions can be made to recruit additional clusters. On the other hand if the ICC is lower than expected the trial duration can be reduced to one year or fewer clusters can be recruited. Both of these scenarios have cost and feasibility implications and detailed analyses needs to be performed to account for best and worst case scenarios.
- vi. **Using composite endpoints** as the primary outcome variable
- vii. **Using surrogate endpoints**, these latter two measures are not likely needed for our trial and will not be discussed in detail.

5.8: Missing data

The approach taken to missing data can be critical to proper interpretation of trial results, especially if loss to follow-up rate is large. There are many strategies to account for missing data and include- imputation techniques that assign numerical estimates for the missing values. These include mean substitution, 'last value forward', or the use of regression methods, none of which correctly accounts for the effect of imputation on the resulting estimates of variance. More sophisticated techniques include using multiple estimates for each missing observation. Since all such strategies are subject to the potential for

bias, performing multiple analyses is advisable, including one in which subjects with missing data are removed from the analyses (per-protocol analyses). The trial will employ a pre-specified 'sensitivity' analyses to account for how the missing data might affect the study conclusions. If the trial conclusions remain consistent across various sets of assumptions, the overall credibility of the trial results will be strengthened (104, 105, 106).

5.9: Sensitivity analyses

Sensitivity analysis is defined as the uncertainty in the results of a statistical model that occurs as a result of variations in the input variables. The technique involves exploring different input variables in the model to determine the effects on the outcomes. With respect to cRCTs this technique involves estimating the required trial size under a number of different clinical scenarios. Studies have shown that even small changes in the number of subjects recruited, effect of the intervention or the ICC can have large effects on the required sample size, thereby altering the power of the study. This is particularly important in cRCTs as investigators often have difficulty in accurately estimating intra and inter cluster variation. A sensitivity analysis for the primary outcome will be performed as part of the statistical analysis plan for the proposed clinical trial (74,107,108).

Chapter 6- Ethical and legal issues

6.1: Introduction

This chapter discusses the broad issues involving trial monitoring and interim reporting adopted by an independent committee to safeguard the interests of the participants.

6.2: Source of funding and potential conflicts of interest

The proposed clinical trial will be submitted to various peer- reviewed and non-peer reviewed funding agencies including the HHS Innovation fund, Canadian Institute of Health Research (CIHR), the Heart and Stroke Foundation of Canada (HSFC) and the Canadian Network and Centre for Trials internationally (CANNeCTIN). The primary investigators of the proposed clinical trial do not serve in these organizations under any capacity. The steering committee members including Drs. Connolly, Worster, Morillo, Stiell and Roberts serve as grant reviewers for the above mentioned organizations and will declare a conflict of interest when the study is submitted to these agencies.

6.3: Ethical issues in cluster randomized trials

In therapeutic trials, informed consent is obtained prior to random assignment, with the physician admitting ignorance regarding optimal therapy. The patient must agree to allow chance to dictate the treatment that they receive.

This often introduces a barrier to patient accrual and randomized trial designs were introduced to lower this barrier and hasten trial completion in cRCTs.

There are two randomized consent designs in use known as the double consent design and the single consent design. In the double-consent design, patients are randomly assigned to an experimental or a standard therapy prior to asking for consent. Patients are then asked for consent to receive the intervention that they have been assigned. If they refuse, they may be offered the alternative therapy under consideration. In the single-consent design, patients are also randomly assigned either to standard or to an experimental therapy. However, consent is sought only from patients randomly assigned to the experimental therapy. These patients have the option of declining their assigned intervention assignment. All analyses are according to the intent to treat' strategy. The potential gain in accrual is expected to offset any loss in efficiency resulting from patients preferring an alternative treatment (109,110).

The decision to omit informed consent, when deemed necessary, merits careful consideration. This approach, though controversial, has been adopted in some trials, where the ethics committee decided that the risk of harm was too slight to warrant informed consent. However, several critics have argued that respect for study subjects should have outweighed these methodological considerations and consent should have been obtained. (74,87).

The randomized consent design has proven to be controversial since first being proposed and has only been adopted in a small number of therapeutic trials. Reluctance to use randomized consent arises because of concern for the ethical implications of randomizing subjects prior to obtaining consent. However, this strategy is typical of most cRCTs (74, 87, 111, 112).

We have decided not obtain consent for inclusion of patients in the proposed cRCT as it is a low risk observational trial without the use of novel pharmacological agents or therapeutic procedures. This will increase the ease of recruitment in our study without compromising patient safety or rights. We will obtain a verbal consent at the time of the 30-day follow up for data collection.

A cluster representation mechanism (CRM) will be established in all participating clusters. This committee will be comprised of the clinical director of the participating ED and a representative of the hospital administration. The CRM will act in good faith with regards to the interests of patients recruited to the cluster. They will function as advocates for patients wanting to withdraw from the trial or seeking alternate treatment strategies. The CRM committee members will be independent of the investigating team to avoid any conflicts of interest. Institutional research ethics board (REB) approval will be obtained at McMaster University and each participating site (113,114,115,116,117,118).

6.4: Health Canada approval and trial registration

The proposed cRCT does not involve the use of a novel off-label medication or medical device. Therefore, a clinical trial or medical device licence application will not be obtained from Health Canada. The study also does not involve long-term storage of tissues or bio-samples. The study will be registered at the National Institutes of Health (Bethesda, the USA) website- www.clinicaltrials.gov, other Canadian federal agencies and the CIHR.

6.5: Trial Monitoring

During the course of a trial an independent committee of experts (the Data and Safety Monitoring Committee) is constituted to safeguard the interests of participants. Efficacy monitoring in accordance with a predetermined plan will be used to terminate the trial in case of early benefits or harm. Interim safety and efficacy analyses will be performed at the 12-month mark. The committee will also monitor the actual accrual of clusters and individuals to the trial, as well as intermittently assess baseline comparability. The interim analyses will also allow investigators an opportunity to reassess the adequacy of trial size on the basis of early information and revise sample size calculations and trial design accordingly (87).

6.6: Patient confidentiality

All data collected during the study including patient demographic and clinical details will be secured in a password-protected server located in the PHRI affiliated to McMaster University. Trial investigators, research personnel, statisticians, DSMB and CRM members will be granted access to review and analyse the confidential data (40,87).

Chapter 7- Summary

7.1: Conclusions

Atrial fibrillation is the most common arrhythmia encountered in clinical practice. The rising prevalence of AF in the community has resulted in frequent ED visits for management of RAFF. Strategies to optimally manage patients with RAFF in the ED and reduce hospital admissions are urgently required to reduce ED visits and health care costs and their effectiveness is at present unproved. Clinical pathways have been shown to reduce inappropriate hospital admissions, provide continuity of care for patients discharged from the ED in a variety of clinical conditions in ED patients. This thesis explores the methodological issues relevant to the design of a cRCT evaluating a clinical pathway in the management of acute onset, low risk AF patients presenting to the ED.

