# REDUCING BLOOD SAMPLING VOLUMES IN THE INTENSIVE CARE UNIT

# SMALL-VOLUME BLOOD COLLECTION TUBES TO REDUCE ANEMIA AND TRANSFUSION IN INTENSIVE CARE UNIT PATIENTS

By DEBORAH M. SIEGAL, MD, BSC, MSC, FRCPC

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree Master of Science

McMaster University MASTER OF SCIENCE (2017) Hamilton, Ontario (Health Research Methodology)

TITLE: Small-Volume Blood Collection Tubes to Reduce Anemia and Transfusion in Intensive Care Unit Patients

AUTHOR: Deborah Siegal, B.Sc., M.Sc., M.D., F.R.C.P.C. (McMaster University)

SUPERVISORS: Professor Mark A. Crowther, Professor Nancy M. Heddle

NUMBER OF PAGES: xii, 85

## LAY ABSTRACT

Blood testing is a preventable cause of blood loss. Patients in the intensive care unit (ICU) have about 41 mL of blood taken per day for testing (like donating 1 unit of blood every 8 days). This contributes to anemia (low red blood cells) and transfusion, which are harmful. About 40% of ICU patients get at least one red blood cell transfusion which is a limited resource with health risks. Most of the blood sent to the laboratory is discarded (up to 90%) suggesting that volumes can be reduced without compromising care.

The goals of this thesis are to (i) summarize the evidence for reducing blood loss for laboratory testing; (ii) discuss cluster randomized trials; (iii) discuss use of health care administrative data for research; (iv) discuss the role of pilot studies; and (v) present a pilot stepped wedge cluster randomized trial of small-volume versus standard-volume blood collection tubes in ICU patients.

**ABSTRACT**

Blood sampling causes significant blood loss in intensive care unit (ICU) patients (up to 41 mL per day). Only 10% of the blood collected is used for testing suggesting that volumes can be reduced without compromising patient care or laboratory processes. Blood loss contributes to anemia which is highly prevalent in the ICU (>90% after 3 days) and is associated with major adverse cardiovascular outcomes and death.

Diagnostic blood loss increases the likelihood of red blood cell (RBC) transfusion which is administered to about 40% of ICU patients (half are given in absence of hemorrhage) and has significant health risks. Small-volume blood collection tubes, which collect about 50% less blood, are available, but rarely used in adults. They have the same cost as standard-volume tubes and are compatible with laboratory equipment. The rationale for the continued use of standard-volume tubes is a theoretical concern about inadequate volume for testing, and the absence of data showing the benefit of small-volume tube use on an important clinical outcome.

A study is needed to show that small-volume tubes reduce blood loss, anemia and RBC transfusion without harms or negative consequences on patient care and hospital procedures compared to standard-volume tubes. If this could be shown, it may lead to practice change regarding blood collection for laboratory testing. A stepped-wedge cluster randomized trial is the ideal study design for this low-risk intervention. By incorporating the small-volume tubes into routine clinical practice and using administrative and hospital electronic medical record data, this study would be a pragmatic, cost-effective way to evaluate effectiveness and implementation. However,

prior to conducting a full-scale trial powered on clinical outcomes, a pilot study is needed to determine whether a larger study will be feasible.

The goals of this thesis are to (i) summarize the existing evidence regarding small-volume tubes; (ii) discuss cluster randomized trial methodology;(iii) discuss the use of health care administrative data for research; (iv) discuss the role of pilot studies; and (v) present the design of a pilot stepped wedge randomized trial of small-volume versus standard-volume blood collection tubes to evaluate the feasibility of a full-scale trial.

**ACKNOWLEDGEMENTS**

I am grateful to my supervisors Dr. Mark Crowther and Prof. Nancy Heddle with whom this idea first emerged and subsequently developed into a research program. Completion of this thesis is a testament to their support and encouragement. I appreciate the insights of my committee members Dr. Deborah Cook and Dr. Adam Cuker. I extend additional gratitude to Dr. Stuart Connolly whose scientific input and encouragement were instrumental in launching the Small-Tubes to Reduce Anemia and Transfusion (STRATUS) study. Finally, I share this accomplishment with Dan Anskaitis and Chris Hillis who patiently supported me throughout this process and without whom it would not have been possible.

**TABLE OF CONTENTS**

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| ACS | acute coronary syndrome |
| BD | Beckton Dickinson |
| CCI | Canadian Classification of Health Interventions |
| CI | confidence interval |
| CIHI | Canadian Institutes of Health Information |
| DAD | Discharge Abstract Database |
| DE | design effect |
| EDTA | ethylenediamenetetraacetic acid |
| EPOC | Effective Practice and Organisation of Care |
| ESS | effective sample size |
| GRACE | Global Registry of Acute Coronary Events |
| ICC | intra-cluster correlation coefficient |
| ICD | International Statistical Classification of Diseases |
| ICH | International Council for Harmonisation |
| ICU | intensive care unit |
| LOS | length of stay |
| LR | likelihood ratio |
| MI | myocardial infarction |
| NIHR | National Institute for Health Research |
| OR | odds ratio |
| PRECIS | Pragmatic Explanatory Continuum Indicator Summary |
| RBC | red blood cell |
| ROBINS-I | Risk of Bias in Non-Randomized Studies of Interventions |
| SD | standard deviation |
| TACO | transfusion-associated circulatory overload |
| TRALI | transfusion-related acute lung injury |
| VAMP | venous arterial blood management |
| VIF | variance inflation factor |
| WHO | World Health Organization |

# CHAPTER 1: INTRODUCTION

## 1.1 Laboratory Testing Results in Unnecessary Blood Loss

Blood sampling can result in significant blood loss, particularly in hospitalized patients who undergo frequent blood testing. For example, in patients with acute myocardial infarction (MI), reported blood loss for diagnostic testing is about 21 mL per day[1]. Intensive care unit (ICU) patients undergo frequent blood testing, often multiple times each day, with sample volumes up to 41 mL per day reported (equivalent to donating a unit of blood every 8 days)[2]. Importantly, only about 10% of the blood collected is actually used for testing procedures with the remainder discarded as waste[3]. For example, in Hamilton, Ontario, Canada the volume of blood collected into each lithium heparin tube for chemistry testing is 4 mL, which yields 2 mL (2000 µL) of plasma. The minimum and maximum volumes of plasma required per chemistry test in our laboratory are 2 µL and 35 µL, respectively (personal communication, Dr. S. Hill, Clinical Biochemist, McMaster University). This suggests that blood sample volumes could be decreased without compromising patient care.

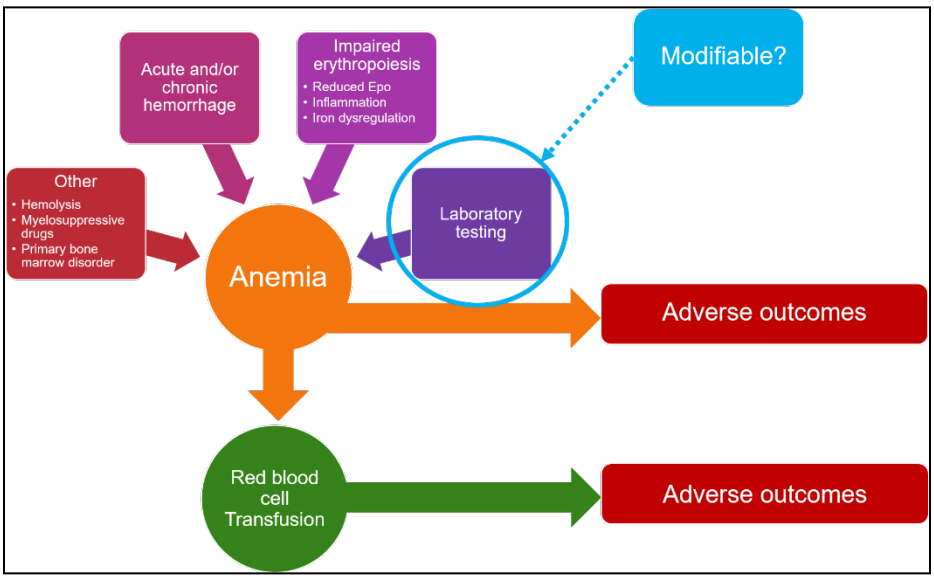
## 1.2 Consequences of Blood Loss from Laboratory Testing

### 1.2.1 Exacerbation of Anemia

Anemia is defined by the World Health Organization [WHO] as hemoglobin less than 120 g/L in women and 130 g/L in men. Up to 50% of hospitalized patients and 75% of the hospitalized elderly experience anemia[4]. Patients in the intensive care unit (ICU) are at particularly high risk for anemia with 60% anemic upon ICU admission, 90% by day 3 and 97% by day 8[2,5,6]. In hospitalized patients, anemia is often considered

multifactorial and generally attributed to acute or chronic hemorrhage and/or reduced erythropoiesis from infection, inflammation, and reduced erythropoietin production (**Figure 1.1**)[7]. Hemolysis, myelosuppressive drugs and primary bone marrow disorders can also contribute, although less commonly. Limited evidence suggests that frequent laboratory testing is a potentially modifiable cause of blood loss and a contributor to anemia. As a proof of concept, phlebotomy of 314 mL in healthy volunteer subjects led to a small but significant reduction in average hematocrit from 44.2% to 39.9%[8]. Patients with acute medical illness are expected to be more susceptible to the impact of blood loss due to a reduced production of and response to erythropoietin. For example, in non-anemic patients with MI, every 50 mL of total blood drawn for testing has been shown to increase the risk of moderate to severe anemia (hemoglobin ≤ 110 g/L) by 15% in adjusted analysis[1]. In hospitalized patients, blood loss of 100 mL was associated with a reduction in hemoglobin of 7 g/L[9].

**Figure 1.1 Conceptual Framework**



*1.2.1.1 Anemia is Associated with Adverse Cardiovascular Outcomes and Death*

Anemia is associated with adverse outcomes in a wide spectrum of patients including those with acute myocardial infarction[10-14], heart failure[15-17], renal failure[18,19], diabetes[20,21], stroke[22] and those undergoing surgery[23-25]. However, anemia remains an under-recognized cause of morbidity and mortality in hospitalized patients compared to competing medical conditions such as cardiovascular events (e.g. myocardial infarction [MI]), nosocomial infections and critical illness. ICU patients with lower hemoglobin levels during admission have been shown to have longer ICU and hospital length of stay [LOS] and higher mortality at 30 days[5]. In the prospective observational CRIT study, nadir (but not baseline) hemoglobin less than 100 g/L was associated with a longer ICU LOS compared to nadir hemoglobin 100 g/L or more (hemoglobin less than 80 g/L: LOS ratio 1.41, 95%CI 1.29-1.53; hemoglobin 80 to 89 g/L: LOS ratio 1.57, 95%CI 1.47-1.69; hemoglobin 90 to 99 g/L: LOS ratio 1.30, 95%CI 1.23-1.39, $p<0.0001$)[5]. Similarly, nadir (but not baseline) hemoglobin less than 100 g/L was associated with longer hospital LOS compared to hemoglobin 100 g/L or more (hemoglobin less than 80 g/L: LOS ratio 1.17, 95%CI 1.08-1.27; hemoglobin 80 to 89 g/L: LOS ratio 1.23, 95%CI 1.16-1.31; hemoglobin 90 to 99 g/L: LOS ratio 1.14, 95%CI 1.08-1.21, $p<0.0001$). Baseline hemoglobin levels were not statistically significantly associated with ICU or hospital length of stay. In multivariate logistic regression analysis, the odds ratio for 30-day mortality was higher in ICU patients with nadir hemoglobin less than 90 g/L compared to those with hemoglobin 100 g/L or more in that study (hemoglobin less than 80 g/L: odds ratio [OR] 1.54, 95%CI 112-2.12, $p<0.009$; hemoglobin 80 to 89 g/L OR 1.49, 95%CI 1.13-1.95, $p<0.004$).

In patients with MI, the development of moderate to severe hospital acquired anemia (defined as hemoglobin concentration ≤ 110 g/L) was associated with higher mortality (8.4% vs. 2.6%; HR 1.82, 95%CI 1.11-2.98) and poorer health status at 1 year compared to patients without anemia after adjustment for known confounding factors such as bleeding and Global Registry of Acute Coronary Events (GRACE) score[26]. In patients with ST elevation MI (STEMI), each 10 g/L decrement in hemoglobin is associated with increased cardiovascular mortality (adjusted odds ratio (OR) 1.21, 95% confidence interval [CI] 1.12-1.30) based on data from 16 Thrombolysis in Myocardial Infarction (TIMI) trials[10]. Similarly, each 10 g/L decrement in hemoglobin was associated with an increased risk of the composite outcome of cardiovascular death, recurrent MI or recurrent ischemia in patients with non-ST elevation acute coronary syndrome (ACS) (adjusted OR 1.45, 95%CI 1.33-1.58). Among all (ACS) patients, baseline hemoglobin <110 g/L increased the risk of the composite outcome by about 2-fold (OR 2.26, 95%CI 1.83-2.79). In patients with acute ischemic stroke, anemia (WHO definition) was associated with an increased odds of death in hospital (OR 1.87, 95% CI 1.57-2.47 for men and OR 1.47, 95% CI 1.08-1.98 for women) and up to 1 year post-stroke (OR 2.90, 95% CI 2.18-3.86 for men and 1.86, 95% CI 1.44-2.41 for women)[22].

### 1.2.2 Red Blood Cell Transfusion

Indications for red blood cell (RBC) transfusions in critically ill patients include hemorrhagic shock, acute hemorrhage with hemodynamic instability or evidence of inadequate oxygen delivery, and hemoglobin concentration of less than 70 g/L or less than 80 g/L depending on the patient population[27,28]. Current guidelines from the AABB

recommend a hemoglobin transfusion threshold of 70 g/L for hemodynamically stable hospitalized adult patients (including critically ill patients) and a threshold of 80 g/L for patients undergoing orthopedic or cardiac surgery, or those with underlying cardiovascular disease[29]. The Transfusion Requirements in Critical Care (TRICC) was a landmark trial that demonstrated reduced hospital mortality in patients randomized to a restrictive RBC transfusion strategy (< 70 g/L) compared to a conservative RBC transfusion strategy (<100 g/L) and, in conjunction with subsequent meta-analyses, established the recommended threshold for RBC transfusion in critical illness.

About 40% of ICU patients receive one or more red blood cell (RBC) transfusions to correct anemia, half of which are administered in the absence of clinically significant hemorrhage[2,5,30,31]. Up to 75% of patients with ICU stays of longer than one week receive RBC transfusion[2]. In a prospective observational study, 17.8% were admitted to ICU for more than 7 days[2]. Diagnostic blood loss is associated with an increased likelihood of RBC transfusion in non-bleeding ICU patients and accounts for 50% of the variation in the amount of RBC transfusion[30,32]. Among ICU patients admitted for 21 days or longer, increases in average diagnostic blood loss of 3.5 mL per day are associated with a 2-fold increase in the odds of transfusion[32].

#### *1.2.2.1 Health Risks of Red Blood Cell Transfusion*

Transfusions are a limited resource and are associated with health risks such as infection, transfusion-related acute lung injury (TRALI), transfusion-associated circulatory overload (TACO), hemolytic transfusion reactions, immunosuppression, allosensitization and allergy[33,34]. RBC transfusion is associated with death, longer ICU

and hospital admissions, infection, prolonged mechanical ventilation and organ dysfunction[2,5,35]. Two observational studies showed that receipt of a RBC transfusion in the ICU increased the odds of dying by 1.37 (95%CI 1.02-1.84) and 1.65 (95% CI 1.35-2.03) in adjusted analyses[2,5]. For example, in the CRIT study, RBC transfusion was significantly associated with an increased risk of death after propensity matching (1059 transfused and 1059 non-transfused patients) with an adjusted mortality ratio of 1.65 (95% confidence interval, 1.35–2.03; log-rank, p<0.001)[5]. The number of RBC units transfused was associated with increased ICU LOS (1-2 units: LOS ratio 1.47, 95%CI 1.35-1.59; 3-4 units: LOS ratio 1.84, 95%CI 1.66-2.03; more than 4 units: LOS ratio 3.20, 95%CI 2.88-3.55; p<0.0001) and hospital LOS (1-2 units: LOS ratio 1.32, 95%CI 1.22-1.42; 3-4 units : LOS ratio 1.61, 95%CI 1.47-1.75; more than 4 units: LOS ratio 2.51, 95%CI 2.28-2.76; p<0.001) compared with patients who did not receive transfusions. The median ICU and hospital LOS in non-transfused patients were 4.6 days and 11.0 days, respectively. Patients who received 1-2, 3-4, or more than 4 RBC units had increases in median ICU LOS of 2.1, 3.8, and 10.1 days, and increases in median hospital LOS of 3.5, 6.7, and 16.6 days, respectively.

*1.2.2.2 Cost of Red Blood Cell Transfusion*

One unit of RBCs costs between $400 and $500 (Canadian) (personal communication Dr. K. Webert, Canadian Blood Services). However, this estimate does not account for the cost of preparing and administering transfusion and managing adverse events. Interventions that reduce RBC transfusion have been shown to decrease cost[36,37]. For example, one blood conservation initiative which utilized an

electronic medical record transfusion trigger alert reduced RBC transfusion by 25% and produced estimated savings of $5.9 million over 4 years by reducing RBC unit acquisition costs[37]. This does not account for additional potential sources of savings to hospitals including reduced nursing workload (administration of transfusions, monitoring and management of adverse reactions), reduced laboratory workload (testing for transfusion and adverse reactions), reduced LOS, and 30-day readmissions.

