Faculty of Health Sciences

School of Public Health

Statistical Modelling of Breastfeeding Data

Jian Zhao

This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University

August 2018

Declaration

To the best of my knowledge and belief, this thesis titled "Statistical modelling of breastfeeding data" contains no material previously published by any other person, except where due acknowledgement has been made. This thesis does not contain material which has been accepted for the award of any other degree or diploma in any university.

This thesis contains five published papers and one manuscript prepared for submission. The statement of primary contribution of the first author and the permission to include the publications in this thesis are presented in the Appendix A. Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.

Human Ethics: The research presented and reported in this thesis was conducted in accordance with the National Health and Medical Research Council National Statement on Ethical Conduct in Human Research (2007) – updated March 2014. The proposed research studies received human research ethics approval from the Curtin University Human Research Ethics Committee (approval number: RDHS-101-15).

Signature: ________________

Jian Zhao

Date: 16/08/2018

Abstract

Background: The benefits of breastfeeding to both infant and maternal health are widely recognised. Breastfeeding health promotion is critical for maternal and child health. Observational studies investigating factors associated with breastfeeding outcomes such as breastfeeding rate and breastfeeding duration are widely implemented in the literature. Various study designs, including cross-sectional, case-control and cohort studies, have been used in observational breastfeeding studies. Among these study designs, prospective cohort, which is able to satisfy the 'temporality criterion' for causality compared to ecological or case-control studies and has less ethical issues, is most commonly used in breastfeeding research.

Breastfeeding data can be sourced as aggregated data from publications or reports and individual participant data (IPD) from original studies. However, in the statistical modelling of breastfeeding data, some key methodological problems still remain unresolved. For example, few studies have accounted for time-dependent exposures in longitudinal analysis of breastfeeding data. Time-dependent bias due to misclassification of time-varying exposures remains in survival analysis of time-to-event breastfeeding outcomes. In the current literature, intra-cluster correlation is often ignored when modelling correlated breastfeeding data. In addition, there is no statistical approach available for modelling clustered time-to-event breastfeeding data with clumping at zero.

Aims: This doctoral thesis aimed to address these key methodological problems in the statistical modelling of breastfeeding data with following specific objectives:

Objective 1: Systematically review and analyse aggregated data from publications to assess the impact of caesarean section and maternal education on breastfeeding practices in China;

Objective 2: Address the time-varying exposure issue and apply a generalized linear mixed model with time-dependent exposures for modelling longitudinal breastfeeding data;

Objective 3: Address time-dependent bias issue in analysing time-to-event breastfeeding data and provide a statistical solution to avoid the bias;

Objective 4: Address the issue of heterogeneity between clustering units in correlated breastfeeding data and develop a novel two-part mixed-effects model to deal with the augmented zero problem in analysing clustered time-to-event breastfeeding data with an extension of the two-part mixed-effects model to accommodate heterogeneity and correlated random effects.

Methods: This thesis developed or used different statistical models for modelling diverse types of breastfeeding data to achieve various research objectives.

Method for Objective 1: Aggregated data from publications were collected and analysed to examine the association between caesarean section and breastfeeding practices postpartum, and the association between maternal education and breastfeeding practices postpartum, in China using systematic review and meta-analysis techniques.

Method for Objective 2: Time-varying exposure analysis was carried out by pooling longitudinal breastfeeding data from two breastfeeding cohort studies in Sichuan Province, China to examine the association between breastfeeding and calcium supplementation postpartum by incorporating time-varying exposures in a generalised linear mixed model (GLMM).

Method for Objective 3: With the use of the time-to-event breastfeeding data collected in Sichuan Province, China between 2010 and 2011 as an example, time-dependent bias introduced by ignoring or misclassification of time-varying exposure in paediatric time-to-event outcomes was illustrated, and the solution, an extended Cox model, to avoid such bias was given.

Method for Objective 4: Data collected from a Nepal breastfeeding cohort study in 2014 were utilised to illustrate two different frailty modelling approaches to account for heterogeneity between clustering units. Furthermore, the same data were used to illustrate how to analyse clustered time-to-event breastfeeding data with clumping at zero using the proposed two-part mixed-effects model.

Results:

Results (Objective 1): Both caesarean section and higher maternal education level were identified as detrimental factors associated with breastfeeding practices postpartum in China. Compared with vaginal birth, the likelihood of mothers exclusively breastfeeding their babies during the early postpartum period was reduced by 47% and 39% reduction was found for the odds of breastfeeding at 4 months postpartum after caesarean section. The odds of breastfeeding of mothers who had ‘more than six years’ or ‘more than twelve years’ education was reduced by 10% or 9% respectively, compared to mothers obtaining ‘six years or less’ or ‘twelve years or less’ education.

Results (Objective 2): Revealed by the GLMM, which took into account of both inherent correlations among repeated measurements and time-varying exposures, the pooled likelihood of taking calcium supplementation postpartum among breastfeeding mothers was 4.02 times (95% Confidence Interval (CI): 2.30, 7.03) higher than that of their non-breastfeeding counterparts.

Results (Objective 3): In the illustrative example, based on the extended Cox model incorporating time-varying covariates, the effect of ‘solid foods introduction’ (Hazard Ratio (HR) 0.61, 95% CI 0.50, 0.75) on breastfeeding cessation from the time-fixed analysis was reversed (HR 1.50, 95% CI 1.17, 1.93). The non-significant association between ‘maternal return to work’ and breastfeeding cessation (HR 0.99, 95% CI 0.73, 1.36) from the time-fixed analysis became significant (HR 1.45, 95% CI 1.06, 2.00). The findings suggested the time-fixed analysis may produce biased estimates with a smaller magnitude.

Results (Objective 4): Using the semiparametric frailty approach, the random effects representing the variation in the hazard of breastfeeding cessation shared by mothers living in the 27 distinct communities were estimated and graphically presented. Compared with the conventional Cox model, Cox frailty model reduced the chance of type I error occurring and provided a better model fit in the presence of correlated survival data. Among candidate parametric frailty approaches, a Weibull proportional hazards (PH) model with a gamma frailty term was chosen to be used for modelling the correlated breastfeeding duration data from the cohort study in Nepal. Furthermore, when considering the augmented zeros existing in the same clustered time-to-event breastfeeding data, the developed two-part mixed-effects model was applied. A significantly positive correlation between the two parts was confirmed, suggesting the baseline prevalence of exclusive breastfeeding (EBF) and

EBF duration were correlated. This correlated two-part mixed-effects model outperformed the independent two-part model with a smaller -2 log-likelihood given that there was a correlation between the two processes.

Conclusion: This thesis found that both caesarean section and higher maternal education had negative impacts on breastfeeding practices postpartum in China. Relevant health policy for improving breastfeeding outcomes in China should take these detrimental factors into consideration. We addressed some key methodological problems with statistical modelling of breastfeeding data. Specifically, to avoid time-dependent bias and to achieve more precise and reliable statistical inference results, this thesis confirmed that the time-varying exposures should be accounted for in the statistical modelling of both time-to-event breastfeeding data and longitudinal breastfeeding. We further proposed a useful approach for modelling clustered time-to-event breastfeeding data with clumping at zero. This approach makes full use of all available information at baseline and during follow-up, and takes into account not only the correlation within each clustering unit but also the correlation between the baseline process and the follow-up process. This approach has a potential to be generalised to other types of health related data in the presence of many zeros.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AF	Any breastfeeding
AFT	Accelerated failure time
AI	Adequate intake
AIC	Akaike information criterion
BCG	Bacillus Calmette-Guerin
CI	Confidence interval
CNKI	China National Knowledge Infrastructure
D+L	DerSimonian and Laird
EAR	Estimated average requirement
EBF	Exclusive breastfeeding
FB	Full breastfeeding
GEE	Generalised estimating equations
GLM	Generalised linear models
GLMM	Generalised linear mixed model
GMM	Generalised method of moments
HR	Hazard Ratio
IPD	Individual participant data
I-V	Inverse variance
MeSH	Medical Subject Headings

ML	Maximum likelihood
MM	Method of moments
MOOSE	Meta-analysis Of Observational Studies in Epidemiology
OR	Odds Ratio
PH	Proportional hazards
PQL	Penalised quasi-likelihood
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT	Randomised control trials
REML	Restricted maximum likelihood
SE	Standard error
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
WHO	World Health Organization
ZINB	Zero-inflated negative binomial
ZIP	Zero-inflated Poisson

Acknowledgements

I would like to express my sincere gratitude and appreciation to my supervisors: Dr Yun Zhao, Professor Andy Lee and Professor Colin Binns. I still remember the moment when we first met each other in Chengdu, China. They provided me such a terrific opportunity to get scientific training in Curtin University, Australia. Throughout my PhD study, they have invested a considerable amount of time and efforts on my PhD research. I am always appreciative of the guidance of Dr Yun Zhao and Professor Andy Lee on statistical methodology and practical implications. I could never have gotten here without you. I am also forever grateful to Professor Colin Binns, who has been inspiring me with his commitment to applying good science to public health research and ultimately to improving the health and wellbeing of our citizens.

I would like to thank others who have helped me complete this thesis through discussions and suggestions. Thanks to Dr. Li Tang for kindly sharing her research experience on breastfeeding and providing me the Jiangyou cohort data. Thanks to Dr. Chuan Yu for providing me the Chengdu cohort data. Thanks to Mengran for helping search and retrieve Chinese breastfeeding literature for systematic review and meta-analysis in the thesis. I would also like to acknowledge Professor Liming Xiang (School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore) for discussing with me on the development of the two-part model. Thanks to Dr. Vishnu Khanal for providing the illustrative cohort data.

I am grateful to the faculty of the School of Public Health, Curtin University. I have improved my knowledge on public health and broadened my horizon through seminars and workshops. Also, I am grateful to fellow PhD students in the HDR Hub for their friendship and insights.

Last but not least, I am eternally grateful to my parents who have constantly encouraged and supported me for the last thirty-one years and to my wife for her infinite love.

Dedication

To my parents,

and to my beloved wife.

Chapter 1: Introduction

1.1 Background

1.1.1 Breastfeeding benefits and status

Breastfeeding is the normal way for feeding infants, and it is widely recognised to offer numerous health benefits over breastmilk substitutes (Fulhan et al., 2003; Lawrence, 1994). Breastmilk provides sufficient nutrients that infants need for the first six months of their life, and the composition of breastmilk varies over time to meet the changing nutrition needs of infants growing in different periods after birth (Binns, 2003; Lawrence, 1994).

Breastfeeding is not only the optimal nutrition resource for infants, but also plays a protective role against infant and maternal adverse health outcomes. Compared to non-breastfed infants, breastfed infants have been found to be less likely to develop adverse health outcomes, such as mortality and morbidity caused by infectious diseases (Labbok et al., 2004; Stuebe & Schwarz, 2010), allergic diseases (Gijsbers et al., 2005; Obihara et al., 2005), hyperlipidaemia (Owen et al., 2008; Owen et al., 2002), hypertension (Martin et al., 2005; Owen et al., 2003; Stuebe et al., 2011), and type 2 diabetes (Owen et al., 2006; Pettitt et al., 1997; Young et al., 2002) in their childhood or adulthood. For mothers, breastfeeding has been confirmed to be able to reduce postpartum haemorrhage (Diaz et al., 1988), risk of postmenopausal hip fracture (Bjornerem et al., 2011) and risk of breast cancer (Collaborative Group on Hormonal Factors in Breast, 2002; do Carmo Franca-Botelho et al., 2012) and ovarian cancer (Li et al., 2014; Su et al., 2013).

Given the extraordinary range of benefits from breastfeeding, the World Health Organization (WHO) recommends that infants should be exclusively breastfed up to 6 months of age followed by continued breastfeeding while complementary foods are introduced (World Health Organization, 2018). Unfortunately, global breastfeeding levels fall short of the WHO recommendations, especially the exclusive breastfeeding rate and duration, which are at a relatively low level worldwide (Centers for Disease Control and Prevention, 2007; Labbok et al., 2006; Victora et al., 2016). It was reported that in most high-income countries, the prevalence rate of breastfeeding is lower than 20% (Victora et al., 2016).

1.1.2 Gaps in knowledge

To understand the substantial impacts of factors on breastfeeding outcomes, studies investigating factors associated with breastfeeding rate and duration are widely conducted and reported in the literature (Garcia et al., 2016; Mangrio et al., 2017; Scott & Binns, 1999). Most researchers adopt quantitative methods to design their studies and guide their data collection followed by a series of statistical analyses. A variety of epidemiological designs including cross-sectional, case-control, prospective/retrospective cohort, and randomised control trials (RCT) have been used in breastfeeding studies. Among these designs, the prospective cohort study, which is able to satisfy the ‘temporality criterion’ for causality compared to ecological or case-control studies and has less ethical issues, is often chosen by breastfeeding researchers. (Rothman et al., 2008).

However, in the statistical analysis of breastfeeding data, including aggregated data from publications and individual participant data (IPD) from original research, there are some key methodological problems and issues remaining unsolved.

In the analysis of aggregated breastfeeding data, meta-analysis has been used to synthesise evidence from publications and quantify weighted overall effect size between exposures (or interventions) and breastfeeding rate or duration in the literature (Buccini et al., 2017; Lau et al., 2016; Schiff et al., 2014). However, inconsistent, even controversial, findings/conclusions were drawn from studies of Western and Asian countries, due to diverse cultural and economic settings between these countries. For example, in China, higher maternal education level was found having an inverse association with breastfeeding practices after birth in a number of studies (Huang et al., 2012; Kang et al., 2013; Liu et al., 2013; Qiu, 2008; Ye et al., 2007). However, the opposite conclusion has been widely reported in Western countries (Heck et al., 2006; Kehler et al., 2009; Scott & Binns, 1999). Thus, Western-China differences in factors associated with breastfeeding rate/duration remain unclear.

In the analysis of breastfeeding IPD collected from prospectively designed longitudinal studies, time-varying exposures (also known as time-dependent exposures) are common during the follow-up. Studies on how to deal with time-varying exposures in longitudinal data and time-to-event data have been sparse in the breastfeeding literature. In addition, breastfeeding data are commonly clustered. That is, lower level units (e.g., mothers) are often nested in higher level units (e.g. hospitals or communities), suggesting that those subjects nested in the same clustering units are correlated to each other due to some similar characteristics. Ignoring this inherent (intra-cluster) correlation within the clustering units may lead to an inflated Type I error, and incorrect statistical inferences. However, there are very few published breastfeeding studies taking the intra-cluster correlation into account.

Furthermore, a notable proportion of mothers are reported not to exclusively breastfeed their babies at hospital discharge (Scott et al., 2006; Vila-Candel et al., 2017) so that the breastfeeding durations for these mothers/infants are noted as zeros. Comprising other positive exclusive breastfeeding (EBF) durations, this type of data is known as time-to-event data with clumping at zero. Because standard survival analysis, which is often performed in the followed-up of an EBF duration study, only makes use of positive observations, in the literature most breastfeeding studies simply exclude mothers/infants with zero duration from their survival analysis. Consequently, the baseline information of these mothers/infants is lost in the standard survival analysis.

A series of updated and rigorous statistical models or approaches for modelling breastfeeding data with taking into account for the above methodological issues is crucial and beneficial to researchers working in the field of infant nutrition and similar areas.

1.2 Aim of this thesis

The aim of this thesis is to address some key methodological problems in the analysis of breastfeeding data.

This thesis has the following specific objectives:

Objective 1: Systematically review and analyse aggregated data from publications to assess the impact of caesarean section and maternal education on breastfeeding practices in China;

Objective 2: Address the time-varying exposure issue and apply a generalized linear mixed model with time-dependent exposures for modelling longitudinal breastfeeding data;

Objective 3: Address time-dependent bias issue in analysing time-to-event breastfeeding data and provide a statistical solution to avoid the bias;

Objective 4: Address the issue of heterogeneity between clustering units in correlated breastfeeding data and develop a novel two-part mixed-effects model to deal with the augmented zero problem in analysing clustered time-to-event breastfeeding data with an extension of the two-part mixed-effects model to accommodate heterogeneity and correlated random effects.

1.3 Significance of the thesis

The two systematic review and meta-analysis studies, namely, on the association between caesarean section and breastfeeding practices and the association between maternal education and breastfeeding practices, gained the knowledge about the relationship between these two important factors and breastfeeding practices under Chinese culture and settings, and cleared the picture about the differences in some degree between Western countries and China. Findings from these studies could provide further evidence to the Chinese government and health administration for future decision making to inform targeted interventions aimed at improving breastfeeding promotion.

Time-varying exposures are common in longitudinal data and time-to-event data. However, the application of time-varying exposures modelling is relatively sparse in the breastfeeding literature, and researchers often ignore the time-varying nature of exposures measured during follow-up in statistical analysis, which may result in incorrect statistical inference and conclusions. It is, therefore, important to account for time-varying exposures in longitudinal and time-to-event breastfeeding data analysis. An example on the importance of accounting for time-varying exposure in longitudinal data analysis can be found in our study on the association between calcium supplementation postpartum and breastfeeding (Zhao et al., 2016). In this study, we treated the breastfeeding status at different time points postpartum as time-varying exposure and incorporated it into a generalised linear mixed model (GLMM). A dynamic effect size with an enhanced strength of statistical inference was derived from the time-varying modelling. In the study of analysing time-to-event data with time-varying exposures, time-dependent bias was illustrated and discussed along with a solution, an extended Cox regression, to avoid it. Approaches used and illustrated in both longitudinal data and time-to-event data can be implemented in future breastfeeding studies.

The proposed novel two-part mixed-effects model for analysing clustered time-to-event data with clumping at zero makes full use of all available information at the baseline and during the follow-up. The proposed approach provides less biased parameter estimation and contributes to the body of knowledge in this area as well. In addition to breastfeeding research, the proposed approach can be applied to other public health research areas.

1.4 Structure and contents of this thesis

This thesis contains five published papers and one manuscript prepared for submission. In Chapter 2, we examine the association between caesarean section and breastfeeding practices in China based on the evidence from a systematic review and meta-analysis. In Chapter 3, we assess the association between maternal education and breastfeeding practices in China using a systematic review and meta-analysis approach as well. In Chapter 4, we examine the calcium supplementation status and identify whether breastfeeding is associated with increased calcium supplementation among Chinese mothers after child birth with incorporating time-varying exposures into GLMM. In Chapter 5, we illustrate time-dependent bias in conventional survival analysis due to ignoring time-varying exposures in time-fixed analysis and provide a statistical solution, an extend Cox model, to avoid such bias. In Chapter 6, we address the heterogeneity issue in the analysis of clustered time-to-event breastfeeding data and propose a two-part mixed-effects model for analysing clustered time-to-event data with clumping at zero and apply this approach to breastfeeding data from a maternal cohort study in Nepal as an illustrative example.

In Chapter 7, we synthesise the results of all the studies and provide an overall discussion, including our conclusions on statistical modelling in breastfeeding studies.

1.5 References

Binns, C. (2003). 4 Encourage and support breastfeeding. *Dietary Guidelines for Australian Adults*, 227.

Bjornerem, A., Ahmed, L.A., Jorgensen, L., Stormer, J., & Joakimsen, R.M. (2011). Breastfeeding protects against hip fracture in postmenopausal women: the Tromso study. *J Bone Miner Res, 26*(12), 2843-2850.

Buccini, G.D.S., Perez-Escamilla, R., Paulino, L.M., Araujo, C.L., & Venancio, S.I. (2017). Pacifier use and interruption of exclusive breastfeeding: Systematic review and meta-analysis. *Matern Child Nutr, 13*(3)

Centers for Disease Control and Prevention. (2007). Breastfeeding trends and updated national health objectives for exclusive breastfeeding--United States, birth years 2000-2004. *MMWR Morb Mortal Wkly Rep, 56*(30), 760-763.

Collaborative Group on Hormonal Factors in Breast, C. (2002). Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease. *Lancet, 360*(9328), 187-195.

Diaz, S., Rodriguez, G., Marshall, G., del Pino, G., Casado, M.E., Miranda, P., et al. (1988). Breastfeeding pattern and the duration of lactational amenorrhea in urban Chilean women. *Contraception, 38*(1), 37-51.

do Carmo Franca-Botelho, A., Ferreira, M.C., Franca, J.L., Franca, E.L., & Honorio-Franca, A.C. (2012). Breastfeeding and its relationship with reduction of breast cancer: a review. *Asian Pac J Cancer Prev, 13*(11), 5327-5332.

Fulhan, J., Collier, S., & Duggan, C. (2003). Update on pediatric nutrition: breastfeeding, infant nutrition, and growth. *Curr Opin Pediatr, 15*(3), 323-332.

Garcia, A.H., Voortman, T., Baena, C.P., Chowdhurry, R., Muka, T., Jaspers, L., et al. (2016). Maternal weight status, diet, and supplement use as determinants of breastfeeding and complementary feeding: a systematic review and meta-analysis. *Nutr Rev, 74*(8), 490-516.

Gijsbers, B., Mesters, I., Andre Knottnerus, J., Legtenberg, A.H., & van Schayck, C.P. (2005). Factors influencing breastfeeding practices and postponement of solid food to prevent allergic disease in high-risk children: results from an explorative study. *Patient Educ Couns, 57*(1), 15-21.

Heck, K.E., Braveman, P., Cubbin, C., Chavez, G.F., & Kiely, J.L. (2006). Socioeconomic status and breastfeeding initiation among California mothers. *Public Health Rep, 121*(1), 51-59.

Huang, H.T., Liang, H., Hu, J.W., Feng, P., Tong, W.J., Zheng, Y.M., et al. (2012). Breastfeeding and associated factors in five areas of China. *Journal of Soochow University Medical Science Edition, 32*(4), 454-458. Chinese.

Kang, Y., Liang, X.H., Liu, Y.X., Chen, J., & Li, T.Y. (2013). Relative factors of breastfeeding cessation in infants before 6-month-old *J Third Mil Med Univ, 35*(16), 1735-1739. Chinese.

Kehler, H.L., Chaput, K.H., & Tough, S.C. (2009). Risk factors for cessation of breastfeeding prior to six months postpartum among a community sample of women in Calgary, Alberta. *Can J Public Health, 100*(5), 376-380.

Labbok, M.H., Clark, D., & Goldman, A.S. (2004). Breastfeeding: maintaining an irreplaceable immunological resource. *Nat Rev Immunol, 4*(7), 565-572.

Labbok, M.H., Wardlaw, T., Blanc, A., Clark, D., & Terreri, N. (2006). Trends in exclusive breastfeeding: findings from the 1990s. *J Hum Lact, 22*(3), 272-276.

Lau, Y., Htun, T.P., Tam, W.S., & Klainin-Yobas, P. (2016). Efficacy of e-technologies in improving breastfeeding outcomes among perinatal women: a meta-analysis. *Matern Child Nutr, 12*(3), 381-401.

Lawrence, P.B. (1994). Breast milk. Best source of nutrition for term and preterm infants. *Pediatr Clin North Am, 41*(5), 925-941.

Li, D.P., Du, C., Zhang, Z.M., Li, G.X., Yu, Z.F., Wang, X., et al. (2014). Breastfeeding and ovarian cancer risk: a systematic review and meta-analysis of 40 epidemiological studies. *Asian Pac J Cancer Prev, 15*(12), 4829-4837.

Liu, J., Shi, Z., Spatz, D., Loh, R., Sun, G., & Grisso, J. (2013). Social and demographic determinants for breastfeeding in a rural, suburban and city area of South East China. *Contemporary Nurse: A Journal for the Australian Nursing Profession, 45*(2), 234-243 210p.

Mangrio, E., Persson, K., & Bramhagen, A.C. (2017). Sociodemographic, physical, mental and social factors in the cessation of breastfeeding before 6 months: a systematic review. *Scand J Caring Sci,*

Martin, R.M., Gunnell, D., & Smith, G.D. (2005). Breastfeeding in infancy and blood pressure in later life: systematic review and meta-analysis. *Am J Epidemiol, 161*(1), 15-26.

Obihara, C.C., Marais, B.J., Gie, R.P., Potter, P., Bateman, E.D., Lombard, C.J., et al. (2005). The association of prolonged breastfeeding and allergic disease in poor urban children. *Eur Respir J, 25*(6), 970-977.

Owen, C.G., Martin, R.M., Whincup, P.H., Smith, G.D., & Cook, D.G. (2006). Does breastfeeding influence risk of type 2 diabetes in later life? A quantitative analysis of published evidence. *Am J Clin Nutr, 84*(5), 1043-1054.

Owen, C.G., Whincup, P.H., Gilg, J.A., & Cook, D.G. (2003). Effect of breast feeding in infancy on blood pressure in later life: systematic review and meta-analysis. *BMJ, 327*(7425), 1189-1195.

Owen, C.G., Whincup, P.H., Kaye, S.J., Martin, R.M., Davey Smith, G., Cook, D.G., et al. (2008). Does initial breastfeeding lead to lower blood cholesterol in adult life? A quantitative review of the evidence. *Am J Clin Nutr, 88*(2), 305-314.

Owen, C.G., Whincup, P.H., Odoki, K., Gilg, J.A., & Cook, D.G. (2002). Infant feeding and blood cholesterol: a study in adolescents and a systematic review. *Pediatrics, 110*(3), 597-608.