We will conduct an intention-to-treat analysis at the individual level using proportions and means according to the variable in question with an alpha level of 0.05 and power of 0.80 for the primary outcome. EDs in Ontario with an annual volume of >140 RAAF visits will be selected for the trial. The proposed cRCT (assuming a 30% RRR) will be conducted over a two year time period. A total of 13 clusters will need to be recruited in each intervention arm (a total of 26 clusters for the trial). A total of 3500 ED visits for RAFF will need to be included from these 26 clusters.

7.2: Limitations of the study

7.2.1: Cluster level limitations

One of the biggest challenges for the investigators of the proposed cRCT will be the ability to recruit between 30-40 (EDs) clusters during a two-year period. In addition each of these clusters need to be able to provide an estimated 140-150 low risk AF patients for inclusion into the trial each year. This is not going to be an easy task and will be possible only if the trial is conducted across all major EDs in Canada. In addition the investigators will need the whole-hearted support of thought leaders in the field of emergency medicine and cardiology to be able to successfully complete the trial. In addition it may be necessary to identify a “reserve” of clusters that can be randomized in case interim analyses indicate that additional clusters may need to be recruited.

Another cluster level factor is the wide variation in institutional practices for management of AF in the ED across the province of Ontario and nationally. This will affect the internal validity of the trial especially if a particular institution has already instituted a detailed, evidence based clinical pathway for management of AF patients. This may reduce the expected effect of the intervention on the primary outcome. It may also be necessary to exclude EDs that have previously participated in pilot trials evaluating clinical pathways for the management of AF.

7.2.2: Patient level limitations

Data from pilot trials evaluating clinical pathways for management of acute onset AF in the ED seem to indicate that 10-15% patients will not agree to participate in the proposed cRCT. This can affect the power calculations and affect the internal validity of the trial. To avoid this the investigators will have to get accurate estimates of the volume of AF patients in each of the participating EDs.

7.2.3: Physician level limitations

A Canadian ED physician practice survey described earlier on has shown wide variation in physician management preferences. It is possible that individual physicians working in EDs randomized to the clinical pathway arm may disagree with the management steps listed in the clinical pathway. This will impact the internal validity of the trial significantly. To offset this the investigators will need to have training and orientation sessions for participating ED physicians to get their “buy-in” for the clinical pathway. In addition periodic audits of data collection forms and patient care records will be performed to ascertain ED physician adherence to the clinical pathway.

7.2.4: Limitations of the intervention

Lastly, one of the biggest challenges facing the investigators is the intervention in this trial- the clinical pathway. Unlike drug trials, where a precise dose of the novel medication is used as an intervention, our trial is evaluating a clinical pathway. Clinical pathways (also known as care pathways, critical pathways or care maps) are tools used to manage quality in healthcare concerning the standardization of health care processes.

Clinical pathways for management of AF have been evaluated in non-randomized pilot trials. However, these trials were conducted in tertiary care centers by groups involved in research and had highly motivated and dedicated physicians implementing the clinical pathway. Though the clinical pathway has been shown to reduce hospital admissions, it has not been evaluated in a randomized pilot trial with clear feasibility endpoints. The data from these pilot trials cannot be used to precisely estimate the effect size of this intervention or the feasibility of patient and cluster recruitment. Therefore, it may be necessary to first perform a randomized, feasibility cRCT and then proceed to the main trial if the pilot trial is promising.

References

1. Go AS, Hayek EM, Phillips KA, et al. Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the Anticoagulation and Risk Factors in Atrial Fibrillation Study (ATRIA). *JAMA*. 2001; 285: 2370-2375.
2. Chugh SS, Blackshear JL, Shen WK, et al. Epidemiology and natural history of atrial fibrillation: clinical implications. *J Am Coll Cardiol*. 2001;37:371-378.
3. Miyasaka Y, Barnes ME, Gersh BJ, et al. Secular trends in incidence of atrial fibrillation in Olmstead County, Minnesota, 1980 to 2000, and implications on the projections for future prevalence. *Circulation* 2006;114:119-125.
4. McDonald AJ, Pelletier AJ, Ellinor PT et al. Increasing US emergency department visit rates and subsequent hospital admissions for atrial fibrillation from 1993 to 2004. *Ann Emerg Med*. 2008;51:58-65.
5. Wattigney WA, Mensah GA, Croft JB: Increasing trends in hospitalization for atrial fibrillation in the United States, 1985 through 1999. *Circulation* 2003;108:711-716.
6. Aztema CL, Yun L and Dorian P. Atrial Fibrillation in the emergency department: A population based description. *Can J Cardiol* 2010; 26 (Suppl.D): 247D
7. Dell'Orfano JT, Patel H, Wolbrette DL, et al. Acute treatment of atrial fibrillation: spontaneous conversion rates and cost of care. *Am J Cardiol*. 1999;83:788-790, A10.
8. Reynolds MR, Lavelle T, Essebag V, et al. Influence of age, sex, and atrial fibrillation recurrence on quality of life outcomes in a population of patients

with new-onset atrial fibrillation: The fibrillation registry assessing costs, therapies, adverse events and lifestyle (FRACTAL) study. *Am Heart J.* 2006;152:1097-103.

9. Reynolds MR, Essebag V, Zimetbaum P, et al. Healthcare Resource utilization and costs associated with recurrent episodes of atrial fibrillation: The FRACTAL registry. *J Cardiovasc Electrophysiol.* 2007;18:628-33.
10. Friberg J, Buch P, Scharling H et al. Rising rates of hospital admissions for atrial fibrillation. *Epidemiology* 2003;14:666-672.
11. Li G, Lau JT, McCarthy ML et al. Emergency department utilization in the United States and Ontario, Canada. *Acad Emer Med* 2007;14:582-584.
12. Divakaramenon SM, Pazsynski T, PB Nery PB et al. Management of atrial fibrillation in the emergency department: are we adhering to guidelines. *Can J Cardiol* 2008 Oct; 24 (Suppl SE 0839).
13. Agoritsas T, Bovier PA, Perneger TV. Patient Reports of Undesirable Events During Hospitalization. *J Gen Intern Med* 2005;20:922-928.
14. Michael JA, Stiell IG, Agarwal S, et al. Cardioversion of paroxysmal atrial fibrillation in the emergency department. *Ann Emerg Med.* 1999;33:379-387.
15. Mulcahy B, Coates WC, Hennerman PL, et al. New-onset atrial fibrillation: When is admission medically justified? *Acad Emerg Med.* 1996;3:114-119.
16. Joseph AP, Ward MR. A prospective, randomized controlled trial comparing the efficacy and safety of sotalol, amiodarone, and digoxin for the reversion of new-onset atrial fibrillation. *Ann Emerg Med.* 2000;36:1-9.

