ACCOUNTING FOR CENTRE IN MULTICENTRE TRIALS

ACCOUNTING FOR CENTRE IN THE EARLY EXTERNAL CEPHALIC VERSION TRIALS: AN EMPIRICAL COMPARISON OF STATISTICAL METHODS TO ACCOUNT FOR CENTRE IN MULTICENTRE RANDOMISED CONTROLLED TRIALS WITH BINARY OUTCOMES

By ANGELA H. REITSMA, BSc, BHSc

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

McMaster University MASTER OF SCIENCE (2012) Hamilton, Ontario (Health Research Methodology)

TITLE: Accounting for Centre in the Early External Cephalic Version Trials: An Empirical Comparison of Statistical Methods to Account for Centre in Multicentre Randomised Controlled Trials with Binary Outcomes AUTHOR: Angela Reitsma, BSc (University of Waterloo), BHSc (McMaster University) SUPERVISOR: Dr. Eileen K. Hutton NUMBER OF PAGES: xi, 55

ABSTRACT

Accounting for centre in the Early External Cephalic Version trials: An empirical comparison of statistical methods to account for centre in multicentre randomised controlled trials with binary outcomes

Master of Science, 2012, Angela H. Reitsma, Department of Health Research Methodology, McMaster University

Background Breech presentation affects 3-4% of term pregnancies and results in increased rates of Caesarean section (CS). External cephalic version (ECV) is an effective intervention to reduce breech presentation and the corresponding CS rate. The Early ECV (EECV) trials were international multicentre randomized controlled trials that compared the timing of ECV (early or delayed) for breech pregnancies on obstetric and neonatal outcomes. In consideration of current reporting recommendations that multicentre trials should account for centre effects in their analysis, a secondary analysis of the EECV trial data was undertaken.

Purpose To analyse the EECV Trial data using statistical methods that account for centre effect and compare the results to standard analysis.

Methods Fisher's exact test was used to provide overall results unadjusted for centre effects. The outcomes of interest were CS, preterm birth, and non-cephalic presentation at birth.

Seven statistical models that accounted for centre effects were applied to the data: i) Mantel-Haenzsel test, ii) fixed effects regression, iii) fixed effects regression with a treatment-by-centre interaction term (weighted and iv) un-weighted by centre size), v) random intercept model, vi) random intercept and random slope model, and vii) generalized estimating equations.

Results For the three outcomes of interest, accounting for centre effect showed a strengthened statistical association with point estimates moving away from the null value.

Conclusion Accounting for centre effects in the EECV trials changes the effect estimates and confidence intervals for three selected outcomes, but does not alter the overall conclusions of the trial. Each method used to account for centre had advantages and disadvantages in relation to the dataset, but for this application, the Mantel-Haenzsel test and the random effects regressions performed the best. This study provides empirical evidence to support recommendations that multicentre trials account for centre in both design and analysis.

ACKNOWLEDGEMENTS

I am grateful for the assistance of many people who have helped me to gather these ideas into a thesis project. Three years ago I embarked upon my Master's degree with the urging of my mentor and supervisor, Dr. Eileen Hutton. Her faith in my abilities kept me going, and each meeting with her pushed me forward. Thank you, Eileen, for your advice, mentorship, and example. I have also been supported by my thesis committee members, Dr. Lehana Thabane and Dr. Sarah McDonald. Their leadership and direction for this project has been much appreciated. I would also like to thank Julia Thorpe and Rachel Chu who have been tremendously generous with their time to teach me the concepts of biostatistics that were vital for this project.

It has been a pleasure to work with the data of a large RCT, and I would therefore like to acknowledge all those who took part in the EECV Trials – from the collaborators at each centre, to the women and the babies who participated.

Finally, I would like to recognise my friends and family for their support. Thank you to my parents and parents-in-law for their unfailing confidence in me, and to Dave and Arie, for helping me to keep it all in perspective.

TABLE OF CONTENTS

Descriptive Note		ii
Abstract		iii
Acknowledgements		iv
List of Tables		vii
List of Figures		vii
List of Appendices		ix
List of Abbreviations		x
Declaration of Academic Achievement		xi
1.0	**BACKGROUND**	**1**
1.1	The Clinical Problem: Breech presentation and Caesarean section	1
1.2	The EECV Trials	2
1.2.1	The theoretical rationale of the EECV Trials	2
1.2.2	The clinical setting during the EECV Trials	2
1.2.3	Overview of the EECV Trials	3
1.2.4	Results of the EECV Pilot Trial and the EECV2 Trial	4
1.2.5	Discussion of findings by trial collaborators	5
1.3	The Methodological Problem: Centre effect in multicentre trials	5
1.3.1	Centre Effect	5
1.3.2	Review of reviews on accounting for centre effect in RCTs	8
1.4	Accounting for centre in the EECV Trials	9
1.4.1	Rationale	9
1.4.2	Trial Characteristics	9
1.4.2.1	Type of RCT	10
1.4.2.2	Type of intervention	10
1.4.2.3	Type of outcomes	10
1.4.2.4	Centres and enrollment	11
1.4.2.5	Randomisation stratified by centre	11
1.4.3	Methods to account for centre effect	12

1.5	Research Objective	12
2.0	**METHODS**	**17**
2.1	Data preparation	17
2.2	Data description	17
2.3	Statistical methods to account for centre effect	17
2.3.1	Conditional methods	18
2.3.1.1	Fixed effects	19
2.3.1.1.1	Mantel-Haenzsel	19
2.3.1.1.2	Fixed-effects regression	20
2.3.1.1.3	Fixed-effects regression with interaction term	21
2.3.1.2	Random effects	22
2.3.1.2.1	Random intercept	24
2.3.1.2.2	Random intercept and random slope	24
2.3.2	Unconditional methods	25
2.3.2.1	Generalized estimating equations	25
3.0	**RESULTS**	**26**
3.1	Characteristics of data from the EECV Trials	26
3.2	Results of statistical methods to account for centre	27
3.2.1	Dealing with small centres	27
3.2.2	Outcome 1: Caesarean section	28
3.2.3	Outcome 2: Preterm birth	29
3.2.4	Outcome 3: Non-cephalic presentation at birth	30
4.0	**DISCUSSION**	**38**
4.1	Summary of findings	38
4.2	Discussion of statistical models to account for centre	38
4.3	Considerations for future research	42
5.0	**CONCLUSION**	**44**
REFERENCE LIST		**45**

LIST OF TABLES

Table 1-1	Summary of literature review on methods used to account for centre in RCTs of non-pharmacologic interventions and binary outcomes	13
Table 1-2	Summary of statistical methods used to account for centre	15
Table 3-1	Summary of results for Outcome 1: Caesarean section	31
Table 3-2	Summary of results for Outcome 2: preterm birth	32
Table 3-3	Summary of results for Outcome 3: non-cephalic presentation at birth	33

LIST OF FIGURES

Figure 3-1	Centre recruitment and balance of stratification	34
Figure 3-2	Forest plot for Outcome 1: Caesarean section	35
Figure 3-3	Forest plot for Outcome 2: preterm birth	36
Figure 3-4	Forest plot for Outcome 3: non-cephalic presentation at delivery	37

LIST OF APPENDICES

Appendix 1	Code for Mantel Haenszel test	49
Appendix 2	Code for Logistic Regression	50
Appendix 3	Code for Generalized Estimating Equations	55

LIST OF ABBREVIATIONS

ANOVA	analysis of variance
CI	confidence interval
CS	Caesarean section
CONSORT	Consolidated Standards of Reporting Trials
ECV	external cephalic version
EECV	early external cephalic version
GEE	generalized estimating equation
GLM	general linear model
MH	Mantel-Haenzsel
n	number
OR	odds ratio
p	probability of significance
PTB	preterm birth
RR	relative risk

DECLARATION OF ACADEMIC ACHIEVEMENT

This thesis project encompasses the work undertaken over 16 months of research and collaboration. Though the project focusses on statistical methods, the project is approachable to clinicians because it is written from the perspective of a clinician.

Since randomised controlled trials (RCTs) are considered the gold standard for biomedical research, continued research into the improvement of RCT design and analysis is necessary. This thesis contributes to the body of knowledge around RCT design and will appeal to the clinician/researcher.

1.0 BACKGROUND

1.1. The Clinical Problem: Breech Presentation and Caesarean Section

Breech presentation is a variation in the polarity of the fetus in a longitudinal lie. Typically, the term fetus will be head-down in a cephalic presentation. When the leading pole is the fetal pelvis, the fetal buttocks and/or feet will be presenting and this is termed breech presentation. (1) Breech presentation complicates three to four percent of pregnancies at term. (1;2) There are some maternal and fetal factors known to be associated with a higher incidence of breech pregnancy such as uterine anomalies, placenta previa, multiple pregnancy, and fetal anomalies such as hydrocephalus, but most often the cause of breech pregnancy is unknown. (1;3) Birth is more complicated for the breech baby regardless of mode of delivery, and breech pregnancy is a major contributing indication for Caesarean section. (4) Approximately eleven percent of Caesarean sections in the developed world are due to breech presentation, and 30 percent are repeat Caesarean sections due to primary Caesarean section. (5)

Both planned and emergency Caesarean sections confer increased risk of severe morbidity and mortality to women compared to vaginal birth. (5;6) Since breech pregnancy contributes to the Caesarean section rate, interventions to reduce the incidence of breech presentation are important. One intervention that has been shown to reduce the chance of breech presentation at term and Caesarean section is external cephalic version (ECV). (7) ECV is an obstetric procedure undertaken prenatally in which the practitioner manually turns a breech-presenting fetus into a cephalic presentation by manipulating the

fetus through the maternal abdominal wall. (4) Despite low complication rates overall, current research still concludes that ECVs be conducted at term gestation in facilities equipped to provide immediate Caesarean section if required. (4;7;8)

The ECV procedure is a hands-on skill that requires practice to perform. One source sites a learning curve of approximately 20 ECVs to become proficient. (9) The overall success rate of ECV is around 60 percent, but success rates between studies vary widely. (8;10)

1.2 The EECV Trials

1.2.1. The theoretical rationale of the EECV Trials

Most ECVs are attempted after a woman has reached full term pregnancy (37 weeks) in order to allow time for the fetus to turn spontaneously, as well as to ensure that any complication from the procedure necessitating emergency delivery results in a full-term birth. The investigators of the Early External Cephalic Version (EECV) trials theorized that starting the procedure a little earlier in the pregnancy could increase success rates by turning babies before they descend into the pelvis, and while the maximal amount of amniotic fluid is present. (11)

1.2.2 The clinical setting during the EECV Trials

The trial was conceived during a period of time when nearly all breech pregnancies were delivered by Caesarean section. An international multicentre randomised control trial had recently been published that concluded that neonatal outcomes were better for term breech babies born by Caesarean section versus vaginal delivery. (12) Clinical practice in

many centres changed after the publication of the trial. (13;14) This clinical setting allowed for the trial protocol to state that Caesarean section would be recommended to all women with a fetus that remained non-cephalic at the time of birth, and that those pregnancies would be delivered by emergency Caesarean section if labour began spontaneously. The protocol also specified that all women with a cephalic presentation at the time of birth and with no other indication for Caesarean section should be delivered vaginally. (11) Therefore, women who were planning trial of labour with breech presentation, or who were planning elective Caesarean section with cephalic presentation, were not eligible for inclusion in the trial. Since Caesarean section was a primary outcome of the trial, this protocol ensured the clinical effect could be maximized by minimizing crossover between presentation at birth and mode of delivery.