## 1.3 Small-Volume Blood Collection Tubes

Small-volume (soft-draw) Vacutainer® tubes for blood collection (Beckton, Dickinson and Company) have the same cost and physical dimensions as standard-draw Vacutainer® tubes (Beckton, Dickinson and Company), but draw less blood (2 to 3 mL vs. 4 to 6 mL) due to lower vacuum inside the tube which fills to a smaller pre-determined volume. The soft-draw tubes are routinely used in children and patients who refuse blood products (e.g. patients who are Jehovah Witness) and are compatible with current laboratory equipment. Their routine use could decrease blood loss from laboratory testing in hospitalized patients.

### 1.3.1 Evidence for Use of Small-Volume Tubes

There is a paucity of studies evaluating the benefits and harms of small-volume blood collection tubes as a strategy to mitigate diagnostic blood loss in adult ICU patients. A systematic review conducted for this thesis work found 3 small observational studies (judged to have high risk of bias) that evaluated small-volume blood collection tubes to reduce diagnostic blood loss in ICU patients (see Chapter 2: Strategies to

Reduce Blood Loss from Laboratory Testing in ICU Patients: A Systematic Review)[38-40]. The volume of blood collected for laboratory testing in these studies was reduced by 42% to 74%. Dolman et al., showed that incident severe anemia (hemoglobin < 70 g/L) was less frequent with small-volume compared to standard-volume tubes (10% vs. 22%, p=0.01)[40]. This was the only study which evaluated RBC transfusion. There was a 27% reduction in RBC transfusion in the small-volume tube group, but this was not statistically significant. ICU LOS was similar in the 2 studies in which it was evaluated[39,40]. Although these studies were limited by small sample size and the likely presence of confounding factors that could influence results, the studies suggested that small-volume tubes may be a feasible and effective strategy for reducing diagnostic blood loss and RBC transfusion.

### 1.3.2 Why Are Small-Volume Tubes Not Used Routinely?

Although small-volume tubes are available for ordering and clinical use throughout Canada (including in hospitals in Hamilton), they are not used routinely in adults. Advances in laboratory equipment technology have led to reductions in the volume required for conducting testing. The reasons clinical practices for blood sampling have not kept pace with changes in laboratory processes are uncertain, but are likely multifactorial. Although some evidence regarding potential blood savings with small-volume tubes has been published since the 1980s, changes in clinical practice are complex, particularly without reliable evidence of benefits and harms[41]. Implementation involves an understanding of the intervention itself, the local context into which it is being introduced, and the behavioural strategies used for implementation[41].

This complexity is reflected by the scheme developed by the Cochrane Effective Practice and Organisation of Care (EPOC) Group to classify determinants of practice (i.e. barriers, facilitators, obstacles, enablers) as follows: information management, clinical uncertainty, competence, liability, patient expectations, standards of practice, financial considerations, administrative constraints and other[42]. For example, with regards to small-volume blood collection tubes, there is a general lack of awareness of their availability and favourable characteristics. There are also legitimate concerns from healthcare providers regarding the impact of blood volumes on patient care, laboratory processes (e.g. test validation), staff workload and validity of laboratory test results which may also limit implementation in the absence of evidence.

# CHAPTER 2: STRATEGIES TO REDUCE BLOOD LOSS FOR LABORATORY TESTING IN ICU PATIENTS: A SYSTEMATIC REVIEW

## 2.1 Objective

A systematic review was conducted to summarize the existing literature regarding: (i) the efficacy of strategies used to mitigate blood loss from diagnostic testing in ICU; and (ii) the impact of conservation strategies on RBC transfusion, mortality, and ICU and hospital LOS.

## 2.2 Methods

### 2.2.1 Search Strategy

An electronic search was conducted in MEDLINE and EMBASE for English-language studies published from inception until Oct 5, 2017 with assistance from Library Services at the Health Sciences Library of McMaster University. The search strategy is shown in **Figure 2.1**. A manual review of reference lists and citations was also conducted.

### 2.2.2 Study Selection

Studies were eligible if they included adult patients in intensive/critical care units and assessed the effect of strategies to reduce blood sampling volume compared to standard practice (or another intervention). Randomized control trials, non-randomized (observational) studies, systematic reviews and meta-analyses were eligible for inclusion. Studies in neonatal and pediatric patients, case reports, case series (n < 20

patients), reports lacking original data (e.g. editorials, narrative reviews), abstracts, and phlebotomy not for laboratory testing (i.e. treatment of polycythemia or iron overload) were excluded.

**Figure 2.1. Search Strategy**

| | |
|---|---|
| 1. Phlebotomy/ | 20. Intensive Care/ |
| 2. (venipuncture* or venesection or phlebotom*).ti,ab. | 21. Intensive Care.ti,ab. |
| 3. Bloodletting/ | 22. Critical Illness/ |
| 4. bloodletting.ti,ab. | 23. critically ill.ti,ab. |
| 5. Blood Specimen Collection/ | 24. Critical Illness.ti,ab. |
| 6. blood specimen collection*.ti,ab. | 25. ICU.ti,ab. |
| 7. blood sampling.ti,ab. | 26. intensive care units/ or burn units/ or coronary care units/ or respiratory care units/ |
| 8. or/1-7 | 27. Ventilators, Mechanical/ |
| 9. (blood adj2 loss).ti,ab. | 28. ventilator*.ti,ab. |
| 10. iatrogenic.mp. | 29. Respiration, Artificial/ |
| 11. 9 and 10 | 30. respirator*.ti,ab. |
| 12. 8 or 11 | 31. or/18-26 |
| 13. blood conservation.mp. | 32. 17 and 31 |
| 14. laboratory testing.mp. | 33. or/27-30 |
| 15. Diagnostic Tests, Routine/ | 34. 17 and 33 |
| 16. diagnostic test*.ti,ab. | 35. 32 or 34 |
| 17. or/12-16 | 36. limit 35 to "all adult (19 plus years)" |
| 18. critical care/ | 37. limit 36 to english language |
| 19. critical care.ti,ab. | |

### 2.2.3 Outcomes

The primary outcome was the volume of blood loss from diagnostic testing during ICU admission. Secondary outcomes included RBC transfusion during ICU admission, LOS in ICU and hospital, and mortality in hospital and ICU.

### 2.2.4 Data Abstraction

Two reviewers independently reviewed titles/abstracts and full-text articles for eligibility, and extracted data for analysis. Disagreements were resolved by consensus. A standardized data collection form was developed, piloted and then used for full-text review and data abstraction. The kappa statistic was used as a measure of inter-rater agreement.

### 2.2.5 Data Synthesis and Analysis

We planned to calculate a pooled estimate (meta-analysis) for the outcomes of interest, but after consultation with a statistician, pooling was judged to be inappropriate due to significant differences in patient populations, interventions, outcome reporting and summary statistics provided in individual studies. When clarification of published materials was required, attempts were made to contact the authors for missing data, but were unsuccessful. Although not meta-analyzing the available data reduces our ability to provide precise estimates of treatment effect, the differences between studies outlined above suggested that the results of this analysis would be inaccurate[43].

### 2.2.6 Quality Assessment

The risk of bias of included studies was assessed independently by 2 reviewers using the Risk of Bias in Non-Randomized Studies of Interventions (ROBINS-I) tool for non-randomized studies and the Cochrane Risk of Bias tool for randomized studies[44,45]. ROBINS-I evaluates the risk of bias in non-randomized studies comparing the effects of 2 or more interventions including quantitative studies assessing the benefits and/or harms of an intervention in which individuals were not randomly allocated to treatment interventions (observational studies). It is based on the Cochrane Risk of Bias tool for randomized studies and includes assessments of 7 bias domains: confounding, selection of participants, classification of interventions, deviations from intended interventions, missing data, measurement of outcomes and outcome reporting. The Cochrane Risk of Bias tool developed in 2005 and updated in 2011 is used to assess

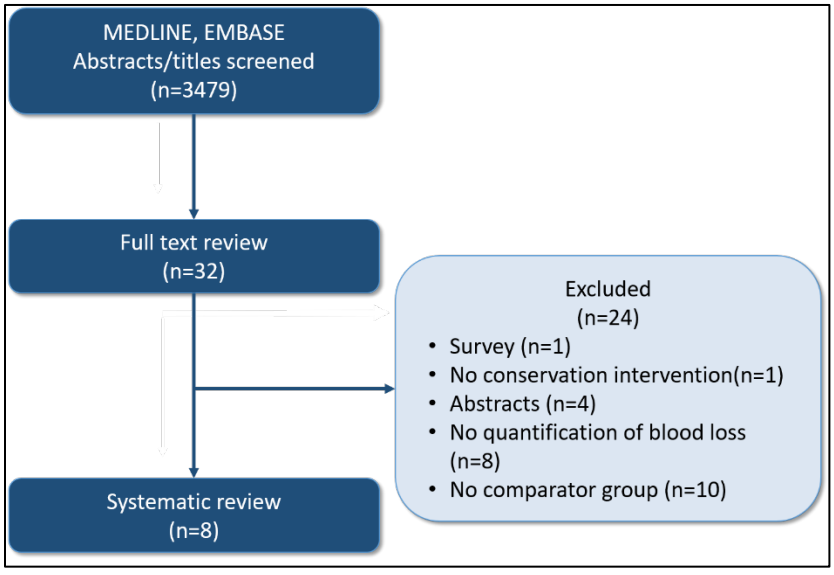
bias in 7 domains: selection bias (allocation sequence generation, concealment of allocation), performance bias (blinding of participants and personnel), detection bias (blinding of outcome assessors), attrition bias (incomplete reporting), reporting bias (selective reporting) and other bias.

## 2.3 Results

### 2.3.1 Search Results

As shown in **Figure 2.2**, the literature search yielded 3479 potentially eligible studies. After screening titles and abstracts, 3447 studies were excluded using predefined inclusion and exclusion criteria. Thirty-two studies were subjected to full-text review with an additional 24 studies excluded for the following reasons: no comparator group (n=10), no quantification of blood loss (n=8), abstract only (n=4), no conservation intervention (n=1), and survey (n=1). In total, 8 studies including 1204 patients were included in the analysis. Kappa was 0.97 (95%CI 0.92-1.0) indicating good agreement.

**Figure 2.2 Flow Diagram**



### 2.3.2 Study Characteristics

Study characteristics are shown in **Table 2.1**. The included studies consisted of 4 randomized controlled trials[46-49] and 4 observational studies[39,40,50,51]. The randomized controlled trials evaluated methods to reduce discarded blood when sampling from arterial catheters. Observational studies evaluated the presence and the absence of arterial catheters (n=1) and small-volume (pediatric) blood collection tubes to standard-volume tubes (n=3). All studies were judged to have significant risk of bias (**Table 2.2 and Figure 2.3**).

**Table 2.1 Study Characteristics.**

| Author (year) | Study Design | Total patients (n) | Population | Intervention | Comparator |
|---|---|---|---|---|---|
| **Blood Collection Tube Interventions** | | | | | |
| Dolman (2015) | Observational (retrospective) | 248 | Admitted to surgical ICU for 48 hours or longer | Small-volume blood collection tubes | Standard-volume blood collection tubes |
| Sanchez-Giron (2008) | Observational (prospective) | 473 | Not specified | Small-volume blood collection tubes | Standard-volume blood collection tubes |
| Smoller (1989) | Observational (prospective cohort with historical controls) | 56 | Admitted to ICU | Small-volume blood collection tubes | Standard-volume blood collection tubes |
| **Arterial Line Interventions** | | | | | |
| Gleason (1992) | RCT | 68 | Admitted to surgical ICU with arterial catheter | Arterial catheter sampling – closed method | Arterial catheter sampling – open method |
| Harber (2006) | RCT | 49 | Admitted to ICU with arterial catheter | Arterial catheter sampling – return blood | Arterial catheter sampling – discard blood |
| Low (1995) | Observational (prospective) | 50 | Admitted to medical/surgical ICU and APACHE score 9-20 | No arterial catheter | Arterial catheter placed within 6 |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | hrs of admission |
| MacIsaac (2003) | RCT | 160 | Admitted to ICU and arterial catheter placed | Arterial catheter sampling – VAMP Plus | Arterial catheter sampling – no VAMP |
| Peruzzi (1993) | RCT | 100 | Admitted to medical ICU and required arterial catheter | Arterial line sampling – VAMP Plus | Arterial line sampling – no VAMP |

APACHE score= acute physiology and chronic health evaluation score; ICU=intensive care unit; RCT=randomized controlled trial; VAMP=venous arterial blood management protection device.

**Figure 2.3 Risk of Bias Assessment for Randomized Studies**

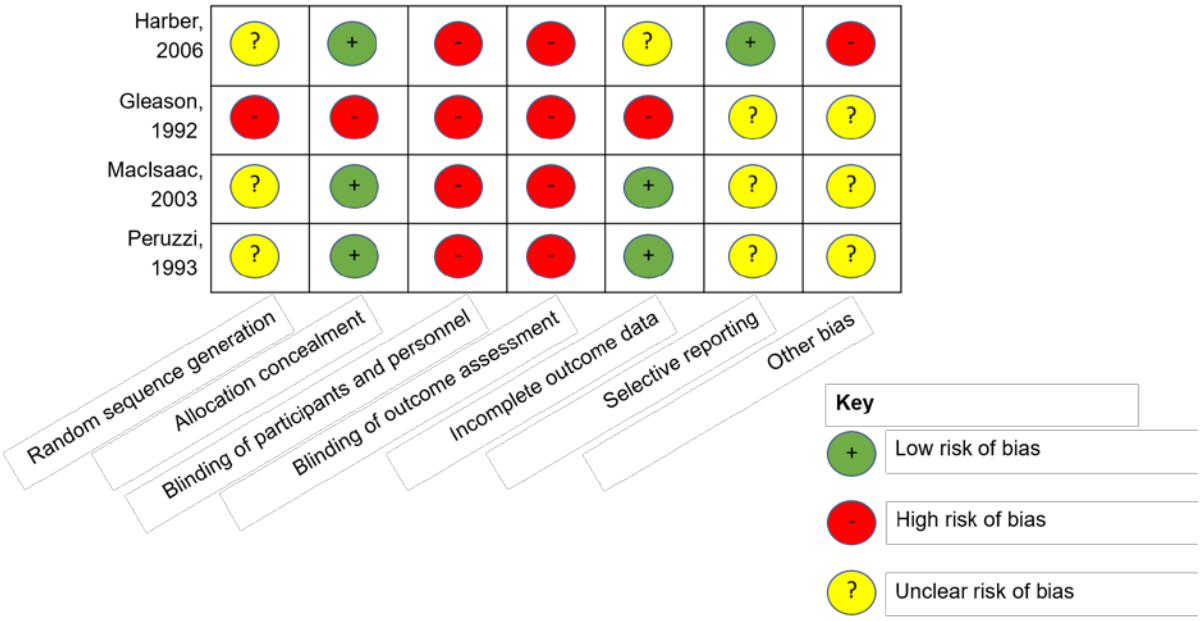


**Table 2.2 Risk of Bias Assessment for Non-Randomized Studies**

| Author (year) | Confounding | Selection | Classification | Deviations | Missing Data | Outcomes | Reporting | Overall |
|---|---|---|---|---|---|---|---|---|
| Dolman (2015) | S | NI | S | S | L | S | M | S |
| Low (1995) | S | S | L | NI | L | S | L | S |
| Sanchez-Giron (2008) | S | NI | S | NI | NI | S | NI | S |
| Smoller (1989) | S | L | L | NI | NI | S | NI | S |

Assessed using ROBINS-I tool for non-randomized studies. L=low risk of bias, M=moderate risk of bias, NA=not applicable, NI=no information, S=severe risk of bias.

### 2.3.3 Volume of Blood Loss

All 8 studies reported a reduction in blood loss from laboratory testing with interventions to reduce diagnostic blood loss compared to standard practice (range 18.9% to 80%) (**Table 2.3**). Compared to standard practice, diagnostic blood loss was significantly lower with conservation strategies in 6 studies which reported statistical analysis ($p<0.05$). All 4 randomized controlled trials reported reduced daily diagnostic blood loss using arterial catheter blood conservation devices compared to standard practice (mean/day 35 mL vs 69 mL, $p<0.01$; median/day 8 mL vs 40 mL, $p<0.001$; median/day 63 mL vs 133 mL, $p=0.001$; mean/day 260.3 mL vs 320.8 mL, p-value not specified). In all 4 observational studies, interventions reduced diagnostic blood loss compared to standard practice (mean/day 63.6 mL vs 114.7 mL, $p<0.001$; median/day 5.1 mL vs 19.9 mL, $p<0.001$; mean/day 32.2 mL vs 55.6 mL, p-value not specified). Use of small-volume blood collection tubes reduced daily blood loss (range 29% to 42%) and total blood loss during ICU admission (range 46.8% to 74.4%) compared to standard-volume blood collection tubes.