Pettitt, D.J., Forman, M.R., Hanson, R.L., Knowler, W.C., & Bennett, P.H. (1997). Breastfeeding and incidence of non-insulin-dependent diabetes mellitus in Pima Indians. *Lancet, 350*(9072), 166-168.

Qiu, L. (2008). *Factors that influence breastfeeding initiation and duration in urban, suburban and rural areas of Zhejiang Province, Peoples Republic of China.* (Ph.D thesis). Curtin University, Perth.

Rothman, K.J., Greenland, S., & Lash, T.L. (2008). *Modern epidemiology*. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins.

Schiff, M., Algert, C.S., Ampt, A., Sywak, M.S., & Roberts, C.L. (2014). The impact of cosmetic breast implants on breastfeeding: a systematic review and meta-analysis. *Int Breastfeed J, 9*, 17.

Scott, J.A., & Binns, C.W. (1999). Factors associated with the initiation and duration of breastfeeding: a review of the literature. *Breastfeed Rev, 7*(1), 5-16.

Scott, J.A., Binns, C.W., Oddy, W.H., & Graham, K.I. (2006). Predictors of breastfeeding duration: evidence from a cohort study. *Pediatrics, 117*(4), e646-655.

Stuebe, A.M., & Schwarz, E.B. (2010). The risks and benefits of infant feeding practices for women and their children. *J Perinatol, 30*(3), 155-162.

Stuebe, A.M., Schwarz, E.B., Grewen, K., Rich-Edwards, J.W., Michels, K.B., Foster, E.M., et al. (2011). Duration of lactation and incidence of maternal hypertension: a longitudinal cohort study. *Am J Epidemiol, 174*(10), 1147-1158.

Su, D., Pasalich, M., Lee, A.H., & Binns, C.W. (2013). Ovarian cancer risk is reduced by prolonged lactation: a case-control study in southern China. *The American journal of clinical nutrition, 97*(2), 354-359.

Victora, C.G., Bahl, R., Barros, A.J., Franca, G.V., Horton, S., Krasevec, J., et al. (2016). Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect. *Lancet, 387*(10017), 475-490.

Vila-Candel, R., Duke, K., Soriano-Vidal, F.J., & Castro-Sanchez, E. (2017). Effect of Early Skin-to-Skin Mother-Infant Contact in the Maintenance of Exclusive Breastfeeding. *J Hum Lact*, 890334416676469.

World Health Organization. (2018). Breastfeeding Retrieved from http://www.who.int/topics/breastfeeding/en/

Ye, J.L., Cui, Y., & Pan, X.P. (2007). Study of the breast feeding and the influence factors in Xinjiang rural areas. *Maternal and Child Health Care of China, 22*(17), 2420-2422.Chinese.

Young, T.K., Martens, P.J., Taback, S.P., Sellers, E.A., Dean, H.J., Cheang, M., et al. (2002). Type 2 diabetes mellitus in children: prenatal and early infancy risk factors among native canadians. *Arch Pediatr Adolesc Med, 156*(7), 651-655.

Zhao, J., Zhao, Y., Binns, C.W., & Lee, A.H. (2016). Increased calcium supplementation postpartum is associated with breastfeeding among Chinese mothers: finding from two prospective cohort studies. *Nutrients*, *8*(10), 622.

Chapter 2: Association between caesarean section and breastfeeding practices in China: a systematic review and meta-analysis

The content of this chapter is covered by a published paper "Zhao, J., Zhao, Y., Du, M., Binns, C. W., & Lee, A. H. (2017). Does Caesarean Section Affect Breastfeeding Practices in China? A Systematic Review and Meta-Analysis. Maternal and child health journal, 21(11), 2008-2024. DOI: https://doi.org/10.1007/s10995-017-2369-x". (see Appendix C.2)

The statement of primary contribution of the first author and the permission to include the publication in this thesis can be found in the Appendix A. The permission to reproduce the material from the publisher can be found in the Appendix C.1.

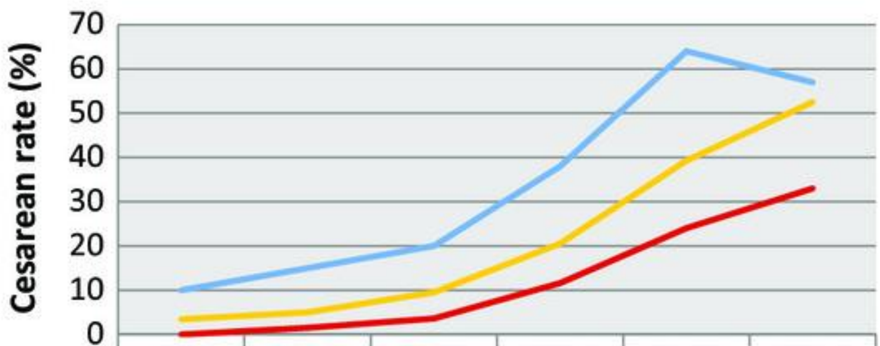
2.1 Caesarean section and breastfeeding

2.1.1 Caesarean section and breastfeeding globally

Given that caesarean section rates continue increasing in most countries during the past decade, for example, 31.9% in the United States (Martin et al., 2018), 45.9% in Brazil (Gibbons et al., 2010), 42% in China (Mi & Liu, 2014), the effect of caesarean section on breastfeeding practices has been studied extensively (Guo et al., 2013; Joshi et al., 2014; Kohlhuber et al., 2008; Patel et al., 2015). However, findings from individual studies are inconsistent in the literature. For example, some studies reported the negative associations between caesarean delivery and breastfeeding outcomes such as breastfeeding initiation, breastfeeding rates at the first week and at 4 months postpartum (Guo et al., 2013; Kohlhuber et al., 2008; Patel et al., 2015), while other studies reported no association between caesarean delivery and breastfeeding rates at 42 days postpartum, and breastfeeding prevalence within 6 months postpartum (Joshi et al., 2014; Patel et al., 2015). A systematic review and meta-analysis on the association between caesarean section and breastfeeding was conducted by Prior et al. (2012) based on English literature searched in PubMed. The review concluded that elective caesarean section was associated with early breastfeeding negatively and no statistically significant association was found with the breastfeeding rate at 6 months postpartum (Prior et al., 2012). However, the majority of breastfeeding studies undertaken in China were published in Chinese and hence excluded from this review (Prior et al., 2012). Some degree of bias may have resulted from this omission.

2.1.2 Caesarean section and breastfeeding in China

The caesarean section rate in China is high at 46.2% in 2007-2008 and 56.1% in 2011 as shown in Figure 2-1, which are nearly double the global rate (Feng et al., 2012; Hellerstein et al., 2015). Understanding the relationship between caesarean section and breastfeeding practices accurately in more depth is important to a country like China, where the caesarean section rate is far beyond the rate of 15% recommended by the World Health Organization (WHO) (World Health Organization, 2015). An updated systematic review and meta-analysis based on both English and Chinese literature would reduce the publication bias and improve the knowledge of the relationship between caesarean delivery and breastfeeding practices in China.



	1988	1993	1998	2003	2008	2010
Overall	3.4	5	9.5	20.5	39.3	52.5
Urban	10	15	20	38	64	57
Rural	0	1.5	3.6	11.6	24	33

Figure 2-1 Temporal trend of caesarean section rate in China (Hellerstein et al., 2015)

2.2 Systematic review and meta-analysis

2.2.1 Systematic review

Search strategy

A systematic electronic search of both Chinese and English language articles on breastfeeding and method of delivery was conducted using the Chinese database China National Knowledge Infrastructure (CNKI) as well as Medline, EMBASE, CINAHL, ProQuest, and Science Direct from January 1990 to June 2015. A two-stage search strategy was adopted following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) and MOOSE (Meta-analysis Of Observational Studies in Epidemiology) guidelines (Moher et al., 2009; Stroup et al., 2000).

Stage 1: The following Medical Subject Headings (MeSH) terms and keywords “breast feeding”, “milk human”, “breastfeeding duration”, “breastfeeding cessation”, “human lactation”, “infant feed*”, “breastfed”, “risk factor*”, “protective factor*”, “determinant*”, “socioeconomic factor*”, “China”, “mainland China”, “Chinese” were used.

Stage 2: The following MeSH terms and keywords “caesarean delivery”, “cesarean delivery”, “caesarean section”, “cesarean section” and “c-section” were further added to the search process.

Corresponding Chinese terms and keywords were used to search in the database of CNKI.

Literature screening and selection criteria

At the initial screening stage, abstracts of eligible publications were retrieved by two independent reviewers (JZ and MRD). Relevant citations were identified after screening the abstracts, and their full-texts were then obtained and evaluated. Discrepancies on relevancy were resolved through consensus or referred to a third investigator (YZ) when necessary. Articles were included if they met the following criteria: (i) published in peer-reviewed journals or theses/dissertations; (ii) observational study design; (iii) reported the association between caesarean delivery and breastfeeding quantitatively; (iv) effect size could be obtained directly or calculated from raw tabulated data. The exclusion criteria were as follows: (i) studies that did not specify sample size; (ii) studies that did not report or define time points of breastfeeding outcomes; (iii) studies that reported inappropriate statistical result (statistical error).

Data extraction

The following information was extracted from each eligible study for qualitative and quantitative synthesis: publication year, name of first author, study design, location of study, sample size, breastfeeding outcomes (including definitions of breastfeeding, types of breastfeeding, time points of measurements), other factors associated with the breastfeeding outcome, and raw tabulated data or effect size (odds ratios) reported via univariate analysis or multivariate analysis. In cases where the relevant results or raw data were missing, the authors were contacted by email.

Odds ratios (ORs), either crude or adjusted, estimated from logistic regression analysis of ‘exclusive breastfeeding during early postpartum period’ and ‘breastfeeding at 4 months postpartum’ for caesarean section versus vaginal delivery and their corresponding 95% confidence intervals (CIs), were extracted from the eligible studies.

The definitions of breastfeeding used in the data extraction followed the World Health Organization (WHO) definitions (World Health Organization, 2003, 2008).

Exclusive breastfeeding: Breastfeeding while giving no other food or liquid, not even water, except drops or syrups consisting of vitamins, mineral supplements or medicines.

Full breastfeeding: Exclusive breastfeeding or predominant breastfeeding (or almost exclusive breastfeeding). Breastmilk is the only source of milk given to the infant regardless of supplementation with other fluids such as water and orange juice.

Any breastfeeding: The child has received breastmilk (direct from the breast or expressed) with or without other drink, formula or other infant food.

Quality assessment

To assess methodological quality of the selected studies, we developed a checklist based on the criteria proposed in Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) (von Elm et al., 2007) and the previous checklist proposed by Tooth et al. (2005). The possible score of our checklist ranges from 0 to 18, with scores above 14, between 11 and 14, and below 11 indicating high, medium and low quality, respectively (see Appendix C.2).

2.2.2 Meta-analysis

The odds ratios of 'exclusive breastfeeding during early postpartum period' (defined as initiation of breastfeeding or exclusive breastfeeding before discharge or exclusive breastfeeding at 42 days post birth) and 'breastfeeding (including exclusive breastfeeding, full breastfeeding and any breastfeeding) at 4 months postpartum', for caesarean section versus vaginal delivery, were the primary outcomes of interest. A meta-analysis was performed to determine the pooled effect size of caesarean delivery on these two breastfeeding outcomes separately. Based on the raw data extraction from the selected studies, the natural logarithmic transformed ORs were used in the meta-analysis.

Fixed-effect (inverse variance (I-V) method of fixed-effect model) meta-analysis was performed initially, and heterogeneity across studies was assessed by the I-square statistic (Higgins & Thompson, 2002; Higgins et al., 2003). A random-effect model (DerSimonian and Laird (D+L) method of random-effect model) was further utilised and presented when the heterogeneity was confirmed statistically significant (DerSimonian & Laird, 1986). A meta-regression was then performed to investigate potential sources contributing to the heterogeneity. Subgroup analysis was undertaken to assess the magnitude of effect on the outcome variables of interest under different stratifications, which included different study designs (cross-sectional, prospective cohort and retrospective cohort), different exclusive breastfeeding time points (initiation, before discharge and at 42 days after birth) and different definitions of breastfeeding (WHO and non-WHO).

To test the dependence of effect size reported in each study, sensitivity analysis using the jackknife approach was performed to assess the robustness of the results (Miller, 1974). Such sensitivity analysis was repeated multiple times with one study removed per cycle.

In order to ascertain publication bias and small sample size bias among the studies, Begg's funnel plot/test and Egger's test were applied (Egger et al., 1997). All calculations and statistical analyses were performed using the Stata package version 14.1 (StataCorp LP, College Station, USA). A p value less than 0.05 was considered as statistically significant.

2.3 Results of the systematic review

As shown in Figure 2-2, a total of 1061 records were identified from both English and Chinese databases, and five more records were obtained from other sources. After removal of duplicates, 1054 articles were screened by manually reading titles and abstracts. As a result, 197 articles were deemed eligible for full-text review. After formal review, 151 of them were excluded (86 qualitative studies, 54 not suitable for analysis, 4 full-text Chinese publications incomplete, 1 sample size not specified, 3 with statistical errors, 3 duplicated the same study). One author was contacted to obtain additional data to calculate the rate of exclusive breastfeeding initiation (Xu, 2008). Finally, 46 articles (38 published in Chinese and 8 published in English, respectively) were included for the qualitative synthesis of systematic review (Table 2-1), among which 13 were suitable for meta-analysis of the association between caesarean delivery and 'exclusive breastfeeding during early postpartum period', and 14 were appropriate for meta-analysis of the effect of caesarean delivery on 'breastfeeding at 4 months postpartum' (Chen et al., 2010; Fang et al., 1996; Gan et al., 2007; Guo et al., 2013; He et al., 1994; Huang & Lu, 2010; Jiang, 2000; Jiang & Li, 2008; Kang et al., 2013; Leng, 2014; Li, 2014; Liu, 2008; Liu & Xing, 1998; Liu et al., 2014; Liu et al., 2012; Liu & Shao, 2009; Ma et al., 2009; Qin & Hua, 2013; Qiu, 2008; Ruan et al., 2012; Tang, 2013; Tang, 2014; Tian et al., 2008; Wang et al., 2006; Wang et al., 2005; Wang, 2010; Wang et al., 2009; Wang et al., 1995; Wang et al., 2013; Wei & Li, 2009; Xu, 2008; Xu & Yu, 2009; Xue et al., 2012; Yang & Feng, 2014; Ye, 2008; Yin et al., 2012; Yu, 2013; Zhang & Wang, 2000; Zhang & Shi, 2013; Zhang & Wu, 1999; Zhang et al., 2006; Zhang, 2012; Zhang et al., 2013; Zheng et al., 2008; Zhu et al., 2013; Zhu et al., 2014). Among these 46 articles, 20 of them were assessed to be high scientific quality, 24 medium quality and two low quality, according to the methodological quality checklist scores in Appendix C.2.

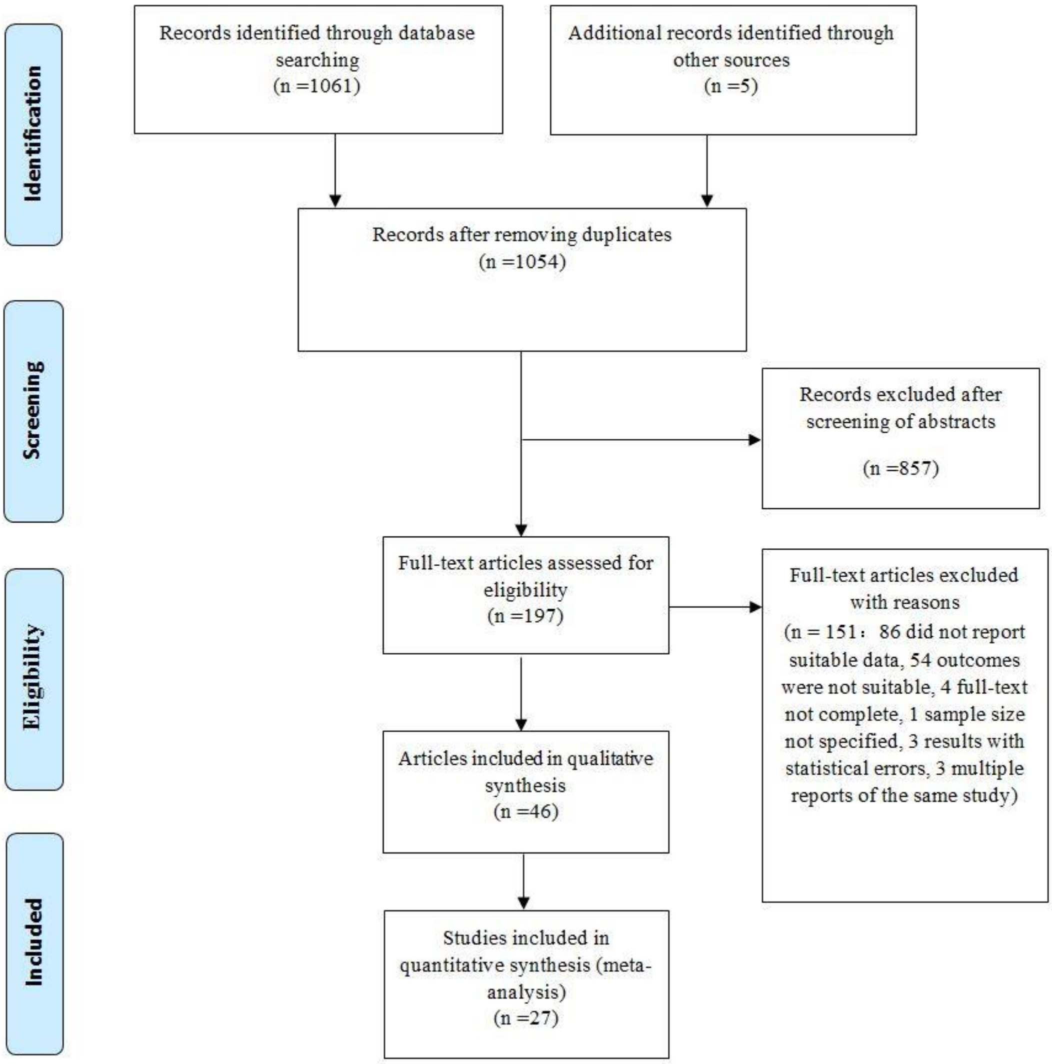


Figure 2-2 PRISMA flow chart of the systematic review process

2.4 Results of the meta-analysis

2.4.1 Effect of caesarean section on 'exclusive breastfeeding during early postpartum period'

The random-effect model meta-analysis of the 13 studies (441,044 subjects: 27,152 in the caesarean delivery group and 413,892 in the vaginal delivery group) showed that the odds of 'exclusive breastfeeding during early postpartum period' was 47%

(pooled OR=0.53, 95% CI: 0.41, 0.68) lower in the caesarean delivery group than the vaginal delivery group (Figure 2-3), with a significant heterogeneity in effect sizes evident across the studies (I^2 =90.6%, p<0.001).

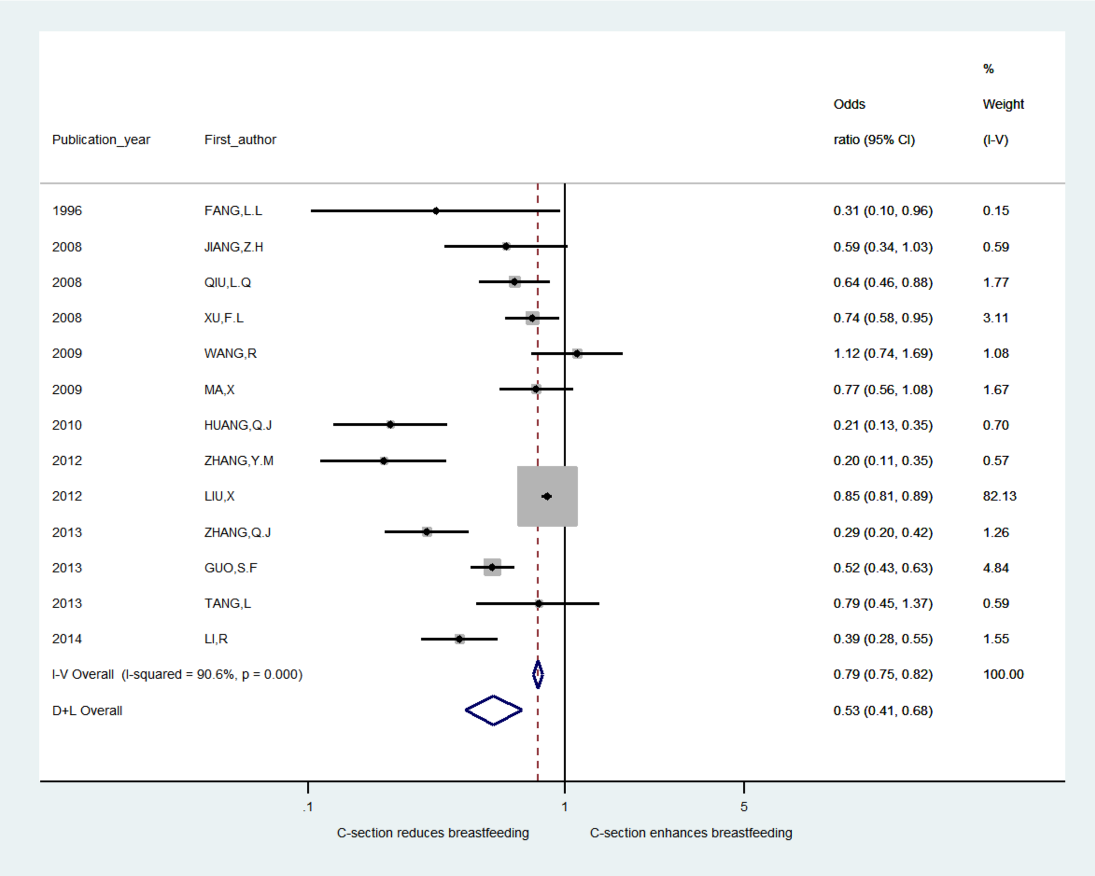


Figure 2-3 Forest plot showing the fixed-effect and random-effect meta-analysis for exclusive breastfeeding during early postpartum period

The stepwise meta-regression analysis revealed no clear main sources for heterogeneity, which may be due to the small number of studies included in this meta-analysis (Higgins & Green, 2011). Exploratory subgroup analyses were further conducted with regard to study design (cross-sectional, prospective cohort and retrospective cohort), exclusive breastfeeding time points (initiation, before discharge and at 42 days after birth) and definitions of breastfeeding (WHO and non-WHO).

Subgroup analysis stratified by study design

Of the 13 studies, six adopted a cross-sectional design, two used retrospective design and five were prospective cohort studies. Considerable heterogeneity remained present in both cross-sectional and prospective cohort subgroups (I^2=86.1%, p<0.001 and I^2=79.9%, p=0.001, respectively) whereas no significant heterogeneity was observed in the retrospective cohort subgroup (I^2=38.5%, p=0.202). Meta-analysis stratified by the study design (Figure 2-4) showed that the adverse effect of caesarean delivery remained on the exclusive breastfeeding prevalence during early postpartum period for the three study designs (cross-sectional: pooled OR=0.42, 95% CI: 0.27, 0.64; retrospective: pooled OR=0.85, 95% CI: 0.81, 0.89; prospective: pooled OR=0.59, 95% CI: 0.41, 0.85).

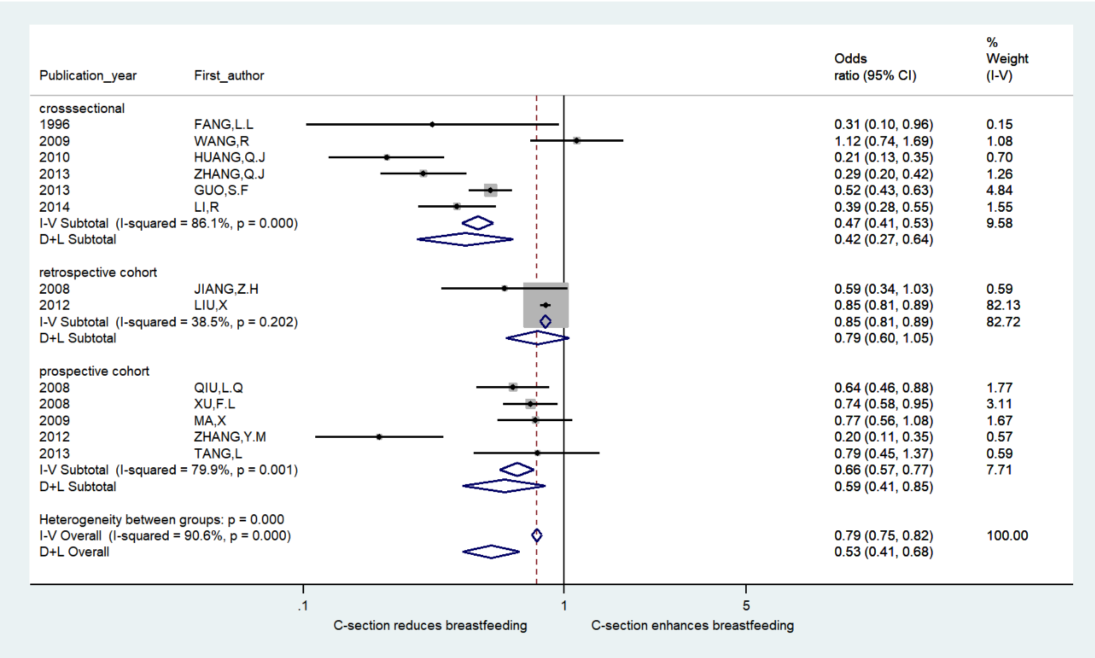


Figure 2-4 Subgroup analysis for exclusive breastfeeding during early postpartum period stratified by study design

Subgroup analysis stratified by exclusive breastfeeding time points

Figure 2-5 presents the subgroup analysis stratified by time points (initiation, before discharge and at 42 days post birth) of exclusive breastfeeding outcomes measured. Again, the negative association between caesarean delivery and the exclusive breastfeeding prevalence during early postpartum period remained persistent (initiation: pooled OR=0.71, 95% CI: 0.55, 0.91; before discharge: pooled OR=0.39, 95% CI: 0.23, 0.67; at 42 days post birth: pooled OR=0.52, 95% CI: 0.32, 0.86). Except the third subgroup (I^2=0.0%, p=0.319), significant heterogeneity was evident for the first two subgroups (initiation: I^2=69.9%, p=0.010 and before discharge: I^2=95.1%, p<0.001, respectively).