1.2.3 Overview of the EECV Trials

In 2003 the EECV Pilot Trial was published, (15) followed by a full-scale trial, the EECV2 Trial in 2011. (16) The EECV Trials were multi-centre randomised controlled trials aimed at investigating the effectiveness of early ECV (conducted between $34^{0/7}$ and $35^{6/7}$ weeks' gestation) compared to the usual timing of ECV at full term ($37^{0/7}$ weeks or beyond) on pregnancy outcomes. The primary outcome was the rate of Caesarean section and the secondary outcome was the rate of preterm birth. The study measured overall success of the ECV procedure by including the rate of non-cephalic presentation at delivery as another outcome.

Eighty-one centres in 22 countries recruited subjects for the EECV Trials. Eligible participants were women with a singleton breech fetus at a gestational age of $33^{6/7}$ weeks' to $35^{6/7}$ weeks. Participants were randomly assigned to have a first ECV procedure early or delayed, with stratification by centre and parity to ensure approximately equal numbers of early and delayed ECVs at each centre, as well as balance in the number of multiparous women (a known predictor of ECV success) in each group at each centre.

Fishers' exact test for binary outcomes was used to quantify the relationship between the exposure (timing of ECV procedure) and the primary and secondary outcomes (Caesarean section and preterm birth). The effects of the intervention were reported using relative risks (RR) and 95% confidence intervals (CI). Subgroup analyses were completed using logistic regression to test for interactions between baseline characteristics and treatment group for the primary and secondary outcomes.

1.2.4 The results of the EECV Pilot Trial and EECV2 Trial

The EECV Pilot Trial recruited 232 women from 25 centres in 7 countries. There were non-significant decreases in the rates of Caesarean section (RR 0.90, 95% CI [0.76, 1.08]; p=0.32) and non-cephalic presentation at birth (RR 0.86, 95% CI [0.70, 1.05] p=0.09) for women in the early ECV group. There was a non-significant increase in the rate of preterm birth for women in the early ECV group (RR 1.42, 95% CI [0.56, 3.59] p=0.31). (15) The clinically important findings of the EECV Pilot Trial supported the funding of a full scale RCT.

The EECV2 Trial recruited 1543 women from 68 centres in 21 countries. Women in the early ECV group were less likely to have a non-cephalic presentation at delivery (RR 0.84, 95% CI [0.75, 0.95] p=0.002), but the decrease in the Caesarean section rate remained not statistically significant (RR 0.93, 95% CI [0.85, 1.02] p=0.12) and the trend for an increase in preterm birth was strengthened though still not statistically significant (RR 1.48, 95% CI [0.97, 2.26] p=0.07). (16)

1.2.5 Discussion of findings by trial collaborators

Upon completion of data collection and data analysis, a collaborators meeting was held to discuss the findings of the EECV2 Trial prior to publication. (17) Investigators from some trial centres were surprised by the inconclusive findings of the trial, convinced that the treatment intervention had made a difference at their centre. Given that the power to detect a difference is much reduced at an individual centre, one can expect considerable chance variation between centres. (18) Basing future clinical practice on one centre's result is not advisable; however, if the centre is so different from the others that their outcomes actually are different, then it is worth exploring. The concepts of between-centre variance in the EECV trials formed the basis of this current study of centre effect.

1.3 The Methodological Problem: Centre effect in multicenter trials

1.3.1 Centre effect

Typical methods that have been used for analyses of multi-centre randomised controlled trials have assumed that outcomes for each study participant are independent of each

other, but this assumption does not hold when factors at centres cause outcomes to be more similar to each other than they are to outcomes at other centres. This so called "centre effect" refers to variance in trial outcomes that can occur between trial centres due to homogeneity (or clustering) of outcomes at each centre. Clustering of outcomes at a centre may occur due to a number of reasons. First, the population-level characteristics such as ethnicity and socio-economic status may be more similar. Further, the centre characteristics such as volume of patients treated and availability of experienced staff and up-to-date equipment could alter outcomes between centres. When the resulting clustering of outcomes is ignored in the analysis of multi-centre trials, incorrect effect estimates, confidence intervals, and p-values can occur. (19) There are several methods available to take centre effect into account in the statistical analyses with the aim of providing an effect estimate that includes between-centre variance. (20) Given that site investigators of the EECV2 Trial were reporting discordance between their centre statistics and the overall trial findings, investigating centre effect using appropriate statistical methods was examined.

The rise of evidence-based medicine has increased the number of randomised controlled trials (RCTs) conducted to test health care interventions. (21) Multicentre trials are often used to accumulate large sample sizes in a short time period, or to meet sample size requirements that would be impossible within one centre. The inclusion of different centres and providers is beneficial in pragmatic trials as it allows for greater generalisability of trial results. However, the variation between centres raises methodological issues. Each centre needs to rigorously follow the study protocol,

particularly around inclusion/exclusion criteria and the application of the intervention to reduce heterogeneity and allow for outcome results to be pooled. Furthermore, accurate reporting of the characteristics of the centres involved in the study can allow readers to assess the risk of bias and the usefulness of the results. (21;22)

As introduced above, the assumption made in many multicentre trials is that participants recruited to the trial are independent of each other. This assumption of independence is necessary to apply routine statistical methods such as the t-test, chi squared test, or Fisher's exact test to analyse data from the trial. However, management of individuals within the same trial centre may be similar, leading to the potential of outcomes from these individuals being correlated with each other. It is not hard to imagine that intervention success rates could differ from one centre to the next due to any number of combinations of practitioner experience, nursing support and expertise, medical equipment, and centre-specific treatment practices. When trial centres are in different international locales, the dissimilarities could be magnified. The correlation of individual outcomes at study centres is termed clustering, and if clustering is overlooked, the conclusions of the trial may be incorrect. (19;22)

Many RCTs test non-pharmacologic treatments such as: surgery, technical procedures, devices, rehabilitation, psychotherapy, behavioural interventions, and complementary and alternative medicine. (23) A review of all RCTs published in 2000 revealed that ten percent of RCTs were testing surgical or procedural interventions. (24) These trials have specific issues compared to pharmacologic trials because treatments are less standard and blinding is more difficult. (23) The CONSORT (Consolidated Standards of Reporting

Trials) Statement is a set of guidelines for researchers to improve the quality of reporting in RCTs. (25) An extension of the CONSORT Statement provides guidelines for improving the quality of RCTs of non-pharmacologic treatments. (23) The Statement identifies centre characteristics such as provider skill and centre volume that could impact patient outcomes. Since clustering of outcomes at study centres reduces statistical power, the CONSORT group recommends accounting for clustering in sample size calculation and statistical analysis. (23)

1.3.2 Review of reviews on accounting for centre effect in RCTs

Despite the development of statistical methods to account for centre effect and the recommendations by trial reporting guidelines, evidence from reviews of the literature indicate that most individually randomised multicentre trials do not account for centre effect. (26;27) Biau *et al* conducted a systematic review of the account of center and provider effects in large surgical and interventional RCTs. (21) Sixty-eight multicentre interventional randomized trials of 200+ patients from the years 2000 – 2005 met the inclusion criteria. They found that stratification by centre was reported in 38 percent of trials, and analysis adjusted for center was reported in 6 percent of trials. (21) Tangri *et al* published a similar systematic review of the literature to assess the extent of adjustment for centre in RCTs of medicinal products. (27) They included 101 multicentre RCTs published in 2007 in four prominent medical journals. Of the 101 trials, 36 percent used random allocation stratified by centre, and 18 percent adjusted for centre in the statistical analysis. (27) Both reviews conclude that improvements to trial reporting regarding centre effects are needed.

1.4 Accounting for centre in the EECV Trials

1.4.1 Rationale

It is recommended that researchers undertaking multicentre trials account for centre in both design and analysis, yet reviews of the literature indicate that most of the time this does not occur. The statistical analysis of the EECV2 trial was done according to the trial protocol written in 2003 and, like many other trials, did not take centre effect into account. Knowledge about centre effect has been gradually disseminated from the biostatistics literature to the clinical research literature over the last decade. (19;23) Because centre effect was not considered during the analysis of the EECV trials, and because issues of centre differences were raised during the EECV2 Trial Collaborator's meeting, a secondary analysis to account for centre effect was undertaken. Several methods to adjust for centre will be applied, and the results of each method will be compared to the other and to the original individual-level analysis of the trial. Clinical implications of altered trial results and considerations for future research will be explored.

1.4.2 Trial characteristics

In order to choose statistical methods with which to conduct a secondary analysis of the EECV trials, an understanding of the trial's characteristics is necessary. The following points outline the characteristics considered when identifying statistical methods.

1.4.2.1 Type of RCT

The EECV trials had individually-randomised parallel group designs where each centre had patients randomised to each intervention group. It is important to differentiate this design from cluster-randomization where the unit of randomization is the centre.

1.4.2.2 Type of intervention

The EECV trials tested an obstetric procedural intervention that was administered by various practitioners at each centre. A study protocol was implemented at each centre to ensure that the procedure was offered at the appropriate time in relation to the intervention group that patient was allocated to (early ECV or delayed ECV). Best attempts were made to ensure that the practitioners involved in providing the intervention were experienced with the procedure by recording approximate numbers of past ECVs completed, and having proficiency signed off by a clinical department head. (11) Blinding to group allocation was not possible due to the type of intervention. As noted in the CONSORT statement, non-pharmacologic interventions such as this one may be more prone to a centre effect. (23)

1.4.2.3 Type of outcomes

The primary and secondary outcomes in these trials were the rates of Caesarean section and preterm birth. The rate of non-cephalic presentation at the time of birth was another outcome of the trial. All three outcomes are binary outcomes, so the statistical analyses chosen to adjust for centre must accommodate binary outcomes.

1.4.2.4 Centres and enrollment

The EECV trials were international multicentre trials with varying numbers of participants at each centre. Having unequal numbers of patients per centre is a common finding in RCTs. Multicentre design is typically chosen to increase recruitment rate, so it is inefficient for an investigator to limit recruitment at one centre and wait for another centre to catch up in order to maintain equal numbers of patients at each centre. (18) When centre size varies, analyses that incorporate weighting by centre size are preferred.