### 2.3.4 RBC Transfusions

RBC transfusion data were reported in 3 randomized studies and 1 observational study (**Table 2.4**). One randomized study showed a statistically significant decrease in the number of patients receiving RBC transfusion with the arterial catheter device intervention vs. control (21% vs. 38%, $p=0.01$)[48].

## Table 2.3 Blood Sampling Volumes

| Author (year) | Type of average and timeframe | Average volume of blood (mL) | | Percent change in volume of blood collected (%) |
|---|---|---|---|---|
| | | Intervention | Comparator | |
| **Blood Collection Tube Interventions** | | | | |
| Dolman (2015) | Mean (SD) per ICU stay | 174.0 (182.0) | 299.0 (355.0) | 41.8** |
| | Mean (SD) per day | 22.5 (17.3) | 31.7 (15.5) | 29.0** |
| Sanchez-Giron (2008) | Median (IQR) per ICU stay | 5.1 (2.3-10.9) | 19.9 (12.0-35.8) | 74.4** |
| Smoller (1989) | Mean per ICU stay | 120.2† | 226.1† | 46.8** |
| | Mean per day | 32.2† | 55.6† | 42.1** |
| **Arterial Line Interventions** | | | | |
| Gleason (1992) | Mean per day | 35.0† | 69.0† | 49.3** |
| Harber (2006) | Median (range) per day | 8.0 (7.0-10.0) | 40.0 (28.0-43.0) | 80.0** |
| MacIsaac (2003) | Median (range) per ICU stay | 63.0 (0-787.0) | 133.0 (7.0-1227.0) | 52.6** |
| Peruzzi (1993) | Mean per ICU stay | 260.3† | 320.8† | 18.9** |

*p<0.05. **p-value not provided. †standard deviation not provided.

## Table 2.4 Red Blood Cell Transfusion

| Author (year) | Patients transfused N (%) | | RBC units per patient Mean (SD) or median [range] | |
|---|---|---|---|---|
| | Intervention | Comparator | Intervention | Comparator |
| Dolman (2015) | Not reported | Not reported | 4.4 (3.6) | 6.0 (9.2) |
| Harber (2006) | 2/25 (8) | 3 /24(12) | Not reported | Not reported |
| MacIsaac (2003) | 17/80 (21)* | 30/80 (38)* | 4.0 [1.0-21.0] | 3.0 [1.0-11.0] |
| Peruzzi (1993) | 16/50 (32) | 13/50 (26) | 0.7 (1.3) | 0.6 (1.1) |

*statistically significant difference, p=0.01.

### 2.3.5 ICU Length of Stay

ICU LOS was reported in 3 randomized controlled trials and 2 observational studies (**Table 2.5**). There was no difference in duration of ICU admission between groups.

**Table 2.5 ICU Length of Stay**

| Author (year) | Descriptive Summary Statistic | ICU Length of Stay (days) | |
|---|---|---|---|
| | | Intervention | Comparator |
| Dolman (2015)<br>Emergency surgery/trauma<br>Medicine | Mean (SD) | 9.2 (10.1)<br>9.7 (8.8) | 10.6 (13.8)<br>6.6 (4.0) |
| Gleason (1992)* | Mean | 6.4 | 6.4 |
| Harber (2006) | Median (range) | 3.0 (3.0-4.8) | 3.0 (2.0-4.0) |
| MacIsaac (2003) | Median (range) | 2.0 (0.2-54.0) | 3.1 (0.2-30.0) |
| Smoller (1988)* | Mean | 3.7 | 4.1 |

*standard deviation not provided.

## 2.4 Discussion

Interventions to reduce diagnostic blood loss were reviewed in 8 studies including 1204 critically ill patients. This is the first systematic review evaluating the effect of diagnostic blood loss reduction techniques on blood loss and clinical outcomes. We identified 2 strategies in the included studies (i) strategies to reduce blood loss from arterial catheters; and (ii) smaller volume blood collection tubes for laboratory testing. The majority of the studies used strategies aimed at minimizing blood loss from arterial catheter sampling. These included (i) changes to practices for arterial catheter blood draws where blood that would normally be discarded is returned to the patient; (ii) a

closed system involving additional pressure tubing and stopcocks in which blood that would normally be discarded does not leave the tubing and is returned to the patient; and (iii) the VAMP and VAMP Plus systems (Baxter Healthcare Systems, Irvine, CA, USA) which are commercially available tubing devices which reinfuse blood that would normally be discarded. The remaining studies used small-volume (pediatric) sized blood collection tubes.

The main finding from our study is that, when used, these strategies significantly reduce blood loss associated with diagnostic testing in critically ill patients. Although the incorporation of these strategies was clearly associated with lower diagnostic blood loss, the impact on RBC transfusion, hospital/ICU LOS and mortality remain unclear. These outcomes were underreported precluding firm conclusions. Further, the studies which reported these outcomes were small and likely underpowered to detect differences between groups.

There were no assessments of potential harms associated with interventions. Therefore, the net clinical benefit of these strategies is unclear as the costs and harms associated with implementation are unknown. Arterial catheter sampling techniques may be associated with higher acquisition costs related to additional tubing or commercial devices, and time for training employees. There are also theoretical concerns regarding a risk of infection associated with intravascular devices. Small-volume blood collection tubes have not been studied on a large scale. The effect of these tubes on sample adequacy, laboratory test validity and staff workload are uncertain and these could affect implementation. Further, it is unclear whether small-

volume tubes perform similarly to standard-volume tubes in critically ill hypotensive patients with compromised vascular access.

### 2.5 Limitations

Our study has some important limitations. There was significant heterogeneity among studies including differences in study design, patient population, interventions used, and outcomes measured. As such, we were unable to pool results between studies. Outcomes of interest were not consistently defined or reported. Additionally, the majority of included studies were small (median 114 patients, range 49-473). Although the results are suggestive of efficacy with regards to reduction in blood loss, firm conclusions cannot be drawn about the magnitude of effect based on these data.

### 2.6 Conclusions

Strategies designed to mitigate diagnostic blood loss appear to be effective in reducing blood sampling volumes in ICU patients although the risk of bias, heterogeneity and lack of reporting of clinical outcomes (including benefits and harms) severely limit conclusions regarding the magnitude of effect. The impact of such strategies on patient-important clinical outcomes such as mortality is unclear. High methodological quality randomized studies evaluating patient-important clinical outcomes are needed to evaluate the efficacy and possible harms associated with interventions that reduce blood loss taken for laboratory testing.

# CHAPTER 3: STUDY DESIGN CONSIDERATIONS

## 3.1 Rationale for a Clinical Trial

Diagnostic blood loss, anemia and RBC transfusion are potentially modifiable complications of critical illness which are associated with poor outcomes[52-54]. A clinical trial which evaluates the impact of reducing the volume of blood collected for testing in critically ill patients is both timely and important, and focuses attention on the risks of anemia and RBC transfusion. By reducing the amount of blood collected per sample, small-volume blood collection tubes represent a simple, cost-neutral way to minimize blood loss from laboratory testing in ICU patients. A randomized trial is needed for an unbiased assessment of whether small-volume blood collection tubes can prevent anemia and avoid RBC transfusions without compromising care due to the need for additional investigations or other unanticipated impacts.

## 3.2 Impact

Such a trial has the potential to lead to a simple, immediate and impactful change in clinical practice. If it can be shown that small-volume tubes can be implemented for blood collection in adult patients without significant adverse consequences, and that their use is associated with improved patient outcomes, this could lead to widespread practice change regarding blood collection for laboratory testing. Further, this study has the potential to influence policy to reduce waste and encourage stewardship of valuable blood products.

### 3.3 What is the Optimal Design for this Randomized Trial?

An individual patient randomized trial design in which ICU patients are randomized to small-volume or standard-volume blood collection tubes would have significant feasibility challenges in a busy ICU setting and would not address whether a policy of using reduced-volume tubes for all adult ICU patients reduces anemia and RBC transfusion. Conversely, cluster randomized trial methodology is well suited to assess the effectiveness of the intervention when applied to a unit of care in clinical practice. By incorporating the research protocol and interventions into routine clinical care and evaluating outcomes that are available from hospital administrative databases and electronic medical records, such a trial would be pragmatic, cost-effective and implementable in a "real world" setting. Cluster trials are discussed in detail in Section 4.5.

### 3.4 Explanatory vs. Pragmatic Clinical Trials

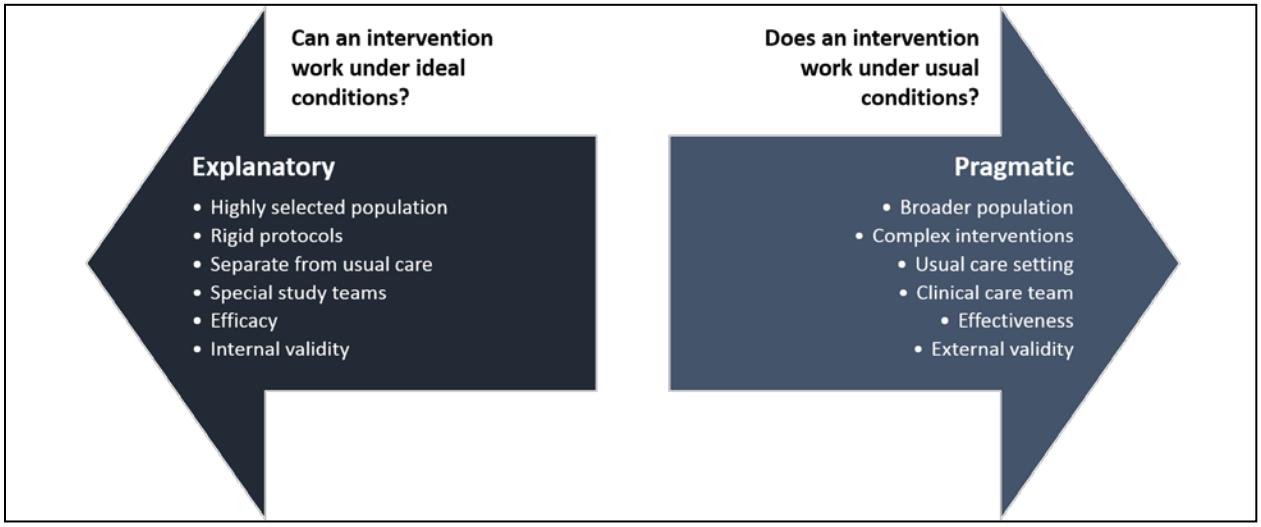
The distinction between explanatory and pragmatic clinical trials was originally proposed by Schwartz and Lellouch in 1967[55]. Although it is useful to consider these concepts dichotomously, very few trials are purely explanatory or pragmatic and current concepts support the existence of a spectrum between these extremes (**Figure 3.1**)[56].

#### 3.4.1 Explanatory Trials

The goal of explanatory trials is to test a hypothesis regarding the efficacy of an intervention under ideal conditions[57]. Explanatory trials seek to reduce the chance that another factor other than the intervention influences the outcome, thereby maximizing

internal validity. They include highly selected participants, and use rigid protocols executed by expert study teams separate from usual care. These characteristics which maximize internal validity also reduce external validity, or generalizability to patients seen in usual care settings. Explanatory trials are generally expensive and implementation of results into clinical practice is slow[58]. However, they provide the best evidence regarding the efficacy of an intervention under ideal circumstances and minimize sample size, so are more efficient.

**Figure 3.1 Spectrum of Explanatory vs. Pragmatic Trials**




## 3.4.2 Pragmatic Trials

Pragmatic trials evaluate the effects of an intervention in usual care settings, which maximizes generalizability (i.e. external validity) at the expense of internal validity. In doing so, pragmatic trials address the effectiveness of the intervention, or how well it works under usual circumstances[59]. Pragmatic trials include a broader (more diverse) population of participants and often evaluate complex interventions or policies applied in the usual care setting by the clinical care team. Cluster (group) randomization

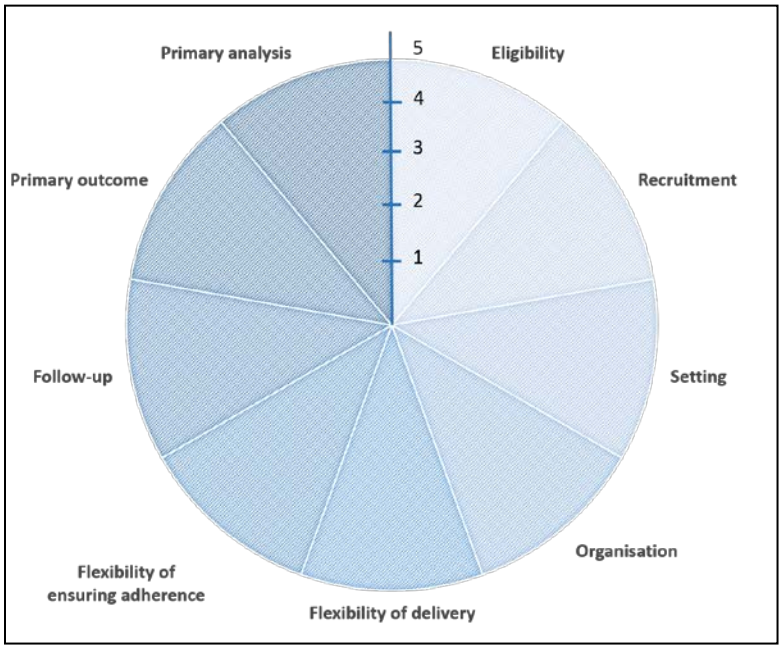
is frequently used such that the intervention is applied and studied as it would be used in practice[60]. An important feature of pragmatic trials is the use of data from hospital administrative databases and electronic medical records which can be repurposed for research after appropriate validation thereby reducing the cost and complexity of studies since such data is generally being collected as a component of patient care. Further, such data is usually collected using automated systems in a codified manner allowing it to be analyzed (in many cases) electronically. The use of administrative data for research is discussed in Chapter 5.

### 3.4.3 The PRECIS-2 Tool for Evaluating Trial Purpose and Design Choices

The Pragmatic Explanatory Continuum Indicator Summary (PRECIS) tool, published in 2009, was developed to assist researchers in designing clinical trials with an explicit purpose and design choices that are consistent with the intended purpose [56]. The new version of the tool, PRECIS-2, maintains the original objectives while addressing weaknesses such as a lack of face validity and inter-rater reliability, lack of scoring system, redundancy in some domains and lack of guidelines for use[61]. The PRECIS-2 tool is intended for use during the design process by following 4 steps: 1) clarifying the intention of the trial (explanatory vs. pragmatic); 2) considering design choices for PRECIS domains; 3) scoring the choices using a 5-point Likert scale from very explanatory (1 point) to very pragmatic (5 points) and creating a PRECIS-2 wheel (**Figure 3.2**); and 4) reviewing the design choices to determine whether the trial is consistent with the intended aim. Within the PRECIS-2 tool, there are 9 domains with which to evaluate the consequences of design decisions: 1) eligibility; 2) recruitment; 3)

setting; 4) organization; 5) flexibility of intervention delivery; 6) flexibility of assessing adherence; 7) follow-up; 8) primary outcome; and 9) primary analysis. By addressing trial applicability and encouraging researchers to address the impact of design decisions on domains 1 to 9 *a priori*, this tool can facilitate the design of studies that address the needs of patients, clinicians and policy-makers.

**Figure 3.2 The PRECIS-2 Wheel**




Adapted from Loudon et al 2015[61].

## 3.5 Cluster Randomized Trials

### 3.5.1 Overview of Cluster Randomized Trials

Whereas individual participants are randomly assigned to interventions in conventional individually randomized trials, groups of participants are randomized to interventions in cluster randomized trials[62]. Cluster randomized trial methodology was first used to conduct public health research in the 1970s because it is well suited to

evaluate complex interventions that target individuals and/or health systems[63]. The unit of randomization in cluster randomized trials can be units of care (e.g. hospital, clinic, hospital ward), communities, schools, churches, or workplaces.

As described by Campbell and Walters, there are 5 main reasons to use a cluster randomized trial design: (i) to avoid contamination which occurs when individuals crossover between treatment groups which reduces the treatment effect using intention-to-treat-analysis; (ii) to represent real-life practice; (iii) to enhance the convenience of research and reduce costs by delivering an intervention to a group instead of to individuals; (iv) to improve the effectiveness of an intervention by applying it to a group; and (v) to avoid ethical concerns that arise when individuals in the same group are treated differently[63-65].

An important challenge in cluster randomized trial design is the loss of efficiency compared to individual participant randomization (i.e. more participants are needed to achieve the same statistical power). This is because participants in the same cluster are more similar to one another than participants from different clusters and observations within a cluster vary less than observations from the overall population. As a result, the estimates of effect have higher variance (standard error) than individual randomized trials with the same number of participants. A detailed discussion regarding the effects of clustering is provided in Section 3.5.3.

### 3.5.2 Types of Cluster Randomized Trials

**Table 3.1** provides a summary of the key features, advantages and disadvantages of the different cluster randomized trial designs discussed below.