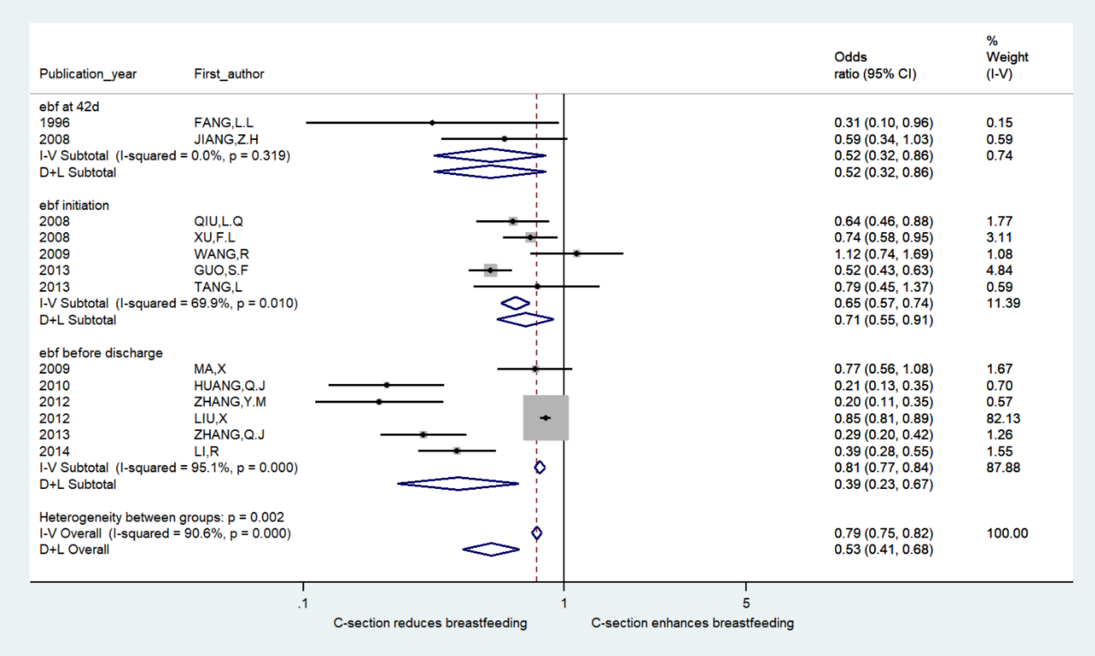


Figure 2-5 Subgroup analysis for exclusive breastfeeding during early postpartum period stratified by time points of breastfeeding outcomes measured

Subgroup analysis stratified by definitions of breastfeeding

Figure 2-6 presents the result of subgroup analysis stratified by the definitions of breastfeeding (WHO and non-WHO). The adverse effect of caesarean delivery on the exclusive breastfeeding prevalence during early postpartum period was confirmed significant for both subgroups (WHO: pooled OR=0.58, 95% CI: 0.43, 0.77; non-WHO: pooled OR=0.48, 95% CI: 0.29, 0.78). Significant heterogeneity was detected as well (WHO: I^2=89.3%, p<0.001; non-WHO: I^2=88.2%, p<0.001).

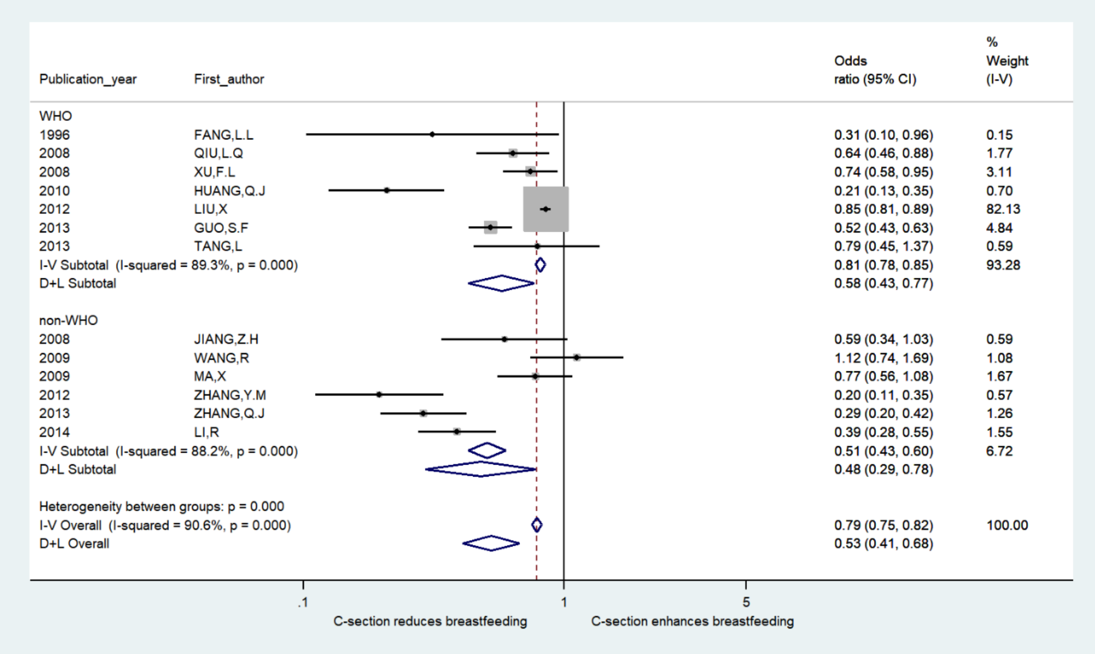


Figure 2-6 Subgroup analysis for exclusive breastfeeding during early postpartum period stratified by definitions

2.4.2 Effect of caesarean section on ‘breastfeeding at four months postpartum’

Fourteen studies were included for this meta-analysis, among which one study (Liu, 2008) had four sub-studies carried out independently in four different periods and consequently these sub-studies were regarded as separate studies in the meta-analysis. Figure 2-7 shows that the prevalence of breastfeeding at 4 months postpartum was significantly lower after caesarean section (pooled OR=0.61, 95% CI: 0.53, 0.71). Significant heterogeneity (I^2=60.6%, p=0.002) was found.

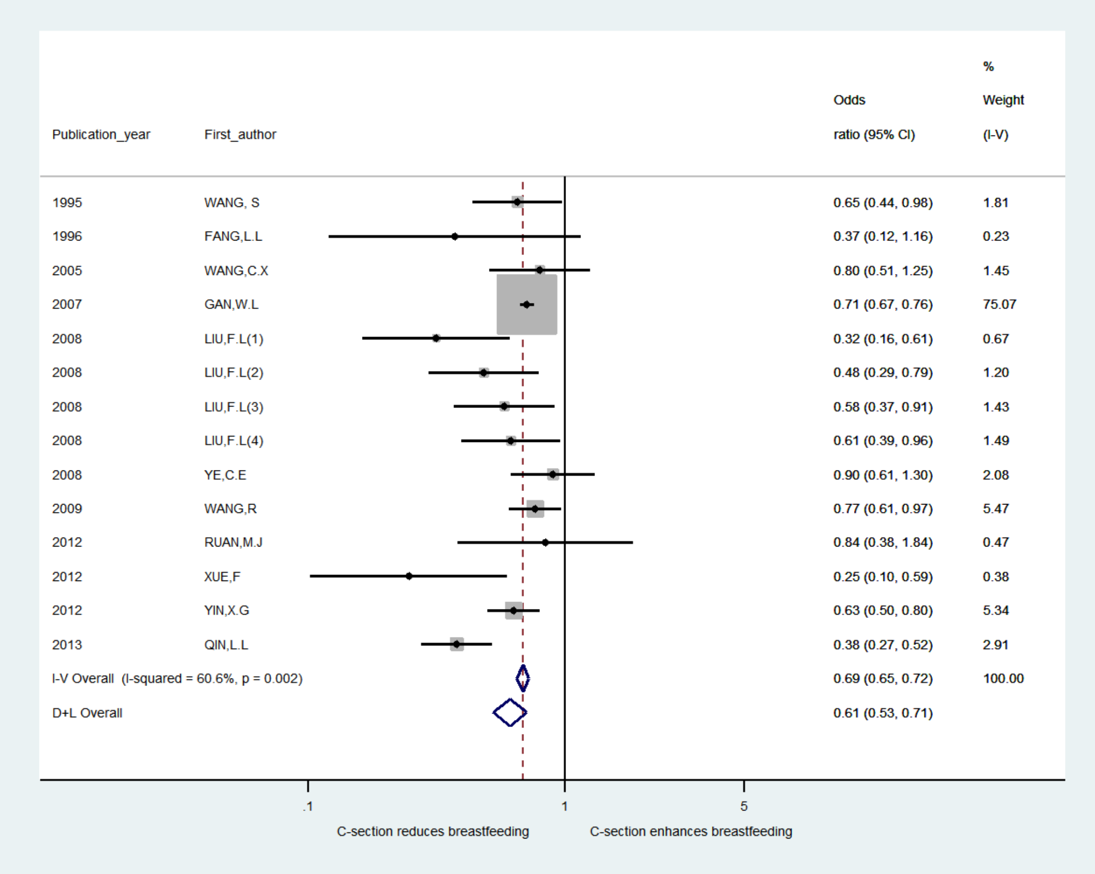


Figure 2-7 Forest plot showing the fixed-effect and random-effect meta-analysis for breastfeeding at 4 months postpartum

Among the 14 studies, four used the WHO definitions. Based on the subgroup analysis stratified by the definitions of breastfeeding, the odds of breastfeeding at 4 months postpartum reduced by 48% (pooled OR=0.52, 95% CI: 0.38, 0.72) when compared with that using non-WHO definitions (pooled OR=0.66, 95% CI: 0.56, 0.77), as indicated in Figure 2-8. Heterogeneity was found marginally significant in both subgroups (WHO: I^2=62.0%, p=0.048; non-WHO: I^2=47.0%, p=0.049).

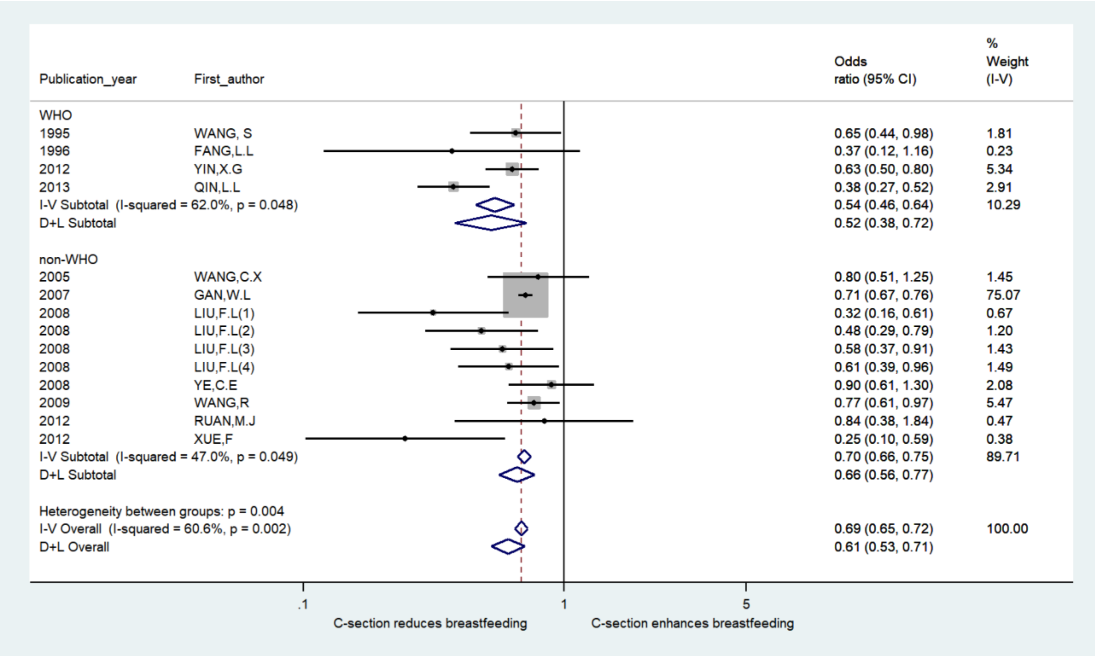


Figure 2-8 Subgroup analysis for breastfeeding at 4 months postpartum stratified by definitions

2.4.3 Sensitivity analysis

The results of sensitivity analysis for both breastfeeding outcomes showed that the pooled ORs remained significant when one study was omitted at a time during recalculation of the pooled ORs (Figure 2-9 and Figure 2-10), suggesting the pooled ORs were not substantially influenced by any individual study so that the results of meta-analysis were robust.

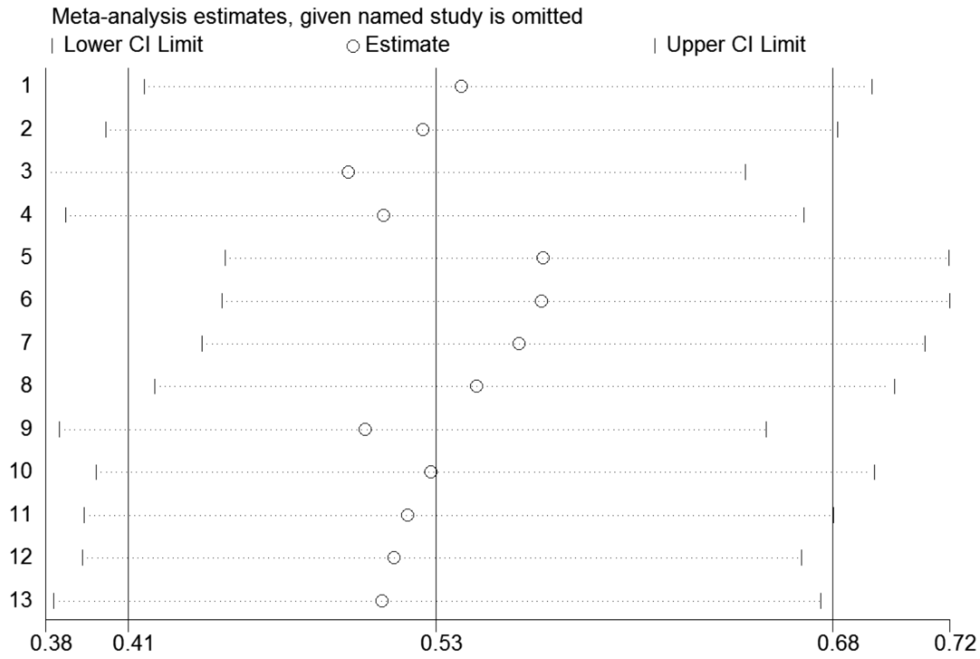


Figure 2-9 Sensitivity analysis for exclusive breastfeeding during early postpartum period

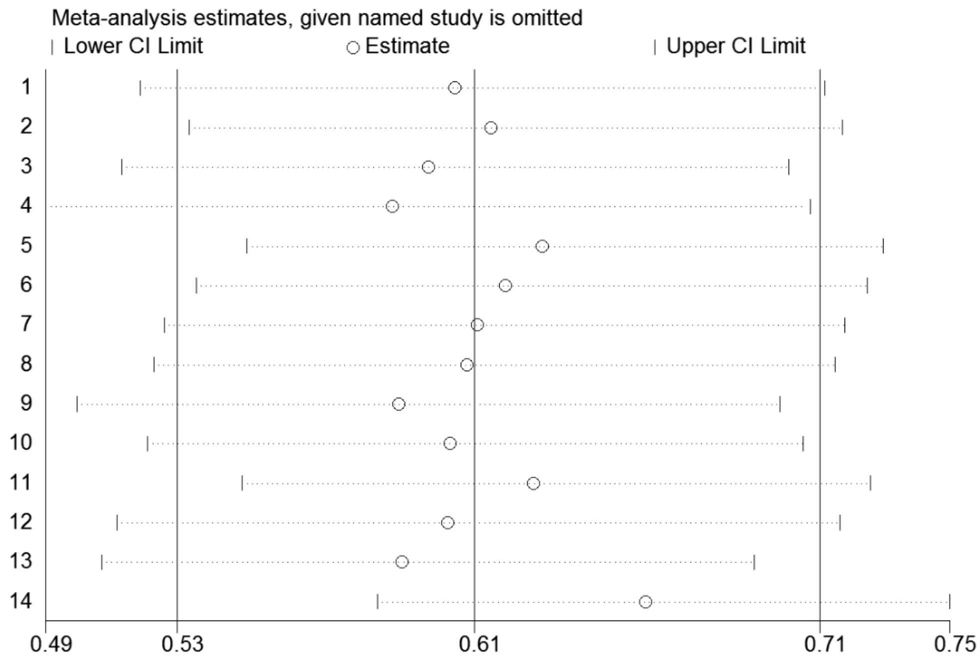


Figure 2-10 Sensitivity analysis for breastfeeding at 4 months postpartum

2.4.4 Publication bias

The distribution of 13 studies involved in the meta-analysis of 'exclusive breastfeeding during early postpartum period' showed in the funnel plot (Figure 2-11) was considered asymmetric, however the Begg's test was not significant (p=0.502), while some evidence of publication bias or small sample size effect was detected by the Egger's test (p=0.007).

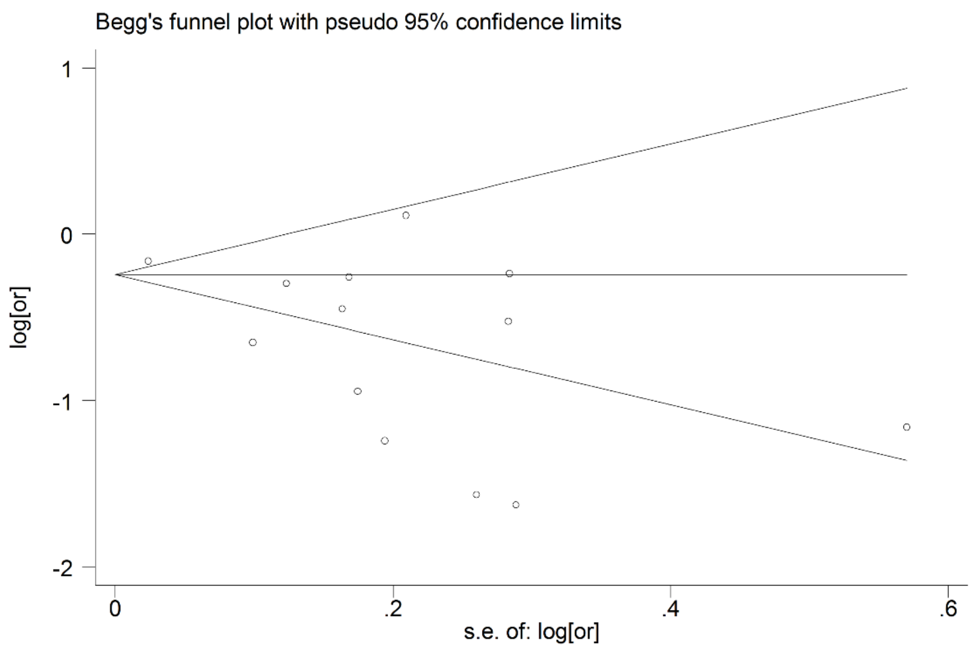


Figure 2-11 Funnel plot exploring publication bias in exclusive breastfeeding during early postpartum period

For the meta-analysis of 'breastfeeding at 4 months postpartum', the funnel plot (Figure 2-12) appeared symmetric (p=0.049 for Begg's test and p=0.059 for Egger's test), suggesting little publication bias or small sample size effect was present.

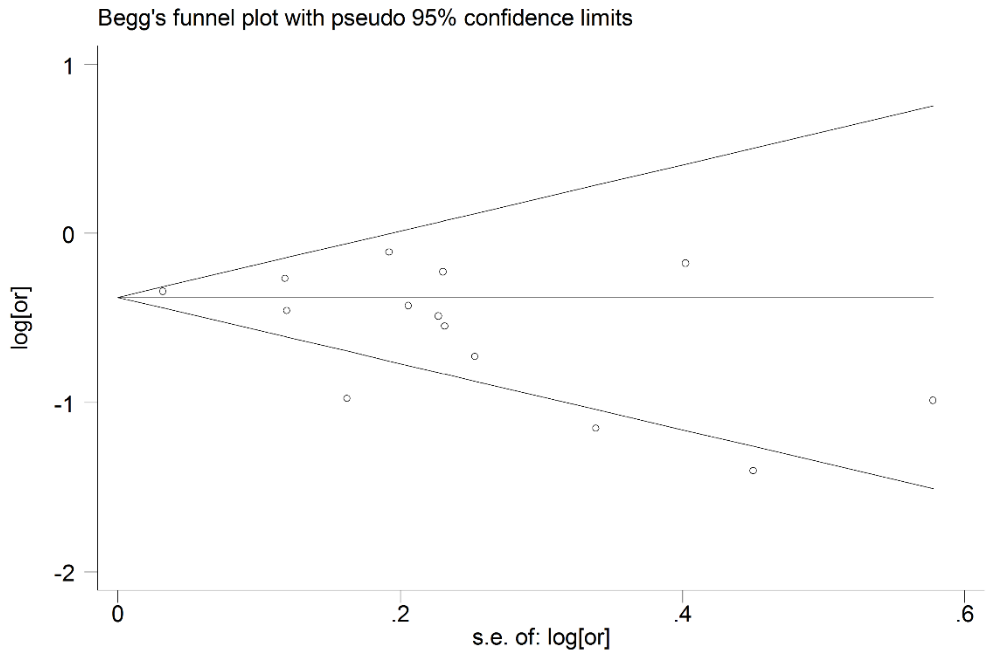


Figure 2-12 Funnel plot exploring publication bias in breastfeeding at 4 months postpartum

2.5 Discussion

The present systematic review incorporated 46 studies to assess the association between caesarean delivery and breastfeeding practices in China. The meta-analysis comprised of 13 studies (441,044 subjects) on ‘exclusive breastfeeding during early postpartum period’ and 14 studies (8,771 subjects) on ‘breastfeeding at 4 months postpartum’ to estimate the pooled relationship between caesarean delivery and breastfeeding outcomes. Based on our findings, caesarean delivery was found to have an adverse effect on the breastfeeding outcomes. More specifically, compared with vaginal birth, the likelihood of mothers exclusively breastfeeding their babies during the early postpartum period was reduced by 47% (pooled OR=0.53) and 39% (pooled OR=0.61) reduction was found for the odds of breastfeeding at 4 months postpartum after caesarean section. Our finding of ‘exclusive breastfeeding during the early postpartum period’ (pooled OR=0.53; 95% CI: 0.41, 0.68) appears to be consistent with a previous review (pooled OR=0.57, 95% CI:0.50, 0.64) (Prior et al., 2012). Our finding of significant impact of caesarean section on ‘breastfeeding at 4 months postpartum’ (pooled OR=0.61, 95% CI: 0.53, 0.71) supports the statistically significant conclusion for any breastfeeding at 6 months postpartum regardless of whether women did or did not initiate breastfeeding (pooled OR=0.91, 95% CI: 0.86, 0.97) reported by Prior et al. (2012). However, it is different from the non-significant result from subgroup analysis restricted to women initiating breastfeeding in that same review (pooled OR=0.95, 95% CI:0.89, 1.01) (Prior et al., 2012). The current knowledge of surgery consequences suggests that postsurgical pain, haemorrhage, infections as well as some hormone issues (like prolactin level postpartum) affect breastfeeding, however these effects lessen in strength over time, indicating the impact of caesarean delivery on breastfeeding may attenuate over time (Chapman & Perez-Escamilla, 1999; Liu et al., 2012; Marcus et al., 2015; Mkontwana & Novikova, 2015; Prior et al., 2012; Wang et al., 2006). Previous studies showed that caesarean section remained an important barrier to early initiation of breastfeeding as well as to the implementation of hospital practices such as delayed skin-to-skin contact between mother-infant pairs (Bramson et al., 2010; Rowe-Murray & Fisher, 2002). Given the negative association between caesarean section and breastfeeding outcomes revealed in the present analysis, it suggests that early initiation of breastfeeding may play a role of mediator variable between caesarean section and premature cessation of exclusive breastfeeding or any breastfeeding.