1.4.2.5 Randomisation stratified by centre

Stratification is a method used to ensure that trial groups are balanced on important prognostic factors. (28) Randomisation alone should result in balanced groups, but can lead to random differences between groups. If chance imbalance occurs for a known prognostic factor, the credibility of the trial is at risk. (28;29) In the EECV trials, the randomisation was stratified by parity and centre. Parity was chosen as a stratification variable because it is a known predictor of ECV success. The choice of centre as a second stratification variable suggests that the investigators believed centre to be an important factor in treatment outcomes.

Although stratification by centre is considered one way of accounting for centre differences by ensuring that each centre contributes approximately equal numbers of patients to the treatment groups, (19) stratification alone is not an adjustment for centre effect. When a variable is chosen as a stratification variable, it still needs to be accounted for in the statistical analysis. (29) In the EECV trials, no further analysis of the centre

variable was undertaken after the initial step of stratification. This is not uncommon as shown by Tangri *et al* in their review of RCTs that concluded stratification by centre occurred in 36 percent of RCTs and adjustment for centre in 18 percent. (27) Ignoring centre in the statistical analysis after stratifying by centre downplays the strength of the design. (30)

1.4.3 Methods to account for centre effect

The biostatistics literature was reviewed to determine which statistical methods would be appropriate to account for centre in the EECV trials. An overview of the literature is shown in Table 1-1. In consultation with a statistician, seven statistical models were chosen to apply to the EECV trial data, and the attributes of each model are described according to the characteristics of the data (Table 1-2).

1.5 Research Objective

Using a combined dataset of the EECV trials, the objective of this project is: i) to estimate the effect of ECV timing (early or delayed) on the outcomes of Caesarean section, preterm birth, and non-cephalic presentation at the time of birth without accounting for centre, and ii) to assess the consistency or robustness of the results under different methods of accounting for centre. The robustness will be assessed relative to the results of part i) in terms of the impact of each method on the magnitude, direction and significance of the effect estimate.

Table 1-1: Summary of literature review on methods used to account for centre in RCTs of non-pharmacologic interventions and binary outcomes

Author/Year	Overview	Types of outcomes	Types of analyses
Agresti 2000 (20)	Reviews strategies for comparing treatments on a binary response with multi-centre data. Problems presented by sparse data are discussed.	Binary response with multi-centre data	Frequentist approaches: -Fixed effects model -Random effects model -Treatment-by-centre interaction -Mantel-Haenszel test
Berlin 1999 (31)	Applies several analytic methods to adjust for centre effect for a cardiac trial.	Binary response data from multicentre trials	- Mantel-Haenszel odds ratio estimate stratified on centre - Random intercept logistic model - Generalized estimating equation
Gould 1998 (32)	Compares fixed and mixed model ANOVAs with empirical and conventional Bayes methods to account for centre effect and treatment-by-centre interaction.	Continuous. "Empirical and conventional Bayes methods … can be applied to …binary data, with appropriate changes in computational details."	- Fixed ANOVA - Mixed model ANOVA - Empirical Bayes - Conventional Bayes The 4 models performed similarly, but the Bayes methods provided more information about effect variability.
Hanley 2003 (33)	Describes generalized estimating equations and their application to complex multivariate data	Binary and quantitative response data	-Generalized Estimating Equations
Hardin 2008 (34)	Outlines an empirical binomial hierarchical Bayesian model for evaluating multisite demonstration and effectiveness studies	The example given has continuous outcomes	Comments on: -Fixed effects ANOVA with terms for the center and the treatment, and an interaction term for treatment by the center. -Un-weighted GLM Type III analysis for multicentre trials when treatment effects differ among centres. Describes: -Empirical binomial hierarchical Bayesian model could more accurately use the data from each of the sites to evaluate the treatments intervention.

Lee 2005 (35)	In multicentre trials, random effect models can be used to model the hierarchical structure of patients within clusters (centres).	Continuous outcomes; similar models can be fitted for binary outcomes using logistic regression.	- Random intercept model - Random intervention effect model (includes a treatment-by-cluster interaction term)
Localio 2001 (19)	Provides examples from the biomedical literature to review the analytic options for adjusting for centre in a multicenter study.	Binary	- Mantel-Haenszel - Conditional logistic regression (fixed or random effects) - Fixed-effect logistic regression with center-by-treatment interaction - Random coefficient model - Generalized estimating equation
Senn 1998 (18)	Discusses controversies in the planning and analysis of multi-centre trials including: fixed vs. random effects models and interaction terms.	Not specified	-GLM Type II analysis -GLM Type III analysis -Fixed vs. Random effects models
Zhang 1997 (36)	Proposes new Mantel-Haenszel test statistics for correlated binary data.	Binary	-Mantel-Haenszel
ANOVA: analysis of variance; GLM: general linear model			

Table 1-2: Summary of statistical models used to account for centre

	Model Name		Description	Comments	Relation to EECV Trial data
Conditional Methods					
A	Mantel Haenzsel		Can be used when the response variables are binary, and there are only two treatment groups (intervention and control). The data are summarized into a series of 2x2 tables based on centre (ie. each centre has its own 2x2 table) (36)	- Has computational simplicity - Is appropriate for studies with many small centres - Assumes homogeneity of treatment effects across centres	The data type is ideal: binary outcome, two groups, and many small centres. Does not account for heterogeneity of treatment effect which is a real possibility in the EECV trials.
B	Fixed effects regression		Odds of outcome conditional on group, with centres as fixed intercepts.	Works well when there are a small number of large centres. The results reflect the centres involved, not all possible centres.	In the EECV trials, small centres need to be removed from the analysis, which may be problematic if the total N is reduced.
C	Fixed effects regression with interaction		Odds of outcome conditional on group, with centre as a fixed intercept, and incorporating treatment-by-centre interaction.	Accounts for heterogeneity of treatment effects across centres.	Small centres and centres with zero counts in the treatment-by-centre interaction term need to be removed, further reducing the size of the dataset and contributing to loss of power.
	C1	Weighted by centre size	An average weighted on the size of the centre takes the amount of information provided by each centre into account.	Information provided by each patient is equally used	
	C2	Unweighted by centre size	A simple average of the treatment effect at each centre.	- Each centre has a true average - Increases variance when there are very small centres	Many small centres holding the same weight as larger centres can bias the estimate.
D	Random intercept		Odds of outcome conditional on group, with centres as random intercepts. Assumes no variation in treatment effect,	Regards variation of treatment effect across centres as random variation (32)	The trial centres are thought to be generalisable to other centres

E	Random intercept and random slope	Odds of outcome conditional on group, with centres as random intercepts and incorporating variation of treatment effects across centre.	A different slope for each centre allows for effect modification	Probably a good model since effect modification is a possibility
Unconditional Methods				
F	Generalized Estimating equations	Estimates the average treatment effect and then adjusts confidence intervals for the correlation of patients within centres.	Need a large number of centres (at least 30)	The model fits with the data; there are enough centres for the model to apply.

2.0 METHODS

2.1 Data preparation

The data from the EECV Pilot Trial and the EECV2 Trial were merged for the outcomes of interest using SPSS. The data was utilised in SAS and R for the statistical analysis. Participants were excluded from analysis if they withdrew from the trial, if there was loss to follow-up, or if there was missing data regarding mode of delivery, gestational age at birth, or presentation.

2.2 Data description

Characteristics of the merged dataset were obtained by running frequencies to determine the number of centres, the number of women enrolled at each centre, and the balance of allocation to treatment groups at each centre as achieved by stratification.

2.3 Statistical methods to account for centre effect

Statistical methods to account for centre were chosen by reviewing the biostatistics literature and relating the functions of each test to the specifics of the combined EECV trial dataset. The literature does not suggest one preferred statistical method to account for centre effects. In fact, there is a lack of evidence on which models perform best in various situations. (37) Each method has strengths and weaknesses given the characteristics of the data (Table 1-2). As outlined in the Background, the methods were chosen to work with data from a multicentre trial that used independent random allocation

to group stratified by centre, with binary outcomes. The methods are broadly grouped as conditional and unconditional methods and are detailed as follows:

2.3.1 Conditional Methods

Conditional methods estimate treatment effects by stratifying (or conditioning) on centre. These methods are appropriate for multicentre trials in which each centre has participants randomly assigned to different treatment arms. (19) Conditional methods can be subdivided into fixed- and random-effects models. A fixed-effects analysis considers a centre to represent only itself, while a random-effects analysis represents the population of centres from which the study sample was drawn. (18;19) Both fixed and random effects models produce a single estimate of treatment effect if the treatment effect is assumed constant across centres. If the treatment effect is suspected to be different across centres, altering the model to include a treatment-by-centre interaction term can improve the model's fit. (19) Trials that test procedural interventions, such as the EECV trials, may be more likely to have heterogeneity of treatment effects across trial centres. (23) This is due to the difficulty of standardising and administering the procedural intervention in a consistent manner across study sites, and is downplayed when a strict study protocol is followed. (23) The interaction that occurs when trial sites have different treatment effects is also known as effect modification.

Six different conditional methods were employed to account for centre in the EECV trials. Four are fixed-effects models, and two are random-effects models. Two models

include treatment-by-centre interaction terms to account for the possibility of effect modification in the EECV trials.

2.3.1.1 Fixed-effects

2.3.1.1.1 Mantel-Haenzsel method

The first method we used to account for centre was the Mantel-Haenzsel method, a fixed-effect analysis that summarizes data into a series of 2×2 tables based on covariates or strata. The Mantel-Haenzsel test is often used when the trial has binary response variables and only two treatment groups; (36) however, it has been generalized to analyse $2(response) \times J(exposure)$ tables. (38) The Mantel-Haenzsel test performs well even with sparse data and it is suitable for studies like the EECV trials that have many centres and few participants per centre. (19;20) The simplicity of the Mantel-Haenzsel test is considered an advantage (20).

To perform the Mantel-Haenzsel test, data from the EECV trials was organized into two columns. The first column represented cells in a 2×2 table, and the second column contained count data for the outcome of interest per cell. For example, for the first outcome (Caesarean Section) the first centre was represented as:

2×2 Table:

		Outcome: Caesarean section	
		1	0
Group	1	a	b
	0	c	d

Data format:

cell	count
a1	4
c1	9

b1	6
d1	3

Centres that had only one participant were removed from the analysis. The statistical software program R Version 2.12.2 was used to run the Mantel-Haenszel test. The code can be found in Appendix 1.

2.3.1.1.2 Fixed-effects regression

The fixed-effects regression estimates only within-centre effects. (19) It achieves this by including a separate intercept for each centre as a fixed effect, restricting the inference of the results to included centres. (37) The model works best when there are many participants spread across few centres. Severely biased estimates occur when each centre has few patients and few events because each centre acts as a covariate – and one should always avoid having too many covariates in a regression. (19) Since the EECV trials had low enrollment at some centres, the deletion of some small centres was expected for this model.