17. Decker WW, Smars PA, Vaidyanathan L et al. A Prospective, Randomized Trial of an Emergency Department Observation Unit for Acute Onset Atrial Fibrillation. *Ann Emerg Med.* 2008;52:322-328.
18. Innes GD, Vertesi L, Dillon EC, et al. Effectiveness of verapamil-quinidine versus digoxin-quinidine in the emergency department treatment of paroxysmal atrial fibrillation. *Ann EmergMed.* 1997;29:126-134.
19. Gomez MA, Jefferey LA, Karagounis LA, et al. An emergency department based protocol for rapidly ruling myocardial ischemia reduces hospital time and expense: results of a randomized study (ROMIO). *J Am Coll Cardiol.* 1996;28:25-33.
20. Koenig BO, Ross MA, Jackson RE. An emergency department observation unit protocol for acute-onset atrial fibrillation is feasible. *Ann Emerg Med.* 2002;39:374-81.
21. Zimetbaum P, Reynolds MR, Ho KKL et al. Impact of a practice guideline for patients with atrial fibrillation on medical resource utilization and costs. *Am J Cardiol* 2003;92:677-681.
22. Fuster V, Ryden LE, Asinger RW, et al. ACC/AHA/ESC guidelines for management of patients with atrial fibrillation: executive summary: a report of the American College of Cardiology/American Heart association Task Force on Practice Guidelines and the European Society of Cardiology Committee for Practice Guidelines and Policy Conferences (Committee to Develop Guidelines for the Management of Patients with Atrial Fibrillation) developed in collaboration with the North American Society of Pacing and Electrophysiology. *Circulation.* 2006;114:700-752.
23. Snow V, Weiss KB, LeFevre M, et al. Management of newly detected atrial fibrillation: A clinical practice guideline from the American Academy of Family Physicians and the American College of Physicians. *Ann Intern Med.* 2003;139:1009-17. Canadian Cardiovascular Society Guidelines for Management of AF 2010.

http://www.ccsguidelineprograms.ca/images/stories/Artrial_Fibrillation_Program/ExecutiveSummary/2010recomendations_exec_summary_6.pdf

24. Stiell IG and Macle L; CCS Atrial Fibrillation Guidelines Committee. Canadian Cardiovascular Society atrial fibrillation guidelines 2010: management of recent-onset atrial fibrillation and flutter in the emergency department. *Can J Cardiol.* 2011 Jan-Feb;27(1):38-46
25. Stiell IG, Clement CM, Brison RJ, et al. Variation in Management of Recent-Onset Atrial Fibrillation and Flutter Among Academic Hospital Emergency Departments. *Ann Emerg Med*; 57:13-21
26. Yealy DM and Delbridge TR. Dysrhythmias. In: Marx JA, Hockberger RS, Walls RM, editors. *Rosen's emergency medicine: concepts and clinical practice*, Vol 2. 5th ed. St. Louis (MO): Mosby; 2002. p. 1083-6.
27. Stiell IG, Clement CM, Perry JJ, et al. Association of the Ottawa Aggressive Protocol with rapid discharge of emergency department patients with recent-onset atrial fibrillation or flutter. *CJEM.* 2010 May;12(3):181-91.
28. Alboni P, Botto GL, Baldi N, et al. Outpatient Treatment of Recent-Onset Atrial Fibrillation with the "Pill-in-the-Pocket" Approach. *N Engl J Med* 2004;351:2384-91.
29. Wattanasuwan N, Khan IA, Mehta NJ, et al. Acute Ventricular Rate Control in Atrial Fibrillation * : IV Combination of Diltiazem and Digoxin vs IV Diltiazem Alone. *Chest* 2001;119;502-506
30. Ellenbogen KA. Role of Calcium Antagonists for Heart Rate Control in Atrial Fibrillation. *Am J Cardiol*; 1992;69:36B-40B

31. Wang HE, O'Connor RE, Megargel RE, et al. The Use of Diltiazem for Treating Rapid Atrial Fibrillation in the Out-of-Hospital Setting. *Ann Emerg Med.* 2001;37:38-45
32. Demircan C, Cikriklar HI, Engindeniz Z, et al. Comparison of the effectiveness of intravenous diltiazem and metoprolol in the management of rapid ventricular rate in atrial fibrillation. *Emerg Med J.* 2005;22:411-414.
33. Borgundvaag B and Ovens H. Cardioversion of uncomplicated paroxysmal atrial fibrillation: a survey of practice by Canadian emergency physicians. *Can J Emerg Med.* 2004;6(3):155-60.
34. Kinsman L, Rotter T, James E, et al. What is a clinical pathway? Development of a definition to inform the debate. *BMC Medicine.* 2010; 8:31-34.
35. De Bleser LDR, De Waele K, Vanhaecht K, et al. Defining pathways. *J Nurs manag.* 2006; 14:553-563
36. Rotter T, Kugler J, Koch R, et al. A systematic review and meta-analysis of the effects of clinical pathways on length of stay, hospital costs and patient outcomes. *BMC Health Serv Res.* 2008; 8: 265-280.
37. Grimshaw J, Campbell M, Eccles M et al. Experimental and quasi-experimental designs for evaluating guideline implementation strategies. *Fam Pract.* 2000; 17:S11-18.
38. Cook TD, Campbell DT. *Quasi-experimentation: Design and analysis issues for field settings.* Chicago: Rand McNally, 1979.

39. Lipsey MW, Wilson DB. The efficacy of psychological, educational and behavioural treatment. Confirmation from meta-analysis. *Am Psychol*. 1993; 48:1181-1209.
40. Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. *BMC Med Res Methodol*. 2006; 6:54.
41. Saint SHT, Hofer TP, Rose JS, et al. Use of critical pathways to improve efficiency: a cautionary tale. *Am J Manag Care*. 2003; 9(11): 758-765.
42. Cornfield J. Randomization by group: a formal analysis. *Am J Epidemiol*. 1978; 108: 100-102.
43. McKinlay SM, Stone EJ and Zucker DM. Research design and analysis issues. *Health Educ Q*. 1989; 16: 307-313.
44. Hayes RJ and Moulton LH. *Cluster Randomised Trials*. 1 Boca Raton: CRC Press, 2009.
45. Campbell MK, Elbourne DR and Altman DG. Consort Statement: Extension to Cluster Randomised Trials. *BMJ* 2004; 328:702-8.
46. Friedman LM, Furberg CD and DeMets DL. *Fundamentals of clinical trials*. 3rd edn. Mosby-Year Book, St. Louis, MO.1996.
47. Gail MH, Byar DP, Pechacek TF, et al. Aspects of statistical design for the community intervention trial for smoking cessation (COMMIT). *Controlled Clinical Trials*. 1992; 13: 6-21.