It is well known that China has a high caesarean section rate where nearly half of the babies born were delivered by caesarean section in 2010 (Hellerstein et al., 2015). The reasons for such a high rate mainly include three important factors: the obstetric care system in China (in hospital births, urbanisation, highly covered New Co-operative Medical Scheme that reduces patient costs and increases revenues for doctors and hospitals for caesarean than for vaginal deliveries, and high volume of deliveries), health care provider factors (insufficient nurses/midwives, some doctors do recommend caesarean section to avoid possible lawsuits in view of the medical malpractice environment), and cultural aspects of patient preference (demand for perfect baby, fear of pain, increasing numbers of macrosomia, increasing pregnancies in older women and delivery date choosing because of luck and belief) (Hellerstein et al., 2015; Long et al., 2012; Mi & Liu, 2014). It was reported that among women with high caesarean section rate of 46.2% in China, 11.7% were cases of caesarean section on maternal request (Lumbiganon et al., 2010). Although the association between caesarean section on maternal request and breastfeeding has been investigated in China (Liu et al., 2012), there is a gap in research on the different impacts of caesarean section on breastfeeding outcomes with respect to medical indication and maternal request (Liu et al., 2015; Tang et al., 2006).

The findings of this study suggest that practices or interventions after caesarean delivery are of benefits to improve breastfeeding behaviours. The practice of maternal-infant skin to skin contact after birth has been increasingly considered as an efficient way to promote breastfeeding status postpartum especially breastfeeding initiation (Moore et al., 2016). The feasibility of an intervention of skin to skin contact after caesarean delivery in the operating room to improve breastfeeding as well as maternal satisfaction and pain perception outcomes was described and evaluated in previous studies (Hung & Berg, 2011; Sundin & Mazac, 2015). However, to our best knowledge, there is no intervention described to improve breastfeeding practice following caesarean section available in the literature. Future intervention studies to promote breastfeeding after caesarean delivery are warranted.

A major strength of this study was the extensive searches conducted in both Chinese and English literature to ensure all relevant articles have been included to reduce reporting bias. Therefore, relative to the previous systematic review work by Prior et al. (2012), this study provided further insights into the association between caesarean section and breastfeeding practice in China, a country with relatively high caesarean section rate. In addition to the associations between caesarean section and breastfeeding practice at two time points, namely, early breastfeeding and 6 months postpartum, which were mainly reported in the previous systematic review (Prior et al., 2012), the study synthesised the effect sizes at 4 months postpartum. Consequently, a temporal trend of the pooled magnitudes can be roughly observed. Furthermore, several studies without clarification of valid time-point for the breastfeeding outcomes had been excluded to enhance the quality of our evaluation. However, publication bias was still detected in the meta-analyses for early postpartum exclusive breastfeeding using the Egger's test (Higgins & Green, 2011). Six months postpartum is the recommended period for exclusive breastfeeding by the WHO, and it would be more comparable with other studies if the impact of caesarean section on 'breastfeeding at 6 months postpartum' could be examined in our study, however most of available reports in the present systematic review comprised the breastfeeding outcomes only for within 6 months postpartum. In view of the small number of studies addressing breastfeeding outcomes at 6 months postpartum, the corresponding quantitative synthesis was not performed. Another limitation concerns that the pooled weighted effect size was estimated based on both crude and adjusted ORs due to the limited number of eligible studies available. Ideally, extraction of data and analysis should be performed for adjusted ORs, as consequence, the impact of caesarean section on breastfeeding practice could be assessed under the controlling for possible confounding effects. Since most of studies retrieved were cross-sectional in this systematic review, a causal relationship between caesarean section and breastfeeding rates could not be concluded. Compared to the previous systematic review (Prior et al., 2012), in which caesarean section by sub-types (elective or emergency) was examined and subgroup analysis by restricting to women who initiated breastfeeding after delivery was performed, unfortunately, these analyses were not performed due to lack of such information available in those studies included in our systematic review.

In conclusion, the present study confirmed that the likelihood of breastfeeding, including 'exclusive breastfeeding during early postpartum period' and 'breastfeeding at 4 months postpartum', was significantly lower after caesarean section in China. Breastfeeding, especially exclusive breastfeeding, provides optimal infant nutrition and benefits both maternal and child health (Victora et al., 2000). Therefore, health policy and measures to improve breastfeeding outcomes should target the reduction of caesarean rate and health intervention after caesarean delivery in China.

2.6 References

Bramson, L., Lee, J.W., Moore, E., Montgomery, S., Neish, C., Bahjri, K., et al. (2010). Effect of early skin-to-skin mother--infant contact during the first 3 hours following birth on exclusive breastfeeding during the maternity hospital stay. *J Hum Lact, 26*(2), 130-137.

Chapman, D.J., & Perez-Escamilla, R. (1999). Identification of risk factors for delayed onset of lactation. *J Am Diet Assoc, 99*(4), 450-454; quiz 455-456.

Chen, Y.F., Ouyang, Y.Q., & Wu, S. (2010). Current situation of breastfeeding of healthcare workers in Wuhan and its countermeasures. *Maternal and Child Health Care of China, 25*(27), 3921-3923. Chinese.

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Control Clin Trials, 7*(3), 177-188.

Egger, M., Smith, D.G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*(7109), 629-634.

Fang, L.L., Zheng, X.X., & Guo, A.B. (1996). Survey on breastfeeding related factors. *Med J Chin Civil Admin 8*(2), 98-101. Chinese.

Feng, X.L., Xu, L., Guo, Y., & Ronsmans, C. (2012). Factors influencing rising caesarean section rates in China between 1988 and 2008. *Bull World Health Organ, 90*(1), 30-39, 39A.

Gan, W.L., Zhao, Y., & Wang, J.L. (2007). Breastfeeding associated factors analysis for infants 0-4 months postpartum in urban area of Chongqing *Maternal and Child Health Care of China, 22*(20), 2805-2806. Chinese.

Gibbons, L., Belizán, J.M., Lauer, J.A., Betrán, A.P., Merialdi, M., & Althabe, F. (2010). The global numbers and costs of additionally needed and unnecessary caesarean sections performed per year: overuse as a barrier to universal coverage. *World health report, 30,* 1-31.

Guo, S., Fu, X., Scherpbier, R.W., Wang, Y., Zhou, H., Wang, X., et al. (2013). Breastfeeding rates in central and western China in 2010: implications for child and population health. *Bull World Health Organ, 91*(5), 322-331.

He, H.L., Wang, M., Zhong, Y.Y., Xu, X.X., & Xu, G.Y. (1994). Breastfeeding status 1 month postpartum and associated factors analysis. *Maternal and Child Health Care of China, 9*(4), 23-24. Chinese.

Hellerstein, S., Feldman, S., & Duan, T. (2015). China's 50% caesarean delivery rate: is it too high? *BJOG, 122*(2), 160-164.

Higgins, J.P., & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions* (Vol. 4): John Wiley & Sons.

Higgins, J.P., & Thompson, S.G. (2002). Quantifying heterogeneity in a meta-analysis. *Stat Med, 21*(11), 1539-1558.

Higgins, J.P., Thompson, S.G., Deeks, J.J., & Altman, D.G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*(7414), 557-560.

Huang, Q.J., & Lu, X. (2010). Investigation on breastfeeding status in baby-friendly hospital in Pudong New Area. *Maternal and Child Health Care of China, 25*(9), 1242-1244. Chinese.

Hung, K.J., & Berg, O. (2011). Early skin-to-skin after cesarean to improve breastfeeding. *MCN Am J Matern Child Nurs, 36*(5), 318-324; quiz 325-316.

Jiang, G.F. (2000). Associated factors analysis for breastfeeding *Jiangsu Prev Med, 11*(4), 49-50. Chinese.

Jiang, Z.H., & Li, S. (2008). Survey on factors influencing breastfeeding of neonates. *Maternal and Child Health Care of China, 23*(24), 3495-3496. Chinese.

Joshi, P.C., Angdembe, M.R., Das, S.K., Ahmed, S., Faruque, A.S.G., & Ahmed, T. (2014). Prevalence of exclusive breastfeeding and associated factors among mothers in rural Bangladesh: a cross-sectional study. *Int Breastfeed J, 9,* 7.

Kang, Y., Liang, X.H., Liu, Y.X., Chen, J., & Li, T.Y. (2013). Relative factors of breastfeeding cessation in infants before 6-month-old *J Third Mil Med Univ, 35*(16), 1735-1739. Chinese.

Kohlhuber, M., Rebhan, B., Schwegler, U., Koletzko, B., & Fromme, H. (2008). Breastfeeding rates and duration in Germany: a Bavarian cohort study. *Br J Nutr, 99*(5), 1127-1132.

Leng, X.L. (2014). Investigation of exclusive breastfeeding rate for infants under 6 months *China Medical Engineering 22*(7), 172-173. Chinese.

Li, R. (2014). Research of factors affecting breastfeeding in baby-area and countermeasures of nursing intervention. *J Gannan Med Univ, 34*(3), 453-455. Chinese.

Liu, F.L. (2008). Survey on status of breastfeeding at 4 months postpartum from 2000 to 2006 in the firtst hospital of people's Xinxiang City *Chin J Misdiagn, 8*(19), 4777-4778. Chinese.

Liu, J., & Xing, E.Z. (1998). Associated factors analysis for breastfeeding *J Baotou Med College, 14*(4), 29-30. Chinese.

Liu, L.F., Hu, J.Y., Jin, H.J., & Lei, J. (2014). Investigation and related factors analysis of breastfeeding of infants under 6 months in Lishui. *Chin J Child Health Care, 22*(10), 1092-1095. Chinese.

Liu, X., Landon, M.B., Cheng, W., & Chen, Y. (2015). Cesarean delivery on maternal request in China: what are the risks and benefits? *Am J Obstet Gynecol, 212*(6), 817 e811-819.

Liu, X., Zhang, J., Liu, Y., Li, Y., & Li, Z. (2012). The association between cesarean delivery on maternal request and method of newborn feeding in China. *PLoS One, 7*(5), e37336.

Liu, X.Q., & Shao, H.M. (2009). Investigation on status of postpartum breastfeeding and its influencing factors *Occup and Health 25*(24), 2746-2748. Chinese.

Long, Q., Klemetti, R., Wang, Y., Tao, F., Yan, H., & Hemminki, E. (2012). High Caesarean section rate in rural China: is it related to health insurance (New Co-operative Medical Scheme)? *Soc Sci Med, 75*(4), 733-737.

Lumbiganon, P., Laopaiboon, M., Gulmezoglu, A.M., Souza, J.P., Taneepanichskul, S., Ruyan, P., et al. (2010). Method of delivery and pregnancy outcomes in Asia: the WHO global survey on maternal and perinatal health 2007-08. *Lancet, 375*(9713), 490-499.

Ma, X., Ma, L., Gan, L., Jia, H., & Wang, B.Z. (2009). Cohort study on the impact of different delivery methods on breastfeeding. *Chin J Child Health Care, 17*(4), 428-429. Chinese.

Marcus, H., Gerbershagen, H.J., Peelen, L.M., Aduckathil, S., Kappen, T.H., Kalkman, C.J., et al. (2015). Quality of pain treatment after caesarean section: Results of a multicentre cohort study. *Eur J Pain*, *19*(7), 929-939.

Martin, J.A., Hamilton, B.E., Osterman, M.J., Driscoll, A.K., & Drake, P. (2018). Births: Final data for 2016.

Mi, J., & Liu, F. (2014). Rate of caesarean section is alarming in China. *Lancet, 383*(9927), 1463-1464.

Miller, R.G. (1974). The jackknife-a review. *Biometrika*, *61*(1), 1-15.

Mkontwana, N., & Novikova, N. (2015). Oral analgesia for relieving post-caesarean pain. *Cochrane Database Syst Rev, 3*, CD010450.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med, 151*(4), 264-269, W264.

Moore, E.R., Bergman, N., Anderson, G.C., & Medley, N. (2016). Early skin-to-skin contact for mothers and their healthy newborn infants. *Cochrane Database Syst Rev*, *11*, CD003519.

Patel, A., Bucher, S., Pusdekar, Y., Esamai, F., Krebs, N.F., Goudar, S.S., et al. (2015). Rates and determinants of early initiation of breastfeeding and exclusive breast feeding at 42 days postnatal in six low and middle-income countries: A prospective cohort study. *Reprod Health, 12 Suppl 2*, S10.

Prior, E., Santhakumaran, S., Gale, C., Philipps, L.H., Modi, N., & Hyde, M.J. (2012). Breastfeeding after cesarean delivery: a systematic review and meta-analysis of world literature. *Am J Clin Nutr, 95*(5), 1113-1135.

Qin, L.L., & Hua, J.F. (2013). Investigtion and analysis of infant feeding within four months among floating population in Suzhou. *Nursing Practices and Research 10*(10), 138-139. Chinese.

Qiu, L. (2008). *Factors that influence breastfeeding initiation and duration in urban, suburban and rural areas of Zhejiang Province, Peoples Republic of China.* (Ph.D thesis). Curtin University, Perth.

Rowe-Murray, H.J., & Fisher, J.R. (2002). Baby friendly hospital practices: cesarean section is a persistent barrier to early initiation of breastfeeding. *Birth*, *29*(2), 124-131.

Ruan, M.J., Ba, L., Wang, Z., & Song, H. (2012). Study on breastfeeding status and the related influencing factors in Chaoyang district of Beijing *Maternal and Child Health Care of China, 27*(22), 3478-3480. Chinese.

Stroup, D.F., Berlin, J.A., Morton, S.C., Olkin, I., Williamson, G.D., Rennie, D., et al. (2000). Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA, 283*(15), 2008-2012.

Sundin, C.S., & Mazac, L.B. (2015). Implementing skin-to-skin care in the operating room after cesarean birth. *MCN Am J Matern Child Nurs, 40*(4), 249-255.

Tang, L. (2013). *A cohort study of infant feeding practices and maternal breastfeeding problems in a rural area of Sichuan Province, PR China.* (Ph.D thesis). Curtin University, Perth.

Tang, S., Li, X., & Wu, Z. (2006). Rising cesarean delivery rate in primiparous women in urban China: evidence from three nationwide household health surveys. *Am J Obstet Gynecol, 195*(6), 1527-1532.

Tang, Z.J. (2014). Investigation on current situation and influencing factors of infant breastfeeding in Liwan District, Guangzhou City *Chin J Women and Children Health 5*(1), 73-76. Chinese.

Tian, J.Z., Tian, X.S., & Li, L. (2008). Analysis of impact factors on breastfeeding of 183 mothers before discharge. *Zhejiang Prev Med, 20*(7), 45-55. Chinese.

Tooth, L., Ware, R., Bain, C., Purdie, D.M., & Dobson, A. (2005). Quality of reporting of observational longitudinal research. *Am J Epidemiol, 161*(3), 280-288.

Victora, C.G., Barros, A.J.D., Fuchs, S.C., de Francisco, A., Morris, J., Hall, A.J., et al. (2000). Effect of breastfeeding on infant and child mortality due to infectious diseases in less developed countries: a pooled analysis. *Lancet, 355*(9202), 451-455.

von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gotzsche, P.C., Vandenbroucke, J.P., et al. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med, 4*(10), e296.

Wang, B.S., Zhou, L.F., Zhu, L.P., Gao, X.L., & Gao, E.S. (2006). Prospective observational study on the effects of caesarean section on breastfeeding. *Zhonghua Fu Chan Ke Za Zhi, 41*(4), 246-248.

Wang, C.X., Chen, H., & Liu, Y.Z. (2005). Breastfeeding failure factors analysis and countermeasures. *J Community Med, 3*(9), 11-12. Chinese.

Wang, H.Z. (2010). Status analysis on breastfeeding in 1296 infants. *Practical Preventive Medcine 17*(3), 519-520. Chinese.

Wang, R., Liu, J., & Liu, J.T. (2009). Effects of delivery mode on breastfeeding rate and child growth *Maternal and Child Health Care of China, 24*(25), 3522-3523. Chinese.

Wang, S., Liang, Y.S., & Li, L. (1995). Survey on breastfeeding associated factors. *Chin J Family Planning 20*(6), 358-359. Chinese.

Wang, Z., Gao.J.K., Zhang, Y., Wang, F.H., Jin, J.M., & Zhu, B.H., et al. (2013). Present status and influencing factors analysis on breastfeeding for infants in Zhejiang Province. *Nursing and Rehabilitation J, 12*(7), 628-630. Chinese.

Wei, X.J., & Li, Q. (2009). Investigation on breastfeeding status of 4537 infants 42 days postpartum from 1996 to 2007. *Chin J Child Health Care, 17*(1), 106-108. Chinese.

World Health Organization. (2003). Infant and young child feeding: a tool for assessing national practices, policies and programmes.

World Health Organization. (2008). Indicators for assessing infant and young child feeding practices: part 1: definitions: conclusions of a consensus meeting held 6-8 November 2007 in Washington DC, USA.

World Health Organization. (2015). WHO statement on caesarean section rates.

Xu, F. (2008). *The Determinants of Breastfeeding Initiation and Duration in Xinjiang, Peoples Republic of China.* (Ph.D thesis). Curtin University, Perth.

Xu, T., & Yu, X.S. (2009). Breastfeeding influencing factors analysis *Chin J Health Statistics 26*(4), 406-410. Chinese.

Xue, F., Gong, J.F., & Luo, Y. (2012). Study on breastfeeding influencing factors in Changshu. *Maternal and Child Health Care of China, 27*(10), 1536-1537. Chinese.

Yang, Y.L., & Feng, G.Z. (2014). Related factors of breastfeeding *J Mathematical Med, 27*(4), 427-428. Chinese.

Ye, C.E. (2008). Analysis of breastfeeding current situation. *Maternal and Child Health Care of China, 23*(19), 2770. Chinese.

Yin, X.G., Gui, J.J., Wang, H.Q., Huang, X.M., Zhou, M.Y., & Zhu, P. (2012). Follow-up survey of risk factors for early cessation of breastfeeding in early infancy *Chin J Neonatology 27*(3), 148-152. Chinese.

Yu, C. (2013). *Breastfeeding and health outcomes in infants who receive continuing care from hospitals or community health centres in Chengdu Sichuan Province, People's Republic of China.* (Ph.D thesis). Curtin University, Perth.

Zhang, G.D., & Wang, J.L. (2000). Obstetric factors and breast-feeding. *Chongqing Med, 29*(2), 137-138. Chinese.

Zhang, Q.J., & Shi, Y.J. (2013). Breastfeeding influencing factors analysis for 612 women after delivery *Maternal and Child Health Care of China, 28*(30), 5095-5096. Chinese.

Zhang, S.J., & Wu, L.X. (1999). Breastfeeding 1 month postpartum and associated factors analysis *J North China Coal Medical College, 1*(1), 61. Chinese.

Zhang, W.K., Hao, B., Chen, L.J., Zhao, G.L., Bao, Y.Q., & Zhou, M., et al. (2006). The effects of method of delivery on rate of breastfeeding. *Chin J Child Health Care, 14*(5), 461-463. Chinese.

Zhang, Y.M. (2012). Impact of different delivery modes on breastfeeding *Seek Medical and Ask the Medicine 10*(10), 423. Chinese.

Zhang, Y.X., Dai, N., Zhang, A.H., Li, P.X., Yan, Y.H., & Li, N. (2013). Breastfeeding influencing factors analysis for 3057 women after delivery in Danyang City *J Qiqihar Univ of Med, 34*(18), 2736-2738. Chinese.

Zheng, K.Y., Zhang, X.H., Li, L., Shu, Y., & Qiu, L.Q. (2008). Factors influencing breastfeeding before discharge for delivery women in Hangzhou City *Chin J Child Health Care, 16*(6), 716-718. Chinese.

Zhu, P., Hao, J., Jiang, X., Huang, K., & Tao, F. (2013). New insight into onset of lactation: mediating the negative effect of multiple perinatal biopsychosocial stress on breastfeeding duration. *Breastfeed Med, 8*, 151-158.

Zhu, X., Tian, J., Chen, G., & Christensson, K. (2014). Predictors of breastfeeding exclusivity in three cities of China. *Breastfeed Med, 9*(2), 103-104.

Table 2-1 Characteristics of studies assessing the association between caesarean delivery and breastfeeding in China

Published year	First author	Study design	Location	Sample size	Other factors associated with breastfeeding outcomes	Breastfeeding outcome measured	Definition	Score
1994	HE, H.L	Cross-sectional	Fuzhou	216	Initiation time, milk bottle using, sleep, alcohol drinks	Breastfeeding at 1mo postpartum	Non-WHO	13
1995	WANG, S	Cross-sectional	Beijing	439	BFHI, maternal age, method of delivery, infant gender, gestational reaction	FB at 4mo postpartum	Reclassified with WHO definition	13
1996	FANG, L.L	Cross-sectional	Beijing	60	Initiation time	EBF at 42d/4mo postpartum	WHO	13
1998	LIU, J	Cross-sectional	Inner Mongolia	374	Maternal health condition, complication, nutrition condition, appetite after delivery, maternal education, initiation time	Breastfeeding 0-90 day postpartum	Non-WHO	11
1999	ZHANG, S.J	Prospective cohort	Hebei	207	Knowledge of breastfeeding, confidence of breastfeeding,	Breastfeeding at 1mo postpartum	Non-WHO	15

					initiation time, breast milk substitute			
2000	ZHANG, G.D	Prospective cohort	Chongqing	627	Postpartum hemorrhage	Breastfeeding 28d/4mo postpartum	Non-WHO	11
2000	JIANG, G.F	Cross-sectional	Liuhe	736	Maternal education, rooming-in	EBF within 6mo postpartum	WHO	15
2005	WANG, C.X	Retrospective cohort	Jinan	853	Maternal education, health education, milk powder promotion, initiation time	EBF at 4mo postpartum	Non-WHO	16
2006	WANG, B.S	Prospective cohort	Shanghai	602	NA	FB at 1,6,12 mo postpartum	Reclassified with WHO definition	13
2006	ZHANG, W.K	Prospective cohort	Beijing	802	Maternal age, early touch time, perception of breastfeeding during gestation, living space	FB 2-5d, 42d postpartum	Reclassified with WHO definition	16
2007	GAN, W.L	Cross-sectional	Chongqing	375	Maternal age, maternal education, occupation, maternal mood, postpartum home visit, income monthly	EBF at 4mo postpartum	Non-WHO	12

2008	XU, F.L	Prospective cohort	Xinjiang	1064	Giving breastmilk as the first feed, feeding on demand, mother feeling given enough information about breastfeeding, minority ethnic group, giving birth in spring or summer, medical staff not recommending formula to parents, prelacteal feeds of water or formula	AF initiation/ EBF initiation	WHO	18
2008	LIU, F.L	Retrospective cohort	Xinxiang	288 392 415 376	NA	AF at 4mo postpartum in year 2000, 2002, 2004 and 2006 separately	Non-WHO	13
2008	ZHENG, K.Y	Cross-sectional	Hangzhou	628	Neonatal disease, initiation time, early feeding	AF before discharge	Reclassified with WHO definition	16

2008	TIAN, J.Z	Cross-sectional	Zhejiang	253	Neonate disease, early sucking, initiation time, breastfeeding confidence, maternal education	FB before discharge	Reclassified with WHO definition	14
2008	YE, C.E	Cross-sectional	Ninghai	931	Maternal education, health education	EBF at 4mo postpartum	Non-WHO	16
2008	JIANG, Z.H	Retrospective cohort	Herbin	310	Initiation time, maternal education	EBF at 42d postpartum	Non-WHO	14
2008	QIU, L.Q	Prospective cohort	Zhejiang	917	Living in the suburb and rural areas, maternal age, mother decides to breastfeed until after birth, prelacteal feeding	EBF initiation/ at discharge	WHO	18
2009	LIU, X.Q	Cross-sectional	Beijing	123	Health education, initiation time	Breastfeeding 42d postpartum	Non-WHO	11
2009	XU, T	Retrospective cohort	Shenyang	1025	Maternal age, family income	FB at 6mo postpartum	Reclassified with WHO definition	17
2009	WEI, X.J	Retrospective cohort	Zhengzhou	501 445	NA	FB 42d postpartum in year 1996-2007 separately		16