**Table 3.1 Features of Cluster Trial Designs**

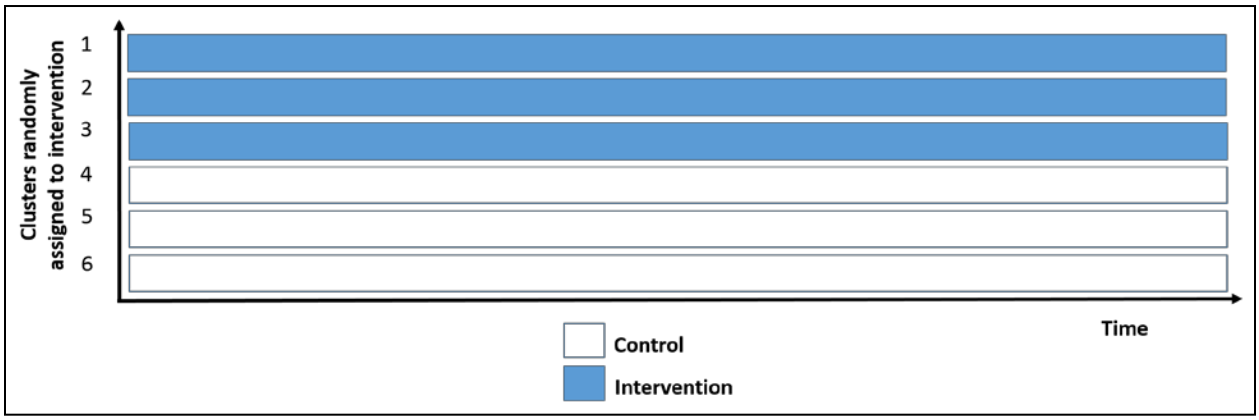
| Cluster Trial Design | Key Features | Advantages | Disadvantages |
|---|---|---|---|
| Parallel Group (Simple) | Clusters are randomly allocated to control or intervention for the duration of the study | Straightforward, easy to implement | Large design effect, possible ethical issues when clusters treated differently within a health system |
| Cluster Crossover | All clusters receive both control and intervention, sequence of treatment is randomly assigned (e.g. A→B, B →A) | More statistically efficient than parallel design, multiple crossovers can account for potential temporal effects | At risk of carryover effect between interventions, multiple crossovers challenging for complex interventions or policies |
| Stepped Wedge | All clusters switch from control to intervention, timing of switch between interventions is randomly assigned at predetermined timepoints (steps) | More statistically efficient than parallel design, useful for interventions with high likelihood of benefit and minimal harm, or where multiple switches between interventions is infeasible, or when the intervention is in limited supply | All sites need to be ready to implement intervention at trial start, at risk of carryover effect, at risk of temporal effects |

*3.5.2.1 Parallel Group (Simple) Cluster Randomized Trials*

In parallel group cluster designs, simple randomization is used to allocate clusters to interventions for the duration of the study (**Figure 3.3**)[64]. This design is straightforward and may be easy to implement. However, the power and sample size of

parallel group cluster trials are particularly affected by the effect of clustering because there are no within-cluster comparisons. As a result, power is reduced and sample size is increased. Another disadvantage of this design is that not all clusters receive the intervention which may be ethically unfavourable within a health system[60].

**Figure 3.3. Parallel Group Cluster Randomized Design**




*3.5.2.2. Cluster Crossover Design*

In cluster crossover designs, each cluster receives both interventions. But, the sequence in which the interventions are introduced is randomly assigned (e.g. A→B vs. B→A)[63]. **Figure 3.4** shows an example of a cluster crossover study with 2 treatments and 2 treatment periods. Cluster crossover trials are more efficient than parallel group designs because each cluster provides data for both treatments (i.e. each cluster acts as its own control)[64]. This results in a gain of precision and power due to removal of between-cluster variability from the variance of the estimates of treatment effect.

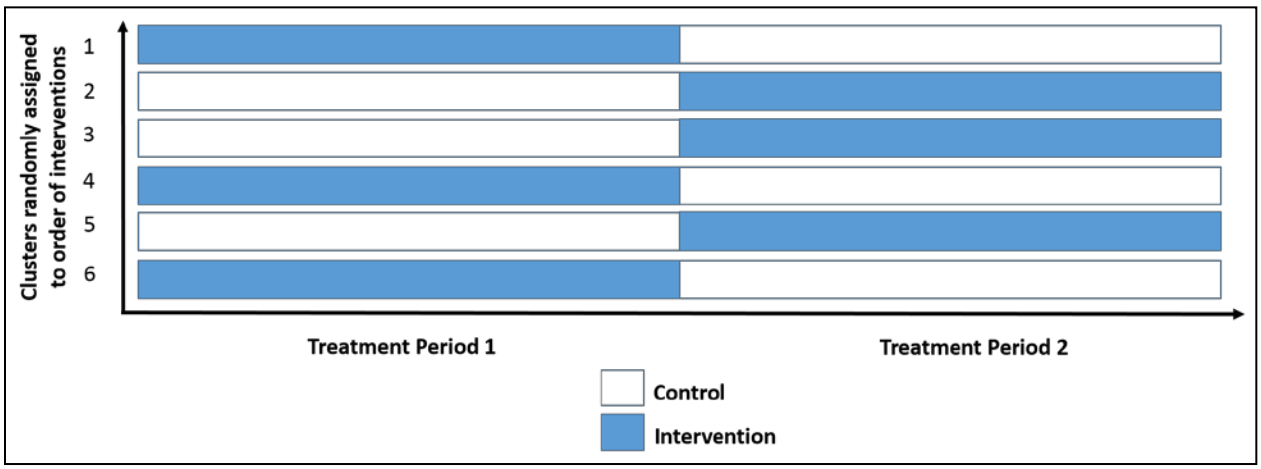
An underlying assumption of cluster crossover trials is that the intervention has no carry-over effect[63]. This may be true of some interventions, but not others. A washout period between treatments reduces the chance that residual effects of the

initial treatment affect the subsequent treatment. The washout period is assumed to be sufficient for the effect of the initial treatment to dissipate. A strategy to account for carry-over effect is the use of designs with multiple crossovers (e.g. A→A→B and B→B→A) such that any carry-over effect can be estimated.

Another limitation of cluster crossover designs is the possibility that temporal (period) effects may affect outcomes independent of a treatment effect[64]. These may occur due to secular trends or changes in the cluster environment. However, the use of multiple crossovers and statistical methods can account for period effects.

An alternative to the cluster crossover design is the partial crossover design in which a baseline observation period is followed by parallel cluster randomization which has the advantage of improving statistical power due to some within-cluster comparisons[60]. This design is useful for studies with feasibility challenges or lack of availability of the intervention as not all clusters receive the intervention.

**Figure 3.4 Cluster Crossover Design**



*3.5.2.3 Stepped Wedge Cluster Randomized Design*

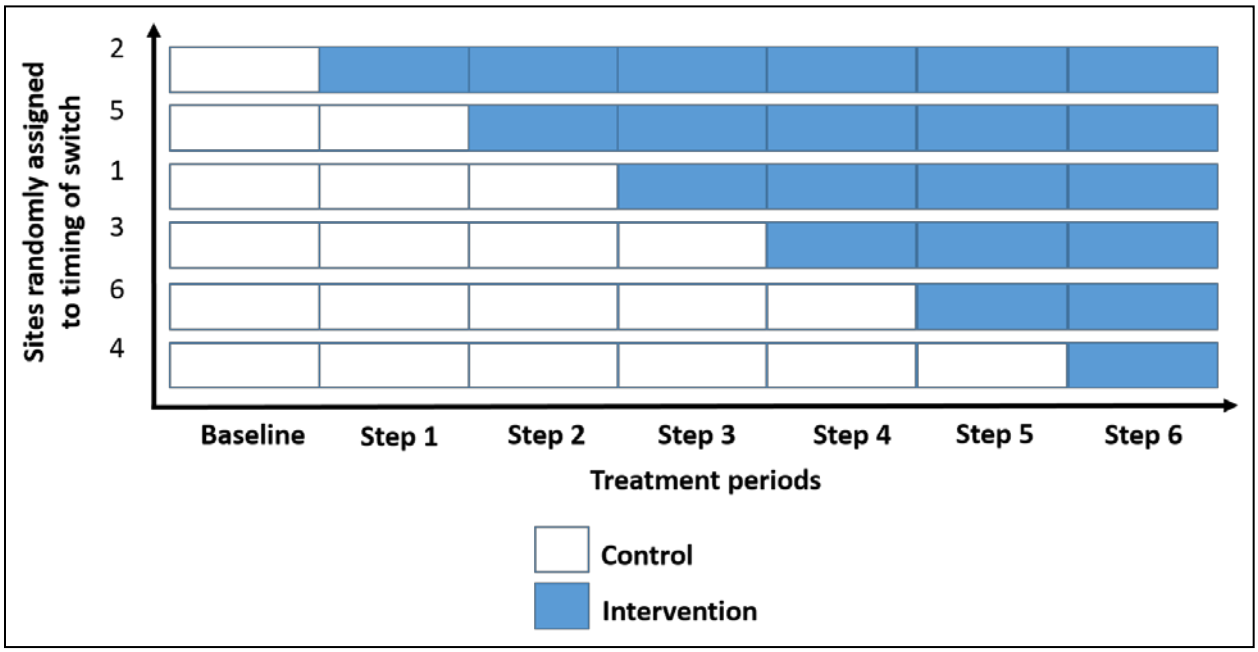
The stepped wedge cluster randomized trial design has been used increasingly to address research questions in a range of areas including infectious disease, cancer, social policy and criminal justice[66]. In this design, clusters switch sequentially from control to intervention, but the timing of the switch is determined randomly during study periods (steps). The study begins with a baseline period in which data are collected, but none of the clusters is exposed to the intervention. By the end of the study, all clusters have received the intervention (**Figure 3.5**). Similar to cluster crossover trials, stepped wedge designs regain some of the power lost due to clustering because of comparisons both between and within clusters (i.e. each cluster acts as its own control).

The stepped wedge design is particularly useful for settings in which there are logistical or political constraints which would interfere with the conduct of a more conventional research study. Because the intervention is implemented at all sites by the end of the study, it is well suited for interventions that have a high likelihood of benefit and minimal harm. An advantage of this design is sequential implementation which can be important for planning and executing complex interventions with logistical constraints or requiring input from multiple stakeholders. Outcome data can be cross-sectional (i.e. single measurements taken from individuals, but different individuals at each step) or longitudinal (repeated measurements on the same individuals throughout the study) in nature[66].

Similar to cluster crossover designs, stepped wedge designs are at risk of carry-over effect. A washout period during which the intervention is implemented but no outcome data is collected can reduce the chance that the effects of the first intervention

influence outcome data after switching. Stepped wedge designs are also susceptible to period effects especially because as the study progresses the proportion of clusters exposed to the intervention increases[66]. Therefore, data are collected prior to implementation of the intervention occur from an earlier calendar time period than observations after implementation of the intervention. Because calendar time could be associated with the intervention and possibly the outcome, it is a potential confounder and needs to be adjusted for during analysis. Because of the randomization of start time, time can be controlled for in the analysis[60].

**Figure 3.5 Stepped Wedge Design**




## 3.5.3 The Consequences of Clustering

### *3.5.3.1 Intracluster Correlation Coefficient*

An important consequence of clustering is that individuals within clusters are correlated, which means that two individuals within a cluster are more similar than two people selected randomly from different clusters[63]. As a result, cluster trials require a

larger sample size compared to corresponding individually randomized trials[66]. For sample size calculation, it is assumed that two observations from the same cluster have a constant correlation between them known as the intra-cluster correlation coefficient (ICC). The ICC is defined as the ratio of the between-cluster variance to the total variance of an outcome and is represented by the following equations[63]:

1. For continuous outcomes: $\rho = \sigma_\alpha^2 / \sigma_\alpha^2 + \sigma_\varepsilon^2$ (where $\sigma_\alpha^2$ is the between-cluster variance and $\sigma_\varepsilon^2$ is the within-cluster variance)
2. For binary outcomes: $\rho = \sigma_\alpha^2 / \pi(1 - \pi)$ (where $\pi$ is the probability of success and $\sigma_\alpha^2$ is the variance)

The ICC has possible values from 0 to 1, although values are typically small in human studies usually ranging from 0.01 to 0.02[67]. A value of 0 indicates that there is no correlation between individuals in the cluster with respect to the measurement such that they are *not* more similar to one another than to the general population[65]. Conversely, a value of 1 means that individuals in the cluster are exactly the same as one another. As a result, a high ICC (greater similarity of individuals within a cluster) decreases the effective sample size and reduces the precision of the estimates of treatment effect and statistical power. Conversely, for a low ICC, the effective sample size is closer to the total number of individuals in the trial. **Table 3.2** shows the effect of ICC on statistical power with different cluster trial designs. As shown, the stepped wedge design is the most efficient (i.e. minimizes the number of clusters and total cluster size) when the ICC is higher[68]. The stepped wedge design also achieves much higher power than the parallel design particularly when the number of clusters is low

and the ICC is high. Put another way, in stepped wedge trials power is determined primarily by within-cluster variability and is less sensitive to between-cluster variability compared to parallel cluster trials[69].

**Table 3.2 Effect of ICC on Statistical Power**

| | ICC 0.01 | | | ICC 0.1 | | |
|---|---|---|---|---|---|---|
| | Parallel cluster trial | Cluster crossover trial | Stepped wedge trial | Simple parallel cluster trial | Cluster crossover trial | Stepped wedge trial |
| Number of clusters = 10, cluster size = 100 | | | | | | |
| Total sample size | 1000 | 1000 | 1020 | 1000 | 1000 | 1020 |
| Power | 61% | 43% | 55% | 16% | 41% | 49% |
| Number of clusters = 10, cluster size = 300 | | | | | | |
| Total sample size | 3000 | 3000 | 3000 | 3000 | 3000 | 3000 |
| Power | 78% | 87% | 91% | 16% | 83% | 90% |

Adapted from[68]. This example relates to a trial requiring 786 observations for an individually randomized trial. The stepped wedge trial assumes 5 steps with 2 clusters randomized per step. ICC=intracluster correlation.

*3.5.3.2 Design Effect*

As a result of clustering, there is a net loss of independent data when accounting for similarities between individuals in a cluster[67]. The increase in sample size needed is determined by calculating the design effect (DE) which depends on the magnitude of the ICC and the number of subjects in each cluster. The DE is defined as the ratio of the variance of an outcome accounting for the effect of clustering to the variance of the outcome without accounting for clustering[63]. For clusters of equal sample size, it is represented by the equation $DE = 1 + (m - 1) / \rho$ (where m is sample size and $\rho$ is the

ICC). Because it measures the amount needed to increase the variance estimate to account for clustering, the DE it is also known as the variance inflation factor (VIF). **Table 3.3** shows the effect of ICC on the DE, total sample size and number of clusters for different cluster trial designs.

**Table 3.3 Effect of ICC on Design Effect**

| | | Parallel cluster trial | | | Cluster crossover trial | | | Stepped wedge trial | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Cluster size** | **ICC** | **DE** | **Total sample size** | **No. clusters** | **DE** | **Total sample size** | **No. clusters** | **DE** | **Total sample size** | **No. clusters** |
| 100 | 0.01 | 1.99 | 1569 | 16 | 2.64 | 2084 | 21 | 2.16 | 1702 | 18 |
| 100 | 0.25 | 25.75 | 20291 | 203 | 2.92 | 2298 | 23 | 2.25 | 1772 | 18 |

Adapted from[68]. This example relates to a trial requiring 788 observations for an individually randomized trial. The stepped wedge trial is assumed to have 9 steps. DE=design effect, ICC=intracluster correlation coefficient.

*3.5.3.3 Effective Sample Size*

The effective sample size (ESS) is the calculation of sample size (total number of clustered subjects) which takes into account the effects of clustering[67]. It is represented by the following equation ESS = mk / DE (where m is the number of subjects in a cluster, k is the number of clusters and DE is the design effect). A large DE effectively reduces the subjects enrolled in the trial from a statistical perspective. **Table 3.4** shows the relationship between DE, ESS and power. A large number of clusters and low number of individuals within a cluster give the smallest DE. Therefore, when designing a study, a high number of clusters and low number of subjects within a cluster give the smallest DE and highest power.

These equations can be used to calculate sample size estimates for a given ICC and desired power by varying the number of subjects per cluster and number of clusters. Sensitivity analysis can also be conducted to explore the impact of various

values of ICC on sample size calculations using the most conservative estimate that is practically feasible[62].

**Table 3.4 Relationship Between Design Effect, Effect Sample Size and Power**

| No. clusters | No. patients | Total no. subjects | DE | ESS | Power |
|---|---|---|---|---|---|
| 4 | 32 | 128 | 1.527 | 84 | 61% |
| 8 | 16 | 128 | 1.255 | 102 | 70% |
| 16 | 8 | 128 | 1.119 | 114 | 75% |
| 32 | 4 | 128 | 1.051 | 122 | 78% |
| 64 | 2 | 128 | 1.017 | 126 | 79% |
| 128 | 1 | 128 | 1.000 | 128 | 80% |

Adapted from[67]. Values shown are for ICC of 0.017 and total sample size of 128.

### 3.5.4 Ethical Issues in Cluster Randomized Trials

The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) contains regulations that outline international standards under which clinical trials are conducted to protect the interests of trial participants[70,71]. Randomized clinical trials in which therapeutic interventions are being tested typically require participants to provide informed consent. This ensures that the expected risks and benefits are communicated to and understood by eligible participants.