The fixed-effects regressions were run using data in aggregate form. Data was organized into a table that included a row for each group in each centre and count data for the number of participants that did not experience an event and the number of participants that did experience an event. For example, for the first outcome (CS), the first centre was represented as:

Data format:

Centre	Group	no CS	CS
1	1	6	4
1	0	3	9

The model for the fixed-effects regression (Equation 1) models the impact of treatment $(X)(1 = early\ ECV, 0 = delayed\ ECV)$ on the odds of having the outcome event $(Y)(1 = CS, 0 = noCS)$ (log of the odds) $logit\ (\pi_{ik}) = log \frac{\pi_{ik}}{1-\pi_{ik}}$ and allows a separate intercept for each centre (k) as a fixed effect. Let $\pi_{ik} = \mathrm{P}\ (Y = 1 \mid x = i, z = k)$, for $i = 1,0, k = 1, \dots, 81$. This is the probability of having the outcome event for someone receiving the i^{th} treatment in the k^{th} centre. β_{0k} represents the log odds for the control group in individual centres, and β_1 represents the log odds of the treatment across all centres.

$$logit(\pi_{ik}) = \beta_{0k} + \beta_1 X \quad (1)$$

2.3.1.1.3 Fixed-effects regression with a "treatment-by-centre" interaction term

As described above, most analyses assume the absence of effect modification. When effect modification is suspected or present, adding an interaction term can estimate treatment effects specific to each center. We used a fixed-effects regression with a centre-by-treatment interaction term to account for the possibility that centres could have varying treatment effects (Equation 2). Here β_{1k} represents the log odds of having the outcome in the treatment group compared to the control group for centre k.

$$logit(\pi_{ik}) = \beta_{0k} + \beta_{1k} X \quad (2)$$

The data format was maintained in the aggregate form, just as it was for the first fixed-effect regression. The model was run twice: once with and once without weighting by the size of the centre (number of women enrolled). In multicentre trials with a large disparity between the size of the centres, as was seen in the EECV trials, weighting by centre size prevents small centres from inflating the variance. (18)

The fixed-effects regressions were fitted using *proc genmod* in SAS Version 9.2 (Cary, NC). The code can be found in Appendix 2.

2.3.1.2 Random-effects

Random effects models are another way to model the hierarchical structure of patients within centres in individually randomized multicentre trials. (35) In contrast to fixed effect models that provide results relevant only to the study sample, random effects models are generalized to the entire population of possible centres by assuming that the trial centres are a random sample of all centres. Although this is not the way that centres are chosen in a pragmatic randomised controlled trial, the underlying notion is that the results of the trial provide inference about patients in general, even those attending centres not included in the trial. (18) A random effects model can be an improvement over a fixed-effects model when there are many centres. (18)

Random effects models have been used for decades for continuous outcomes, but model interpretation and fitting is more difficult with binary data. Random-effects models are

also known by various other names such as: centre-specific, mixed-effects, variance component, hierarchical, multistage, or empirical Bayes regressions. (19)

These models assume an absence of association between the random effect (centre) and the chance of being treated. In the EECV trials, this assumption is satisfied because participants were allocated in equal proportions to treatment and control within each centre. Observations from centres that have no "events" do not contribute to effect estimation in the random effects models. (19)

The original dataset of individual level data was used to run the random effects regressions. For example, for the first outcome (CS), the first centre was represented as:

ID	Centre	Group	CS
1	1	1	1
2	1	1	1
3	1	1	1
4	1	1	1
5	1	1	0
6	1	1	0
7	1	1	0
8	1	1	0
9	1	1	0
10	1	1	0
11	1	0	1
12	1	0	1
13	1	0	1
14	1	0	1
15	1	0	1
16	1	0	1
17	1	0	1
18	1	0	1
19	1	0	1
20	1	0	0
21	1	0	0
22	1	0	0

Two random effects models were used to adjust for centre in the EECV trials. The statistical software program R Version 2.12.2 was used to run both models and the code can be found in Appendix 1.

2.3.1.2.1 Random intercept

When the random effect is the intercept $(\beta_0 + b_{0k})$ we are referring to the log odds of having the outcome in the control group at each centre. This model, reflected in Equation 3, adjusts for a centre effect under the assumption that though there may be clustering of outcomes across both treatment arms in each centre, variation in the treatment effect is unlikely (35). β_1 represents a single treatment effect over participating centres.

$$logit(\pi_{ik}) = (\beta_0 + b_{0k}) + \beta_1 X \quad (3)$$

2.3.1.2.2 Random intercept and random slope

This model includes an additional random effect $(\beta_1 + b_{1k})$, often known as the random slope, to refer to the random treatment effect at each centre. (37) By including random intercept and random slope in the model (Equation 4), we account for variation in treatment effects across centres. The random intercept, $(\beta_0 + b_{0k})$, and the random slope, $(\beta_1 + b_{1k})$, both follow the assumptions of the normal distribution. This model has also been described as a "random intervention effects" model. (35)

$$logit(\pi_{ik}) = (\beta_0 + b_{0k}) + (\beta_1 + b_{1k}) X \quad (4)$$

2.3.2 Unconditional Methods

Unconditional methods include marginal or population-averaged models that estimate an average treatment effect across all centres and then adjust for correlation of outcomes at centres. (19) One unconditional method, generalized estimating equations, was applied to the EECV trial data.

2.3.2.1 Generalized estimating equations

Generalized estimating equations (GEE) model the marginal population treatment effects averaged across centres in two steps. First a model similar to ordinary logistic regression without regard to the centre is fitted. (19;37) Then the model is refitted to adjust the standard error and confidence intervals for within-centre dependence. (37) By using weighted combinations of observations, the GEE approach extracts the appropriate amount of information from correlated data. (33) Large numbers of centres (at least 30) are required for the underlying theory of the GEE model to apply. (19)

Application of the GEE model to the EECV trial data was achieved using data in the same individual patient-level format as was used for the random-effects regressions. The GEE model was run using *proc genmod* in SAS Version 9.2 (Cary, NC) and the code is provided in Appendix 3.

3.0 RESULTS

3.1 Characteristics of data from the EECV Trials

The data from the EECV Pilot Trial and the EECV2 Trial were merged to create a dataset containing 834 variables and 1775 participants. Eleven participants were removed from the analysis due to withdrawal (1), loss to follow-up (8), or missing significant data (2). This resulted in a dataset containing information on 1764 participants at 81 centres.

The 81 trial centres were located in 22 countries: Canada (22), UK (6), USA (5), Australia (13), Israel (5), South Africa (2), Argentina (6), New Zealand (1), Chili (5), Denmark (2), Germany (1), Ireland (1), Jordan (1), The Netherlands (3), Poland (1), Spain (1), Brazil (1), Egypt (1), Portugal (1), Hungary (1), Estonia (1), and Oman (1). Centre sizes were unequal with the number of women recruited at each centre varying from 1 to 117. There were a small number of centres that recruited large numbers of participants, and a large number of centres that recruited small numbers of participants. The mean centre size was 45 and the median centre size was 13. Random block sizes ensured that approximately equal numbers of patients were randomized to the intervention and control groups at each centre. (11) Overall, 881 were randomized to the early ECV group and 883 to the delayed ECV group. The recruitment rates and balance of stratification are presented in Figure 3-1.

3.2 Results of Statistical methods to account for centre

The seven statistical models described in the methods were applied to the EECV trial data to adjust for centre effect, and the results for three selected outcomes are reported in Tables 3-1, 3-2, and 3-3.

3.2.1 Dealing with small centres

Different methodological approaches required different ways of handling low-recruiting or "small" centres in the analysis. The Mantel-Haenszel test used data from 78 of 81 centres after the 3 centres that enrolled only one woman were removed from the analysis. For the three fixed-effects regressions, centres had to be removed from analysis if all the participants at that centre were in one treatment group, or if all the participants experienced the same outcome. The removal of these centres was necessary because the statistical model is constructed only with centres that provide sufficient statistics. (20) Differing numbers of centres were removed from the analysis for the three outcomes of interest. Seventy-one centres representing 1739 participants were included for the outcomes of caesarean section and non-cephalic presentation at birth; 46 centres representing 1434 women were included for the outcome of preterm birth. Further centres were removed from analysis when the treatment-by-centre interaction term was added to the fixed-effects regression model due to zero counts in the interaction term. For the outcome of Caesarean section, 57 centres representing 1655 women were included, for preterm birth 14 centres representing 646 women were included, and for non-cephalic presentation at birth, 57 centres representing 1649 women were included in the analysis.

The random-effects models and the generalized estimating equation were run with the entire dataset. Sparse centre data does not cause problems for the random-effects approach because parameter space does not increase with the number of centres. (20)

3.2.2 Outcome 1: Caesarean Section

ECV has been shown to reduce the risk of Caesarean section at term, (7) but the EECV2 trial did not demonstrate a statistically significant difference in CS rates when comparing early ECV with the routine practice of conducting ECV at term. (16) The merged dataset (EECV Pilot Trial and EECV2 Trial) was analysed using Fisher's exact test to provide an individual-level baseline analysis to which the methods that account for centre could be compared. Unadjusted for centre, the odds ratio for CS for those in the early ECV group was 0.92, 95% CI [0.85, 1.00] p=0.07.

Using methods to adjust for centre effect, the odds ratio changed by 2 to 11 percent toward an increased treatment effect. The fixed-effect regression was the only model that indicated statistical significance favouring treatment (OR 0.81, 95% CI [0.67,0.99], p=0.04). The Mantel-Haenzsel test, random-effects regressions and GEE provided similar results with effect estimates varying from 0.82 to 0.86, and upper limit 95% CI varying from 1.00 to 1.06. The fixed-effect regression with interaction term un-weighted by centre size produced the widest 95% CI (0.68, 1.20).

Overall, adjusting for centre effect strengthens the association between early ECV and the reduction of CS. The results are reported in Table 3-1 and Figure 3-2.

3.2.3 Outcome 2: Preterm Birth

Preterm birth was selected as a secondary outcome for the EECV trials because it impacts neonatal morbidity and health care utilisation. (11) As an outcome it is simple to measure, and acts as an important surrogate outcome for serious perinatal or neonatal morbidity. (11)

The odds of preterm birth for women in the early ECV group, unadjusted for centre, was 1.44, 95% [CI 0.98, 2.13] p=0.06. Methods to adjust for centre strengthened the association indicated in the unadjusted results. The fixed-effect model was once again the only model to show a statistically significant result with the odds of preterm birth equalling 1.57, 95% CI [1.02, 2.43] p=0.04. The Mantel-Haenzsel test also bordered significance with a p-value of 0.05. The random-effects regressions and GEE model provided similar odds ratios and confidence intervals that were close to the Mantel-Haenzsel test and fixed-effect regression results. The fixed-effect regression with the interaction term, both weighted and un-weighted for centre size gave different results than the 5 other models. The confidence intervals were wider and the p-values higher. The fact that these models utilised less than half of the trial data is the likely cause.