48. Gail MH, Mark SD, Carroll RJ, et al. On design considerations and randomization-based inference for community intervention trials. *Stat Med*. 1996; 15: 1069-1092.
49. Susser M. Editorial: The tribulations of trials – intervention in communities. *Am J Public Health*. 1995; 85:156-158.
50. Fishbein M. Editorial: great expectation, or do we ask too much from community interventions? *Am J Public Health*. 1996; 86: 1075-1076
51. Donner A, Brown KS and Brasher P. A methodological review of non-therapeutic intervention trials employing cluster randomization, 1979-1989. *Int J Epidemiol*. 1990; 19:795-800.
52. Simpson JM, Klar N and Donner A. Accounting for cluster randomization: a review of primary prevention trials, 1990 through 1993. *Am J Public Health*. 1995; 85: 1378-1382.
53. Whiting-O'Keefe QE, Henke C and Simborg DW. Choosing the correct unit of analysis in medical care experiments. *Medical Care*. 1984; 22: 1101-1114.
54. Kelder SH, Jacobs DR Jr., Jeffery RW, et al. The worksite component of variance: design effects and the Healthy Worker Project. *Health Educ Res*. 1993; 8: 555-566
55. Koepsell TD, Martin DC, Diehr PH, et al. Data analysis and sample size issues in evaluations of community-based health promotion and disease prevention programs: a mixed-model analysis of variance approach. *J Clin Epidemiol*. 1991; 44:701-713.

56. Bland JM and Kerry SM. Statistics notes. Trials randomized in clusters. *BMJ*. 1997; 315:600.
57. Smith PJ and Morrow RH. *Methods for field trials of interventions against tropical diseases: a 'toolbox'*. Oxford university press, Oxford:1991.
58. Koepsell TD. Epidemiologic issues in the design of community intervention trials. In Brownson RC and Petitti DB (eds), *Applied epidemiology: theory to practice*. Oxford University Press, New York; 1998: 177-211.
59. Murray DM, Rooney BL, Hannan PJ, et al. Intraclass correlation among common measures of adolescent smoking: estimates, correlates and applications in smoking prevention studies. *American Journal of Epidemiology* 1994; 140:1038 {1050.
60. Hannan PJ, Murray DM, Jacobs DR, et al. Parameters to aid in the design and analysis of community trials: intraclass correlations from the Minnesota Heart Health Program. *Epidemiology* 1994; 5:88 {95.
61. Siddiqui O, Hedeker D, Flay BR, et al. Intraclass correlation estimates in school-based smoking prevention study: outcome and mediating variables by gender and ethnicity. *American Journal of Epidemiology* 1996; 144:425{433.
62. Feng Z, Diehr P, Yasui Y, et al. Explaining community-level variance in group randomized trials. *Statistics in Medicine* 1999; 18:539{556.
63. Campbell MK, Mollison J and Grimshaw JM. Cluster trials in implementation research: estimation of intracluster correlation coefficients and sample size. *Stat Med*. 2001; 20: 391-399
64. Verma V and Le T. An analysis of sampling errors for the demographic and health surveys. *International Statistical Review*. 1996; 64: 265-294

65. Hsieh FY. Sample size formulae for intervention studies with the cluster as unit of randomization. *Stat Med.* 1988; 8:1195-1201.
66. Hayes R. Design of human immunodeficiency virus intervention trials in developing countries. *J R Stat Soc Ser A.* 1998; 161:251-263.
67. Feldman HA and McKinlay SM. Cohort versus cross-sectional design in large field trials: precision, sample size, and a unifying model. *Stat Med.* 1994; 13:61-78.
68. Thompson SG, Pyke SDM and Hardy RJ. The design and analysis of paired cluster randomized trails: an application of meta-analysis techniques. *Stat Med.* 1997; 16: 2063-2980.
69. Cohen J. The cost of dichotomization. *Appl Psychol Meas.* 1983; 7: 249-253.
70. Fleming TR and DeMets DL. Surrogate endpoints in clinical trials: are we being misled? *Ann Intern Med.* 1996; 125: 605-613.
71. Bland JM and Kerry SM. Statistics notes. Weighted comparison of means. *Br Med J.* 1998; 316: 129.
72. Shao J. Ordinary and weighted least-squares estimators. *Can J Stat.* 1990; 18: 327-336.
73. Raudenbush SW. Hierarchical linear models and experimental design. In Edwards LK. (ed.), *Applied analysis of variance in behavioural science.* 1993: Marcel Dekker, Ch.13.

74. Kreft IGG. An illustration of item homogeneity scaling and multilevel analysis techniques in the evaluation of drug prevention programs. *Eval Rev.* 1998; 22: 46-77.
75. Murray DM and Short BJ. Intraclass correlation among measures related to tobacco use by adolescents: estimates, correlates, and applications in intervention studies. *Addict Behav.* 1997; 22:1-12.
76. Austin P. A comparison of the statistical power of different methods for the analysis of cluster randomization trials with binary outcomes. *Stat Med.* 2007; 26: 3550-3565.
77. Armitage P and Berry G. *Statistical methods in medical research.* 3rd edn. Oxford: Blackwell Scientific Publications.
78. Heeren T and D'Agostino. Robustness of the two-independent samples t-test when applied to ordinal scaled data. *Stat Med.* 1987; 6: 79-90.
79. Donner A and Klar N. Statistical considerations in the design and analysis of community intervention trials. *J Clin Epidemiol.* 1996; 49: 435-439.
80. Conover WJ. *Practical nonparametric statistics.* 2nd edn. New York: John Wiley; 1980.
81. Rosner B. *Fundamentals of biostatistics.* 4th edn. Blemont, CA: Wadsworth;1995.
82. Donner A and Donald A. The statistical analysis of multiple binary mesasurements. *J Chronic Dis.* 1988; 41: 899-905.