				458				
				503				
				541				
				399				
				340			Reclassified with WHO definition	
				223				
				347				
				284				
				250				
				246				
2009	WANG, R	Cross-sectional	Kunming	1328	NA	EBF initiation, at 4mo postpartum	Non-WHO	15
2009	MA, X	Prospective cohort	Shaanxi	605	Initiation time, perception of breastfeeding, supplementary	EBF before discharge and 1mo postpartum	Non-WHO	13

					feeding, confidence of breastfeeding, milk bottle using			
2010	CHEN, Y.F	Cross-sectional	Wuhan	445	Prenatal preparation, initiation time, right feeding method, feeding confidence, maternal nutrition, feeding setting	Breastfeeding within 4mo-6mo postpartum	Non-WHO	13
2010	WANG, H.Z	Cross-sectional	Changli	1296	Health education, maternal education	EBF at 6mo postpartum	WHO	13
2010	HUANG, Q.J	Cross-sectional	Shanghai	350	Early feeding, milk bottle using	EBF before discharge	WHO	13
2012	RUAN, M.J	Cross-sectional	Beijing	103	Maternal education, fixed term job, maternal age	EBF at 4mo postpartum	Non-WHO	13
2012	ZHANG, Y.M	Prospective cohort	Chengdu	268	NA	EBF at discharge	Non-WHO	15
2012	XUE, F	Cross-sectional	Changshu	126	Area (rural or urban)	EBF at 4mo postpartum	Non-WHO	13
2012	YIN, X.G	Prospective cohort	Hefei	2522	Maternal education, family income, premature birth	AF at 2mo, 4mo postpartum	Reclassified with WHO definition	18

2012	LIU, X	Retrospective cohort	27 sites	431,704	NA	EBF before discharge	WHO	17
2013	ZHANG, Q.J	Cross-sectional	Zhengzhou	612	NA	EBF 24h before discharge	Non-WHO	8
2013	ZHANG, Y.X	Cross-sectional	Danyang	3057	First breastfeeding duration, perception of breastmilk amount, psychosocial factors	FB before discharge and after discharge	WHO	15
2013	KANG, Y	Cross-sectional	Chongqing	939	Infant gender, maternal education, income monthly, birth weight, duration of maternal leave, perception of breastmilk amount, prelacteal feeding	Breastfeeding at 6mo	Non-WHO	14
2013	WANG, Z	Cross-sectional	Zhejiang	528	Infant age, infant gender, early feeding, perception of breastmilk amount	EBF within 6mo postpartum	WHO	16
2013	GUO, S.F	Cross-sectional	26 counties	2293	Maternal antenatal clinic visit, child’s age	EBF initiation /within 6mo postpartum	WHO	14

2013	ZHU, P	Prospective cohort	Hefei	1602	Preterm birth, breastfeeding frequency on Day 1, life events in the third trimester, onset of lactation	AF at 2mo postpartum	Reclassified with WHO definition	14
2013	QIN, L.L	Retrospective cohort	Suzhou	1212	Maternal occupation, maternal age, birth region, breastfeeding professional instruction	EBF at 4mo postpartum	WHO	17
2013	TANG, L	Prospective cohort	Jiangyou	693	Father's attitude towards breastfeeding, early breastfeeding initiation	EBF initiation within 1h post birth / AF at discharge/ FB at discharge	WHO	18
2013	YU, C	Prospective cohort	Chengdu	845	Maternal occupation, paternal education, intention of going back to work, first feeding, mothers' friends breastfeed their babies, paternal job, staff encouragement, father's attitude, maternal	AF within 15 d postpartum/ FB within 15d postpartum	WHO	18

					grandmother's breastfeeding history			
2014	TANG, Z.J	Cross-sectional	Guangzhou	315	NA	BF at 6mo postpartum	Reclassified with WHO definition	13
2014	LIU, L.F	Cross-sectional	Lishui	675	Family income monthly, early initiation, health education, neonatal disease	EBF within 6mo	WHO	16
2014	LENG, X.L	Cross-sectional	Shenzhen	1200	Maternal age, initiation time	FB within 6mo postpartum	Reclassified with WHO definition	13
2014	YANG, Y.L	Cross-sectional	Wuhan	513	Health education, maternal education, prenatal high risk factors	EBF within 6mo postpartum	Non-WHO	8
2014	LI, R	Cross-sectional	Shenzhen	840	Initiation time, nipples condition	EBF before discharge	Non-WHO	13
2014	ZHU, X	Cross-sectional	Three cities	151	Occupation status, baby's first sucking time, times of sucking when back home per day	EBF at 6mo postpartum	Non-WHO	14

BFHI: Baby Friendly Hospital Initiative; FB: full breastfeeding; EBF: exclusive breastfeeding AF: any breastfeeding

Chapter 3: Maternal education and breastfeeding practices in China: a systematic review and meta-analysis

The content of this chapter is covered by a published paper "Zhao, J., Zhao, Y., Du, M., Binns, C. W., & Lee, A. H. (2017). Maternal education and breastfeeding practices in China: A systematic review and meta-analysis. Midwifery, 50, 62-71. DOI: https://doi.org/10.1016/j.midw.2017.03.011". (see Appendix D.2)

The statement of primary contribution of the first author and the permission to include the publication in this thesis can be found in the Appendix A. The permission to reproduce the material from the publisher can be found in the Appendix D.1.

3.1 Introduction

Breastfeeding is the optimal feeding method for infants as breastmilk provides all nutrients that infants need for healthy growth and development up to six months of age (Gillman et al., 2001; Schanler et al., 1999). The benefits of breastfeeding to both infant and maternal health, either short term or long term, are well documented in the literature (Bhandari et al., 2003; Howie et al., 1990; Ip et al., 2007). Because of the public health benefits of breastfeeding, many countries have initiated health promotion interventions, prioritizing recognized factors affecting breastfeeding outcomes, to support and promote breastfeeding practices, especially exclusive breastfeeding (Bernaix et al., 2010; Howell et al., 2014; Kramer et al., 2001; Oken et al., 2013; Pisacane et al., 2012). Among the factors affecting breastfeeding practices, maternal education has been investigated widely as maternal intellectual maturity has been speculated as one of the most important factors influencing a woman's decision to breastfeed (Chin et al., 2008; Li et al., 2005). However, the findings and conclusions are inconsistent. The effect of maternal education has often been concluded differently in studies carried out in China compared to those in Western countries (Heck et al., 2006; Kehler et al., 2009; Ludvigsson & Ludvigsson, 2005; Scott & Binns, 1999). In Western countries, including the U.S., Australia, New Zealand, Canada, the Netherlands, Sweden and Italy, studies report that higher maternal education is associated with higher odds for breastfeeding (Dubois &

Girard, 2003; Heath et al., 2002; Heck et al., 2006; Hornell et al., 1999; Lanting et al., 2005; Riva et al., 1999; Scott & Binns, 1999).

Some studies in China agreed well with the above findings from those Western countries (Jiang, 2000; Ma et al., 2014), however, others have reached the opposite conclusions (Huang et al., 2012; Qiu, 2008; Ye et al., 2007). There has been no published systematic review of the impact of maternal education on breastfeeding prevalence in China. In this thesis, a systematic review and meta-analysis of breastfeeding studies carried out in China was undertaken to examine the association between maternal education and breastfeeding prevalence. The findings of this study have the potential to guide the provision of focused and evidence-based interventions aimed to promote breastfeeding in China.

3.2 Systematic review and meta-analysis

3.2.1 Systematic review

Search strategy

Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009), a systematic electronic search was conducted to retrieve literature published in Chinese or English using the Chinese database China National Knowledge Infrastructure (CNKI) and English databases Medline, Embase, CINAHL, ProQuest and Science Direct from January 1990 to June 2015. The search strategy adopted included the following two stages:

Stage 1: The following Medical Subject Headings (MeSH) terms and keywords “breast feeding”, “ human milk”, “breastfeeding duration”, “breastfeeding cessation”, “human lactation”, “infant feed*”, “breastfed”, “risk factor*”, “protective factor*”, “determinant*”, “socioeconomic factor*”, “China”, “Mainland China”, “Chinese” were used.

Stage 2: The following MeSH terms and keywords “maternal education”, “education” and “education status” were further added to the search process.

Literature screening and selection criteria

The abstracts of the identified records were screened by two independent reviewers (JZ and MRD) to check whether they were appropriate to be included in the present study. Full texts were then extracted and evaluated after the abstract screening confirmed the appropriateness. The inclusion criteria for the studies were: (i) published in peer-reviewed journals or theses/dissertations; (ii) observational study design; (iii) reported and evaluated the relationship between maternal education and breastfeeding prevalence quantitatively; (iv) effect size with 95% confidence intervals (CIs) could be obtained directly or calculated from the raw tabulated data. The exclusion criteria were: (i) sample size not specified; (ii) time points of breastfeeding outcomes not reported or defined; (iii) inappropriate statistical result reported (statistical error); (iv) unspecified category of maternal education. Discrepancies were discussed between JZ and MRD until a consensus could be reached or referred to a third investigator (YZ) when necessary.

Data extraction

Information, including published year, first author name, study design, location of study, sample size, breastfeeding outcomes measured, maternal education status, other factors associated with breastfeeding outcomes, results of tabulated data or effect size (odds ratio: OR) with their corresponding 95% CIs reported by univariate analysis (crude ORs) or multivariable analysis with adjustment for confounders (adjusted ORs), was retrieved from each eligible study for qualitative and quantitative synthesis. In cases in which essential information was not provided in articles, the authors were contacted by email.

We followed the World Health Organization (WHO) definitions in the data extraction (World Health Organization, 2003, 2008):

Exclusive breastfeeding: Breastfeeding while giving no other food or liquid, not even water, except drops or syrups consisting of vitamins, mineral supplements or medicines.

Full breastfeeding: Exclusive breastfeeding or predominant breastfeeding (or almost exclusive breastfeeding). Breastmilk is the only source of milk given to the infant regardless of supplementation with other fluids such as water and orange juice.

Any breastfeeding: The child has received breastmilk (direct from the breast or expressed) with or without other drink, formula or other infant food.

Quality assessment

Each selected study had been assessed its methodological quality using a formal checklist 'the Standard Quality Assessment Criteria for Evaluating Primary Research Papers from a Variety of Fields' (Kmet et al., 2004). Two independent researchers (JZ and MRD) performed the scoring process according to each criterion of the checklist. Each item of each eligible study was given a score (0-No, 1-Partial, 2-Yes) based on the degree to which the criteria were met. The summary score was then calculated by summing the total score obtained across relevant items then divided by the total possible score (22) for the study. Therefore, study quality scores ranged from 0 to 1, where a higher score corresponds to higher quality. Any discrepancies between the assessors were resolved by consensus.

Assessment of risk of bias

Based on the Cochrane Collaboration's tool for assessing the risk of bias and GRADE guidelines of assessing study limitations (risk of bias) in observational studies (Guyatt et al., 2011; Higgins et al., 2011), we summarized key criteria to assess the risk of bias across studies included for five domains: incomplete outcome data (attrition bias), selective reporting (reporting bias), lack of adjustment for baseline characteristics, failure to develop and apply appropriate eligibility criteria and flawed measurement of both exposure and outcome. For each domain, we classified studies into low, unknown, and high risk of bias.

3.2.2 Meta-analysis

In this study, maternal education (in years) refers to the formal education mothers receive in primary schools, high schools or tertiary institutes and was coded into two binary categorical variables. All eligible studies were classified into two groups based on different cut-off values used in the coding of maternal education. In Group 1, maternal education was recoded into two categories 'more than 6 years (>6 years) education' versus '6 years or less (<=6 years) education', while in Group 2 the coding criteria was 'more than 12 years (>12 years) education' versus '12 years or less (<=12 years) education', respectively. The ORs of breastfeeding (including exclusive breastfeeding, full breastfeeding and any breastfeeding within 12 months postpartum) comparing higher educational status to lower educational status were the primary statistical measures and were transformed into the logarithmic scale for meta-analysis. Considering the effect of confounding, crude ORs and adjusted ORs were pooled separately in both groups.

A fixed-effect meta-analysis by the inverse variance (I-V) method was conducted to pool the OR between maternal education and breastfeeding prevalence (Higgins & Green, 2011). Visual inspection of a Galbraith plot was conducted to detect the heterogeneity due to individual studies (Bax et al., 2009). The I^2 statistic was used to assess the heterogeneity across the studies (Higgins & Thompson, 2002; Higgins et al., 2003). A random-effect meta-analysis by DerSimonian and Laird (D+L) method was applied when significant heterogeneity was present (DerSimonian & Laird, 1986). A meta-regression was conducted to investigate the potential sources of heterogeneity. Subgroup analysis was carried out if some factors were detected being significant in meta-regression. Considering the difference between fixed time point prevalence of breastfeeding for cohort study and period average prevalence of breastfeeding for cross-sectional study, subgroup analysis based on the design of the study (cohort study or cross-sectional study) was performed for both Group 1 and Group 2. To test the dependence of effect size reported in each study, a sensitivity analysis using the Jack-knife method was performed to assess the robustness of the results (Miller, 1974). Begg's test and Egger's test were conducted to investigate the publication bias or small sample size bias across studies (Egger et al., 1997). All statistical analyses were performed using the Stata package version 14.1 (StataCorp LP, College Station, USA). A p value less than 0.05 was considered statistically significant.

3.3 Results of the systematic review

As shown in Figure 3-1, a total of 181 studies were identified from English and Chinese databases and 4 additional records were obtained via ProQuest Dissertations and Theses. After removal of duplicate articles, 54 articles were retrieved to assess eligibility and ultimately, 31 studies were included in the systematic review, as shown in Table 3-1. The 31 studies included 23 cross-sectional studies (Guo et al., 2013; Huang et al., 2012; Jiang, 2000; Jiang et al., 2013; Kang et al., 2013; Ke, 1993; Li et al., 2003; Lin et al., 1997; Liu et al., 2013; Liu & Zhou, 2013; Ruan et al., 2012; Tang, 2014; Tian et al., 2008; Wang, 2010; Wang et al., 2013; Wu & Qiu, 2015; Xiao, 2001; Xiong et al., 2006; Yang & Feng, 2014; Ye, 2008; Ye et al., 2007; Yu & Song, 2000; Zhang et al., 1998), 4 retrospective studies (Jiang & Li, 2008; Ma et al., 2014; Qin & Hua, 2013; Wang et al., 2005) and 4 prospective studies (Qiu, 2008; Tang, 2013; Xu, 2008; Yu, 2013). The time points of breastfeeding outcomes measured included ‘before hospital discharge’, 15 days, 28 days, 1 month, 42 days, 2 months, 3 months, 4 months and 6 months postpartum. The quality score of each study included was calculated and reported in Table 3-1 with a mean of 0.80 and standard deviation of 0.15, ranging from 0.50 to 1.00. Assessment of risk of bias across studies is shown in Table 3-2. High risk of attrition bias due to incomplete outcome data was found in only 6 (19%) studies, high risk of reporting bias due to selective reporting of outcomes was rated in only 4 (13%) studies, and high risk of bias due to lack of adjustment for baseline characteristics was identified in 18 (58%) studies. No high risk of bias due to failure to develop and apply appropriate eligibility criteria and flawed measurement of both exposure and outcome was found in studies, which were rated as either “low risk of bias” or “unknown risk of bias”.

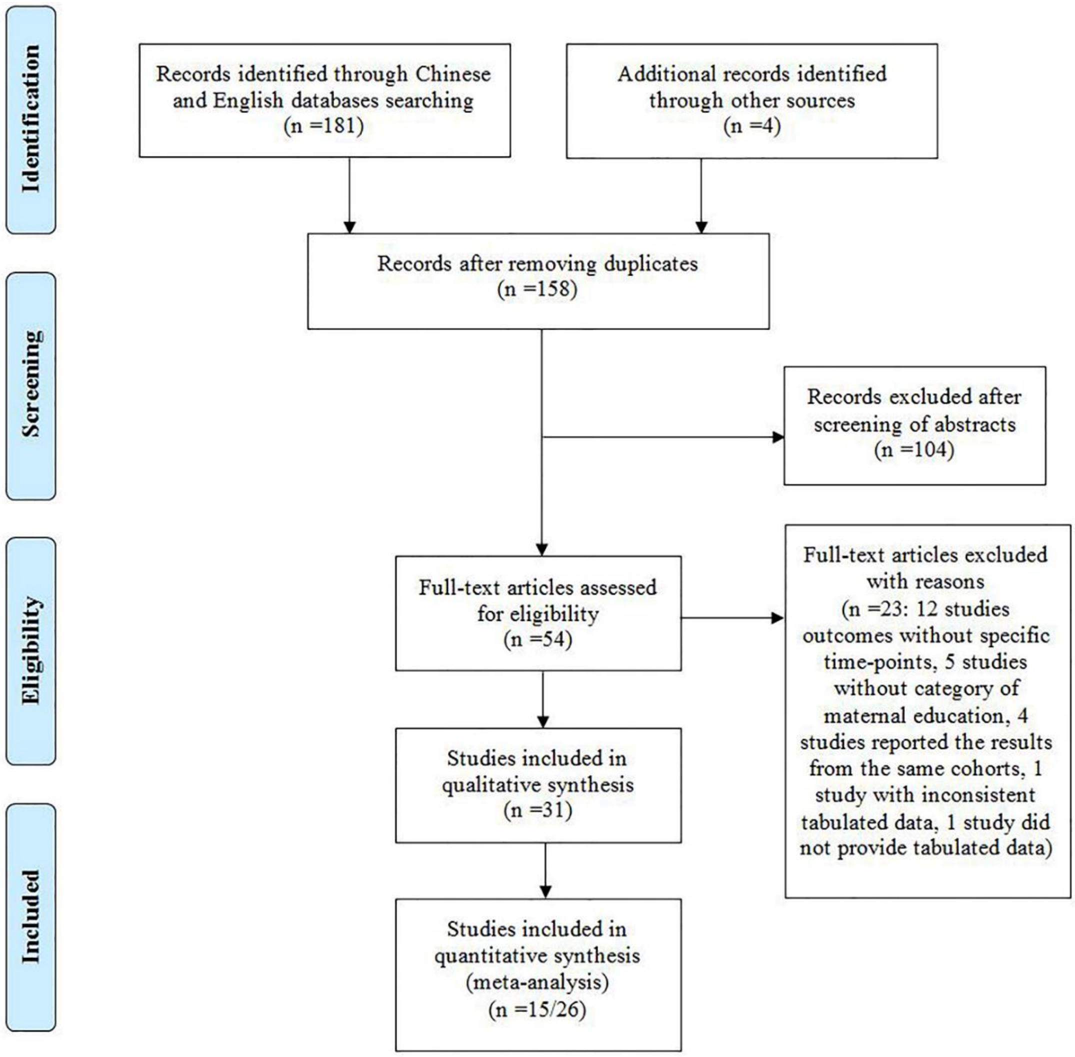


Figure 3-1 PRISMA flow chart of the systematic review process

3.4 Results of the meta-analysis

Fifteen studies were included in Group 1 for meta-analysis (Jiang, 2000; Jiang et al., 2013; Kang et al., 2013; Ke, 1993; Liu et al., 2013; Liu & Zhou, 2013; Ma et al., 2014; Qiu, 2008; Ruan et al., 2012; Wang et al., 2005; Wang, 2010; Wu & Qiu, 2015; Xu, 2008; Ye, 2008; Yu, 2013), and 26 studies were involved in Group 2 for meta-analysis (Jiang & Li, 2008; Kang et al., 2013; Li et al., 2003; Lin et al., 1997; Liu et al., 2013; Liu & Zhou, 2013; Ma et al., 2014; Qin & Hua, 2013; Qiu, 2008; Ruan et al., 2012; Tang, 2013; Tang, 2014; Tian et al., 2008; Wang et al., 2005; Wang et al., 2013; Wu & Qiu, 2015; Xiao, 2001; Xiong et al., 2006; Xu, 2008; Yang & Feng, 2014; Ye, 2008; Yu, 2013; Yu & Song, 2000; Zhang et al., 1998), among which 2 individual studies were classified into 4 sub-studies due to the various breastfeeding outcomes measured (Tang, 2013; Yu, 2013). Note that there were 11 studies commonly shared by both groups as they provided results suitable for both cut-off values (Kang et al., 2013; Liu et al., 2013; Liu & Zhou, 2013; Ma et al., 2014; Qiu, 2008; Ruan et al., 2012; Wang et al., 2005; Wu & Qiu, 2015; Xu, 2008; Ye, 2008; Yu, 2013) and 3 studies excluded in either group as they were not suitable to be quantitatively synthesized using either cut-off values (Guo et al., 2013; Huang et al., 2012; Ye et al., 2007). Four studies reported appropriate adjusted odds ratios from multiple logistic regressions which controlled for potential confounders (Guo et al., 2013; Huang et al., 2012; Li et al., 2003; Ye et al., 2007), but only one study could be used in the quantitative synthesis based on the preselected cut-off value (Li et al., 2003). The remaining studies reported crude (unadjusted) ORs or provided tabulated data to calculate the crude ORs for the meta-analysis.

3.4.1 Meta-analysis for Group 1: '> 6 years' versus '<=6 years'

In Group1, where the maternal education was categorised using the cut-off of '6-year education', 15 studies (13395 subjects) were included in the meta-analysis of the association between maternal education and breastfeeding. The random-effect meta-analysis indicated that the odds of breastfeeding were 10% (pooled OR=0.90, 95% CI: 0.83, 0.97) lower in mothers who had been educated for 'more than 6 years' compared to mothers who had '6 years or less' education. There was substantial heterogeneity across the studies (I^2 =92.5%, p<0.001) as shown in Figure 3-2. A visual inspection of Galbraith plot suggested several individual studies (Jiang, 2000; Jiang et al., 2013; Liu & Zhou, 2013; Ma et al., 2014; Qiu, 2008; Wu & Qiu, 2015; Xu, 2008; Ye, 2008; Yu, 2013) were the potential sources of the heterogeneity (Figure 3-3).

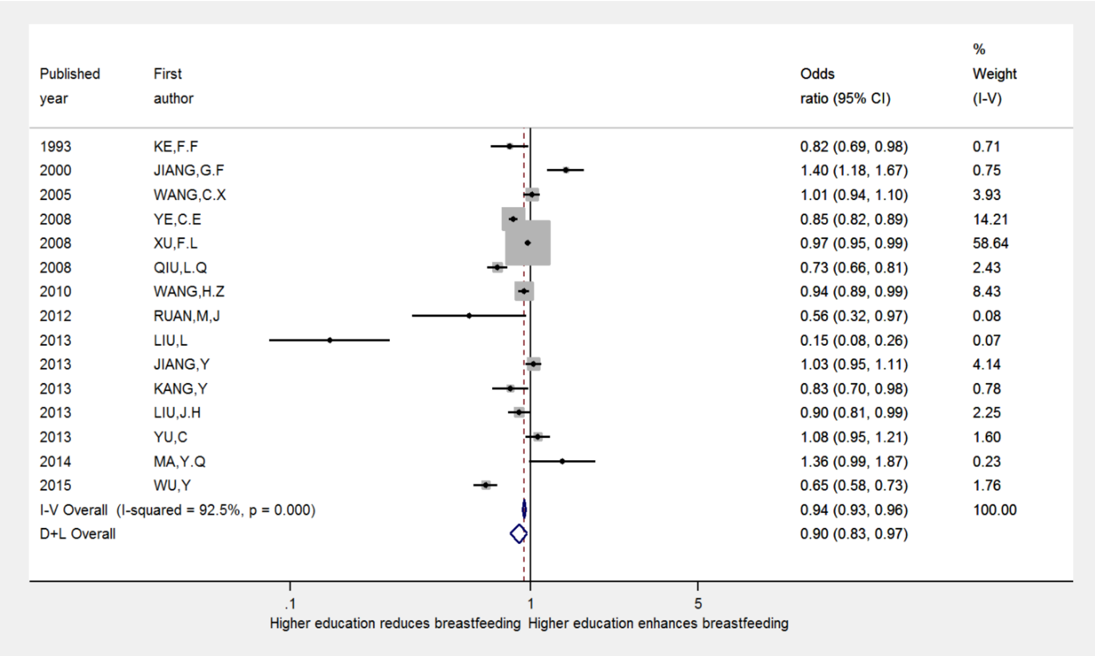


Figure 3-2 Forest plot of the association between maternal education and breastfeeding in Group 1

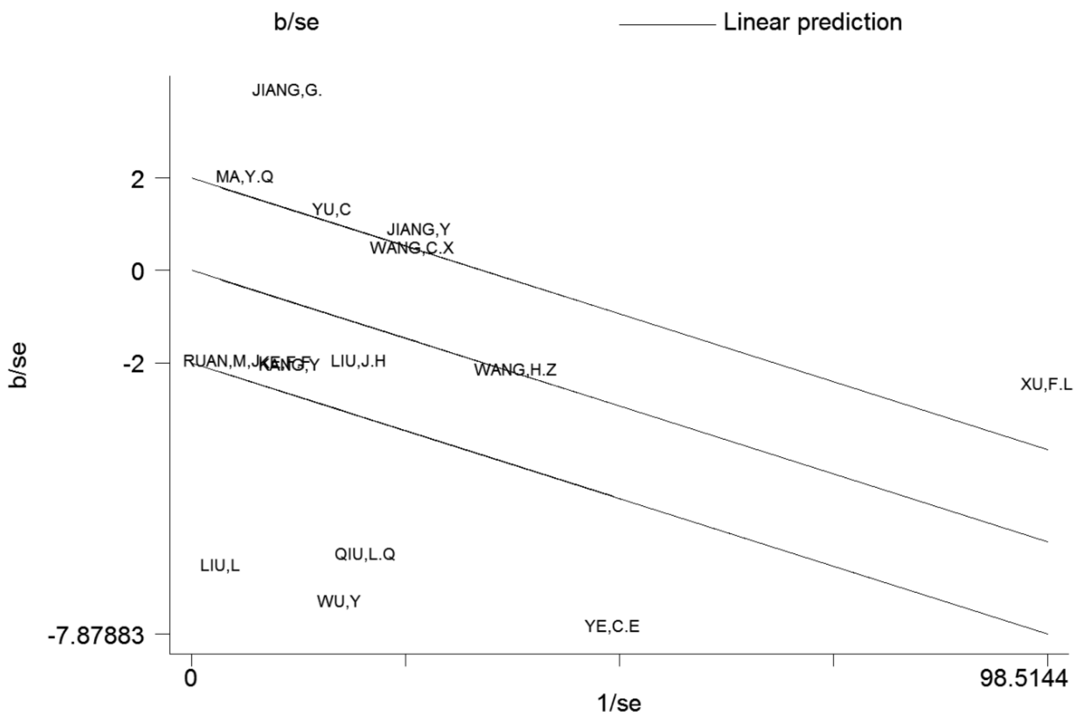


Figure 3-3 Visual inspection of heterogeneity sources via a Galbraith plot in Group 1

A meta-regression analysis including the study level covariates such as published year, sample size, study design, definitions of breastfeeding, and quality assessment scores of studies was performed to investigate the sources of heterogeneity. None of the above covariates were found contributing to the heterogeneity between studies.