In general, adjusting for centre strengthens the association between early ECV and odds of preterm birth suggesting that preterm birth is a risk of performing ECV prior to term gestation.

3.2.4 Outcome 3: Non-cephalic presentation at birth

The presentation of the fetus at the time of birth was collected as one of a list of other outcomes in the EECV trials. Having a non-cephalic presentation at the time of birth represents all women for whom: ECV was unsuccessful, the fetus did not turn spontaneously, or the baby reverted to a non-cephalic presentation prior to birth. For this analysis, the outcome is a way of denoting "success" of the procedure, and may represent centre differences related to practitioner skill and success with the procedure. Many women will undergo CS despite cephalic presentation; therefore, it is important to separate the success of the procedure from the eventual outcome of interest.

Unadjusted for centre, the odds of having a baby in a non-cephalic presentation at the time of birth was 0.84, 95% CI [0.76, 0.93] p=0.001 for women in the early ECV group compared to the delayed ECV group.

When adjusted for centre effect, the association was further strengthened with the OR varying from 0.70 to 0.72 among six out of seven statistical models. The fixed-effect model with interaction term un-weighted for centre resulted in much wider confidence intervals and a non-significant result. Once again, this aberration is likely due to the weighting of small centres equally with larger centres.

The results confirm that those in the early ECV group are more likely to have a cephalic presentation at the time of birth.

Table 3-1: Results for Outcome 1: Caesarean section

Outcome 1: Caesarean Section					
UNADJUSTED FOR CENTRE					
EECV2 Trial – Fisher's exact test*		n=1533	RR 0.93	CI: [0.85, 1.02]	p=0.12
Merged Dataset – Fisher's exact test		n=1764	OR 0.92	CI: [0.85, 1.00]	p=0.07
ADJUSTED FOR CENTRE					
Mantel-Haenszel	78 centres	n=1761	OR 0.82	CI: [0.68, 1.00]	p=0.05
Fixed-effects regression	71 centres	n=1739	OR 0.81	CI: [0.67, 0.99]	p=0.04
Fixed-effects regression with interaction term	57 centres weighted by centre size	n=1655	OR 0.86	CI: [0.70, 1.06]	p=0.16
	57 centres un-weighted	n=1655	OR 0.90	CI: [0.68, 1.20]	p=0.47
Random intercept	81 centres	n=1764	OR 0.83	CI: [0.69, 1.00]	p=0.05
Random intercept and random slope	81 centres	n=1764	OR 0.86	CI: [0.70, 1.05]	p=0.14
Generalized Estimating Equation	81 centres	n=1764	OR 0.83	CI: [0.68, 1.02]	p=0.08

EECV2: Early External Cephalic Version 2 Trial; RR: relative risk; OR: Odds Ratio; CI: Confidence Interval

* as reported in the EECV2 Trial publication using RR (16)

Table 3-2: Results for Outcome 2: Preterm birth

Outcome 2: Preterm Birth					
UNADJUSTED FOR CENTRE					
EECV2 Trial – Fisher's exact test*		n=1533	RR 1.48	CI: [0.97, 2.26]	p=0.07
Merged Dataset – Fisher's exact test		n=1764	OR 1.44	CI: [0.98, 2.13]	p=0.06
ADJUSTED FOR CENTRE					
Mantel-Haenszel	78 centres	n=1761	OR 1.55	CI: [1.01, 2.37]	p=0.05
Fixed-effects regression	46 centres	n=1434	OR 1.59	CI: [1.03, 2.45]	p=0.04
Fixed-effects regression with interaction term	14 centres weighted for centre size	n=646	OR 1.04	CI: [0.49, 2.24]	p=0.91
	14 centres un-weighted	n=646	OR 1.20	CI: [0.61, 2.38]	p=0.59
Random intercept	81 centres	n=1764	OR 1.49	CI: [0.98, 2.29]	p=0.06
Random intercept and random slope	81 centres	n=1764	OR 1.58	CI: [0.97, 2.58]	p=0.07
Generalized Estimating Equation	81 centres	n=1764	OR 1.46	CI: [0.97, 2.19]	p=0.07

EECV2: Early External Cephalic Version 2 Trial; RR: relative risk; OR: Odds Ratio; CI: Confidence Interval

* as reported in the EECV2 Trial publication using RR (16)

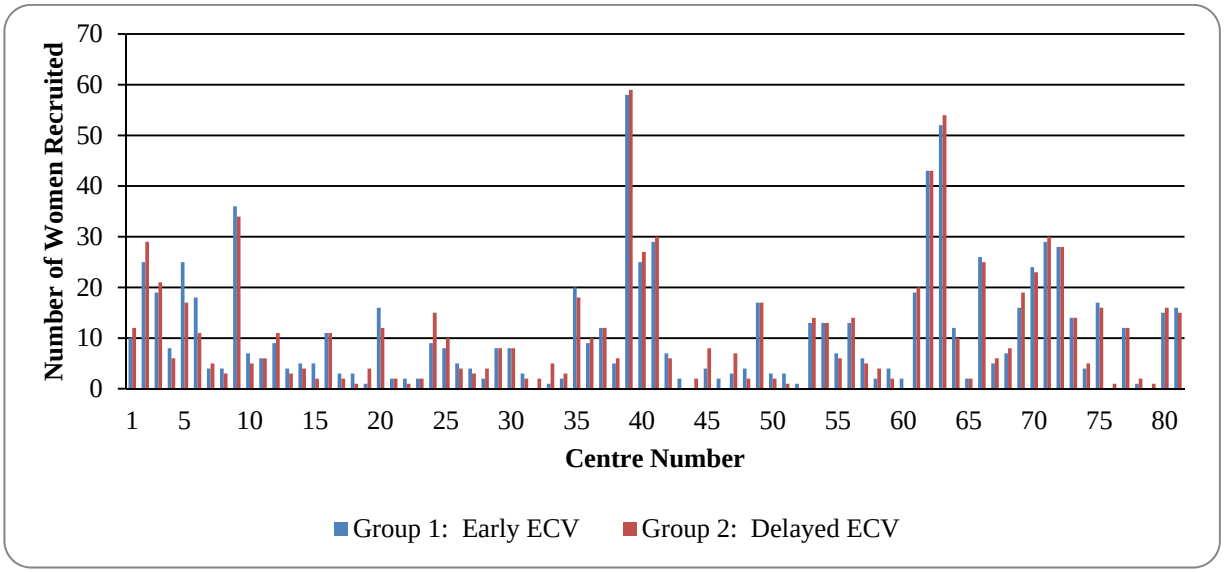
Table 3-3: Results for Outcome 3: Non-cephalic presentation at birth

Outcome 3: Non-cephalic presentation at birth					
UNADJUSTED FOR CENTRE					
EECV2 Trial – Fisher's exact test*		n=1533	RR 0.84	CI [0.75, 0.94]	p=0.002
Merged Dataset – Fisher's exact test		n=1764	OR 0.84	CI [0.76, 0.93]	p=0.001
ADJUSTED FOR CENTRE					
Mantel-Haenszel	78 centres	n=1761	OR 0.71	CI [0.59, 0.87]	p<0.001
Fixed-effects regression	71 centres	n=1739	OR 0.70	CI: [0.58, 0.85]	p<0.001
Fixed-effects regression with interaction term	57 centres weighted for centre size	n=1649	OR 0.72	CI: [0.59, 0.89]	p=0.002
	57 centres un-weighted	n=1649	OR 0.84	CI: [0.63, 1.12]	p=0.23
Random intercept	81 centres	n=1764	OR 0.71	CI: [0.59, 0.86]	p<0.001
Random intercept and random slope	81 centres	n=1764	OR 0.72	CI: [0.59, 0.90]	p=0.003
Generalized Estimating Equation	81 centres	n=1764	OR 0.72	CI: [0.59, 0.89]	P=0.002

EECV2: Early External Cephalic Version 2 Trial; RR: relative risk; OR: Odds Ratio; CI: Confidence Interval

* as reported in the EECV2 Trial publication using RR (16)

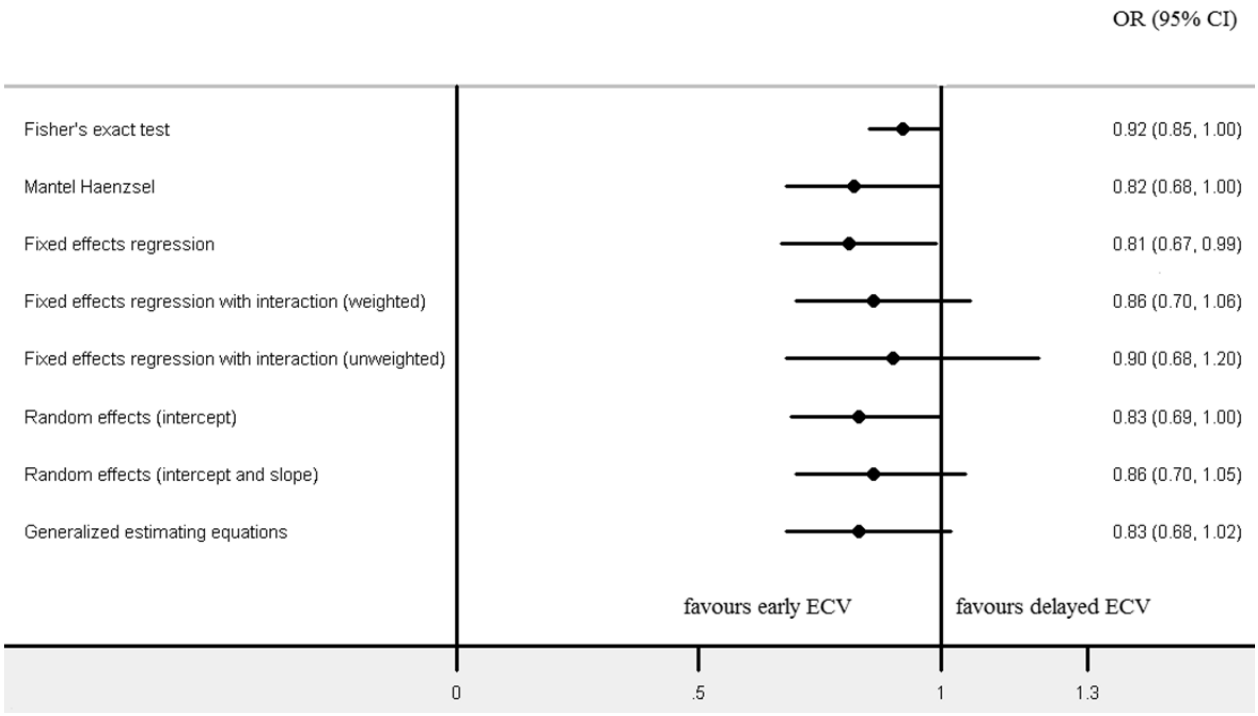
Figure 3-1: Centre recruitment and balance of stratification to groups