83. Donner A. Statistical methods in ophthalmology: an adjusted chi-square approach. *Biometrics*. 1989; 45: 511-605
84. Klar N. Stratified analysis of correlated binary outcome: a comparison of model dependent and robust tests of significance. *Commun Stat Theory Methods*. 1996; 25: 2431-2458.
85. Rao JNK and Scott AJ. A simple method for the analysis of clustered binary data. *Biometrics*. 1992; 48: 577-585.
86. Donner A, Eliasziw M and Klar N. A comparison of methods for testing homogeneity of proportions in teratologic studies. *Stat Med*. 1994; 13: 1253-1264.
87. Fung KY, Krewski D, Rao JN, et al. Tests for trend in development toxicity experiments with correlated binary data. *Risk Analysis*: 1994; 14: 639-648.
88. Hedeker D, McMahon SD, Jason LA, et al. Analysis of clustered data in community psychology: with an example from a worksite smoking cessation project. *Am J Community Psychol*. 1994; 22: 595-615.
89. Neuhaus JM. Statistical methods for longitudinal and clustered designs with binary responses. *Stat Methods Med Res*. 1992; 1: 249-273.
90. Liang KY and Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika*. 1986; 73:13-22.
91. Boos DD. On generalized score tests. *Am Stat*. 1992; 46: 327-333.
92. Donner A and Klar N. Design and analysis of cluster randomization trials in health research. 1st edn. Sussex, UK. John Wiley & sons Ltd: 2000.

93. Neuhaus JM and Kalbfleisch JD. Between- and within-cluster covariate effects in the analysis of clustered data. *Biometrics*. 1998; 54: 638-645.
94. Koepsell TD, Wagner EH, Cheadle AC, et al. Selected methodological issues in evaluating community-based health promotion and disease prevention programs. *Annu Rev Public Health*. 1992; 13: 31-57.
95. Hayes RJ and Bennett S. Simple sample size calculation for cluster-randomized trials. *Int J Epidemiol*. 1999; 28: 319-326.
96. Chakraborty H, Moore J, Carlo WA, et al. A simulation based technique to estimate intracluster correlation for a binary variable. *Contemp Clin Trials*. 2009; 30: 71-80.
97. Candel MJJM and Van Breukelen GJP. Sample size adjustments for varying cluster sizes in cluster randomized trials with binary outcomes analyzed with second-order PQL mixed logistic regression. *Stat Med*. 2010; 29: 1488-1501.
98. Eldridge S, Cryer C, Feder G, et al. Sample size calculations for intervention trials in primary care randomizing by primary care group: an empirical illustration from one proposed intervention trial. *Stat Med*. 2001; 20: 367-376.
99. Eldridge SM, Ashby D, Kerry S. Sample size for cluster randomized trials: effect of coefficient of variation of cluster size and analysis method. *Int J Epidemiol*. 2006; 35:1292-1300.
100. Calkins H, Brugada J, Packer DL, et al. HRS/EHRA/ECAS expert consensus statement on catheter and surgical ablation of atrial fibrillation: Recommendations for personnel, policy, procedures and follow-up. A

report of the Heart Rhythm Society (HRS) task force on catheter and surgical ablation of atrial fibrillation. *Heart Rhythm*. 2007; 4(6): 816-861.

101. Eldridge S, Ashby D, Bennett C, et al. Internal and external validity of cluster randomised trials: systematic review of recent trials. *BMJ* 2008; 336:876-80.
102. Puffer S, Torgerson D and Watson J. Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. *BMJ* 2003; 327:785-9.
103. Schulz KF and Grimes DA. Blinding in randomised trials: hiding who got what. *Lancet* 2002; 359:696-700.
104. Glasgow RE, Vogt TM and Boles SM. Evaluating the public health impact of health promotion interventions: the re-aim framework. *Am J Public Health*. 1999; 89:1322-7.
105. Wu AW, Cavanaugh TA, McPhee SJ, et al. To tell the truth: ethical and practical issues in disclosing medical mistakes to patients. *J Gen Intern Med* 1997; 12:770-5.
106. Weingart SN, Wilson RM, Gibberd RW, et al. Epidemiology of Medical Error. *BMJ* 2000; 320:774-7.
107. Benjamin DM. Reducing Medication Errors and Increasing Patient Safety: Case Studies in Clinical Pharmacology. *J Clin Pharmacol*. 2003; 43:768-83.
108. Thabane L, Ma J, Chu R, et al. A tutorial on pilot studies: The what, why and how. *BMC Med Res Methodol*. 2010; 10:1.

109. LaPrelle J, Bauman KE and Koch GC. High intercommunity variation in adolescent cigarette smoking a 10-community field experiment. *Evaluation Review*. 1992; 16: 115-130.
110. Hollis S and Campbell F. What is meant by intention to treat analysis? survey of published randomised controlled trials. *BMJ* 1999; 319:670-4.
111. Heitjan DF. Annotation: what can be done about missing data? Approaches to imputation. *Am J Public Health*. 1997; 87: 548-550.
112. Zucker DM, Lakatos E, Webber LS, et al. Statistical design of the child and adolescent trial for cardiovascular health (CATCH). Implications of cluster randomization. *Controlled Clinical Trials*. 1995; 16: 96-118.
113. Saltelli A. *Global sensitivity analysis : The primer*. Chichester, England ; Hoboken, NJ: John Wiley, 2008.
114. Ellenberg SS. Informed consent: protection or obstacle? Some emerging issues. *Controlled Clinical Trials*. 1997; 18: 628-636.
115. Zelen M. Randomized consent designs for clinical trials: an update. *Stat Med*. 1990; 9: 645-656.
116. Altman DG, Whitehead J, Parmar MK, et al. Randomised consent designs in cancer clinical trials. *Eur J Cancer*. 1995; 31 A (12): 1934-1944.
117. Villar J, Bakketeig L, Donner A, et al. The WHO antenatal care randomised trial: rationale and study design. *Pediatr Perinat Epidemiol*. 1998; 12 (Suppl.2): 27-58.

118. Taljaard M, Weijer C, Grimshaw JM, et al. Ethical and policy issues in cluster randomized trials: rationale and design of a mixed methods research study. *Trials*. 2009; 10: 61.
119. Gage BF, Waterman AD, Shannon W, et al. Validation of clinical classification schemes for predicting stroke: Results from the national registry of Atrial Fibrillation. *JAMA*. 2001; 285: 2864-2870.