Ethical issues arise with regards to cluster trials because they differ from individual randomized trials in important ways such that standard research ethics guidelines are difficult to apply[72]. In cluster trials, the identification of individuals considered to be research participants and, therefore, entitled to ethical and regulatory protections is complicated. For example, interventions are allocated to groups of individuals (e.g. hospital, community), and may include health professionals in addition to patients[72]. The assessment of benefits and harms is also more challenging in cluster

trials because they may have consequences for groups as well as individuals[72]. Furthermore, it may be logistically difficult or even impossible to obtain informed consent from individuals especially for complex interventions involving large units of care or communities[62].

Some of the main ethical issues raised regarding cluster trials are summarized as follows[63]:

i) Determining whether the research subject is the cluster or the patient

ii) Determining whether informed consent must be obtained and, if so, from whom, how and when. If consent is required, determining how non-consenting patients will be handled.

iii) Determining whether clinical equipoise applies to cluster randomized trials. For example, exposing some clusters to the potential benefits of interventions and/or delaying potential benefits may have ethical implications.

iv) Determining whether the benefits outweigh the risks of cluster randomized trials. For example, in individual randomized trials may be discontinued when there is evidence of harm. This is more difficult in cluster trials when harm may be apparent in one cluster but not another receiving the same intervention.

v) Protecting vulnerable groups. This is particularly relevant for cluster trials that are conducted in developing countries where subjects enrolled in trials may receive a higher standard of medical care than the population.

vi) Identifying the “gatekeepers” who represent the interests of clusters and cluster members and determining their responsibilities.

The Ottawa Statement is a series of 15 recommendations addressing 7 ethical issues developed by a team of 15 investigators from Canada, the United Kingdom and the United States to guide researchers and research ethics committees in the ethical conduct of cluster randomized trials (**Table 3.5**)[72,73].

**Table 3.5 Summary Points from Ottawa Statement Checklist**

| Ethical Issue | Recommendation | Summary Description |
|---|---|---|
| Justifying cluster randomized trial design | 1 | • Clear rationale and use of appropriate statistical methods |
| Research ethics committee review | 2 | • Review and approval by research ethics committee before commencing |
| Identifying research participants | 3 | • Research participants are individual whose interests may be affected by study interventions or data collection |
| Informed consent | 4 | • Informed consent is required unless a waiver of consent is granted by a research ethics committee |
| | 5 | • When participant informed consent is required but not possible before randomization of clusters it must be obtained as soon as possible after randomization and before commencing study interventions or data collection |
| | 6 | • A waiver of consent or alteration of consent requirements may be granted by research ethics committees when the research is not otherwise feasible and the interventions and data collection procedures pose minimal risk |
| | 7 | • Informed consent is required from professionals/health service providers who are research participants unless a waiver of consent is granted |
| Gatekeepers | 8 | • Should not provide proxy consent on behalf of clustered individuals |

| Ethical Issue | Recommendation | Summary Description |
|---|---|---|
| | 9 | • Gatekeeper permission is required to enrol the cluster or organization in the trial (does not replace informed consent process) |
| | 10 | • Mechanisms to protect cluster interests should be in place with cluster consultation regarding study design, conduct, and reporting |
| Assessing benefits and harms | 11 | • Adequate justification of the intervention such that benefits and harms are consistent with practice in the field of study |
| | 12 | • Adequate justification of the choice of control such that individuals in the control arm are not deprived of care that they would have access to in the absence of a trial |
| | 13 | • Adequate justification of data collection procedures such that risks are minimized |
| Protecting vulnerable participants | 14 | • Determine whether additional protections are needed in consultation with research ethics committees |
| | 15 | • If individual informed consent is required research ethics committees should take additional care to recruitment, privacy and consent for participants who may be less able to freely choose participation |

Adapted from[73]

# CHAPTER 4. USE OF HEALTH CARE ADMINISTRATIVE DATA FOR RESEARCH

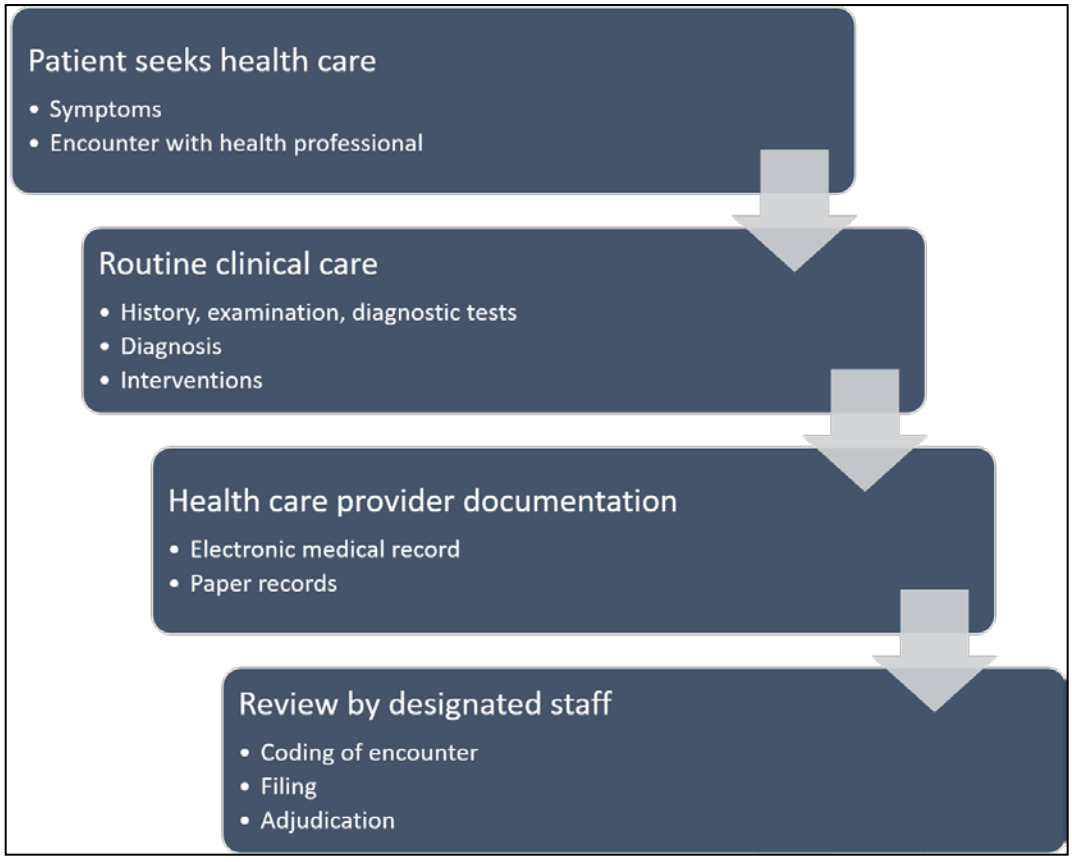
Administrative data refers to data collected for purposes other than research. Because it is available through pre-existing electronic mechanisms, administrative data can be collected with increased efficiency and lower cost compared to traditional manual prospective data collection conducted for prospective observational studies or randomized trials. This allows the completion of less expensive and less administratively complex (and thus more pragmatic) clinical research studies. This chapter describes the nature of administrative data including a discussion of the strengths and limitations when used for clinical research purposes.

## 4.1 What is Health Care Administrative Data?

Health information is increasingly digitized and collected by physicians, pharmacists, hospitals and health insurers (including government). Administrative health data include individual patient information regarding diagnoses, procedures/surgeries, medications, laboratory tests, radiological investigations, visits to the emergency department and hospitalizations[74]. Frequently, these data are derived from physician services claims, drug prescriptions and hospitalization abstraction. Because this information represents a patient's course through the health care system, it can be used to conduct research including epidemiologic, comparative effectiveness and health care utilization studies. Administrative data are being incorporated into innovative prospective studies including randomized trial designs by leveraging information that is already being collected for other databases thereby reducing the cost and increasing efficiency[75].

Unlike primary data collection (e.g. prospective observational or randomized studies) for which the timing and nature of data collected is determined by the investigator, administrative data (i.e. secondary data collection) are generated only if there is an encounter with the health care system associated with a diagnosis, procedure or prescription[76]. Information regarding the encounter must be documented, filed electronically, reviewed and coded accurately. **Figure 4.1** shows a schematic representation the process of data collection for administrative purposes.

**Figure 4.1 Generation of Administrative Data in Health Care**




Adapted from[76].

When using administrative data, it is important to understand the nature of the data collected and describe the methods used to create the analytical dataset (e.g. sampling frame, sampling methods, inclusion and exclusion criteria) which have an

impact on internal and external validity[74]. This includes an understanding of a database's intended purpose (i.e. why was it created) and the methods used to collect data (i.e. how was it created) which influence the completeness and accuracy of the data collected. Datasets may be more complete when those providing information benefit from providing data[74]. For example, databases derived from physician services claims may be more complete for physicians who work on a fee-for-service basis than they are for physicians who are salaried, as the claims data are used for reimbursement in a fee-for-service system[77]. Similarly, databases containing information regarding hospital laboratory tests which are downloaded from the laboratory information system are likely to be highly accurate because the data are collected automatically from the same source that provides information for real-time clinical care and the data does not undergo additional modification or adjudication. Conversely, data regarding diagnoses may be less accurate because diagnostic codes require a disease to be diagnosed, appropriate codes to be available, documented legibly in the hospital record, recognized and interpreted by data abstractors and coded accurately. Data quality checks such as logical checks and random chart re-abstractions ensure completeness and accuracy of administrative data.

### 4.1.1 Randomized Trial Data vs. Administrative Data

The purpose of randomized controlled trials is to compare the effect of one or more interventions on an outcome. Randomization reduces the chance that the observed effect is due to factors other than the interventions being compared. In addition to randomization, randomized controlled trials include highly selected

participants which reduces the generalizability of results to patients in real-world practice. This results from the application of strict inclusion and exclusion criteria which may relate to disease severity, comorbidities, medication use and likelihood of compliance with the study protocol[78]. In addition, randomized controlled trial data are gathered by specialized teams outside routine clinical care and require significant financial, logistical and administrative resources. Because of the logistical and financial challenges of non-pragmatic randomized controlled trials, they are usually of relatively short duration.

Unlike randomized controlled trials, administrative database research is observational in nature leading to important challenges from bias and confounding (see discussion below). However, information linked between physician, hospital and pharmacy information is comprehensive and allows researchers to study associations between diseases, exposures and outcomes[78]. Using administrative data, researchers can conduct large, longitudinal studies of demographically and geographically diverse patients which affords considerable statistical power. Such research is less costly than randomized controlled trials because it leverages the use of data that are already being collected for other purposes.

### 4.1.2 Canadian Institute for Health Information Discharge Abstract Database: An Example of A Health Care Administrative Database

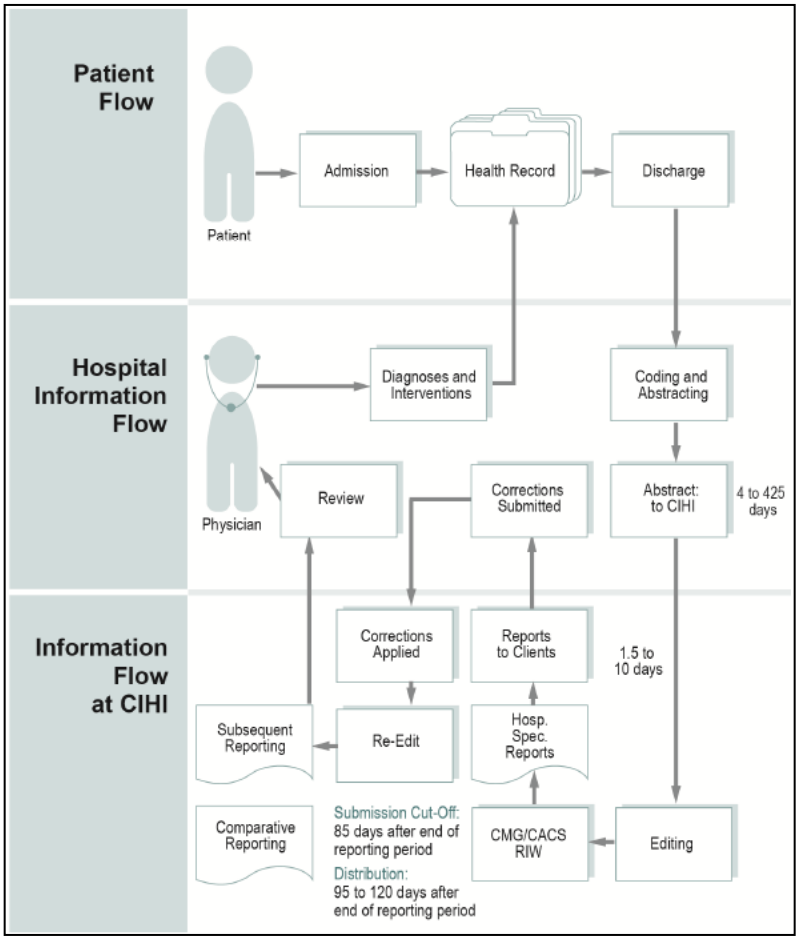
The Canadian Institute of Health Information (CIHI) collects information about Canada's health care system and the health of Canadians[79]. The CIHI discharge abstract database (DAD) is a national database which contains information regarding

discharges, deaths, sign-outs and transfers from acute care institutions in Canada (except those in Quebec) [80]. Other types of care are also captured in the DAD including day surgery procedures, long-term care, and rehabilitation. Data submitted to the DAD represent all "separations from acute inpatient care and day surgery institutions in Canada (excluding stillbirths and cadaveric donor cases from April 1 to March 31" (except those in Quebec) [80]. The DAD contains an inventory of institutions which is validated by individual provinces and territories. The flow of information into the DAD is depicted in **Figure 4.2**.

Data regarding each hospital separation (discharge, death, sign-out, transfer) include coded information regarding diagnoses and interventions, and patient demographics[80]. Data are coded using the International Statistical Classification of Diseases and Related Health Problems 10th Revision, Canada (ICD-10-CA) classification and the Canadian Classification of Health Interventions (CCI) which was developed by CIHI and is the Canadian standard for classifying health care interventions. An abstraction manual containing definitions, instructions, and validation rules in addition to education sessions and client support ensure that data submitted to the DAD data accurately reflect an institution's activities. Data completeness is evaluated on an ongoing basis through analysis of the number of reports generated. An editing and correction process verifies individual data elements. Re-abstraction studies conducted by CIHI evaluate data quality by reviewing original sources of information (e.g. patient chart) and comparing the information obtained with that contained within the DAD. Health information professionals external to participating hospitals conduct the re-abstraction by reviewing hospital charts for acute care diagnosis, interventions and

other data elements submitted to CIHI. In 2009-2010 data from 581 acute care facilities (9 provinces, 3 territories) was received by CIHI[81]. For the 2009-2010 Data Quality Study, 19 re-abstractors reviewed approximately 14,000 acute care abstracts from 85 hospitals and compared the data to previously collected DAD data. The database elements evaluated in the 2009-2010 Data Quality Study are listed in Table 4.1.

**Figure 4.2 Information Flow to CIHI Discharge Abstract Database**




From: Data Quality Documentation, Discharge Abstract Database – Multi-Year Information, Standards and Data Submission, Canadian Institute for Health Information.

**Table 4.1 CIHI Discharge Abstract Database Data Quality Study 2009-2010**

| Database Element Evaluated | Details |
|---|---|
| Coding of interventions | |
| Coding of significant diagnoses | |
| Coding of most responsible diagnosis | |
| Consistency of diagnosis typing | Assignment of significance<br>Most responsible diagnosis<br>Pre-admit comorbidities<br>Post-admit comorbidities |
| Coding for selected health conditions | Infections due to multi-drug resistant organisms<br>Palliative care<br>Pneumonia<br>Post-admit comorbidities<br>Obstetrical trauma<br>Birth trauma<br>Post-intervention conditions<br>Flagged interventions (non-invasive biopsy, per orifice endoscopy, tracheostomy, feeding tube, mechanical ventilation)<br>Intervention pre-admit flag (thrombolytic therapy, induction of labor) |
| Quality of case-mix grouping variables | Major clinical category and case mix group<br>Comorbidity level<br>Expected length of stay<br>Resource intensity weight |

## 4.2 Advantages of Using Administrative Health Data for Research

Health care administrative data contain information collected routinely during health care utilization in daily clinical practice. Such data can be accessed by researchers from individual hospitals, clinics, pharmacies, health systems and/or insurers and has several advantages when applied to clinical research[76,82]. Electronic information which is automatically stored for routine clinical practice provides enormous

potential for answering a variety of research questions in a relatively timely manner and at low cost. Using administrative data, researchers gain access to health information of a large number of patients which allows the study of rare events. The data obtained represent longitudinal routine clinical care and can therefore be used to address additional research questions regarding real-world effectiveness (e.g. of interventions or policies), health care utilization, health care quality, predictive risk modelling, and disease or adverse event surveillance over time[83].