Legend of Centre Numbers:

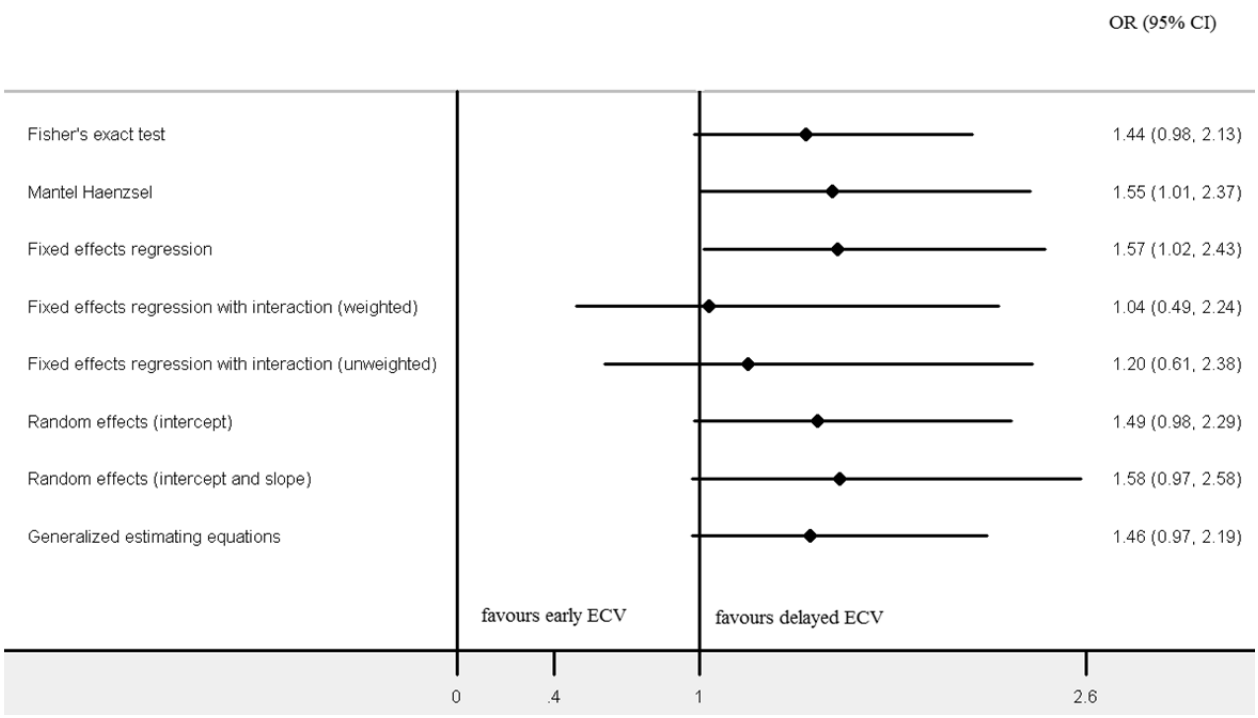
1-22	Canada
23-28	UK
29-33	USA
34-46	Australia
47-51	Israel
52-53	South Africa
54-59	Argentina
60	New Zealand
61-65	Chile
66-67	Denmark
68	Germany
69	Ireland
70	Jordan
71-73	The Netherlands
74	Poland
75	Spain
76	Brazil
77	Egypt
78	Portugal
79	Hungary
80	Estonia
81	Oman

Figure 3-2: Forest Plot for Outcome 1: Caesarean Section



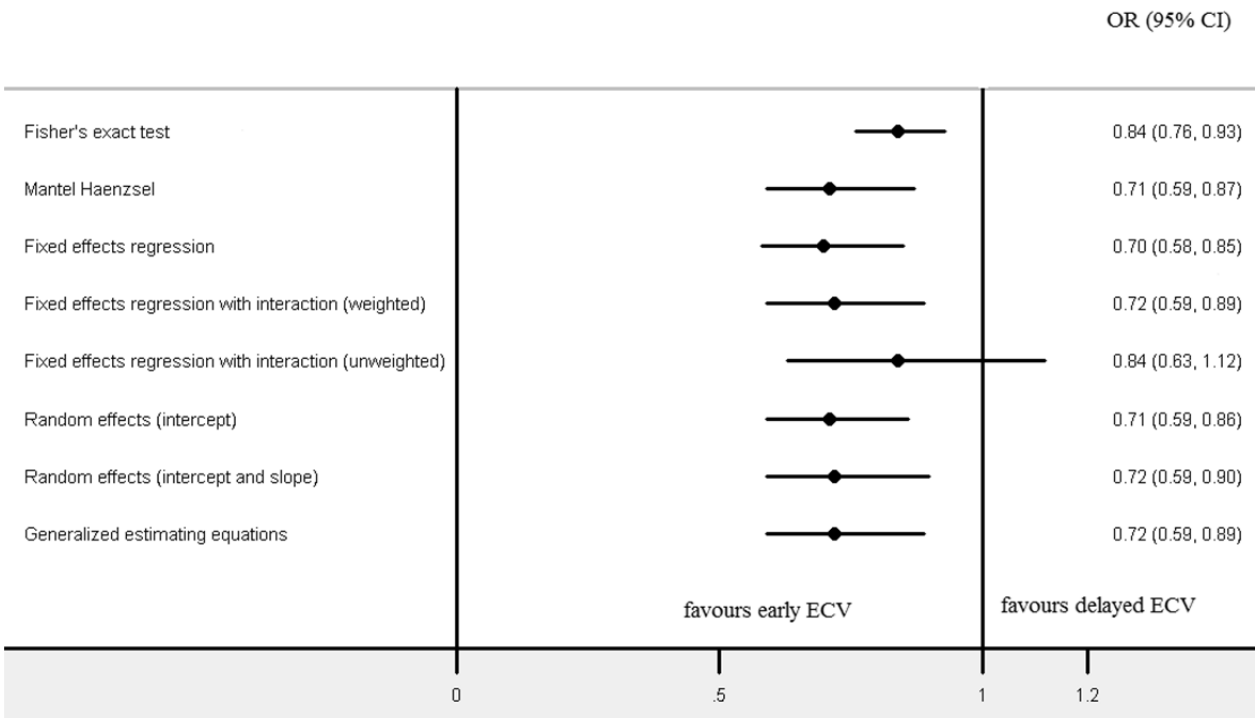
Mantel-Haenzsel test included 78 centres and 1761 participants. The Fixed-effects regression included 71 centres and 1739 participants. The Fixed-effects regression with interaction term, both weighted and un-weighted for centre size, included 57 centres and 1655 participants. The random intercept, random intercept and slope, and the generalized estimating equations used all 81 centres and 1764 participants.

Figure 3-3: Forest Plot for Outcome 2: Preterm birth



Mantel-Haenzsel test included 78 centres and 1761 participants. The Fixed-effects regression included 46 centres and 1434 participants. The Fixed-effects regression with interaction term, both weighted and un-weighted for centre size, included 14 centres and 646 participants. The random intercept, random intercept and slope, and the generalized estimating equations used all 81 centres and 1764 participants.

Figure 3-4: Forest plot for Outcome 3: Noncephalic presentation at birth



Mantel-Haenzsel test included 78 centres and 1761 participants. The Fixed-effects regression included 71 centres and 1739 participants. The Fixed-effects regression with interaction term, both weighted and un-weighted for centre size, included 57 centres and 1649 participants. The random intercept, random intercept and slope, and the generalized estimating equations used all 81 centres and 1764 participants.

4.0 DISCUSSION

4.1 Summary of findings

The results obtained from the statistical models that incorporated centre effect were different than the trial's original analysis. For all three outcomes, point estimates and confidence intervals moved away from the null value when centre was taken into account. This indicates strengthened associations between the intervention and the outcomes. The odds of having a cephalic presentation at delivery was significantly increased for women randomized to the early ECV group, and this translated to a trend toward reduction in Caesarean section rates. The trend of increased preterm birth for women randomized to the early ECV group was also strengthened after accounting for centre.

4.2 Discussion of statistical models that account for centre effect

The seven models used to account for centre have advantages and disadvantages in their application to the EECV trial dataset. Based on background information from the biostatistics literature, several models were expected to perform better than others when applied to the EECV trial data. Four models seemed to be particularly suitable; these were the Mantel-Haenzsel test, the two random effects regressions, and the GEE model. Simulation studies provide valuable information for comparing the performance of statistical models under known conditions; however, this analysis offers a real-world example of how the statistical models function under conditions specific to the EECV trial.

The Mantel-Haenzsel test was applicable to the dataset and simple to apply because the data included two treatment groups and a binary outcome, stratified by centre. The Mantel-Haenzsel test is known to perform well even with sparse data, so the fact that the EECV trials had many low-recruiting centres was not an issue. The one drawback of the Mantel-Haenzsel test was that it does not allow for variation in treatment effect across centres. Based on the nature of the intervention in the EECV trials, a non-pharmacologic, hands-on procedure, a variation in treatment effect between centres is plausible.

The use of fixed-effect regression models to account for centre effects in the EECV trials had unique issues. Fixed-effect regression models are known to perform better when there are few centres with many patients at each centre. Small centres are problematic because centres containing only patients of one treatment group or experiencing only one type of outcome event need to be removed. If the statistical software program does not automatically drop such centres, warning messages of extremely large standard errors indicating unreliability of the results will appear. Due to this issue, the fixed effects regression was not expected to perform well with the EECV trial data. However, as one source comments, patients at small centres provide little information to the overall treatment effect, so removing them often does not cause a problem. (20) This was confirmed by the results of the fixed effect model which, when applied to the EECV trial data, showed effect estimates and confidence intervals similar to the Mantel-Haenzsel test. In fact, the fixed-effects regression performed better than expected, showing that the removal of the lowest recruiting centres was not harmful to the process.

In the next model, a treatment-by-centre interaction term was added to the fixed-effects model to account for possible heterogeneity of treatment effects across centers. When zero-counts occurred within the interaction term at a centre, these centres had to be dropped in addition to the centres already removed from the first fixed-effects analysis. This issue was particularly problematic for the outcome of preterm birth due to the low event rate. Only 14 centres representing 37 percent of the data (646/1764) could be included in the model and the resulting confidence intervals are the widest of any method. One way to avoid the problem of zero counts is to add a small positive constant such as 0.05 to all cells or to all empty cells. This procedure is not positively regarded by many statisticians as the constant chosen could change the treatment effects, (20) and was not attempted in this analysis.