Figures

Figure 1: Clinical pathway for management of low risk AF patients in the ED
(21,22,23)

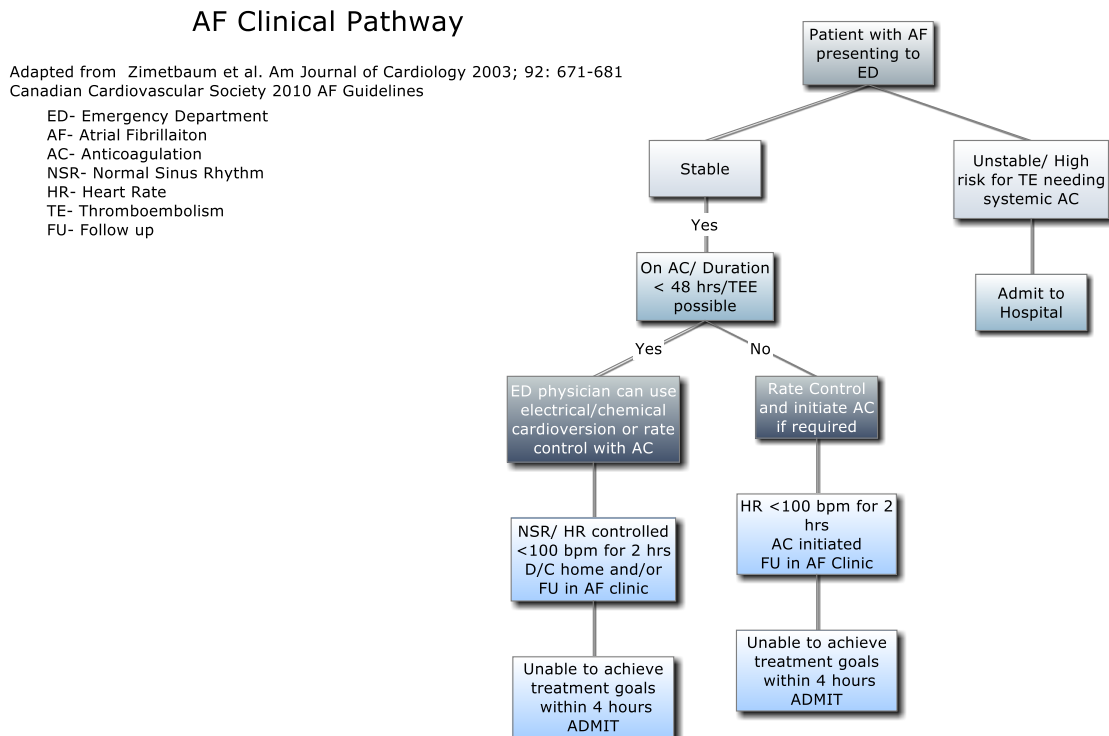


Figure 2: The five-step algorithm for management of AF in the ED:

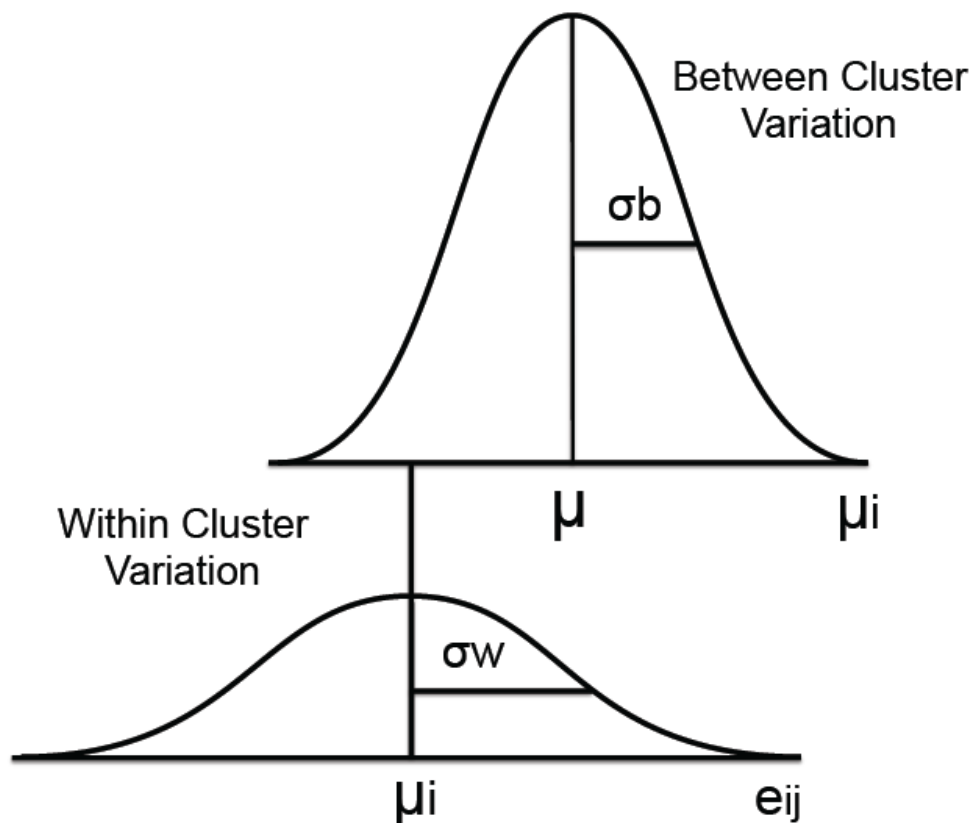
<p>Step 1</p>	<p>Confirm AF/Atrial Flutter on EKG</p> <p>Check Inclusion/Exclusion criteria</p> <p>Patient eligible for study proceed to next step</p>
<p>Step 2</p>	<p>Obtain consent for study-</p> <p>Yes- Proceed to next step</p> <p>No- Continue with institutional practice for management of AF</p>
<p>Step 3</p>	<p>Establish duration of AF/Atrial flutter</p> <p>>48 hours; <48 hours</p>
<p>Step 4</p>	<p><u>AF < 48 hours-</u></p> <p><u>Pharmacological Cardioversion –</u></p> <p>iv Procainamide</p> <p><u>Direct Current Cardioversion</u></p> <p><u>(DCCV)</u> as first line if unable to tolerate procainamide, if it fails or if electrical cardioversion is preferred</p>

	as first line treatment
Step 5	<p><u>AF >48 hours/Contraindication to procainamide or DCCV/Cardioversion strategy fails-</u></p> <p>Rate control with iv Diltiazem and if not tolerated with Metoprolol- Target ventricular rate <100 bpm; for at least one hour at rest before discharge</p>
Admission/Discharge	<p>If successful in achieving above goals in 5 hours- Discharge</p> <p>Start Dabigatran or Coumadin if CHADS score >1</p> <p>If unsuccessful- Admit</p>

Figure 3: This figure shows the mean variance μ and the between cluster variation σ_b of patient outcome y from multiple patients (indexed by the subscript j) and from multiple clusters (indexed by the subscript i). We can think of the j^{th} observation in the i^{th} cluster as:

$$y_{ij} = \mu_i + e_{ij}$$

where, μ_i ($i = 1$ to k , the number of clusters) is the true mean for the i^{th} cluster and e_{ij} ($j = 1$ to m_i , the number of patients in cluster i) is part of the measurement, which causes one patient's observation to be different from another's in the same cluster (ie. within-cluster variation or σ_w) and is simply the variance of e_{ij} .