### 4.3 Challenges of Using Administrative Health Data for Research

There a number of important challenges when using administrative data for clinical research[76,82]. Because administrative data are collected for clinical care and/or reimbursement, data are generated only when there is an encounter with the health care system. The encounter needs to be documented, filed and coded appropriately which raises concerns about data completeness, accuracy and precision. It is possible that diseases, exposures and outcomes are misclassified or omitted. Further, variables for analysis are limited to data which are collected routinely and there may be missing data elements and/or unmeasured confounders. Even though it is possible to link databases to widen the scope of information, data quality and integrity vary among health care databases.

#### 4.3.1 Internal Validity

Internal validity refers to the appropriateness of inferences made about the relationships between variables in a study, in particular a causal relationship between

an exposure and an outcome[78]. In other words, internal validity is the ability of a study to measure what it is designed to measure. Bias threatens internal validity because it undermines the ability to infer causality. Bias refers to systematic error or systematic deviation from the truth. In randomized trials, the process of randomization serves to eliminate bias such that any differences detected in outcomes between groups can be attributed to the effect of the intervention. Observational studies, including administrative database studies, are at risk of bias because study participants are not randomized to interventions resulting in the presence of both known and unknown differences between groups which can affect outcomes. Biases have been classified using several schemes but are often grouped into three general categories: selection bias, information bias and confounding[84]. Examples of bias encountered in administrative database research are included in **Table 4.2**.

#### *4.3.1.1 Selection Bias*

Selection bias addresses the comparability between the groups under study[84]. In cohort studies, selection bias occurs when the exposed and unexposed groups differ in some way besides the exposure of interest. In case-control studies, selection bias is present when there are differences between cases and controls other than the disease under study.

#### *4.3.1.2 Information Bias*

Information bias arises from incorrect determination of the exposure or outcome (or both)[84]. It is also known as ascertainment, observation, classification or

measurement bias. To minimize information bias, information regarding outcomes and/or exposures should be gathered in the same way for comparison groups.

#### *4.3.1.3 Confounding*

Confounding refers to the presence of a variable that is related to the exposure and the outcome, but is not an intermediate in the causal association between the two[84]. Methods to control for confounding are designed to achieve homogeneity between study groups and include restriction, matching, stratification and multivariate regression. These can be applied before (restriction, matching, stratification) or after the study is completed (regression). Confounding can be corrected only if it anticipated and the necessary information is available.

#### *4.3.1.4 Time-Dependent Bias*

Time-dependent bias arises when a patient's outcome influences the value of a time-dependent variable[74]. For example, individuals experiencing an outcome early in the observation period are less likely to have had the exposure of interest. This can result in misleading results with higher outcome rates in unexposed compared to exposed individuals. This is also referred to as "immortal time bias". Immortal time refers to the period of follow-up during which an outcome cannot occur[85]. For example, after a patient is discharged from hospital, there may be a period of time during which follow-up is occurring, but there is a delay in determining treatment status because a prescription has not been dispensed. This period of delay is referred to as "immortal" because individuals must survive (alive, event-free) in order to be assigned to the

exposed or unexposed group. Immortal time in observational studies can bias the results in favour of the treatment. To address time-dependent bias, time-dependent exposures (variables) should be analyzed as time-dependent covariates which allows the variables to change over time.

**Table 4.2 Types and Examples of Bias in Administrative Database Research**

| Type of Bias | Assessment | Examples |
|---|---|---|
| Selection bias | Are the participants similar except for exposure (cohort studies) or disease (case-control studies)? | • Differing compliance with therapeutic strategies<br>• Referral pattern related to exposure status |
| Information bias | Is information obtained in the same way for exposed and unexposed (cohort studies) or cases and controls (case-control studies)? | • Over-coding or under-coding of diagnoses<br>• Differences in diagnostic code reliability<br>• Combining different databases with different sensitivity/specificity for outcomes<br>• Unexposed individuals considered exposed, or vice versa (e.g. prescription vs. over-the-counter medications, drug samples, drug plan) |
| Confounding | Are the findings related to the presence of a factor which is not along the causal pathway, but which is associated with both the exposure and the outcome? | • Baseline risk factors<br>• Indication<br>• Illness severity |

Adapted from[77,84]

*4.3.1.5 Propensity Score Analysis*

Propensity score analysis is increasingly used to mitigate the effect of bias and confounding in observational studies by balancing confounding variables across treatments and groups[82]. The propensity score, which is commonly calculated using logistic regression models, is the conditional probability of an individual being exposed (e.g. to a treatment) given their covariates. The propensity score is then applied in three main ways to balance confounders between groups: i) matching; ii) stratification; and iii)

regression. However, propensity scoring does not address bias arising from the omission of variables from the propensity score estimation, for example because they are unmeasurable or unavailable. Although propensity scoring is an attractive approach to balance potential confounders, it does not adjust for confounding by time-dependent variables, which are variables that are not constant throughout the duration of follow-up.

### 4.3.2 Construct Validity

Construct validity refers to the degree to which a variable accurately measures what it is supposed to measure[78]. Because diagnostic codes are essentially surrogate measures of a disease or procedure, an evaluation of the association between the code and the real variable using code validation studies is important for documenting reliability[74]. The crudest approach is to compare disease incidence or event rates measured using the code or more reliable methods (ecological study). Another approach is to re-abstract information from the medical record to evaluate the integrity of the abstraction process which is susceptible to missed cases or misidentified cases (re-abstraction study). The strongest validation approach is to compare the code with a gold standard determination of the disease or procedure by using: i) standard clinical and laboratory criteria; ii) panel review; or iii) second data set with an accurate measure of disease status.

Positive and negative likelihood ratios can then be used to report code accuracy. Likelihood ratios express the likelihood of a given test result in a patient with a disease compared to the likelihood of the same test result in a patient without the disease. When likelihood ratios are combined with the baseline odds of disease, the probability that an

individual with a code truly has the disease can be calculated using the following equation: $O \times LR^+ / O \times LR^+ + 1 - O$ (where O is the odds of disease and $LR^+$ is the positive likelihood ratio)[74]. Likelihood ratios are preferred for this purpose because unlike positive predictive value they account for disease prevalence. Positive predictive value refers to the probability that an individual with a code truly has the disease and varies significantly with disease prevalence. Although sensitivity and specificity vary less with changes in disease prevalence, they do not describe the probability that people with the code truly have the disease (or vice versa). Ideally, a gold standard method would be used to measure disease prevalence, but likelihood ratios provide a method for estimating the probability that someone with a given code truly has the disease of interest.

### 4.3.3 External Validity

External validity refers to the generalizability of the inferred causal relationship across patient types and settings, and over time[78]. Compared to randomized controlled trials, database studies are less likely to generate concerns regarding external validity. However, the study population, practice setting, practice patterns and costs can influence external validity in database studies. Study population characteristics such as race, socioeconomic status, sex and age can affect any causal relationships found. Similarly, generalizability can be affected by differences in practice settings (i.e urban versus rural, academic versus community) in which access to resources may be variable. The comprehensiveness of health and prescription coverage may vary across regions thereby influencing overall health care costs and access.

4.3.4 Statistical vs. Clinical Significance

Because administrative databases contain such large sample sizes, statistically significant differences between groups which have little or no clinical significance will be detected[74]. The interpretation of such differences should incorporate absolute and relative differences between groups as opposed to solely relying on the *P*-value. **Table 4.3** demonstrates the effect of sample size on *P*-values for statistical testing between two hypothetical groups with similar baseline prevalence of a characteristic (49.9% and 50.1%). As the sample size increases above 250,000, *P*-values less than 0.05 are achieved. However, such a difference may (or may not) be clinically significant. Confidence intervals can also help distinguish between clinical and statistical significance because they are generated around absolute or relative differences between groups. Confidence intervals provide an estimated range of values which contains the true difference between populations allowing writers and readers to reflect on the differences between populations.

## 4.4 Summary

Administrative data can be used to conduct large studies with long duration of follow-up at low cost to answer research questions. Novel approaches include the combination of administrative data with randomized trial methodology. There are important threats to the validity of administrative database studies which must be accounted for during study design and analysis. As with all observational studies, bias and confounding which threaten internal validity are particularly important as they affect the ability to make causal inferences about exposures and outcomes.

**Table 4.3 Effect of Sample Size on *P*-Values**

| Sample Size | *P*-Value |
|---|---|
| 10 | 0.49495 |
| 100 | 0.48405 |
| 1,000 | 0.44967 |
| 10,000 | 0.34458 |
| 100,000 | 0.10295 |
| 500,000 | 0.02275 |
| 1,000,000 | 0.00003 |
| 10,000,0000 | < 0.00001 |

Adapted from[74]. P-values for two independent samples with proportions 0.499 and 0.501.

# CHAPTER 5: SMALL-VOLUME TUBES TO REDUCE ANEMIA AND TRANSFUSION IN ICU PATIENTS: A PILOT STUDY

This chapter outlines the design of a pilot study to assess the feasibility of a full-scale stepped wedge cluster randomized trial to evaluate whether small-volume blood collection tubes reduce RBC transfusion in adult ICU patients compared to standard-volume blood collection tubes. A discussion of the definition and purpose of pilot (feasibility) studies is also provided.

## 5.1 Rationale for Conducting a Pilot Study

### 5.1.1 What is a Pilot Study?

The National Institute for Health Research (NIHR) describes a pilot study as "a smaller version of the main study used to test whether the components of the main study can all work together. It is focused on the processes of the main study…"[86]. Although both 'pilot' and 'feasibility' are used to describe these types of studies, there is a lack of consensus regarding the distinction (if any) between them[87]. Eldridge and colleagues recently published a consensus framework to define pilot and feasibility studies in which feasibility studies are defined as studies that evaluate whether something can be done, if it should be done, and how it should be done[87,88]. Pilot studies, a subset of feasibility studies, are considered by these authors as stand-alone studies conducted prior to definitive randomized trials which inform their design by addressing issues of feasibility.

### 5.1.2 Objectives of Pilot Studies

The objectives of pilot studies address issues of uncertainty which need to be explored in preparation for a future full-scale trial and differ from the objectives of the full-scale trial[87]. Therefore, the purpose of pilot studies is to assess the feasibility of larger full-scale trials[89]. Broadly, pilot studies address the following aspects of study design: (i) process, (ii) resources, (iii) management and, (iv) science[89,90]. Process refers to aspects of the study protocol that are important for success such as eligibility rates, recruitment rates, adherence rates, and retention. Resource issues include determining the time required for processes (e.g. for survey or data collection form completion), availability of equipment, and capacity (e.g. adequate physical space to conduct assessments). Management issues include those related to human resources (e.g. availability and training of study personnel) and data collection (e.g. adequacy of data collection tools, ensuring all relevant data point collected, confirming ability to collect data points). Scientific issues such as evaluating estimates of treatment effects and variance in addition to exploring other outcomes which inform the design of a larger study can also be assessed in a feasibility study.

### 5.1.3 Misconceptions About Pilot Studies

Pilot studies are not simply small versions of randomized controlled trials[89]. Although they are often conducted to generate preliminary estimates of treatment effect to inform sample size calculations, such estimates may be unrealistic or biased due to limited sample sizes[89]. They are not powered to detect differences in treatment effects between groups and can be potentially misleading in this regard. Pilot studies should not be done because there are insufficient resources with which to conduct a large

multi-center trial; they should only be conducted because they will provide information which will be useful for helping researchers plan a definitive trial. Because research studies expose participants to health and/or privacy risks it is unethical to conduct a research study which is unable to achieve its goals and, therefore, exposes participants to these risks unnecessarily.

## 5.2 Pilot Study Design

### 5.2.1 Rationale

Small-volume (soft-draw or reduced-volume) blood collection tubes have the same physical dimensions, blood collection technique and cost as standard-volume tubes. Further, they are clinically available and compatible with a range of existing laboratory equipment. Their use could reduce the volume of blood collected for laboratory testing which may impact the incidence/severity of anemia and/or number of RBC transfusions in ICU patients who undergo frequent laboratory testing, have a high incidence of anemia and receive frequent RBC transfusions. There are limited existing data regarding the use of small-volume blood collection tubes and their impact on blood loss, anemia and RBC transfusion. Further, there are no studies which address potential harms to patients, hospitals or laboratories. Thus, there is equipoise about the overall net benefit of small-volume blood collection tubes.

Prior to embarking on a full-scale stepped wedge cluster randomized trial evaluating small-volume blood collection tubes powered on clinical outcomes (RBC transfusion, incidence and severity of anemia), a pilot study is needed to determine whether a larger study will be feasible. A stepped wedge cluster randomized trial design

was chosen because (i) the intervention (small-volume tubes) has a high likelihood of benefit with minimal harm and (ii) the logistics of widespread introduction of small-volume tubes are more suitable for a one-way switch followed by implementation.

Although the intervention seems straightforward and cost-neutral, implementation is complex and requires input from multiple stakeholders. The pilot study will address concerns that small-volume tubes could increase workload for nursing and/or laboratory staff, compromise test procedures, limit the number of tests that can be performed, and ultimately lead to the need for additional blood collection and subsequent delays in obtaining critical results. Preliminary informal discussions with clinical and laboratory stakeholders identified the following concerns about the use of small-volume tubes: (i) potential difficulty collecting blood from some patients (e.g. hypotensive, dialysis, central venous catheter devices); (ii) lack of experience and education; (iii) increased workload for blood collectors and laboratory technicians; (iv) reduced number of tests that can be performed; (v) increased number of samples with insufficient volume for testing; and (vi) potential need for additional blood collection and subsequent delays in obtaining critical results. These discussions also identified a need for further exploration of acceptability and potential barriers and facilitators of implementation to enhance the conduct of the full-scale trial and facilitate incorporation of the intervention into routine practice. The pilot study will inform a future large stepped wedge cluster randomized trial designed to evaluate the clinical effectiveness of small-volume tubes compared to standard-volume tubes.

### 5.2.2 Pilot Study Objectives

#### *5.2.2.1 Primary Objective*

The primary objective of the pilot study is to evaluate the feasibility of a full-scale trial with emphasis on the following:

1. Evaluating the study protocol and logistics
2. Assessing potential harms
3. Assessing implementation issues
4. Evaluating data collection procedures
5. Informing estimates for the full-scale trial

#### *5.2.2.2 Secondary Objectives*

The secondary objectives are to prospectively evaluate the following clinical outcomes which will inform planning of the larger trial:

1. Change in hemoglobin level from ICU admission to ICU discharge (or death) adjusted for RBC transfusion
2. Blood loss prevented using small-volume tubes
3. Number of RBC transfusions and variance (ICC)
4. ICU and hospital length of stay
5. ICU and hospital mortality

### 5.2.3 Pilot Study Outcomes

#### *5.2.3.1 Primary Outcome*

The primary outcome of the pilot study will be feasibility. The pilot study will be considered a success and to have demonstrated feasibility if the following criteria are fulfilled:

1. Successful switch from standard-volume to small-volume tubes by the end of the 2-week washout period. This will be defined as 95% correct tubes collected during an audit of 100 blood specimens collected in the ICU at the end of the washout period.
2. Adherence to the correct tube size during the intervention period. This will be defined as 95% adherence to allocated tube size evaluated during 2 audits of 100 blood specimens collected in the ICU during the intervention period.
3. Sufficient volume for testing with small-volume tubes. This will be defined as less than 3% of blood specimens collected reported as inadequate volume for testing (NSQ). This estimate is based on discussions with local laboratory testing experts (laboratory staff and Dr. Stephen Hill, Medical Biochemist).
4. Acceptability of the intervention by end-users including individuals taking blood samples (e.g. nurses, phlebotomy technicians and laboratory staff). Acceptability will be evaluated qualitatively during structured focus group discussions during the intervention period and at the end of the study.
5. Barriers and facilitators of implementation will be assessed qualitatively during structured focus group discussion with end-users at the end of the study.
6. Complete primary data collection. This will be defined as 95% of patients with complete data collected. Data will be collected from hospital administrative

databases and electronic medical records. We will evaluate completeness of data collection with respect to the data points shown in **Table 5.2**.

In addition to the criteria listed above, the pilot study may also reveal potential unforeseen problems in conducting the trial and provide an opportunity to find solutions to any such problems. There will be ongoing communication between investigators, ICU staff and laboratory staff at weekly meetings during the study.

*5.2.3.2 Secondary Outcomes*

In addition, the following preliminary prospective data will be collected and described:

1. Change in hemoglobin level from ICU admission to ICU discharge (or death) adjusted for RBC transfusion (subtract 10 g/L for each unit of RBCs transfused)
2. Prospective assessment of the reduction in blood loss from routine hematology, chemistry, and coagulation testing using small-volume tubes. This will be calculated using the total number of blood specimens collected for hematology, chemistry and coagulation testing multiplied by known sample volumes. This information will allow the preliminary sample size calculations to be refined for the full-scale trial. Volumes will be compared between control and intervention periods. The actual volume of reduction achieved will depend on the type of tube and frequency with which it is used. For example, ethylenediamenetetraacetic acid (EDTA) and lithium heparin tubes for complete blood count and chemistry (50% reduction) are expected to be used more frequently than fluoride or serum tubes (25% reduction). Based on data from a cohort of 10,260 medical-surgical ICU patients in Hamilton, Ontario (2012-

2015), mean (SD) diagnostic blood loss during ICU admission was 243 mL (328 mL) (unpublished data). It is hypothesized that a 40% reduction in blood volume obtained for laboratory analysis (approximately 96 mL) can be achieved for an average patient enrolled in the study based on an informal assessment of likely tube mix and duration of ICU stay.