Subgroup analysis stratified by sample size

Studies in Group 1 were divided into two subgroups: one subgroup of studies with a sample size smaller than the mean of sample sizes and the other subgroup of studies with a sample size larger than the mean of sample sizes.

A subgroup meta-analysis stratified by the sample size was performed as shown in Figure 3-4. In the subgroup of studies with a sample size smaller than the average sample size, the pooled OR is 0.77 (95% CI: 0.57, 1.03), suggesting a nonsignificant association between maternal education and breastfeeding prevalence, while in the other subgroup with a sample size larger than the average sample size, the pooled OR is 0.91 (95% CI: 0.85, 0.98), revealing a negative impact of maternal education on breastfeeding prevalence.

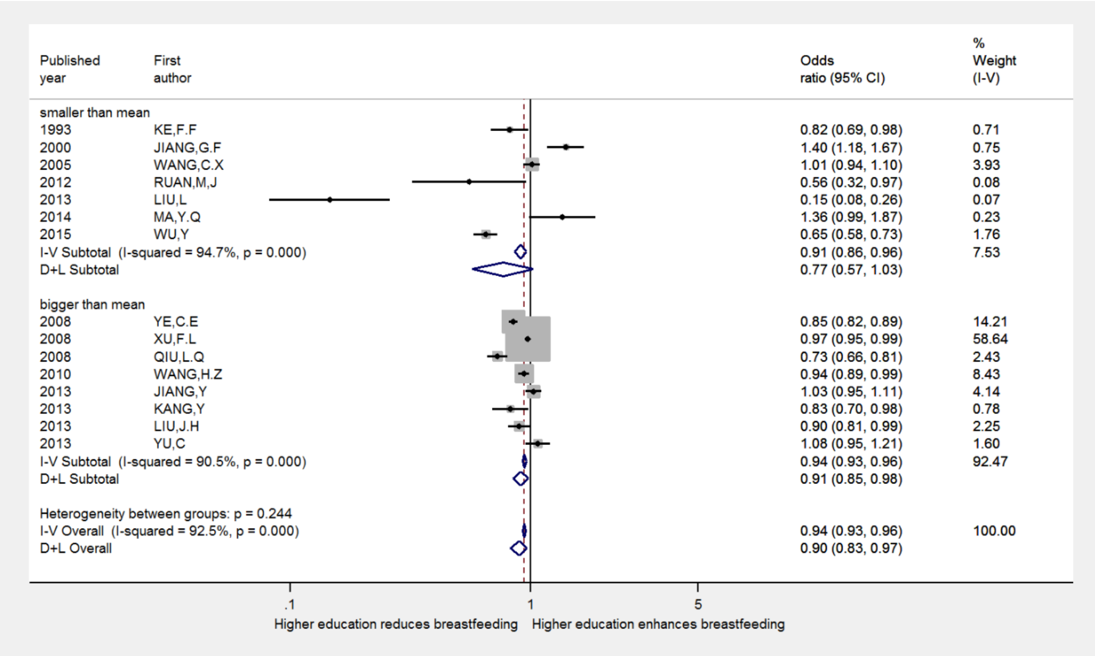


Figure 3-4 Subgroup analysis stratified by sample size in Group 1

Subgroup analysis stratified by study design

Studies in Group 1 were further classified into two subgroups according to their study designs: cross-sectional studies and cohort studies as shown in Figure 3-5. In the subgroup of cross-sectional studies, the pooled OR is 0.84 (95% CI: 0.74, 0.95) showing a significant negative impact of maternal education on breastfeeding, while in the cohort studies subgroup, the pooled OR is 0.98 (95% CI: 0.84, 1.14), indicating no significant association between maternal education and breastfeeding.

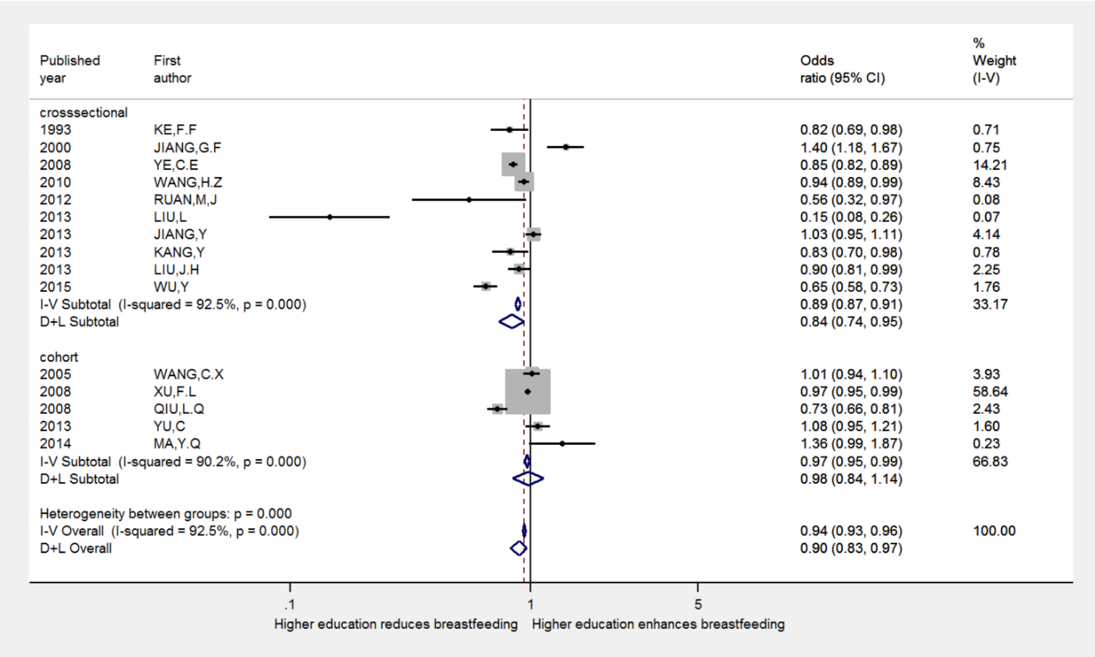


Figure 3-5 Subgroup analysis stratified by study design in Group 1

3.4.2 Meta-analysis for Group 2: '> 12 years' versus '<=12 years'

A total of 26 studies (15366 subjects) were included in Group 2 for the meta-analysis to investigate the pooled effect of maternal education on breastfeeding when the maternal education was categorised using the cut-off of '12-year education'. As Figure 3-6 shows, a random-effect meta-analysis revealed that compared to the mothers who had '12 years or less' education, the odds of breastfeeding in mothers who attained 'more than 12 years' education was significantly lower by 9% (pooled OR=0.91, 95% CI: 0.86, 0.96). An evident heterogeneity was present (I^2 =85.2%, p<0.001) and the corresponding Galbraith plot as shown in Figure 3-7 identified that the individual studies (Jiang & Li, 2008; Lin et al., 1997; Liu & Zhou, 2013; Qin & Hua, 2013; Qiu, 2008; Tang, 2013; Tang, 2014; Wu & Qiu, 2015; Xiong et al., 2006; Yang & Feng, 2014; Ye, 2008; Yu, 2013) were the potential source of the heterogeneity.

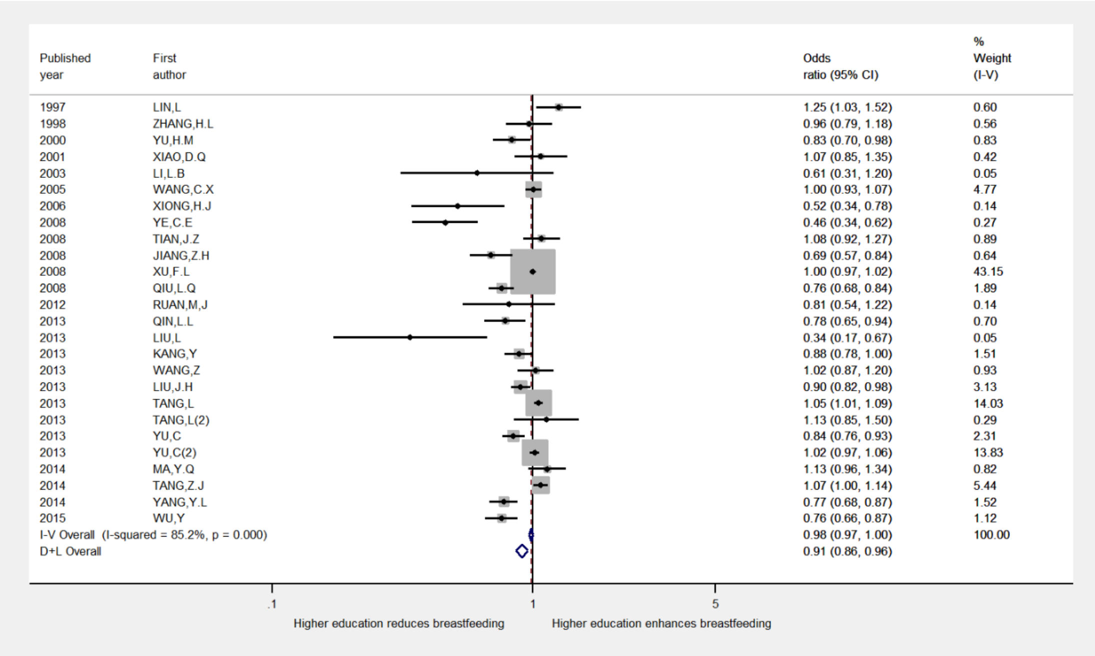


Figure 3-6 Forest plot of the association between maternal education and breastfeeding in Group 2

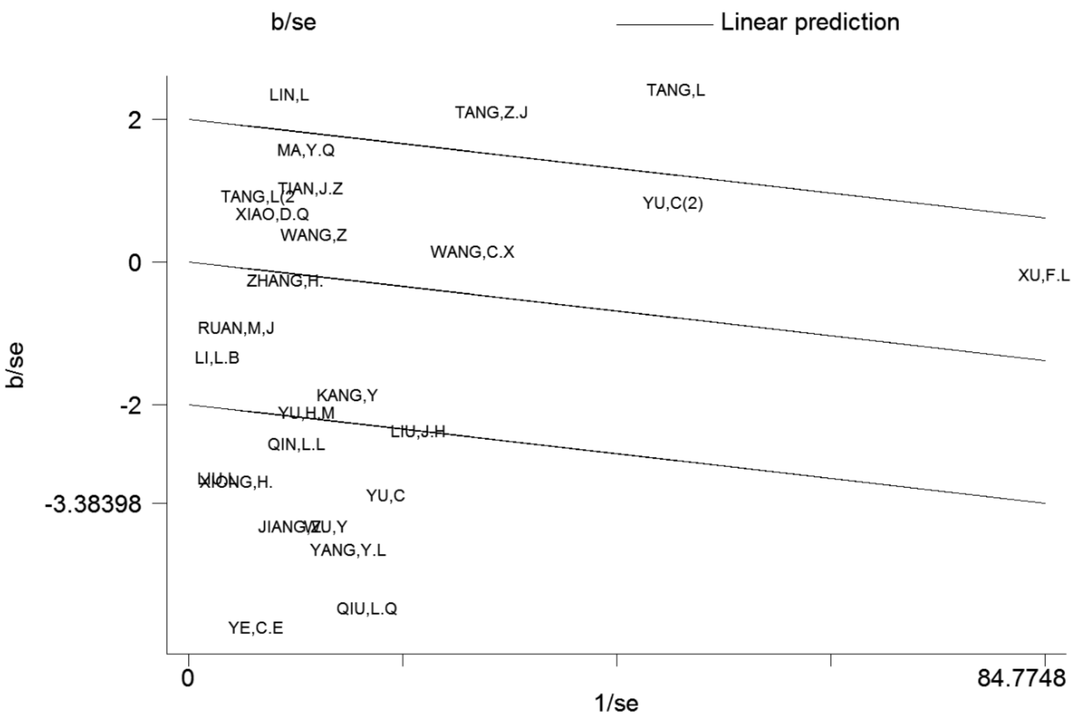


Figure 3-7 Visual inspection of heterogeneity sources via a Galbraith plot in Group 2

A meta-regression analysis was performed to investigate whether study level factors (published year, sample size, study design, definitions of breastfeeding, and quality assessment scores of studies) contribute to the heterogeneity; however, none of them were found to be a source of the heterogeneity.

Subgroup analysis stratified by study design

Subgroup analysis stratified by study design (cross-sectional study and cohort study) was performed as shown in Figure 3-8. The pooled OR in the subgroup of cross-sectional studies is 0.82 (95% CI: 0.73, 0.93), indicating a significant negative association between maternal education and breastfeeding, however, in the subgroup of cohort studies the pooled OR is 0.97 (95% CI: 0.92, 1.02), suggesting nonsignificant impact of maternal education.

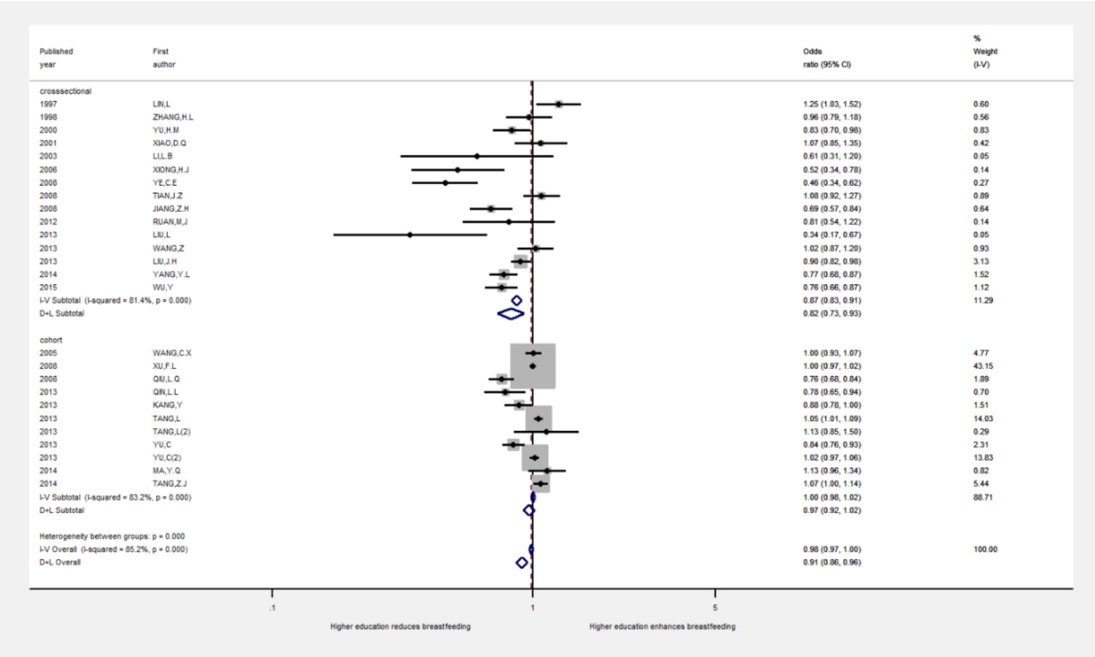


Figure 3-8 Subgroup analysis stratified by study design in Group 2

3.4.3 Sensitivity analysis

The sensitivity analysis (Figure 3-9 and Figure 3-10) indicated that the pooled effect size remained significant when each study was omitted in turn, which suggested that the results of the meta-analysis were robust.

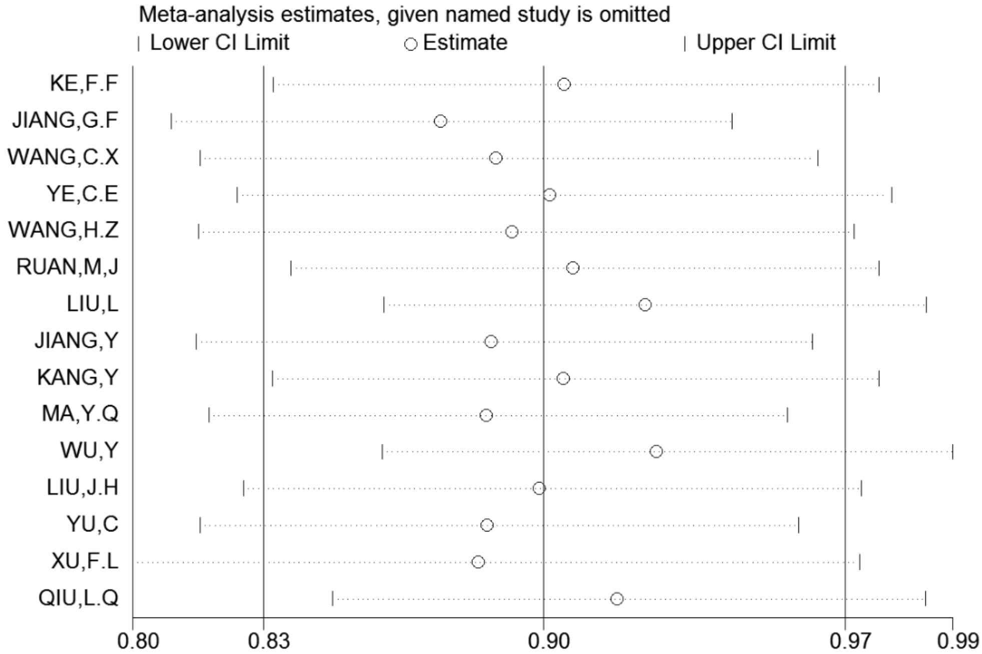


Figure 3-9 Sensitivity analysis assessing the influence of individual study on the pooled effect size in Group 1

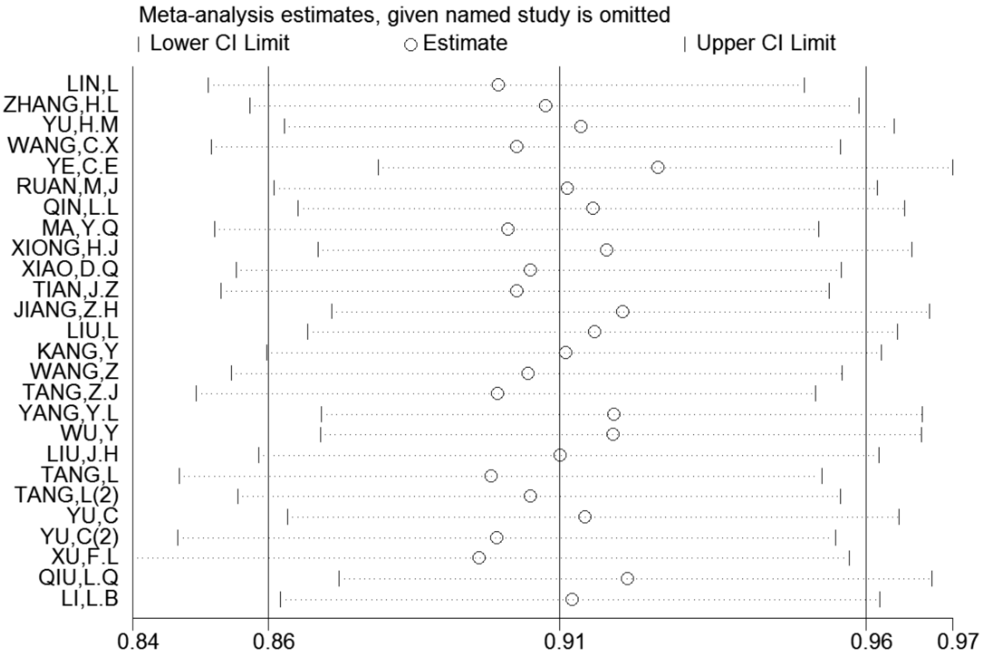


Figure 3-10 Sensitivity analysis assessing the influence of individual study on the pooled effect size in Group 2

3.4.4 Publication bias

For the meta-analysis of 15 studies in Group1, the funnel plot showed symmetric, consistent with the Begg's test (p=0.373) and Egger's test (p=0.236), suggesting that no publication bias or small study effect was present.

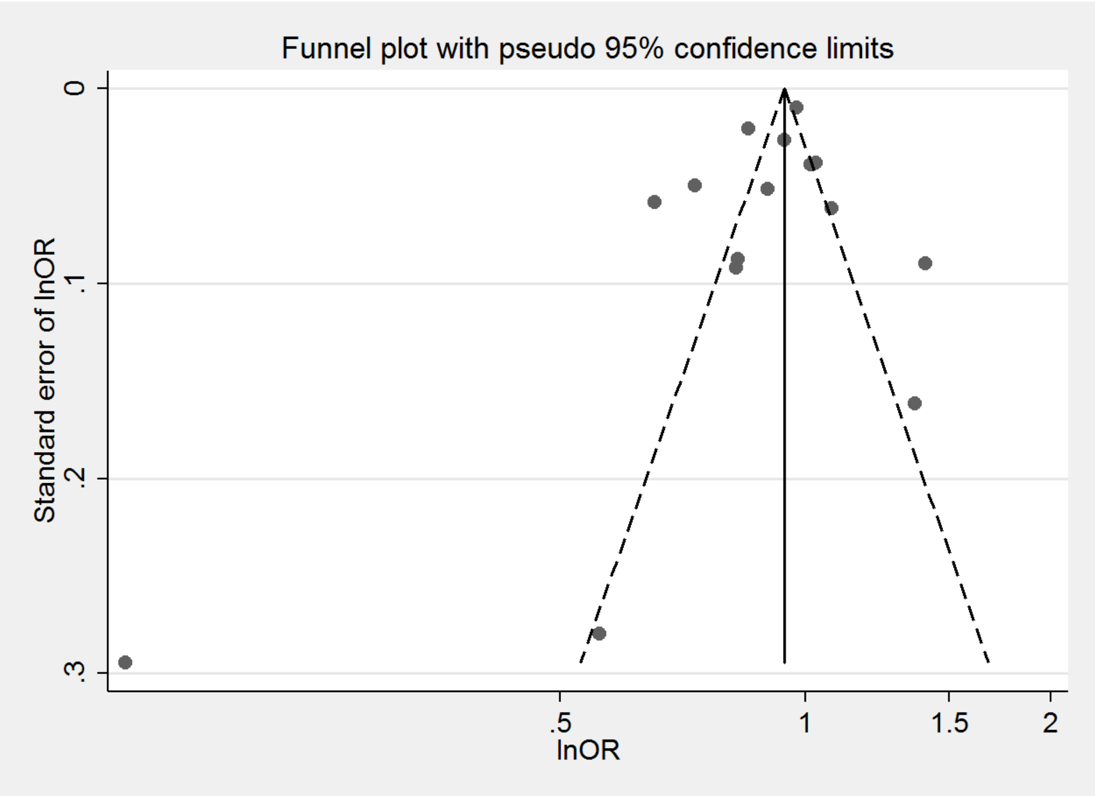


Figure 3-11 Funnel plot visualizing the publication bias in Group 1

The funnel plot of 26 studies in the meta-analysis of Group 2 appeared asymmetric. Although the Begg's test showed no significant publication bias or small study effect (p=0.146), the Egger's test suggested significant publication bias or small study effect (p=0.005).