A further consideration for the fixed-effect model with the treatment-by-centre interaction term was the issue of weighting by centre size. The interaction term is a reflection of the treatment effect at each centre. Each centre's treatment effect can be equally weighted and averaged to provide an average centre treatment effect, or the centre's treatment effect can be weighted according to the number of patients at each centre. Weighting by centre size ensures that the amount of information provided by each center is reflected in the overall treatment effect; not weighting allows small centers (with a correspondingly less precise treatment effect) to add too much information to the final result. In cases where some centers are very small, using the information from that centre can actually increase the variation around the estimate. (18) In fact, this is exactly what the results of the analysis indicate. The results of the un-weighted model have the widest confidence

intervals and the most different effect estimates than any other model. Given that there were many small recruiting centres and few large recruiting centres participating in the EECV trials, weighting by centre size is a logical step and the preferred method for this analysis. Running the model both ways (weighted and un-weighted for centre size) illustrates the effect the variation of small centres can add to the overall result.

The random-effects regressions performed well given the characteristics of the EECV trial data. The results of the random intercept model, where the random-intercept is the centre, were similar to the Mantel-Haenzsel test. The second random-effects regression model included random intercept and random slope, where the random slope represents the treatment effect. Incorporating a different slope for each centre allows for effect modification, or variation in treatment effect at each centre. The results differ slightly from the random-intercept model in that the confidence intervals are wider. Including random-slope in the model acknowledges differences in treatment effect at centre, and this increased variance is reflected in the standard errors and CI calculations.

The generalized estimating equation was the one unconditional method applied to the dataset. Generalized estimating equations use matrices to estimate correlation within centre that are inherently weighted to centre size, so this works well with the EECV trial data where centre size varies widely. Generalized estimating equations are particularly suited to trials with many centres (at least 20-30), so the EECV trials easily meet this criterion. The results given by the generalized estimating equation model are in line with the conditional methods applied and show the most similarity to the random-effects models.

Since this secondary analysis is an empirical study and not a simulation study, it is impossible to choose one statistical method as the best method to account for centre effects. However, as stated in the research objective, one can evaluate the robustness of the results by assessing the impact of the methods on the magnitude, direction and significance of the effect estimate. The results of the analyses that account for centre effects are in line with the unadjusted results, and increase the robustness of the results.

Given the characteristics of the EECV Trials, the Mantel-Haenzsel test and the random-effects regressions worked well with the dataset and would be good choices for future research with similar datasets.

4.3 Considerations for future research

This secondary analysis adds to the accumulating knowledge about centre effect in multicentre RCTs. As researchers move forward, trial protocols ought to include plans to account for centre effects in the trial design and statistical analysis. Centre effect should even be accounted for in sample size calculation as the magnitude of centre effect can change the sample size required. (39)

Anticipating the type of statistical analysis that fits with the dataset will depend on the number of recruiting centres (a small number of centres will suit a fixed-effect method, whereas a large number will fit with the random-effects and GEE methods). Tactics to promote higher recruitment at trial centres or implementing minimum contribution numbers at each centre for trial involvement may help to avoid the pitfalls associated with small centre size.

The statistical method of choice should also be informed by the likelihood that a variation in treatment effect exists between centres. Five of the seven models described in this study do not take this heterogeneity of treatment effect into account, and the one that addresses the differences in treatment effect most directly, the fixed-effect regression with a treatment-by-centre interaction term, requires high recruiting centres.

5.0 CONCLUSION

A secondary analysis of the EECV trials was undertaken to account for centre effect in light of current recommendations that multicentre RCTs consider the centre effect in design and analysis. Seven statistical models to account for centre were applied to three outcomes of interest from the EECV trials. The results of each model were compared to the unadjusted analysis and to each other.

When centre was taken into account, trends identified in the EECV2 Trial were strengthened. The results indicate that women who receive early ECV are more likely to a have a cephalic presentation at the time of delivery, and may be less likely to have a Caesarean section. Their risk of preterm birth, however, is increased.

With the exception of one model (fixed effect regression with an interaction term un-weighted for centre size), the statistical models performed well with the dataset, and provided similar treatment effects and confidence intervals. The Mantel-Haenzsel test and the random effects regressions had advantages over the other models for this application. The results contribute to a growing body of evidence about the consequences of centre effect in RCTs, and provide a concrete example of how to account for centre in the analysis.

REFERENCES

(1) Oxorn H. Breech Presentation. Human Labor & Birth. 5 ed. East Norwalk, CN: Appleton & Lange; 1986. p. 221-67.

(2) Hill L. Prevalence of breech presentation by gestational age. American Journal of Perinatology 1990;7:92-3.

(3) Krebs L. Breech at term: early and late consequences of mode of delivery. Danish Medical Bulletin 2005;52(4):234-52.

(4) Hutton E, Reitsma A. A comprehensive review of the research literature on external cephalic version (ECV). Canadian Journal of Midwifery Research and Practice 2008;7(1):4-16.

(5) Penn Z, Ghaem-Maghami S. Indications for caesarean section. Best Practice & Research Clinical Obstetrics & Gynaecology 2001;15(1):1-15.

(6) Liu S, Liston R, Joseph K, Heaman M, Sauve R, Kramer M. Maternal mortality and severe morbidity associated with low-risk planned cesarean delivery versus planned vaginal delivery at term. Canadian Medical Association Journal 2007;176(4):455-60.

(7) Hofmeyr G, Kulier R. External cephalic version for breech presentation at term. Cochrane Database of Systematic Reviews 1996;(1):Art. No.: CD000083.

(8) Grootscholten K, Kok M, Oei G, Mol B, van der Post J. External cephalic version - related risks: a meta analysis. Obstetrics & Gynecology 2008;112(5):1143-51.

(9) Teoh T. Effect of learning curve on the outcome of external cephalic version. Singapore Medical Journal 1997;38(8):323-5.

(10) Zhang J, Bowes WJr, Fortney J. Efficacy of external cephalic version: a review. Obstetrics and Gynecology 1993;82(2):306-12.

(11) Hutton E. Early external cephalic version 2 trial protocol. 2005.

(12) Hannah M, Hannah W, Hewson S, Hodnett E, Saigal S, Willan A. Planned caesarean section versus planned vaginal birth for breech presentation at term: a randomised multicentre trial. Term Breech Trial Collaborative Group. Lancet 2000;356(9239):1375-83.

(13) Hogle K, Kilburn L, Hewson S, Gafni A, Wall R, Hannah M. Impact of the international term breech trial on clinical practice and concerns: a survey of centre collaborators. J Obstet Gynaecol Can 2003;25(1):14-6.

(14) Rietberg C, Elferink-Stinkens P, Visser G. The effect of the Term Breech Trial on medical intervention behaviour and neonatal outcome in The Netherlands: an analysis of 35,453 term breech infants. BJOG: An International Journal of Obstetrics and Gynaecology 2005;112:205-9.

(15) Hutton EK, Kaufman K, Hodnett E, Amankwah K, Hewson S, McKay D, et al. External cephalic version beginning at 34 weeks' gestation versus 37 weeks' gestation: A randomized multicenter trial. American Journal of Obstetrics and Gynecology 2003;189:245-54.

(16) Hutton E, Hannah M, Ross S, Delisle M-F, Carson G, Windrim R, et al. The Early External Cephalic Version (ECV) 2 Trial: an international multicentre randomised controlled trial of timing of ECV for breech pregnancies. BJOG: An International Journal of Obstetrics and Gynaecology 2011;118(5):564-77.

(17) EECV2 Trial Collaborator's Meeting. 2008 May 5; Hamilton, ON 2008.

(18) Senn S. Some controversies in planning and analysing multi-centre trials. Statistics in Medicine 1998;17(1753):1765.

(19) Localio AR, Berlin JA, Ten Have TR, Kimmel SE. Adjustments for center in multicenter studies: an overview. Annals of Internal Medicine 2001;135(2):112-23.

(20) Agresti A, Hartzel J. Tutorial in biostatistics: strategies for comparing treatments on a binary response with multi-centre data. Statistics in Medicine 2000;19:1115-39.

(21) Biau DJ, Porcher R, Boutron I. The account for provider and center effects in multicenter interventional and surgical randomized controlled trials is in need of improvement: a review. Journal of Clinical Epidemiology 2008;61:435-9.

(22) Petrinco M, Pagano E, Desideri A, Bigi R, Ghidina M, Ferrando A, et al. Information on centre characteristics as costs' determinants in multicenter clinical trials: is modeling center effect worth the effort? Value in Health 2009;12(2):325-30.

(23) Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P, . Extending the CONSORT Statement to Randomized Trials of Nonpharmacologic Treatment: Explanation and Elaboration. Annals of Internal Medicine 2008;148:295-309.

(24) Chan A, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. Lancet 2005;365:1159-62.

(25) Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Annals of Internal Medicine 2001;134:663-94.

(26) Biau DJ, Halm JA, Ahmadieh H, Capello WN, Jeekel J, outron I, et al. Provider and centre effect in multicenter randomized controlled trials of surgical specialties: an analysis on patient-level data. Annals of Surgery 2008;247(5):892-8.

(27) Tangri N, Kitsios GD, Su SH, Kent DM. Accounting for center effects in multicenter trials. Epidemiology 2010;21(6):912-3.

(28) Sackett D. The tactics of performing therapeutic trials. In: Haynes R, Sackett D, Guyatt G, Tugwell P, editors. Clinical Epidemiology. 3 ed. Philadelphia: Lippincott Williams & Wilkins; 2006. p. 66-172.

(29) Kernan W, Viscoli C, Makuah R, Brass L, Horwitz R. Stratified randomization for clinical trials. Journal of Clinical Epidemiology 1999;52(1):19-26.

(30) Thabane L. 2012. Personal Communication.

(31) Berlin JA, Kimmel SE, Ten Have TR, Sammel M. An empirical comparison of several clustered data approaches under confounding due to cluster effects in the analysis of complications of coronary angioplasty. Biometrics 1999;55(2):470-6.

(32) Gould AL. Multi-centre trial analysis revisited. Statistics in Medicine 1998;17(1779):1797.

(33) Hanley J, Negassa A, Edwardes M, Forrester J. Statistical analysis of correlated data using generalized estimating equations: an orientation. American Journal of Epidemiology 2003;157(4):364-75.

(34) Hardin J, Anderson B, Woodby L, Crawford M, Russell T. Using an empirical binomial hierarchical Bayesian model as an alternative to analyzing data from multisite studies. Evaluation Review 2011;32(2):143-56.

(35) Lee K, Thompson S. The use of random effects models to allow for clustering in individually randomized trials. Clinical Trials 2005;2:163-73.

(36) Zhang J, Boos D. Mantel-Haenszel test statistic for correlated binary data. Biometrics 1997;53(4):1185-98.

(37) Chu R, Thabane L, Ma J, Holbrook A, Pullenayegum E, Devereaux P. Comparing methods to estimate treatment effects on a continuous outcome in multicentre randomized controlled trials: a simulation study. BMC Medical Research Methodology 2011;11(21):1-15.