3. Number of RBC units transfused per patient-day in the ICU and variance (ICC)
4. ICU and hospital length of stay
5. ICU and hospital mortality

### 5.2.4 Study Population

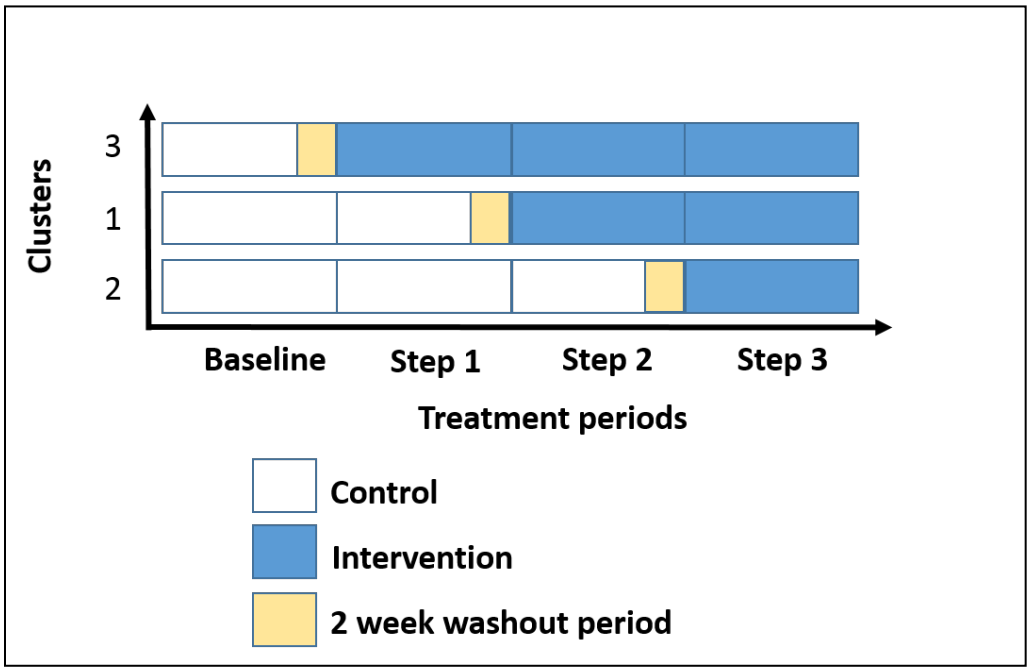
All adult patients admitted to 3 medical-surgical ICUs in Hamilton, Ontario (St. Joseph's Healthcare Hamilton, Juravinski Hospital, and Hamilton General Hospital ICU-East) during the study period. Patients will be followed until hospital discharge, 30 days, or death, whichever is earliest.

### 5.2.5 Study Design

The proposed pilot study will be a stepped wedge cluster randomized trial (**Figure 5.1**). Using a stepped wedge cluster randomized trial design, 3 ICUs (ICU-East, Juravinski Hospital, St. Joseph's Healthcare Hamilton) (clusters) will switch from using standard-volume tubes (control) to small-volume tubes (intervention) for all bloodwork as determined by the clinical care team. The timing at which individual ICUs switch to the intervention will be randomly assigned using a random numbers table. During each time period (step), one site will switch to using small-volume tubes. Therefore, each ICU

provides “before” and “after” observations and switches from control to intervention, but at different times. The randomization schedule will be maintained by the research staff and concealed from individual sites.

**Figure 5.1 Pilot Stepped Wedge Randomized Study**




5.2.6 Study Interventions

The study will involve switching from standard-volume (4 – 6 mL; current practice) to small-volume (2 – 3 mL; intervention) EDTA, lithium-heparin, citrate, fluoride and silica tubes (Vacutainer®, Becton, Dickinson and Company) (**Table 5.1**). These are the types of blood collection tubes used for routine bloodwork (e.g. hematology, chemistry, coagulation) in Hamilton, Ontario. The small-volume tubes have similar physical dimensions as the standard-volume tubes, but have reduced vacuum and draw less blood during collection. The additives present in the tube are adjusted to reflect the reduced volume of blood drawn by the vacuum. Therefore, both tubes can be used on

the same standard laboratory instruments without process modification. They are both available for ordering and distribution through usual processes and have the same unit cost. Thus, from a nursing and laboratory perspective, handling requirements are identical. Education will be required for blood collectors (ICU nurses) and laboratory staff because they may perceive tubes to be inadequately filled when using the small-volume tubes.

Although very small volume capillary blood collection tubes (e.g. Microtainer®) are commercially available, they are more expensive and require manual collection and aliquoting at the bedside and in the laboratory, thereby introducing safety concerns and substantially increasing workload to unacceptable levels. Thus, although using Microtainer® tubes in one arm of the study may appear to be a “better” test of the hypothesis that smaller blood volume tubes reduce iatrogenic anemia and need for transfusion, complexity associated with their use make such an evaluation infeasible. However, should Microtainer® tubes be clinically indicated at a centre participating in the study their use will be allowed. Patients having blood drawn into Microtainer® will not be excluded from the analysis; rather, they will be included with the arm currently being run at that centre. It is anticipated that the number of Microtainer® blood draws to be very small and predominantly limited to neonates and other highly selected populations, few of whom will be receiving care in the intensive care units which are the major participants in the study.

**Table 5.1 Study Interventions: Vacutainer® Tubes for Blood Collection**

| Description | Current practice (standard-volume) mL | Intervention (small-volume) mL |
|---|---|---|
| EDTA (lavender) | 4.0 | 2.0 |
| Fluoride (grey) | 4.0 | 2.0 |
| Serum (red) | 4.0 | 3.0 |
| Heparin (green) | 4.0 | 2.0 |
| Citrate (light blue) | 2.7 | 1.8 |

Source: BD Life Sciences – Preanalytical Systems Product Catalogue 2016

5.2.7 Study Conduct

The pilot stepped wedge cluster randomized study will consist of a baseline period (current practice of standard-volume tubes) and 3 study periods (steps) each lasting 4 weeks for a total duration of 16 weeks. The timing of switching to small-volume tubes will be randomly assigned prior to the start of the study and a supply of small-volume tubes will be available. Sites will be notified 2 weeks prior to the switch which will allow a total of 4 weeks for sites to complete the change from standard to small-volume tubes prior to the start of data collection. Based on discussions with nursing and laboratory staff, this was agreed to be sufficient time to gather experience with the intervention during the study for the purposes of assessing feasibility. There will be a 2-week washout period during which the small-volume tubes will be supplied, but no data collected to prevent the likelihood of carryover effect from previous blood collection that could affect results. An audit of 150 tubes in ICU blood collection tube storage areas will be conducted at the end of the wash-out period to assess the success of switching from standard-volume to small-volume tubes. After switching, only the small-volume tubes

will be stocked and used unless specially requested. A sufficient quantity of tubes will be ordered and available (confirmed and monitored) during the observation period. The research coordinator will conduct an audit of 100 tubes (randomly chosen) in the ICU storage areas during each of the three study periods to assess adherence to tube size.

During the 4-week period preceding the switch, educational sessions will be conducted for ICU nurses and laboratory staff to familiarize them with the intervention and address questions or concerns. This will include in-kind support from an experienced Clinical Practice Consultant from Becton Dickinson Canada (personal communication, Ms. Susan Csatari, National Clinical Practice Consultant, Vascular Access Blood Collection, BD Preanalytical Systems). Educational sessions will include demonstration and feedback and will be performed at various times to ensure maximal coverage of staff who currently provide care on a 24/7 basis determined in collaboration with nursing and laboratory managers. Because Vacutainer® tubes are already used in the ICU and the laboratory, educational efforts will be focused on reinforcing existing blood collection standard operating procedures and emphasizing the key differences between standard- and small-volume tubes such as reduced fill volume and lower vacuum. Additional educational resources such as posters and quick-reference cards will also be provided in-kind by Becton Dickinson Canada.

### 5.2.8 Focus Group Discussion

One of the pilot study objectives is to explore the experiences of individuals using small-volume tubes and identify themes of barriers and facilitators to implementation of small-volume tubes. This will be accomplished using in-depth focus groups (2 per site)

with individuals who work with small-volume blood collection tubes including blood collectors (nurses) and laboratory staff (technologists, technicians, managers). Approximately 4 to 6 individuals will be included in each focus group. This number was proposed by nursing and laboratory managers such that the focus groups will not disrupt clinical responsibilities. Participants will be eligible if they had hands-on experience using the small-volume tubes during the study period and will be selected with the assistance of the ICU nurse managers and laboratory managers by reviewing work schedules. Each focus group will take place in a private meeting room at the most convenient time of day for the group identified by nursing and laboratory managers.

The experiences and barriers/facilitators identified by blood collectors and laboratory staff will be summarized with the qualitative description method using focus groups. Qualitative description allows for a comprehensive summary in everyday terms[91]. The output of qualitative description reflects the participant's experience similar to their own words[92]. This method was chosen because focus groups facilitate in-depth exploration of the knowledge, experiences and insights of participants in a way that would be more difficult with individual interviews[93]. Further, focus groups encourage discussion among participants.

Two members of the research team will conduct the focus group sessions as facilitator and assistant/note-taker. When participants arrive, a consent form will be completed with time to discuss questions or concerns. If potential participants do not attend the focus group, this will be recorded in the study log. The objectives of the focus group and the overall pilot study will be provided to each focus group. Confidentiality and anonymity will be assured and ground rules will be described. Prior to the start of

the meetings, the facilitator and the assistant will introduce themselves to the participants and have some informal discussions to make the participants feel comfortable, and familiarize them with the facilitator.

A note-taking form will be used by the assistant/note-taker for all focus group discussions. In addition, an audio recording will be made of the discussion. Each participant will be given a study ID number by the note-taker who will draw a map of participants and record which participant made which contributions, to match up with the transcript of the focus group afterwards. The note-taker will document non-verbal behaviour and describe the interview setting and atmosphere of the interview. The facilitator and note-taker will meet to discuss the findings of the interview including themes, atmosphere and group dynamics.

All study documents will be maintained in a secure, locked office and/or on a password-protected computer on a secure server at the Population Health Research Institute and will only be shared within the study team. Audio recordings will be transcribed into Microsoft Word and then reviewed and revised by members of the research team prior to data analysis. Participant data will be de-identified. The transcripts will then be exported to NVivo (v10.0 QSR International, Australia) for coding and analysis. Details of the de-briefing session between the facilitator and note-taker after each focus group will be typed into Microsoft Word and shared within the research team.

Data analysis will be conducted using NVivo (v10.0 QSR International, Australia). Transcripts from audio-recorded focus group sessions will be typed in a Microsoft Word document and saved in NVivo. Quantitative data (e.g. demographics) will be stored in a

password-protected Microsoft Excel file. This will be imported into NVivo and linked to transcription files. One or two transcripts will be selected randomly to develop a coding structure. Coding will be generated empirically from the data as per the qualitative description method[94].

### 5.2.9 Data Collection and Management

Except for data noted elsewhere in this thesis (for example, results of tube audits), all data will be collected electronically from hospital administrative databases (e.g. census data, Discharge Abstract Database) and hospital electronic medical records (including the Laboratory Information System) at Hamilton Health Sciences and St. Joseph's Healthcare Hamilton. Understanding the process of data collection is an important component of this pilot study to assist with planning data collection for the full-scale trial. Information will be collected regarding patient demographics, laboratory tests, RBC transfusions administered, mortality, duration of ICU admission, duration of hospital admission, most responsible diagnosis, pre- and post-admission comorbidities and post-admission interventions (ICD-10 codes, CCI codes) (**Table 5.2**). Additional data will include use of mechanical ventilation and renal replacement therapy, and surgeries during hospitalization as these are known covariates which are expected to affect RBC transfusion. Encryption of patient identifiers will be used to ensure confidentiality. Data will be stored in a de-identified database on a secure server at the Population Health Research Institute, McMaster University. All of this data is currently electronically housed and accessible through Decision Support and Laboratory

Information Systems at both Hamilton Health Sciences and St. Joseph's Healthcare Hamilton.

**Table 5.2 Data Points to be Collected**

| Demographic/Clinical | Laboratory | Transfusion |
|---|---|---|
| • Age<br>• Sex<br>• Date of ICU admission<br>• Date of ICU discharge<br>• Date of hospital discharge<br>• Date of death (if applicable)<br>• Most responsible diagnosis<br>• Pre-admission comorbidities<br>• Post-admission comorbidities<br>• Interventions during admission<br>• Use of dialysis during admission<br>• Use of mechanical ventilation<br>• Multi-organ dysfunction score [MODS] | • Number of samples with inadequate volume for testing ("NSQ" samples)<br>• Hemoglobin level at ICU admission<br>• Hemoglobin level at ICU discharge<br>• Hemoglobin level at hospital discharge<br>• Serum creatinine at ICU admission and ICU discharge<br>• Highest serum creatinine during ICU admission<br>• Number and type of blood specimens collected for testing during ICU admission | • RBC units transfused |

5.2.10 Sample Size

Because the pilot trial is designed to assess feasibility, it will not be powered to detect differences in clinical outcomes, although these will be documented. The results of the pilot trial will be used to test the assumptions and inform projections for the sample size and study duration for the full-scale trial. The sample size calculations

described below were conducted using the confidence-interval method for a single proportion (Power and Sample Size, NCSS Statistical Software).

1. Successful switch to small-volume tubes. This will be assessed based on the proportion of correct size tubes collected during an audit of 100 blood specimens collected in the ICU at the end of the wash-out period. With this sample size (n=300), the 95% confidence interval (CI) of the proportion of correct tubes with expected proportion of 95% is 91.9% to 97.2%. Therefore, the successful switch outcome will be considered as accomplished if 300 tubes are collected and the proportion of correct tubes is greater than 91.9%.
2. Adherence to small-volume tubes. This will be assessed based on the proportion of correct size tubes collected during an audit of 100 blood specimens collected in the ICU 4 weeks after the switch. With this sample size (n=300), the 95% CI of the proportion of correct tubes with expected proportion of 95% is 91.9% to 97.2%. Therefore, the adherence feasibility outcome will be considered as accomplished if 300 tubes are collected and the proportion of correct tubes is greater than 91.9%.
3. 95% complete primary data collection. The average ICU admission rate in Hamilton was approximately 58 patients per month during 2012-2015. Therefore, we expect to register approximately 252 patients during the study. With this sample size, the 95% CI of the proportion of patients with complete data collected with expected proportion of 95% is 92.30% to 97.70%. Therefore, data collection will be considered as accomplished if the proportion of patients with complete data is greater than 92%.

### 5.4.11 Statistical Analysis

Pilot trial feasibility outcomes will be presented descriptively. Univariate analyses will be conducted to compare the patient characteristics (age, sex, most responsible diagnosis, pre-admit comorbidities, hemoglobin level at ICU admission, creatinine level at ICU admission, MODS score, use of mechanical ventilation, use of dialysis during admission) in each treatment group. Categorical data will be reported as counts and proportions, and compared using Chi-Square, or Fisher's Exact Test if the number of observations is small. Continuous data will be reported as means with standard deviations (SD) or medians with range and will be compared using student t-test or Wilcoxon rank sum test. Proportions will be reported with 95% confidence intervals (CI).

### 5.2.12 Consent

Because both interventions (small-volume and standard-volume blood collection tubes) fall within current standard of care and are already in use in our hospital system, we will seek a waiver of individual patient consent from the Hamilton Integrated Research Ethics Board (HiREB) according to criteria proposed by the Tri-Council Policy Statement (TCPS 2): Ethical Conduct for Research Involving Humans[70]. This study fulfills TCPS criteria in that: (i) the study poses minimal risk to patients; (ii) waiver of consent will not adversely affect patient rights and welfare; and (iii) it would be impracticable to carry out the research if prior consent is required. To support a request for a waiver of consent it is important to note that these tubes are already used in selected clinical units in Hamilton that vary between sites; as a result, patients receiving care in Hamilton prior to initiation of the study could receive testing results based on an analysis of blood drawn into either standard or reduced volume tubes.

## CHAPTER 6: CONCLUSION AND FUTURE DIRECTIONS

Blood sampling for laboratory testing causes significant blood loss, especially in ICU patients who undergo frequent blood testing. ICU patients are at high-risk for anemia, which is associated with poor outcomes and is frequently corrected with RBC transfusion, a scarce resource that is also associated with harm. Importantly, only 10% of the blood collected is used for testing procedures with the remainder discarded as waste. This suggests that sample volumes can be decreased without compromising patient care or hospital procedures.

Small-volume (soft-draw) blood collection tubes have the same cost as standard-volume tubes and are compatible with laboratory equipment. They draw less blood (2 to 3 mL versus 4 to 6 mL) due to lower vacuum inside the tube which fills to a smaller-volume. There is a paucity of evidence regarding the benefits and harms of small-volume tubes. A randomized trial is needed for an unbiased assessment of whether small-volume tubes reduce blood loss, anemia and RBC transfusion without concomitant harms or negative impact on patient care and hospital procedures. If this could be shown, it may lead to practice change regarding blood collection for laboratory testing.