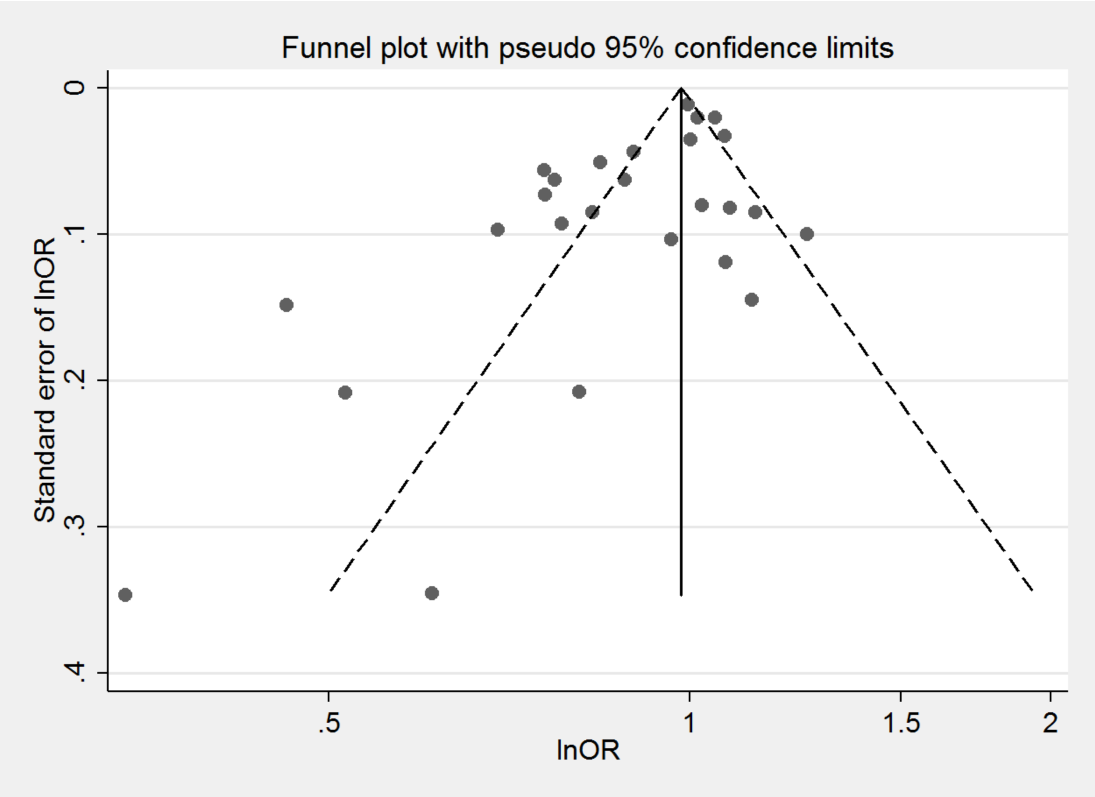


Figure 3-12 Funnel plot visualizing the publication bias in Group 2

3.5 Discussion

To our knowledge, this is the first systematic review and meta-analysis examining the impact of maternal education level on breastfeeding prevalence in China. The results of the present study showed that the odds of breastfeeding of mothers who had 'more than six years' or 'more than twelve years' education was reduced by 10% or 9% respectively, compared to mothers obtaining 'six years or less' or 'twelve years or less' education, respectively.

These results differ from published results of studies carried out in other countries, where a positive association between maternal education and breastfeeding prevalence (Al-Sahab et al., 2010; Barria et al., 2008; Baxter et al., 2009; Esteves et al., 2014) has frequently been reported, indicating higher maternal education is associated with higher odds of breastfeeding. By contrast the majority of studies carried out in China reported a negative or non-significant association between maternal education and breastfeeding prevalence regardless the association was assessed independently (Jiang et al., 2013; Wu & Qiu, 2015; Ye, 2008) with no adjustment of confounders or with an adjustment for potential confounders (Guo et al., 2013; Huang et al., 2012; Tang, 2013; Wang et al., 2013; Xu et al., 2007; Ye et al., 2007). Esteves et al. (2014) discussed the findings of one Brazilian study which also reported that higher schooling was a risk factor for delayed breastfeeding initiation within the first hour after birth in Brazil (Silveira et al., 2008). They speculated that this 'negative' association was caused by not adjusting for the potential impact of other variables on breastfeeding. However, based on our present systematic review, several studies from China concluded that the negative association between maternal education and breastfeeding remained even after adjusting potential confounders (Huang et al., 2012; Ye et al., 2007), similar to the findings of the Brazilian study.

Maternal occupation and maternity leave are two commonly recognised important factors, which may potentially confound or mediate the effect of the maternal education on breastfeeding prevalence (Baxter, 2008; Jiang et al., 2012; Tan, 2011). Mothers who have higher education normally have better paid but heavy workload full-time job in China. As these mothers are paid well and then have the financial resource to afford purchasing foreign imported infant formula, which is much demanding and expensive after the melamine contamination affair in 2008, for their babies rather feeding breastmilk (Gong & Jackson, 2012; Guan et al., 2009; Parry, 2008). Furthermore, highly educated mothers are usually working full time in a professional or executive position and tend to have shorter breastfeeding duration (Kang et al., 2013; Wu & Qiu, 2015). In China, if workplaces do not provide a supportive environment, breastfeeding in public spaces is doubtful to be encouraged or even accepted and full-time working mothers may choose ceasing breastfeeding. At present, female employees in China are entitled to 98 days of maternity leave and among the 98 days, 15 may be taken before giving birth, which means that mothers have to return to work around 3 months after delivery (Liu et al., 2009). In that case, mothers who attained higher education are more likely to return to formal employment after birth compared to the mothers with lower education level.

In addition, to some extent, some incorrect traditional perceptions in China also have strong adverse influence on breastfeeding practices (Xu et al., 2009). One popular typical incorrect traditional perception about breastfeeding among mothers is that breastfeeding could impact on mothers' health or change mothers' body figure (Chen & Ji, 1993). Therefore, some mothers, especially those who have higher education or higher income and pay more attention to their health or body figure after giving birth, are willing to choose infant formula to feed their babies instead of breast milk.

With the short maternity leave that is available, a busy and heavy workload, and a less supportive, even unaccepting attitude to breastfeeding in public, as well as some incorrect traditional perceptions on breastfeeding, all these together might cause a number of better educated mothers to choose to rely on infant formula rather than breastfeeding.

Generally, it is recommended that breastfeeding women have access to midwives who are capable of providing information and support for breastfeeding (Australian Nursing & Midwifery Federation, 2015). However, in China, there is a severe shortage of midwives, with on average only around 3 midwives available for every 100,000 people, which will be exacerbated now that the one child policy has been cancelled (Wang, 2015). The mothers who have attained high education levels usually have a better paid job which allows them the possibility of access to midwives' information and support. A skilled midwives training program has been launched by The National Health and Family Planning Commission in 2014 (Wang, 2015), which may become a novel and effective breastfeeding promotion measure aiming at mothers with high educational status but low breastfeeding prevalence. In addition to the midwife workforce, primary health care which provides maternal care from prenatal to postpartum period needs to ensure the efficient delivery of high-quality breastfeeding promotion to highly educated mothers with suboptimal breastfeeding practice.

The present systematic review retrieved both English and Chinese literature on associations between maternal education and breastfeeding prevalence to reduce reporting bias. In addition, different cut-off points were used to improve the sensitivity of pooled effect size estimate.

However, there are several limitations of this study should be taken into account when considering the results of the study. Firstly, the majority of studies eligible for our systematic review only reported crude (unadjusted) odds ratios without adjustment for any confounders. Due to the limited number of studies that provided adjusted odds ratio, the crude odds ratios were also included in the meta-analysis, which might lead to some concerns about the accuracy and precision of the pooled effect size estimation due to the missing of adjustment for possible confounding effects. If future studies conducted in China could address the possible confounding effects, new investigations controlling for confounders such as other socioeconomic status, health status, ethnicity, residence area and maternity leave are recommended for uncovering the more accurate adjusted relationship between maternal education and breastfeeding. Secondly, no evidence showed that publication bias or small study effect was present in the meta-analysis for Group 1, however, Egger's test and funnel plot both suggested publication bias or small study effect were present in the analysis for Group 2.

In conclusion, the results of our study show a negative association between the maternal education and breastfeeding prevalence within 12 months postpartum. Mothers who have attained a higher level of education are less likely to breastfeed their babies compared to mothers with lower education levels in the context of the Chinese culture and employment environment. Intervention programs with a major focus on the factors related to the effect of maternal education on breastfeeding discussed in this study should be trialled to improve breastfeeding practices in China. These programs would need to promote breastfeeding in general but would give some priority to special groups, including better educated mothers. A further priority is to promote breastfeeding supportive workplaces. The role of the midwife in breastfeeding is essential in promoting and supporting breastfeeding mothers, especially help mothers to practice breastfeeding in a professional manner. Therefore, enhancing community acceptability of breastfeeding integrated with the supportive role of midwives is also an important aspect. Future research should evaluate intervention studies to promote breastfeeding by mothers of all levels of education, including those who are better educated.

3.6 References

Al-Sahab, B., Lanes, A., Feldman, M., & Tamim, H. (2010). Prevalence and predictors of 6-month exclusive breastfeeding among Canadian women: a national survey. *BMC Pediatr*, *10*, 20.

Australian Nursing & Midwifery Federation. (2015). Promoting breastfeeding Retrieved from http://anmf.org.au/documents/policies/P_Promoting_Breastfeeding.pdf

Barria, R.M., Santander, G., & Victoriano, T. (2008). Factors associated with exclusive breastfeeding at 3 months postpartum in Valdivia, Chile. *J Hum Lact, 24*(4), 439-445.

Bax, L., Ikeda, N., Fukui, N., Yaju, Y., Tsuruta, H., & Moons, K.G.M. (2009). More than numbers: the power of graphs in meta-analysis. *American Journal of Epidemiology, 169*(2), 249-255.

Baxter, J. (2008). Breastfeeding, employment and leave: An analysis of mothers in growing up in Australia. *Family Matters, 80*, 17-26.

Baxter, J., Cooklin, A.R., & Smith, J. (2009). Which mothers wean their babies prematurely from full breastfeeding? An Australian cohort study. *Acta Paediatr, 98*(8), 1274-1277.

Bernaix, L.W., Beaman, M.L., Schmidt, C.A., Harris, J.K., & Miller, L.M. (2010). Success of an educational intervention on maternal/newborn nurses' breastfeeding knowledge and attitudes. *J Obstet Gynecol Neonatal Nurs, 39*(6), 658-666.

Bhandari, N., Bahl, R., Mazumdar, S., Martines, J., Black, R.E., Bhan, M.K., et al. (2003). Effect of community-based promotion of exclusive breastfeeding on diarrhoeal illness and growth: a cluster randomised controlled trial. *Lancet, 361*(9367), 1418-1423.

Chen, J., & Ji, P. (1993). The process to promote breastfeeding in Beijing. *Maternal and Child Health Care of China,*(3), 27-28. Chinese.

Chin, A.C., Myers, L., & Magnus, J.H. (2008). Race, education, and breastfeeding initiation in Louisiana, 2000-2004. *J Hum Lact, 24*(2), 175-185.

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Control Clin Trials, 7*(3), 177-188.

Dubois, L., & Girard, M. (2003). Social determinants of initiation, duration and exclusivity of breastfeeding at the population level: the results of the Longitudinal Study of Child Development in Quebec (ELDEQ 1998-2002). *Can J Public Health, 94*(4), 300-305.

Egger, M., Smith, D.G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*(7109), 629-634.

Esteves, T.M., Daumas, R.P., Oliveira, M.I., Andrade, C.A., & Leite, I.C. (2014). Factors associated to breastfeeding in the first hour of life: systematic review. *Rev Saude Publica, 48*(4), 697-708.

Gillman, M.W., Rifas-Shiman, S.L., Camargo, C.A., Berkey, C.S., Frazier, A.L., Rockett, H.R.H., et al. (2001). Risk of overweight among adolescents who were breastfed as infants. *Jama-Journal of the American Medical Association*, *285*(19), 2461-2467.

Gong, Q., & Jackson, P. (2012). Consuming anxiety? *Food, Culture & Society, 15*(4), 557-578.

Guan, N., Fan, Q., Ding, J., Zhao, Y., Lu, J., Ai, Y., et al. (2009). Melamine-contaminated powdered formula and urolithiasis in young children. *N Engl J Med, 360*(11), 1067-1074.

Guo, S., Fu, X., Scherpbier, R.W., Wang, Y., Zhou, H., Wang, X., et al. (2013). Breastfeeding rates in central and western China in 2010: implications for child and population health. *Bull World Health Organ, 91*(5), 322-331.

Guyatt, G.H., Oxman, A.D., Vist, G., Kunz, R., Brozek, J., Alonso-Coello, P., et al. (2011). GRADE guidelines: 4. Rating the quality of evidence-study limitations (risk of bias). *Journal of Clinical Epidemiology, 64*(4), 407-415.

Heath, A.L., Tuttle, C.R., Simons, M.S., Cleghorn, C.L., & Parnell, W.R. (2002). A longitudinal study of breastfeeding and weaning practices during the first year of life in Dunedin, New Zealand. *J Am Diet Assoc, 102*(7), 937-943.

Heck, K.E., Braveman, P., Cubbin, C., Chavez, G.F., & Kiely, J.L. (2006). Socioeconomic status and breastfeeding initiation among California mothers. *Public Health Rep, 121*(1), 51-59.

Higgins, J.P., Altman, D.G., Gotzsche, P.C., Juni, P., Moher, D., Oxman, A.D., et al. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ, 343*, d5928.

Higgins, J.P., & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions* (Vol. 4): John Wiley & Sons.

Higgins, J.P., & Thompson, S.G. (2002). Quantifying heterogeneity in a meta-analysis. *Stat Med, 21*(11), 1539-1558.

Higgins, J.P., Thompson, S.G., Deeks, J.J., & Altman, D.G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*(7414), 557-560.

Hornell, A., Aarts, C., Kylberg, E., Hofvander, Y., & Gebre-Medhin, M. (1999). Breastfeeding patterns in exclusively breastfed infants: a longitudinal prospective study in Uppsala, Sweden. *Acta Paediatrica, 88*(2), 203-211.

Howell, E.A., Bodnar-Deren, S., Balbierz, A., Parides, M., & Bickell, N. (2014). An intervention to extend breastfeeding among black and Latina mothers after delivery. *Am J Obstet Gynecol, 210*(3), 239 e231-235.

Howie, P.W., Forsyth, J.S., Ogston, S.A., Clark, A., & Florey, C.D. (1990). Protective effect of breast feeding against infection. *BMJ, 300*(6716), 11-16.

Huang, H.T., Liang, H., Hu, J.W., Feng, P., Tong, W.J., Zheng, Y.M., et al. (2012). Breastfeeding and associated factors in five areas of China. *Journal of Soochow University Medical Science Edition, 32*(4), 454-458. Chinese.

Ip, S., Chung, M., Raman, G., Chew, P., Magula, N., DeVine, D., et al. (2007). Breastfeeding and maternal and infant health outcomes in developed countries. *Evid Rep Technol Assess (Full Rep),*(153), 1-186.

Jiang, G.F. (2000). Associated factors analysis for breastfeeding *Jiangsu Prev Med, 11*(4), 49-50. Chinese.

Jiang, H., Li, M., Yang, D.L., Wen, L.M., Hunter, C., He, G.S., et al. (2012). Awareness, intention, and needs regarding breastfeeding: findings from first-time mothers in Shanghai, China. *Breastfeeding Medicine, 7*(6), 526-534.

Jiang, Y., Guo, L.N., Zhang, L., Wang, C., & Cui, Y. (2013). Analysis of feeding and its influence factors among rural infants in 4 provinces of Midwest China. *Chin J Health Education, 29*(5), 394-397. Chinese.

Jiang, Z.H., & Li, S. (2008). Survey on factors influencing breastfeeding of neonates. *Maternal and Child Health Care of China, 23*(24), 3495-3496. Chinese.

Kang, Y., Liang, X.H., Liu, Y.X., Chen, J., & Li, T.Y. (2013). Relative factors of breastfeeding cessation in infants before 6-month-old *J Third Mil Med Univ, 35*(16), 1735-1739. Chinese.

Ke, F.F. (1993). A survey on breastfeeding status in a rural area. *Maternal and Child Health Care of China, 8*(2), 54-58. Chinese.

Kehler, H.L., Chaput, K.H., & Tough, S.C. (2009). Risk factors for cessation of breastfeeding prior to six months postpartum among a community sample of women in Calgary, Alberta. *Can J Public Health, 100*(5), 376-380.

Kmet, L.M., Lee, R.C., & Cook, L.S. (2004). Standard quality assessment criteria for evaluating primary research papers from a variety of fields. (Alberta Heritage Foundation for Medical Research Edmonton).

Kramer, M.S., Chalmers, B., Hodnett, E.D., Sevkovskaya, Z., Dzikovich, I., Shapiro, S., et al. (2001). Promotion of Breastfeeding Intervention Trial (PROBIT): a randomized trial in the Republic of Belarus. *JAMA, 285*(4), 413-420.

Lanting, C.I., Van Wouwe, J.P., & Reijneveld, S.A. (2005). Infant milk feeding practices in the Netherlands and associated factors. *Acta Paediatr, 94*(7), 935-942.

Li, L., Li, S., Ali, M., & Ushijima, H. (2003). Feeding practice of infants and their correlates in urban areas of Beijing, China. *Pediatrics International, 45*(4), 400-406.

Li, R., Darling, N., Maurice, E., Barker, L., & Grummer-Strawn, L.M. (2005). Breastfeeding rates in the United States by characteristics of the child, mother, or family: the 2002 National Immunization Survey. *Pediatrics, 115*(1), e31-37.

Lin, L., Ma, R.J., Peng, X.H., & Wen, J.Y. (1997). Investigation of factors influencing exclusive breastfeeding of 4-month infants *Maternal and Child Health Care of China, 12*(2), 93-95. Chinese.

Liu, B., Zhang, Y., & Li, Y. (2009). *Reconciling work and family: issues and policies in China*.

Liu, J., Shi, Z., Spatz, D., Loh, R., Sun, G., & Grisso, J. (2013). Social and demographic determinants for breastfeeding in a rural, suburban and city area of South East China. *Contemporary Nurse: A Journal for the Australian Nursing Profession, 45*(2), 234-243 210p.

Liu, L., & Zhou, W.L. (2013). Present situation investigation of baby breastfeeding within six months in Chengjiaqiao community *Shanghai Medicine, 34*(12), 62-64. Chinese.

Ludvigsson, J.F., & Ludvigsson, J. (2005). Socio-economic determinants, maternal smoking and coffee consumption, and exclusive breastfeeding in 10205 children. *Acta Paediatr, 94*(9), 1310-1319.

Ma, Y.Q., Xu, D., Wang, L., & Wang, Y.P. (2014). Breast feeding studies in Hui nationality infants in Xining and Xi'an areas. *Chin J Child Health Care, 22*(8), 871-896. Chinese.

Miller, R.G. (1974). The jackknife-a review. *Biometrika, 61*(1), 1-15.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med, 151*(4), 264-269, W264.

Oken, E., Patel, R., Guthrie, L.B., Vilchuck, K., Bogdanovich, N., Sergeichick, N., et al. (2013). Effects of an intervention to promote breastfeeding on maternal adiposity and blood pressure at 11.5 y postpartum: results from the Promotion of Breastfeeding Intervention Trial, a cluster-randomized controlled trial. *Am J Clin Nutr, 98*(4), 1048-1056.

Parry, J. (2008). Contaminated infant formula sickens 6200 babies in China. *BMJ, 337*, a1738.

Pisacane, A., Continisio, P., Filosa, C., Tagliamonte, V., & Continisio, G.I. (2012). Use of baby carriers to increase breastfeeding duration among term infants: the effects of an educational intervention in Italy. *Acta Paediatr, 101*(10), e434-438.

Qin, L.L., & Hua, J.F. (2013). Investigtion and analysis of infant feeding within four months among floating population in Suzhou. *Nursing Practices and Research 10*(10), 138-139. Chinese.

Qiu, L.Q. (2008). *Factors that influence breastfeeding initiation and duration in urban, suburban and rural areas of Zhejiang Province, Peoples Republic of China.* (PhD thesis). Curtin University, Perth.

Riva, E., Banderali, G., Agostoni, C., Silano, M., Radaelli, G., & Giovannini, M. (1999). Factors associated with initiation and duration of breastfeeding in Italy. *Acta Paediatr, 88*(4), 411-415.

Ruan, M.J., Ba, L., Wang, Z., & Song, H. (2012). Study on breastfeeding status and the related influencing factors in Chaoyang district of Beijing *Maternal and Child Health Care of China, 27*(22), 3478-3480. Chinese.

Schanler, R.J., Shulman, R.J., & Lau, C. (1999). Feeding strategies for premature infants: beneficial outcomes of feeding fortified human milk versus preterm formula. *Pediatrics, 103*(6 Pt 1), 1150-1157.

Scott, J.A., & Binns, C.W. (1999). Factors associated with the initiation and duration of breastfeeding: a review of the literature. *Breastfeed Rev, 7*(1), 5-16.

Silveira, R.B.d., Albernaz, E., & Zuccheto, L.M. (2008). Fatores associados ao início da amamentação em uma cidade do sul do Brasil. *Rev. bras. saúde matern. infant, 8*(1), 35-43.

Tan, K.L. (2011). Factors associated with exclusive breastfeeding among infants under six months of age in peninsular Malaysia. *International Breastfeeding Journal, 6*, 2-2.

Tang, L. (2013). *A cohort study of infant feeding practices and maternal breastfeeding problems in a rural area of Sichuan Province, PR China.* (PhD thesis). Curtin University, Perth.

Tang, Z.J. (2014). Investigation on current situation and influencing factors of infant breastfeeding in Liwan District, Guangzhou City *Chin J Women and Children Health 5*(1), 73-76. Chinese.

Tian, J.Z., Tian, X.S., & Li, L. (2008). Analysis of impact factors on breastfeeding of 183 mothers before discharge. *Zhejiang Prev Med, 20*(7), 45-55. Chinese.

Wang, C.X., Chen, H., & Liu, Y.Z. (2005). Breastfeeding failure factors analysis and countermeasures. *J Community Med, 3*(9), 11-12. Chinese.

Wang, H.Z. (2010). Status analysis on breastfeeding in 1296 infants. *Practical Preventive Medcine 17*(3), 519-520. Chinese.

Wang, X.D. (2015). Midwife shortage may grow crucial. 2015/05/06. *China Daily*.Retrieved from http://www.chinadaily.com.cn/china/2015-05/06/content_20631508.htm

Wang, Z., Gao.J.K., Zhang, Y., Wang, F.H., Jin, J.M., & Zhu, B.H., et al. (2013). Present status and influencing factors analysis on breastfeeding for infants in Zhejiang Province. *Nursing and Rehabilitation J, 12*(7), 628-630. Chinese.

World Health Organization. (2003). Infant and young child feeding: a tool for assessing national practices, policies and programmes.

World Health Organization. (2008). Indicators for assessing infant and young child feeding practices: part 1: definitions: conclusions of a consensus meeting held 6-8 November 2007 in Washington DC, USA.

Wu, Y., & Qiu, L.Q. (2015). An analysis on the status and influencing factors of breastfeeding among infants wihin six months of age. *Zhejiang Prev Med, 27*(3), 245-248. Chinese.

Xiao, D.Q. (2001). The analysis of the condition on breast-feeding of newborn infants. *Chin J Nat Med, 3*(2), 97-98. Chinese.

Xiong, H.J., Yang, W.P., & Jin, L.N. (2006). Investigation of breastfeeding status. *Chin J Gen Pract, 5*(12), 747-749. Chinese.

Xu, F., Binns, C., Yu, P., & Bai, Y. (2007). Determinants of breastfeeding initiation in Xinjiang, PR China, 2003-2004. *Acta Paediatr, 96*(2), 257-260.

Xu, F., Qiu, L., Binns, C.W., & Liu, X. (2009). Breastfeeding in China: a review. *Int Breastfeed J, 4*, 6.

Xu, F.L. (2008). *The determinants of breastfeeding initiation and duration in Xinjiang, Peoples Republic of China.* (PhD thesis). Curtin University, Perth.

Yang, Y.L., & Feng, G.Z. (2014). Related factors of breastfeeding *J Mathematical Med, 27*(4), 427-428. Chinese.

Ye, C.E. (2008). Analysis of breastfeeding current situation. *Maternal and Child Health Care of China, 23*(19), 2770. Chinese.

Ye, J.L., Cui, Y., & Pan, X.P. (2007). Study of the breast feeding and the influence factors in Xinjiang rural areas. *Maternal and Child Health Care of China, 22*(17), 2420-2422.Chinese.

Yu, C. (2013). *Breastfeeding and health outcomes in infants who receive continuing care from hospitals or community health centres in Chengdu Sichuan Province, People's Republic of China.* (PhD thesis). Curtin University, Perth.

Yu, H.M., & Song, J.G. (2000). Breastfeeding among 368 infants in Karamay City, Xinjiang, and influential factors thereof *Chin J Health Education, 16*(10), 581-583. Chinese.

Zhang, H.L., Wang, F., & Pu, C.X. (1998). The current situation of breast-fed infants in the urban community in Wuhan and the affected factors. *Medicine and Society 11*(3), 28-30. Chinese.