Tables

Table 1: Recommended drugs for pharmacological cardioversion of AF in the ED
(21,22,23)

Drug	Dose	Efficacy	Risks
Class IA Procainamide	15-17 mg/kg IV over 60 min	++	5% hypotension
Class IC*			
Propafenone	450-600 mg PO	+++	Hypotension, 1:1 flutter, bradycardia
Flecainide	300-400 mg PO	+++	Hypotension, 1:1 flutter, bradycardia
Class III Ibutilide	1-2 mg IV over 10-20 min Pre-treat with MgSO ₄ 1-2 mg IV	++	2-3% Torsades de pointes

*Class IC drugs should be used in combination with AV nodal blocking agents (beta-blockers or calcium-channel inhibitors).

Table 2: Intravenous & orally administered pharmacological agents for heart rate control in patients with AF without accessory pathways (21,22,23)

Drug	Class/LOE	Loading Dose	Onset	Maintenance Dose	Major Side Effects
Esmolol*†	Class I, LOE C	500 mcg/kg IV over 1 min	5 min	60 to 200 mcg/kg/min IV	↓ BP, HB, ↓ HR, asthma, HF
Metoprolol†	Class I, LOE C	2.5 to 5 mg IV bolus over 2 min; up to 3 doses	5 min	NA	↓ BP, HB, ↓ HR, asthma, HF
Propranolol†	Class I, LOE C	0.15 mg/kg IV	5 min	NA	↓ BP, HB, ↓ HR, asthma, HF
Diltiazem	Class I, LOE B	0.25 mg/kg IV over 2 min	2 to 7 min	5 to 15 mg/h IV	↓ BP, HB, HF
Verapamil	Class I, LOE B	0.075 to 0.15 mg/kg IV over 2 min	3 to 5 min	NA	↓ BP, HB, HF

Table 3: Intravenous & orally administered pharmacological agents for heart rate control in patients with AF with accessory pathways (21,22,23)

Drug	Class/LOE	Loading Dose	Onset	Maintenance Dose	Major Side Effects
Esmolol*†	Class I, LOE C	500 mcg/kg IV over 1 min	5 min	60 to 200 mcg/kg/min IV	↓ BP, HB, ↓ HR, asthma, HF
Metoprolol†	Class I, LOE C	2.5 to 5 mg IV bolus over 2 min; up to 3 doses	5 min	NA	↓ BP, HB, ↓ HR, asthma,

					HF
Propranolol†	Class I, LOE C	0.15 mg/kg IV	5 min	NA	↓ BP, HB, ↓ HR, asthma, HF
Diltiazem	Class I, LOE B	0.25 mg/kg IV over 2 min	2 to 7 min	5 to 15 mg/h IV	↓ BP, HB, HF
Verapamil	Class I, LOE B	0.075 to 0.15 mg/kg IV over 2 min	3 to 5 min	NA	↓ BP, HB, HF

Table 4: Intravenous & orally administered pharmacological agents for heart rate control in patients with AF and congestive heart failure, without accessory pathways (21,22,23)

Drug	Class/LOE	Loading Dose	Onset	Maintenance Dose	Major Side Effects
Esmolol*†	Class I, LOE C	500 mcg/kg IV over 1 min	5 min	60 to 200 mcg/kg/min IV	↓ BP, HB, ↓ HR, asthma, HF
Metoprolol†	Class I, LOE C	2.5 to 5 mg IV bolus over 2 min; up to 3 doses	5 min	NA	↓ BP, HB, ↓ HR, asthma, HF
Propranolol†	Class I, LOE C	0.15 mg/kg IV	5 min	NA	↓ BP, HB, ↓ HR, asthma, HF
Diltiazem	Class I, LOE B	0.25 mg/kg IV over 2 min	2 to 7 min	5 to 15 mg/h IV	↓ BP, HB, HF
Verapamil	Class I, LOE B	0.075 to 0.15 mg/kg IV over 2 min	3 to 5 min	NA	↓ BP, HB, HF

* Onset is variable and some effect occurs earlier.

† Only representative members of the type of beta-adrenergic antagonist drugs are included in the table, but other, similar agents could be used for this indication

in appropriate doses. Beta blockers are grouped in an order preceding the alphabetical listing of drugs.

‡ Amiodarone can be useful to control the heart rate in patients with atrial fibrillation (AF) when other measures are unsuccessful or contraindicated.

§ Conversion to sinus rhythm and catheter ablation of the accessory pathway are generally recommended; pharmacological therapy for rate control may be appropriate

in certain patients.

½½ If rhythm cannot be converted or ablated and rate control is needed, intravenous (IV) amiodarone is recommended.

¶ Adequacy of heart rate control should be assessed during physical activity as well as at rest.

– BP indicates hypotension; – HR, bradycardia; HB, heart block; HF, heart failure; LOE, level of evidence; and NA, not applicable.

Abbreviations

1. AF: Atrial Fibrillation
2. AFI: Atrial Flutter
3. RAFF: Recent onset atrial fibrillation or flutter (usually within 48-72 hrs)
4. A
5. CHF: Congestive Heart Failure
6. COPD: Chronic Obstructive Pulmonary Disease
7. DCCV: Direct Current Cardioversion
8. ED: Emergency Department
9. EDOU: Emergency Department Observation Unit
10. FY: Fiscal Year
11. HHS: Hamilton Health Sciences
12. ICC: Intra-cluster Correlation Coefficient
13. LVH: Left Ventricular Hypertrophy
14. TEE: Trans-Esophageal Echocardiogram

Appendix

Appendix 1:

Emergency Department Management of Acute Onset AF

A Survey of Current Practice

This survey is trying to gather information on the management of patients with acute onset AF (48-72hrs; Lone AF or AF with minimal comorbidities such as uncomplicated Hypertension or Diabetes Mellitus) and without significant comorbidities (eg. acute coronary syndrome, congestive heart failure, severe COPD, pneumonia etc.) in the ED. You are free to select more than one response and to enter in your own responses if it does not feature in the list of choices.