A stepped-wedge cluster randomized trial is the ideal study design for this low-risk intervention. By incorporating the small-volume tubes into routine clinical practice and using administrative and hospital electronic medical record data, this study would be a pragmatic, cost-effective way to evaluate effectiveness and implementation. However, prior to conducting a full-scale trial powered on clinical outcomes, a pilot study is needed to determine whether a larger study will be feasible.

The proposed pilot study is a stepped-wedge cluster randomized trial at 3 medical-surgical ICUs in Hamilton, Ontario. Sites will switch from standard-volume (4 to 6 mL) to small-volume (2 to 3 mL) blood collection tubes at a time which is randomly allocated. The primary objective is to evaluate the feasibility of a full-scale stepped-wedge cluster randomized trial. The primary outcome will be feasibility as assessed by: (i) successful switch from standard-volume to small-volume tubes; (ii) adherence to the correct tube size; (iii) sufficient volume for testing; (iv) acceptability of the intervention by end-users; (v) identifying barriers and facilitators of implementation; and (vi) complete primary data collection. The pilot study may also identify potential unforeseen problems and provide an opportunity to find solutions. If the pilot study is successful, a full-scale stepped wedge cluster randomized trial designed to evaluate the clinical effectiveness of small-volume tubes compared to standard-volume tubes will be conducted.

## REFERENCES

1. Salisbury AC, Reid KJ, Alexander KP, et al. Diagnostic blood loss from phlebotomy and hospital-acquired anemia during acute myocardial infarction. Archives of Internal Medicine 2011;171:1646-53.

2. Vincent JL, Baron JF, Reinhart K, et al. Anemia and blood transfusion in critically ill patients. JAMA 2002;288:1499-507.

3. Dale JC, Ruby SG. Specimen collection volumes for laboratory tests. Arch Pathol Lab Med 2003;127:162-8.

4. Ania BJ, Suman VJ, Fairbanks VF, Rademacher DM, Melton LJ, 3rd. Incidence of anemia in older people: an epidemiologic study in a well defined population. J Am Geriatr Soc 1997;45:825-31.

5. Corwin HL, Gettinger A, Pearl RG, et al. The CRIT Study: Anemia and blood transfusion in the critically ill--current clinical practice in the United States. Crit Care Med 2004;32:39-52.

6. Thomas J, Martinez A. Blood conservation in the critically ill. American Journal of Health-System Pharmacy 2007;64:S11-8.

7. Siegal D, Verhovsek M. Evaluation of Anemia. In: McKean SC, Ross JJ, Dressler DD, Brotman DJ, Ginsberg JS, eds. Principles and Practice of Hospital Medicine. 2 ed: McGraw-Hill; 2017.

8. Cortinez LI, Somma J, Robertson KM, et al. Changes in hematocrit based on incremental blood sampling: mathematical models perform poorly. Can J Anaesth 2005;52:374-8.

9. Thavendiranathan P, Bagai A, Ebidia A, Detsky AS, Choudhry NK. Do blood tests cause anemia in hospitalized patients? The effect of diagnostic phlebotomy on hemoglobin and hematocrit levels. Journal of General Internal Medicine 2005;20:520-4.

10. Sabatine MS, Morrow DA, Giugliano RP, et al. Association of hemoglobin levels with clinical outcomes in acute coronary syndromes. Circulation 2005;111:2042-9.

11. McKechnie RS, Smith D, Montoye C, et al. Prognostic implication of anemia on in-hospital outcomes after percutaneous coronary intervention. Circulation 2004;110:271-7.

12. Nikolsky E, Mehran R, Aymong ED, et al. Impact of anemia on outcomes of patients undergoing percutaneous coronary interventions. Am J Cardiol 2004;94:1023-7.

13. Al Falluji N, Lawrence-Nelson J, Kostis JB, et al. Effect of anemia on 1-year mortality in patients with acute myocardial infarction. Am Heart J 2002;144:636-41.

14. Salisbury AC, Amin AP, Reid KJ, et al. Hospital-acquired anemia and in-hospital mortality in patients with acute myocardial infarction. Am Heart J 2011;162:300-9 e3.

15. Androne AS, Katz SD, Lund L, et al. Hemodilution is common in patients with advanced heart failure. Circulation 2003;107:226-9.

16. McClellan WM, Flanders WD, Langston RD, Jurkovitz C, Presley R. Anemia and renal insufficiency are independent risk factors for death among patients with congestive heart failure admitted to community hospitals: a population-based study. J Am Soc Nephrol 2002;13:1928-36.

17. Silverberg DS, Wexler D, Iaina A, Steinbruch S, Wollman Y, Schwartz D. Anemia, chronic renal disease and congestive heart failure--the cardio renal anemia

syndrome: the need for cooperation between cardiologists and nephrologists. Int Urol Nephrol 2006;38:295-310.

18. Keough-Ryan TM, Kiberd BA, Dipchand CS, et al. Outcomes of acute coronary syndrome in a large Canadian cohort: impact of chronic renal insufficiency, cardiac interventions, and anemia. Am J Kidney Dis 2005;46:845-55.

19. Tong PC, Kong AP, So WY, et al. Hematocrit, independent of chronic kidney disease, predicts adverse cardiovascular outcomes in chinese patients with type 2 diabetes. Diabetes Care 2006;29:2439-44.

20. Craig KJ, Williams JD, Riley SG, et al. Anemia and diabetes in the absence of nephropathy. Diabetes Care 2005;28:1118-23.

21. Shu DH, Ransom TP, O'Connell CM, et al. Anemia is an independent risk for mortality after acute myocardial infarction in patients with and without diabetes. Cardiovasc Diabetol 2006;5:8.

22. Barlas RS, Honney K, Loke YK, et al. Impact of Hemoglobin Levels and Anemia on Mortality in Acute Stroke: Analysis of UK Regional Registry Data, Systematic Review, and Meta-Analysis. J Am Heart Assoc 2016;5.

23. Carson JL, Duff A, Poses RM, et al. Effect of anaemia and cardiovascular disease on surgical mortality and morbidity. Lancet 1996;348:1055-60.

24. Kulier A, Levin J, Moser R, et al. Impact of preoperative anemia on outcome in patients undergoing coronary artery bypass graft surgery. Circulation 2007;116:471-9.

25. Karkouti K, Wijeysundera DN, Beattie WS, Reducing Bleeding in Cardiac Surgery I. Risk associated with preoperative anemia in cardiac surgery: a multicenter cohort study. Circulation 2008;117:478-84.

26. Salisbury AC, Alexander KP, Reid KJ, et al. Incidence, correlates, and outcomes of acute, hospital-acquired anemia in patients with acute myocardial infarction. Circ Cardiovasc Qual Outcomes 2010;3:337-46.

27. Napolitano LM, Kurek S, Luchette FA, et al. Clinical practice guideline: red blood cell transfusion in adult trauma and critical care. J Trauma 2009;67:1439-42.

28. Retter A, Wyncoll D, Pearse R, et al. Guidelines on the management of anaemia and red cell transfusion in adult critically ill patients. Br J Haematol 2013;160:445-64.

29. Carson JL, Guyatt G, Heddle NM, et al. Clinical Practice Guidelines From the AABB: Red Blood Cell Transfusion Thresholds and Storage. JAMA 2016;316:2025-35.

30. Corwin HL, Parsonnet KC, Gettinger A. RBC transfusion in the ICU. Is there a reason? Chest 1995;108:767-71.

31. Walsh TS, Garrioch M, Maciver C, et al. Red cell requirements for intensive care units adhering to evidence-based transfusion guidelines. Transfusion 2004;44:1405-11.

32. Chant C, Wilson G, Friedrich JO. Anemia, transfusion, and phlebotomy practices in critically ill patients with prolonged ICU length of stay: a cohort study. Critical Care (London, England) 2006;10:R140.

33. Eder AF, Chambers LA. Noninfectious complications of blood transfusion. Arch Pathol Lab Med 2007;131:708-18.

34. Bihl F, Castelli D, Marincola F, Dodd RY, Brander C. Transfusion-transmitted infections. J Transl Med 2007;5:25.

35. Marik PE, Corwin HL. Efficacy of red blood cell transfusion in the critically ill: a systematic review of the literature. Crit Care Med 2008;36:2667-74.

36. Earley AS, Gracias VH, Haut E, et al. Anemia management program reduces transfusion volumes, incidence of ventilator-associated pneumonia, and cost in trauma patients. J Trauma 2006;61:1-5; discussion -7.

37. Goodnough LT, Maggio P, Hadhazy E, et al. Restrictive blood transfusion practices are associated with improved patient outcomes. Transfusion 2014;54:2753-9.

38. Sanchez-Giron F, Alvarez-Mora F. Reduction of blood loss from laboratory testing in hospitalized adult patients using small-volume (pediatric) tubes. Arch Pathol Lab Med 2008;132:1916-9.

39. Smoller BR, Kruskall MS, Horowitz GL. Reducing adult phlebotomy blood loss with the use of pediatric-sized blood collection tubes. Am J Clin Pathol 1989;91:701-3.

40. Dolman HS, Evans K, Zimmerman LH, et al. Impact of minimizing diagnostic blood loss in the critically ill. Surgery 2015;158:1083-7; discussion 7-8.

41. Chaudoir SR, Dugan AG, Barr CH. Measuring factors affecting implementation of health innovations: a systematic review of structural, organizational, provider, patient, and innovation level measures. Implement Sci 2013;8:22.

42. Cochrane Effective Practice and Organisation of Care Group: Data Collection Checklist. EPOC resources for review authors. Revised 2002. http://www.epoc.cochrane.org/en/handsearchers.html (accessed October 15, 2017). 2002.

43. Higgins JPT, Green S. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.handbook.cochrane.org. .

44. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.

45. Sterne JA, Hernan MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016;355:i4919.

46. Gleason EG, S.; Campbell, C. Minimizing diagnostic blood loss in critically ill patients. American Journal of Critical Care 1992;1:85-90.

47. Harber CR, Sosnowski KJ, Hegde RM. Highly conservative phlebotomy in adult intensive care--a prospective randomized controlled trial. Anaesthesia & Intensive Care 2006;34:434-7.

48. MacIsaac CM, Presneill JJ, Boyce CA, Byron KL, Cade JF. The influence of a blood conserving device on anaemia in intensive care patients. Anaesthesia & Intensive Care 2003;31:653-7.

49. Peruzzi WTP, M. A.; Lichtenthal, P. R.; Cochran-Zull, C.; Toth, B.; Blake, M. A clinical evaluation of a blood conservation device in medical intensive care unit patients. Critical Care Medicine 1993;21:501-6.

50. Sanchez-Giron FA-M, F. Reduction of blood loss from laboratory testing in hospitalized adult patients using small-volume (pediatric) tubes. Archives of Pathology and Laboratory Medicine 2008;132:1916-9.

51. Low LL, Harrington GR, Stoltzfus DP. The effect of arterial lines on blood-drawing practices and costs in intensive care units. Chest 1995;108:216-9.

52. Levi M. Twenty-five million liters of blood into the sewer. J Thromb Haemost 2014;12:1592.

53. Ranasinghe T, Freeman WD. 'ICU vampirism' - time for judicious blood draws in critically ill patients. Br J Haematol 2014;164:302-3.

54. Stefanini M. Iatrogenic anemia (can it be prevented?). J Thromb Haemost 2014;12:1591.

55. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. J Chronic Dis 1967;20:637-48.

56. Thorpe KE, Zwarenstein M, Oxman AD, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. CMAJ 2009;180:E47-57.

57. Sox HC, Lewis RJ. Pragmatic Trials: Practical Answers to "Real World" Questions. JAMA 2016;316:1205-6.

58. Zwarenstein M, Oxman A, Pragmatic Trials in Health Care S. Why are so few randomized trials useful, and what can we do about it? J Clin Epidemiol 2006;59:1125-6.

59. Sackett D. The Principles Behind the Tactics of Performing Therapeutic Trials. In: Haynes RB, ed. Clinical Epidemiology: How to Do Clinical Practice Research: Lippincott Williams & Wilkins; 2006.

60. Cook AJ, Delong E, Murray DM, Vollmer WM, Heagerty PJ. Statistical lessons learned for designing cluster randomized pragmatic clinical trials from the NIH Health Care Systems Collaboratory Biostatistics and Design Core. Clin Trials 2016;13:504-12.

61. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. BMJ 2015;350:h2147.

62. Donner A, Klar N. Pitfalls of and controversies in cluster randomization trials. Am J Public Health 2004;94:416-22.

63. Campbell MJ, Walters SJ. How to Design, Analyse and Report Cluster Randomised Trials in Medicine and Health Related Research: John Wiley & Sons, Ltd; 2014.

64. Crespi CM. Improved Designs for Cluster Randomized Trials. Annu Rev Public Health 2016;37:1-16.

65. Meurer WJ, Lewis RJ. Cluster randomized trials: evaluating treatments applied to groups. JAMA 2015;313:2068-9.

66. Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. BMJ 2015;350:h391.

67. Killip S, Mahfoud Z, Pearce K. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. Ann Fam Med 2004;2:204-8.

68. Hemming K, Taljaard M. Sample size calculations for stepped wedge and cluster randomised trials: a unified approach. J Clin Epidemiol 2016;69:137-46.

69. Zhan Z, van den Heuvel ER, Doornbos PM, et al. Strengths and weaknesses of a stepped wedge cluster randomized design: its application in a colorectal cancer follow-up study. J Clin Epidemiol 2014;67:454-61.

70. Canadian Institutes of Health Research NSERC, Social Sciences and Humanities Research Council of Canada. Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2) 2014. http://www.pre.ethics.gc.ca/eng/policy-politique/initiatives/tcps2-eptc2/chapter3-chapitre3/#toc03-1b/. (Accessed March 26, 2016)

71. International Conference on Harmonisation. Integrated Addendum to ICH E6(R1): Guideline for Good Clinical Practice E6(R1) from http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R2__Step_4.pdf (Accessed Oct 25, 2017)

72. Taljaard M, Weijer C, Grimshaw JM, Eccles MP, Ottawa Ethics of Cluster Randomised Trials Consensus G. The Ottawa Statement on the ethical design and conduct of cluster randomised trials: precis for researchers and research ethics committees. BMJ 2013;346:f2838.

73. Weijer C, Grimshaw JM, Eccles MP, et al. The Ottawa Statement on the Ethical Design and Conduct of Cluster Randomized Trials. PLoS Med 2012;9:e1001346.

74. van Walraven C, Austin P. Administrative database research has unique characteristics that can risk biased results. J Clin Epidemiol 2012;65:126-31.

75. Heddle NM, Cook RJ, Arnold DM, et al. Effect of Short-Term vs. Long-Term Blood Storage on Mortality after Transfusion. N Engl J Med 2016;375:1937-45.

76. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. J Clin Epidemiol 2005;58:323-37.

77. Cox E, Martin BC, Van Staa T, Garbe E, Siebert U, Johnson ML. Good research practices for comparative effectiveness research: approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources: the International Society for Pharmacoeconomics and Outcomes Research Good Research Practices for Retrospective Database Analysis Task Force Report--Part II. Value Health 2009;12:1053-61.

78. Motheral BR, Fairman KA. The use of claims databases for outcomes research: rationale, challenges, and strategies. Clin Ther 1997;19:346-66.

79. Canadian Institute for Health Information website https://www.cihi.ca/en (Accessed October 15, 2017).

80. Canadian Institute for Health Information Data Quality Documentation, Discharge Abstract Database - Multi-Year Information from https://www.cihi.ca/en/dad_multi-year_en.pdf (Accessed October 15, 2017).

81. Canadian Institute for Health Information Data Quality Study of the 2009-2010 Discharge Abstract Database https://secure.cihi.ca/free_products/Reabstraction_june19revised_09_10_en.pdf (accessed December 23, 2017).

82. Takahashi Y, Nishida Y, Asai S. Utilization of health care databases for pharmacoepidemiology. Eur J Clin Pharmacol 2012;68:123-9.

83. Deeny SR, Steventon A. Making sense of the shadows: priorities for creating a learning healthcare system based on routinely collected data. BMJ Qual Saf 2015;24:505-15.

84. Grimes DA, Schulz KF. Bias and causal associations in observational research. Lancet 2002;359:248-52.

85. Levesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. BMJ 2010;340:b5087.

86. National Institute for Health Rearch Glossary https://www.nihr.ac.uk/glossary (Accessed October 15, 2017).

87. Lancaster GA. Pilot and feasibility studies come of age! Pilot and Feasibility Studies 2015;1:1.

88. Eldridge SM, Chan CL, Campbell MJ, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. Pilot Feasibility Stud 2016;2:64.

89. 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.

90. van Teijlingen E, Hundley V. The importance of pilot studies. Nurs Stand 2002;16:33-6.

91. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000;23:334-40.

92. Neergaard MA, Olesen F, Andersen RS, Sondergaard J. Qualitative description - the poor cousin of health research? BMC Med Res Methodol 2009;9:52.

93. Kitzinger J. Qualitative research. Introducing focus groups. BMJ 1995;311:299-302.

94. Sandelowski M. Combining qualitative and quantitative sampling, data collection, and analysis techniques in mixed-method studies. Res Nurs Health 2000;23:246-55.