(38) Yanagawa T, Yoshinori F, Mastuoka J. Generalized Mantel-Haenszel procedures for 2 x J tables. Environmental Health Perspectives 1994;102(Suppl 8):57-60.

(39) Vierron E, Giraudeau B. Sample size calculation for multicenter randomized trial: taking the center effect into account. Contemporary Clinical Trials 2007;28(4):451-8.

Appendix 1: Code for Mantel-Haenzsel method

Run in "R"

Code:

```
dataset<-read.csv("cmh_data_mar22_78_final.csv")
summary(dataset)
cmh.data2<-array(c(dataset$cmh), dim=c(2,2,78))
cmh.data2
mantelhaen.test(cmh.data2)
```

Appendix 2: Code for Logistic Regression

Fixed effects (SAS)

Caesarean section	**data** data2a; set angela.logistic_june_3; if centre not in (**28**, **32**, **43**, **44**, **46**, **52**, **58**, **60**, **76**, **79**); /* 20 groups in 10 centres*/ **run**; **proc genmod** data=data2a; class centre (ref='1')/param=ref ; model cs/tot = group centre / dist = bin link = logit; **run**;
Preterm birth	**data** data2a; set angela.logistic_june_3; if centre not in (**7**, **10**, **11**, **13**, **15**, **16**, **17**, **18**, **19**, **22**, **23**, **27**, **30**, **32**, **33**, **34**, **36**, **37**, **40**, **42**, **44**, **46**, **48**, **51**, **52**, **60**, **65**, **67**, **68**, **74**, **75**, **76**, **78**, **79**); /* 34 centres with large SEs*/ if centre not in (**32**, **43**, **44**, **46**, **52**, **60**, **76**, **79**); /*additional centre with zero patient in one arm*/ **run**; **proc genmod** data=data2a; class centre; model ptb/tot = group centre / dist = bin link = logit; **run**;
Non-cephalic presentation at delivery	**data** data2a; set angela.logistic_june_3; if centre not in (**28**, **32**, **43**, **44**, **46**, **52**, **58**, **60**, **76**, **79**); /* 20 groups in 10 centres*/ **run**; **proc genmod** data=data2a; class centre; model bpd/tot = group centre / dist = bin link = logit; **run**;

Fixed effects with interaction term (SAS)

Caesarean section

```
data data3;
set angela.logistic_june_3;
if centre not in (4,15,17,18,19,22,23,26, 28, 32, 33,
43,46, 47, 48, 51,52, 58, 60, 65,76,78,79,44 );
run;

*1/57 = 0.01754386;
proc genmod data=data3;
 class centre (ref='1')/param=ref ;
 model cs/tot = group centre group*centre / dist = bin
                          link = logit covb;
 estimate 'logOR(c1)' group 1 /e exp;
 estimate 'logOR(c2)' group 1
                                    group*centre   1 /e
exp;
 estimate 'logOR(unwt ave)' group 1
            group*centre   0.01754386 0.01754386
0.01754386 0.01754386 0.01754386 0.01754386 0.01754386
0.01754386 0.01754386 0.01754386

                         0.01754386 0.01754386 0.01754386
0.01754386 0.01754386 0.01754386 0.01754386 0.01754386
0.01754386 0.01754386

                         0.01754386 0.01754386 0.01754386
0.01754386 0.01754386 0.01754386 0.01754386 0.01754386
0.01754386 0.01754386

                         0.01754386 0.01754386 0.01754386
0.01754386 0.01754386 0.01754386 0.01754386 0.01754386
0.01754386 0.01754386

                         0.01754386 0.01754386 0.01754386
0.01754386 0.01754386 0.01754386 0.01754386 0.01754386
0.01754386 0.01754386

                         0.01754386 0.01754386 0.01754386
0.01754386 0.01754386 0.01754386/e exp;
estimate 'logOR(wt ave)' group 1
group*centre   0.032628399 0.024169184 0.025377644
0.017522659 0.005438066 0.004229607 0.042296073
0.007250755 0.007250755 0.012084592
                0.004229607 0.005438066 0.013293051
0.016918429 0.002416918 0.014501511 0.010876133
0.004229607 0.009667674 0.009667674
                0.003021148 0.003021148 0.022960725
0.011480363 0.014501511 0.006646526 0.070694864
```

	0.03141994 0.035649547 0.007854985 0.007250755 0.020543807 0.003021148 0.016314199 0.01570997 0.007854985 0.016314199 0.006646526 0.003625378 0.023564955 0.051963746 0.064048338 0.013293051 0.03081571 0.006646526 0.009063444 0.021148036 0.028398792 0.035649547 0.033836858 0.016918429 0.005438066 0.019939577 0.014501511 0.018731118 0.018731118/e exp; **run**;

Preterm birth	**data** data2a;
	set angela.logistic_june_3;
	if centre not in (7, 10, 11, 13, 15, 16, 17, 18, 19, 22, 23,
	27, 30, 32, 33, 34, 36, 37, 40, 42, 44, 46, 48, 51, 52, 60, 65,67, 68, 74, 75, 76, 78, 79); /* 34 centres with large SEs*/
	if centre not in (32, 43, 44, 46, 52, 60, 76, 79);
	/*additional centre with zero patient in one arm*/
	run;
	
	data data3;
	set data2a;
	if centre not in (3, 4, 5, 6, 8, 12, 14, 20, 24, 25, 26, 28, 29, 31, 35, 38, 41, 47, 50, 53, 57, 58, 59, 61, 64, 69, 70, 71, 72, 73, 80, 81);
	run;
	*Only 14 centres left;
	*1/14=0.071428571;
	proc genmod data=data3;
	class centre (ref='1')/param=ref ;
	model ptb/tot = group centre group*centre / dist = bin
	link = logit covb;
	estimate 'logOR(c1)' group 1 /e exp;
	estimate 'logOR(c2)' group 1
	group*centre 1 /e exp;
	estimate 'logOR(unwt ave)' group 1
	group*centre 0.071428571 0.071428571 0.071428571
	0.071428571 0.071428571 0.071428571 0.071428571
	0.071428571 0.071428571 0.071428571
	
	0.071428571 0.071428571
	0.071428571/e exp;
	estimate 'logOR(wt ave)' group 1
	group*centre 0.083591331 0.108359133 0.00619195
	0.181114551 0.018575851 0.052631579 0.040247678
	0.020123839 0.041795666
	
	0.133126935
	0.164086687 0.078947368 0.037151703/e exp;

<table>
<tr><td></td><td><pre>run;</pre></td></tr>
<tr><td>Non-cephalic presentation at delivery</td><td><pre>
data data2a;
set angela.logistic_june_3;
if centre not in (28, 32, 43, 44, 46, 52, 58, 60, 76,
79); /* 20 groups in 10 centres*/
run;

data data3;
set data2a;
if centre not in (4, 15, 17, 18, 19, 22, 23, 33, 47, 51,
59, 65, 68, 78);
run;

*Only 57 centres left;
*1/57=0.01754386;
proc genmod data=data3;
 class centre (ref='1')/param=ref ;
 model bpd/tot = group centre group*centre / dist = bin
                       link = logit covb;
 estimate 'logOR(c1)' group 1 /e exp;
 estimate 'logOR(c2)' group 1
                                  group*centre   1 /e
exp;
estimate 'logOR(unwt ave)' group 1
            group*centre   0.01754386 0.01754386
0.01754386 0.01754386 0.01754386 0.01754386 0.01754386
0.01754386 0.01754386 0.01754386

                       0.01754386 0.01754386 0.01754386
0.01754386 0.01754386 0.01754386 0.01754386 0.01754386
0.01754386 0.01754386

                       0.01754386 0.01754386 0.01754386
0.01754386 0.01754386 0.01754386 0.01754386 0.01754386
0.01754386 0.01754386

                       0.01754386 0.01754386 0.01754386
0.01754386 0.01754386 0.01754386 0.01754386 0.01754386
0.01754386 0.01754386

                       0.01754386 0.01754386 0.01754386
0.01754386 0.01754386 0.01754386 0.01754386 0.01754386
0.01754386 0.01754386

                       0.01754386 0.01754386 0.01754386
0.01754386 0.01754386 0.01754386/e exp;
estimate 'logOR(wt ave)' group 1
group*centre   0.032747119 0.024257126 0.025469982
0.017586416 0.005457853 0.004244997 0.04244997
0.007277138 0.007277138 0.012128563
0.004244997 0.005457853 0.013341419 0.016979988
0.002425713 0.014554275 0.010915706 0.005457853
</pre></td></tr>
</table>

	0.004244997 0.00970285 0.00970285 0.003032141 0.003032141 0.023044269 0.011522135 0.014554275 0.00667071 0.070952092 0.031534263 0.03577926 0.007883566 0.007277138 0.003638569 0.020618557 0.003032141 0.01637356 0.015767132 0.007883566 0.01637356 0.00667071 0.023650697 0.05215282 0.064281383 0.013341419 0.030927835 0.00667071 0.021224985 0.028502122 0.03577926 0.033959976 0.016979988 0.005457853 0.020012129 0.014554275 0.018799272 0.018799272/e exp; run;

Random effects (R)

Caesarean section	m1a <- lmer(cs ~ group + (1|centre), data=data.frame(cs.ind), family=binomial)
Preterm birth	m1a <- lmer(ptb ~ group + (1|centre), data=data.frame(ptb.ind), family=binomial)
Non-cephalic presentation at delivery	m1a <- lmer(bpd ~ group + (1|centre), data=data.frame(bpd.ind), family=binomial)

Random intercept and random slope (R)

Caesarean section	m2a <- lmer(cs ~ group + (group|centre), data=data.frame(cs.ind), family=binomial)
Preterm birth	m2a <- lmer(ptb ~ group + (group|centre), data=data.frame(ptb.ind), family=binomial)
Non-cephalic presentation at delivery	m2a <- lmer(bpd ~ group + (group|centre), data=data.frame(bpd.ind), family=binomial)

Appendix 3: Code for generalized estimating equation

These three datasets contain individual patient-level binary outcome data (1 or 0) for outcomes: Caesarean section, preterm birth, and non-cephalic at time of birth

Run in SAS

Caesarean section	`proc genmod data=angela.cs_ind descending;`
	` class centre;`
	` model cs = group / dist = bin`
	`  link = logit;`
	` repeated subject=centre/type= exch corrw;`
	`run;`

Preterm birth	`proc genmod data=angela.ptb_ind descending;`
	` class centre;`
	` model ptb = group / dist = bin`
	`  link = logit;`
	` repeated subject=centre/type= exch corrw;`
	`run;`

Non-cephalic at the time of delivery	`proc genmod data=angela.bpd_ind descending;`
	` class centre;`
	` model bpd = group / dist = bin`
	`  link = logit;`
	` repeated subject=centre/type= exch corrw;`
	`run;`