1. How do you manage acute onset AF patients presenting to the ED?
 - (a) Routinely refer all patients to the on call internal medicine/cardiology team in hospital
 - (b) Attempt rate or rhythm control and if unsuccessful refer to the internal medicine/cardiology team in hospital
 - (c) Send patients home once rhythm/rate control has been achieved and refer to internist/cardiologist for outpatient consultation
 - (d) Other- please elaborate.....
2. Do you prefer a rhythm or rate control strategy in such patients?
 - (a) Rate control strategy
 - (b) Rhythm control strategy
 - (c) Please fill in if you wish to elaborate.....
3. Would you prefer rate or rhythm control strategy in the following subgroup of patients? Please circle your choice.
 - (a) Patients aged over 65 years- Rate/ Rhythm

- (b) Persistent or long lasting AF- Rate/ Rhythm
 - (c) Unsure of duration of AF- Rate/ Rhythm
 - (d) CHADS score over ≥ 2 - Rate/ Rhythm
 - (e) Other. Please print.....
4. Which of the following medications do you use for rate control in the ED? Please list in order of preference
- (a) Beta Blockers- oral/iv
 - (b) Calcium Blockers- oral/iv
 - (c) Digoxin oral/iv
 - (d) Amiodarone iv
 - (e) Other (including combination therapy); please elaborate...
5. How do you define adequate rate control during AF?
- (a) Resting heart rate < 80 bpm
 - (b) Resting heart rate < 100 bpm
 - (c) Resting heart rate < 80 bpm and heart rate after 6 minute walk < 120 bpm
 - (d) Other. Please print....
6. Do any of the following situations hamper you while treating acute AF patients in the ED? Please circle your responses-
- (a) Side effects of rate and rhythm control medications- rarely/sometimes/frequently
 - (b) Lack of monitoring facilities - rarely/sometimes/frequently
 - (c) Inability to keep patient in ED long enough to control the ventricular rate - rarely/sometimes/frequently
 - (d) Inability to arrange follow up care for ensuring rate/rhythm control and systemic anticoagulation if required

- (e) Concern that oral medications may not control ventricular rate as effectively as iv medications after discharge
 - (f) Other; please print....
7. Which of the following reasons do you use to adopt a rhythm control strategy?
- (a) Younger patients (< 65 years)
 - (b) Personal Preference
 - (c) AF with duration < 48 hrs
 - (d) Patient preference
 - (e) Symptomatic patients (Shortness of breath)
 - (f) No major comorbidities
 - (g) Patient has only very infrequent symptomatic episodes of AF
 - (h) Other. Please elaborate.....
8. What antiarrhythmic medications do you prefer to use in the ED for rhythm control? Please rank them in order of preference.
- (a) Amiodarone..... preference
 - (b) Flecainide..... preference
 - (c) Propafenone..... preference
 - (d) Sotalol..... preference
 - (e) Procainamide.....preference
 - (f) Other. Please list....
9. Have you opted not to attempt pharmacological cardioversion or prescribe antiarrhythmic medications (propafenone/flecainide) for lack of information regarding the patient's left ventricular function?
- (a) Yes
 - (b) No

10. What proportion of patients with acute AF do you cardiovert electrically or pharmacologically?
- (a) Seldom
 - (b) 25%
 - (c) 50%
11. What are the hurdles you face in using cardioversion as a rhythm control strategy in this subgroup of patients?
- (a) Lack of certainty regarding duration of AF
 - (b) Inability to get transesophageal echo done in the ED to rule out LA/LAA clot
 - (c) Uncomfortable using Flecainide, Propafenone or Amiodarone for this indication
 - (d) Personal Preference
 - (e) Patient preference
 - (f) Lack of follow up care
 - (g) Inability to ensure therapeutic systemic anticoagulation following cardioversion
 - (h) Other. Please print....
12. How do you deal with acute AF patients that need systemic anticoagulation?
- (a) Refer to the internal medicine/thrombo team in hospital
 - (b) Discharge if patient has a family physician/internist/cardiologist has accepted follow up care
 - (c) Other. Please print....
13. What are the criteria you use to discharge acute AF patients directly from the ED?
- (a) Spontaneous reversion to sinus rhythm

- (b) Successful DC cardioversion and follow up anticoagulation with family physician or other physician
- (c) Adequate heart rate control of AF while at rest and a physician has accepted the responsibility for follow up care on an outpatient basis
- (d) Other. Please print...

14. Do you refer patients for follow up outpatient care after discharging them from the ED?

- (a) No
- (b) Yes; to an
 - 1. Internist
 - 2. Cardiologist
 - 3. Arrhythmia Specialist

Appendix 2: Definitions:

- i. **Acute Coronary Syndrome-** will be considered to have occurred if either of the three following criteria are met: 1) typical rise and gradual fall (troponin) or more rapid rise and fall (CK-MB) of biochemical markers of myocardial necrosis with at least one of the four followings: (a) ischemic symptoms, (b) development of pathological Q-waves on the ECG, (c) ECG changes indicative of ischemia, (d) coronary artery intervention, or 2) pathological findings of an acute myocardial infarction.
- ii. **Major Bleeding-** Major bleeding must satisfy one or more of the following criteria-

- Bleeding associated with a reduction in haemoglobin of at least 20 grams per litre or leading to a transfusion of at least 2 units of blood or packed cells
- Symptomatic bleeding in a critical area or organ: Intraocular, intracranial, intraspinal or intramuscular with compartment syndrome, retroperitoneal bleeding, intra-articular bleeding or pericardial bleeding

Major bleed are to be classified as life-threatening if they meet one or more of the following criteria:

- Fatal, symptomatic intracranial bleed; reduction in haemoglobin of at least 50 grams per litre; transfusion of at least 4 units of blood or packed cells, associated with hypotension requiring the use of intravenous inotropic agents; necessitated surgical intervention

iii. Congestive Heart Failure (CHF) and hemodynamic compromise-

Hospitalization for CHF will be considered to have occurred if the following criteria are met: signs and symptoms of CHF, radiological evidence of HF, patients with systolic blood pressure < 80mm Hg. and/or needing inotropic agents OR intravenous diuretic therapy, admission to hospital for HF (overnight stay).

Appendix 3: Description of the CHADS₂ score:

CHADS score or CHADS₂ score is a clinical prediction rule for estimating the risk of stroke in patients with non-rheumatic atrial fibrillation (AF; 119).

	Condition	Points
C	Congestive heart failure	1
H	Hypertension: blood pressure consistently above 140/90 mmHg (or treated hypertension on medication)	1
A	Age >75 years	1
D	Diabetes Mellitus	1
S ₂	Prior Stroke or TIA	2

Adding together the points that correspond to the conditions that a patient has will result in the CHADS₂ score. This score is used in the next section to estimate stroke risk.

Risk of stroke

According to the findings of the validation study, the risk of stroke as a percentage per year is:

Annual Stroke Risk		
CHADS ₂ Score	Stroke Risk %	95% CI
0	1.9	1.2–3.0
1	2.8	2.0–3.8
2	4.0	3.1–5.1
3	5.9	4.6–7.3
4	8.5	6.3–11.1
5	12.5	8.2–17.5
6	18.2	10.5–27.4