Table 3-1 Characteristics of studies (n=31) assessing the association between maternal education and breastfeeding in China

Published year	First author	Study design	Location	Sample size	Other factors associated with breastfeeding outcomes	Breastfeeding outcome measured	Definition	Score
1993	KE, F.F	Cross-sectional	Quanzhou	459	Rooming in, breastmilk initiation, income, nutrition during lactation, gender, disease, reproductive plan, menstruation resumed	EBF at 4mo postpartum	WHO	0.73
1997	LIN, L	Cross-sectional	Guangdong	418	Fatigue after delivery, rooming in, lack of breastfeeding knowledge and skills, income per capita, maternity leave	EBF at 4mo postpartum	Non-WHO	0.64
1998	ZHANG, H.L	Cross-sectional	Wuhan	344	Maternal occupation, rooming in, breastfeeding on demand	FB at 4mo postpartum	WHO	0.68
2000	YU, H.M	Cross-sectional	Karamay	368	Maternity leave, ethnicity	FB within 4mo postpartum	WHO	0.82
2000	JIANG, G.F	Cross-sectional	Liuhe	736	Rooming in, delivery method	EBF within 6mo postpartum	WHO	0.77

2001	XIAO, D.Q	Cross-sectional	Beijing	286	Delivery method	EBF within 28 days postpartum	WHO	0.68
2003	LI, L.B	Cross-sectional	Beijing	251	Maternal antenatal feeding plan	Breastfeeding during 6-12mo postpartum	WHO	0.95
2005	WANG, C.X	Retrospective cohort	Jinan	622	Delivery method, health education, milk powder promotion, initiation time	EBF at 4mo postpartum	Non-WHO	0.77
2006	XIONG, H.J	Cross-sectional	Beijing	146	NA	EBF at 1,2,3,4mo postpartum	WHO	0.59
2007	YE, J.L	Cross-sectional	Xinjiang	2076	Ethnicity, maternal age, boiled water drinking	EBF at 4mo postpartum	WHO	0.95
2008	XU, F.L	Prospective cohort	Xinjiang	1136	Giving breastmilk as the first feed, feeding on demand, mother feeling given enough information about breastfeeding, minority ethnic group, giving birth in spring or summer, medical staff not recommending formula to parents,	AF initiation/ EBF initiation	WHO	1.00

					prelacteal feeds of water or formula			
2008	TIAN, J.Z	Cross-sectional	Zhejiang	253	Neonate disease, early sucking, initiation time, breastfeeding confidence, delivery method	FB before discharge	WHO	0.73
2008	YE, C.E	Cross-sectional	Ninghai	975	Delivery method, health education	EBF at 4mo postpartum	Non-WHO	0.68
2008	JIANG, Z.H	Retrospective cohort	Harbin	310	Initiation time, delivery method	EBF at 42d postpartum	Non-WHO	0.59
2008	QIU, L.Q	Prospective cohort	Zhejiang	1511	Living in the suburb and rural areas, maternal age, mother decides to breastfeed until after birth, prelacteal feeding	EBF initiation/ at discharge	WHO	1.00
2010	WANG, H.Z	Cross-sectional	Changli	1296	Health education, delivery method	EBF at 6mo postpartum	WHO	0.64
2012	RUAN, M.J	Cross-sectional	Beijing	101	Delivery method, fixed-term job, maternal age	EBF at 4mo postpartum	Non-WHO	0.73

2012	HUANG, H.T	Cross-sectional	5 provinces	2522	Infant gender, ethnicity, region, premature birth, low birth mass, pregnancy disease, maternal age, history of abortion, history of induced labour, parity	EBF within 4mo postpartum	Non-WHO	0.95
2013	JIANG, Y	Cross-sectional	4 provinces	1272	Health education, ethnicity, low birth weight	EBF within 6mo postpartum	WHO	0.95
2013	LIU, L	Cross-sectional	Shanghai	210	Household income monthly, maternal occupation	EBF within 6mo postpartum	WHO	0.73
2013	KANG, Y	Cross-sectional	Chongqing	938	Infant gender, delivery method, income monthly, birth weight, duration of maternal leave, perception of breastmilk amount, prelacteal feeding	Breastfeeding at 6mo postpartum	Non-WHO	0.91
2013	WANG, Z	Cross-sectional	Zhejiang	528	Infant age, infant gender, early feeding, perception of breastmilk amount, delivery method	EBF within 6mo postpartum	WHO	0.86
2013	GUO, S.F	Cross-sectional	26 counties	2293	Maternal antenatal clinic visit, child's age, delivery method	EBF within 6mo postpartum	WHO	0.91

2013	LIU, J.H	Cross-sectional	Jintan	1280	Delivery method, gender, grandmother's education, maternal occupation, main caregiver, birth order	EBF within 6mo postpartum	Non-WHO	0.91
2013	QIN, L.L	Retrospective cohort	Suzhou	612	Maternal occupation, maternal age, birth region, breastfeeding professional instruction, delivery method	EBF at 4mo postpartum	WHO	0.91
2013	TANG, L	Prospective cohort	Jiangyou	695	Encouragement from facility staff, paternal feeding preference, time of deciding feeding method, maternal grandmother's feeding preference	AF at discharge/ FB at discharge	WHO	1.00
2013	YU, C	Prospective cohort	Chengdu	845	Maternal occupation, paternal education, intention of going back to work, first feeding, mothers' friends breastfeed their babies, paternal job, staff encouragement, father's attitude, maternal	AF within 15 d postpartum/ FB within 15d postpartum	WHO	1.00

					grandmother's breastfeeding history, delivery method			
2014	TANG, Z.J	Cross-sectional	Guangzhou	315	Delivery method, household income per capita	BF at 6mo postpartum	WHO	0.59
2014	MA, Y.Q	Retrospective cohort	Xining & Xiaan	502	NA	EBF at 4mo postpartum	Non-WHO	0.86
2014	YANG, Y.L	Cross-sectional	Wuhan	513	Health education, delivery method, prenatal high risk factors	EBF within 6mo postpartum	Non-WHO	0.50
2015	WU, Y	Cross-sectional	Yongkang	667	Maternal age, maternal occupation, household income monthly	FB at 3mo postpartum	WHO	0.91

FB: full breastfeeding; EBF: exclusive breastfeeding; AF: any breastfeeding

Score: quality assessment score according to the formal checklist developed by Kmet et al. (2004)

Table 3-2 Assessment of risk of bias across studies (n=31) included in the systematic review

Published year	Study	Incomplete outcome data(Attrition bias)	Selective reporting (Reporting bias)	Lack of adjustment for baseline characteristics	Failure to develop and apply appropriate eligibility criteria	Flawed measurement of both exposure and outcome
1993	KE,F.F	+	+	-	?	+
1997	LIN,L	+	-	-	+	+
1998	ZHANG,H.L	+	+	-	?	+
2000	YU,H.M	+	-	-	+	+
2000	JIANG,G.F	+	+	-	+	+
2001	XIAO,D.Q	+	+	-	?	?
2003	LI,L.B	+	+	+	+	+
2005	WANG,C.X	+	+	-	+	+
2006	XIONG,H.J	+	-	-	+	+
2007	YE,J.L	?	+	+	+	+
2008	XU,F.L	+	+	+	+	+
2008	TIAN,J.Z	+	+	-	?	?
2008	YE,C.E	-	+	-	?	?
2008	JIANG,Z.H	+	?	-	?	?
2008	QIU,L.Q	+	+	+	+	+
2010	WANG,H.Z	+	+	-	?	+
2012	RUAN,M.J	-	+	-	+	+
2012	HUANG,H.T	+	+	+	+	+
2013	JIANG,Y	+	+	+	+	+
2013	LIU,L	-	+	-	?	?
2013	KANG,Y	+	+	+	+	+
2013	WANG,Z	-	+	+	?	?
2013	GUO,S.F	?	+	+	+	+
2013	LIU,J.H	-	?	+	+	+
2013	QIN,L.L	-	?	-	?	?
2013	TANG,L	+	+	+	+	+
2013	YU,C	+	+	+	+	+
2014	TANG,Z.J	+	?	-	?	?
2014	MA,Y.Q	+	-	-	+	+
2014	YANG,Y.L	+	?	-	?	?
2015	WU,Y	+	?	+	+	+

+	Low risk of bias
-	High risk of bias
?	Unknown risk of bias

Chapter 4: Modelling longitudinal breastfeeding data with time-dependent covariates

The content of this chapter is covered by a published paper "Zhao, J., Zhao, Y., Binns, C. W., & Lee, A. H. (2016). Increased Calcium Supplementation Postpartum Is Associated with Breastfeeding among Chinese Mothers: Finding from Two Prospective Cohort Studies. Nutrients, 8(10), 622.". DOI: https://doi.org/10.3390/nu8100622. (see Appendix E.1)

The statement of primary contribution of the first author and the permission to include the publication in this thesis can be found in the Appendix A. This paper was published in Nutrients under an open access license called "CC-BY", thus permissions to reuse or reproduce the paper from the publisher are not required.

4.1 Time-dependent nature in longitudinal breastfeeding studies

Longitudinal studies refer to studies where observations of participants/ subjects are repeatedly measured over a follow-up period. Longitudinal data that arise from the repeated measurements are commonly used for assessing outcome trends over time or comparing outcomes between different populations accounting for time effects (Diggle et al., 2002; Fitzmaurice et al., 2011; Zeger & Liang, 1992). Such repeated measurements cause correlations within subjects, and thus when modelling longitudinal data to examine the association between exposures and health outcomes, it is necessary to properly account for the intra-subject correlation to obtain statistically valid analysis and inference (Diggle et al., 2002; Fitzmaurice et al., 2011).

In longitudinal studies, exposure variables are generally measured with the outcomes at baseline and different time points during the follow-up. As a result, except the intra-subject correlation between measurements, another issue associated with longitudinal data is the variation of the values of the exposure variables over time. Some exposure variables change their values over the study period while some of them may keep unchanged in values. Accordingly, in the analysis of longitudinal data, covariates are classified into two categories, namely, time-independent covariates (also known as time-invariant covariates) and time-dependent covariates (also known as time-varying covariates). Time-independent covariates refer to those covariates which values do not change over time, such as gender (in most instances), birth delivery method, born region, blood type and genotype. By contrast, values of time-dependent covariates usually vary over time, for instance, lifestyle related factors including physical activity status, smoking status, alcohol consumption, and dietary intake.

In longitudinal breastfeeding studies, time-dependent covariates arise often. For example, in a longitudinal study where breastfeeding rates over 12 months postpartum are evaluated as outcomes, the values of maternal body mass index, maternal smoking status and maternal medicine use postpartum could vary over time. As another example, in a study where breastfeeding is assessed its association with short-term outcomes such as infant hospitalization or infant growth, the status of breastfeeding can change over time and hence it can be defined as a time-dependent covariate. Unfortunately, there are very few longitudinal breastfeeding studies taking the time-dependent nature into account in the analysis.

4.2 Modelling longitudinal data with time-dependent covariates

Most longitudinal data can be analysed using two classes of models: conditional models and marginal models (Fitzmaurice et al., 2011; Zeger & Liang, 1992). Conditional models are typically referred to as subject-specific models or mixed-effects models, in which estimated coefficients have subject-specific interpretations. Generalised linear mixed-effects model (GLMM) is widely used as a conditional approach in longitudinal data analysis. Marginal models are population-averaged models, where estimated coefficients have population-averaged interpretations. The generalised estimating equations (GEE) and the generalised method of moments (GMM) are two most commonly used marginal models in longitudinal data analysis. For a continuous response, the coefficient estimates of conditional models and marginal models are numerically equivalent (Fitzmaurice et al., 2011). Therefore, in the following section, we focus on modelling binary longitudinal data with time-dependent covariates.

4.2.1 Subject-specific modelling

Let Y_{ij} denote a binary outcome variable corresponding to the jth observation of the ith subject in a subject-specific longitudinal model (random-intercept logistic model) as follows:

$$\text{logit}[P(Y_{ij}=1 \mid \mathbf{X}_{ij}, \mathbf{Z}_i, u_i)] = \beta_0 + \mathbf{X}_{ij}{}^T \boldsymbol{\beta}_1 + \mathbf{Z}_i{}^T \boldsymbol{\beta}_2 + u_i$$

where $\mathbf{X}_{ij}$ indicates a vector of time-dependent covariates collected from the jth observation of the ith subject, $\mathbf{Z}_i$ is a vector of time-independent covariates collected from the ith subject, and u_i is the random effect of the ith subject. In general, the random effects u_i are assumed to follow a normal distribution with mean zero and variance σ^2. Regression coefficients $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ represent the influence of covariates (time-dependent and time-independent covariates, respectively) on the mean response of a specific subject.

Unlike linear mixed-effects models, where a closed form of the likelihood function can be specified, the maximum likelihood (ML) and restricted maximum likelihood (REML) are two common methods for the parameter estimation, numerous approximation methods have been proposed for the parameter estimation of generalised linear mixed-effects model (GLMM). Wolfinger and O'Connell (1993) proposed a pseudo-likelihood approach, Breslow and Clayton (1993) developed a penalised quasi-likelihood approach (PQL) and Pinheiro and Bates (1995) proposed adaptive Gaussian quadrature. Thorough reviews of the statistical inference of GLMM can be found in McCulloch and Searle (2001), Hedeker and Gibbons (2006) and Lee et al. (2006).

4.2.2 Population-averaged modelling

Generalised estimating equations (GEE) and generalised method of moments (GMM) are two most commonly used population-averaged models. GEE approach was proposed by Liang and Zeger (1986) and Zeger and Liang (1986). It is an extension of generalised linear models (GLM) with accounting for the dependency between the repeated measurements in longitudinal studies (McCulloch & Searle, 2001). A logistic regression model using GEE approach is as follows:

$$\text{logit}[P(Y_{ij}=1\,|\,\mathbf{X}_{ij},\mathbf{Z}_i)]=\beta_0+\mathbf{X}_{ij}{}^T\boldsymbol{\beta}_1+\mathbf{Z}_i{}^T\boldsymbol{\beta}_2$$

where Y_{ij} denotes a binary outcome variable corresponding to the jth observation of the ith subject, $\mathbf{X}_{ij}$ is a vector of time-dependent covariates measured at the jth time point for the ith subject and $\mathbf{Z}_i$ indicates a vector of time-independent covariates for the ith subject. In GEE approach, a "working" correlation matrix is specified to account for the correlation between repeated measurements. Different "working" correlation structures can be specified and typically used are independent, exchangeable, autoregressive and unstructured.

When repeated measurements are assumed uncorrelated, an independent working correlation structure is used:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Uniform correlations are assumed by an exchangeable structure:

$$\begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix}$$

A first order autoregressive (AR) process is assumed by an autoregressive structure (AR1):

$$\begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix}$$

Unconstrained pairwise correlations are assumed by an unstructured matrix:

$$\begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix}$$

where $\rho_{ij} = corr(Y_{ij}, Y_{ik})$ for outcomes of the ith subject measured at the jth time point and the kth time point.

The quasi-likelihood estimates of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are obtained from robust or sandwich inference while maximum likelihood inference is not available (Liang & Zeger, 1986). GEE approach produces efficient estimates for time-independent covariates if the working correlation structure is specified correctly. In addition, it provides consistent estimates for time-independent covariates even if the assumed working correlation is misspecified (Liang & Zeger, 1986). Therefore, GEE can be used to model correlated outcomes with time-independent covariates appropriately with arbitrary working correlation structures. However, for time-dependent covariates, the consistency of parameter estimates using GEE is not assured with misspecified working correlation structures (Hu et al., 1998; Pepe & Anderson, 1994).

The generalised method of moments (GMM), on the other hand, generalises the standard method of moments (MM) by allowing the number of moment conditions to be greater than the number of parameters. GMM uses an assumption about specific moments of random variables instead of the assumption about the entire distribution (which would be required in the maximum likelihood framework) (Hansen, 1982). Thus, GMM is more robust than the maximum likelihood (ML) and can be adapted to a wide variety of applications (Hansen, 2010). Lai and Small (2007) developed a new marginal regression approach using GMM to analyse longitudinal data with time-dependent covariates and showed substantial gains in efficiency over GEE with the independent working correlation.

4.3 Modelling binary longitudinal breastfeeding data with time-dependent covariates: the association between calcium supplementation postpartum and breastfeeding

4.3.1 Background

The mineral accretion rate of a neonate reaches about 30–40 mg/kg per day, while calcium transfer between mothers and infants is on average 210 mg per day (Kalkwarf et al., 1999; Olausson et al., 2012; Trotter & Hixon, 1974). For babies who are breastfed exclusively through the first 6 months, the amount of mineral demand from the mothers is four times greater than that during 9 months of pregnancy (Kovacs & Ralston, 2015). The calcium requirement of mothers during lactation has been the subject of much discussion (Kovacs, 2011; Prentice, 2000; Thomas & Weisman, 2006). In 2011, the Institute of Medicine published the calcium dietary reference intakes by life stage, in which Estimated Average Requirement (EAR) of calcium for pregnant and lactating adult women is recommended as 800 mg (Institute of Medicine, 2011).

Compared to western countries, the lower consumption of dairy products in China results in most Chinese residents having a calcium intake lower than the adequate intake (AI) (Chen et al., 2007; Ma et al., 2007; Wang & Li, 2008). In a prospective cohort study of women’s health from Shanghai, the median intake of calcium was 485 mg/day, 60% of calcium from plant sources, and only 20% from milk, which was lower than the age group specific AI (800 mg/day for 18–49 years group and 1000 mg/day for over 50 years group) (Shin et al., 2006; Wang & Li, 2008). Only 6.25% of perimenopausal women reached the standard of calcium intake in Changsha (Deng et al., 2008). The average intake of calcium of Beijing elderly was 505 mg/day, which was about one half of the recommended adequate intake for the elderly (Liu et al., 2004). In the National Nutrition and Health Survey of 2002, fewer than 5% reached the adequate intake levels of calcium for all age groups and the prevalence of calcium supplementation during pregnancy was 41.4% (He et al., 2007; Lai et al., 2007). Besides cultural preferences, the lower consumption of dairy products in China is attributed to the high rate of lactose intolerance, which is around 80% to 95% (De Vrese et al., 2001; Wang et al., 1984).

The Chinese National Health and Family Planning Commission recommends that pregnant women should have a dietary calcium intake of 1000 mg per day from the second trimester and increase to 1200 mg per day from the third trimester until the end of lactation (National Health and Family Planning Commission of the People's Republic of China, 2014). However, low dietary calcium intake in lactating women has been reported in different regions of China, as shown in Table 4-1. This suggests that calcium supplementation for lactating women is an important public health issue to mothers in China based on the current evidence about the benefits of calcium intake during lactation on reducing maternal bone loss (Chan et al., 1987; O'Brien et al., 2012; O'Brien et al., 2006; Yoneyama & Ikeda, 2004).

Table 4-1 Dietary calcium intake of lactating women in different regions of China

Study Location	**Study Design**	**Study Period**	**Average Daily Dietary Calcium Intake (Postpartum)**
Guangzhou (Li et al., 2004)	Prospective cohort	2002	786.45 mg (12 weeks)
Hunan (Huang, 2014)	Cross-sectional	2011–2012	426 mg
Beijing, Suzhou & Guangzhou (Yang et al., 2014)	Cross-sectional	2011–2012	401.4 mg (0–1 month) 585.3 mg (1–2 months) 591.2 mg (2–4 months) 649.0 mg (4–8 months)
Fujian (Chen et al., 2012)	Prospective cohort	2012	428 mg (2 days) 454 mg (7 days) 595 mg (30 days) 544 mg (90 days)
Shanghai (Kong et al., 2016)	Prospective cohort	2014–2015	749.3 mg (1–3 days) 781.1 mg (7–9 days) 762.3 mg (14–17 days) 768.4 mg (25–27 days) 678.5 mg (39–41 days)

The calcium supplementation status during the postpartum period among Chinese lactating women is still unclear. The objective of the present study is to utilise data from two population-based prospective cohort studies to examine the calcium supplementation status and to identify whether breastfeeding is associated with increased calcium supplementation among Chinese mothers after childbirth.

4.3.2 Method

Data source

Two prospective cohort studies were conducted in an urban area, Chengdu (capital city) and a rural area, Jiangyou (county-level city), Sichuan Province, China between 2010 and 2012. Mothers who gave birth to a healthy singleton infant were invited to participate before discharge. These two studies used the same methodology based on same questionnaires, which had been used in Australia and China (Qiu et al., 2008; Scott et al., 2001; Xu et al., 2006) previously, to interview all consented women face-to-face at discharge, and followed up the participants at one, three and six months postpartum by telephone interviews. The baseline interview collected detailed information on mothers and newborns, including socio-demographic, obstetric characteristics and dietary supplements during pregnancy. The follow-up interviews collected detailed information on lactation patterns, durations and dietary supplements during the postpartum period. The World Health Organization (WHO) standard definition of any breastfeeding was used in these two studies; ‘Any breastfeeding’ is defined as the infant has received breast milk (direct from the breast or expressed) with or without other drink, formula or other infant food (World Health Organization, 2008).

In both studies, as shown in Figure 4-1, calcium supplementation and any-breastfeeding status were measured at different time points, namely, 1 month, 3 months and 6 months postpartum. Figure 4-2 shows that any-breastfeeding status varies over the follow-up course. Therefore, the variable of interest, any-breastfeeding status, is a time-dependent covariate.

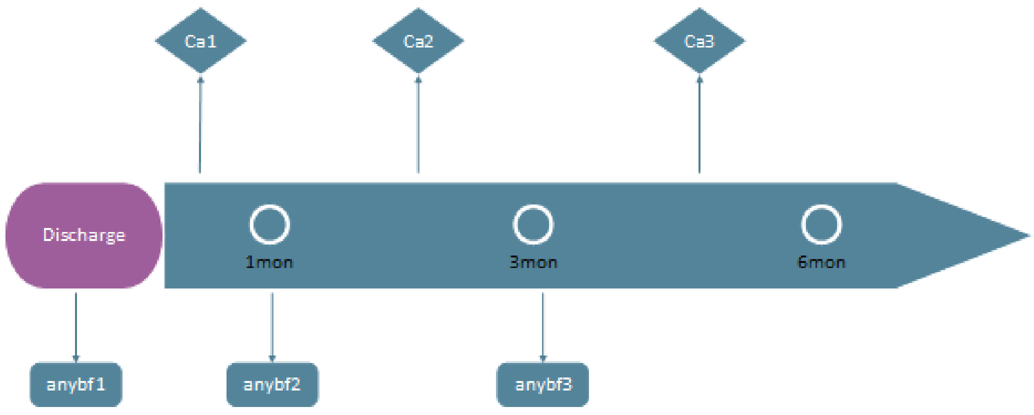


Figure 4-1 Time-dependent covariate (any-breastfeeding) and outcome (calcium supplementation) measurement during the follow-up period

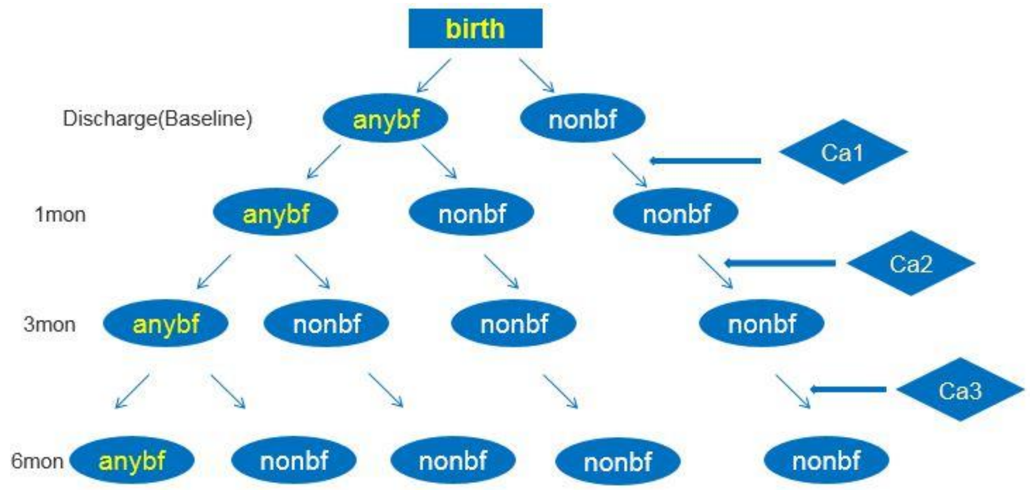


Figure 4-2 Any-breastfeeding status transition over the follow-up course

Ethical approval

The two cohort studies were approved by the Human Research Ethics Committee of Curtin University, Perth, Western Australia (approval numbers: HR169/2009 and HR168/2009, respectively). The present study was also approved by the Human Research Ethics Committee of Curtin University (approval number: RDHS-101-15). The data used in this study were de-identified.

Statistical analysis

As shown in Table 4-2, the outcome of the present study is maternal calcium supplementation status (yes or no) measured longitudinally during three different postpartum periods (from discharge to 1 month, from 1 month to 3 months, and from 3 months to 6 months, respectively) at three follow-up time points (namely, 1 month, 3 months and 6 months postpartum). The main study variable of interest, any breastfeeding status, was measured longitudinally at three different postpartum time points (discharge, 1 month and 3 months postpartum). Descriptive statistics of mothers' socio-demographic status, obstetric characteristics, calcium supplementation during pregnancy and the three postpartum periods, and any breastfeeding status at the three postpartum time points were obtained and reported. Chi-square test was conducted to compare the calcium supplementation rates between breastfeeding group and non-breastfeeding group at the different follow-up time points. Generalised linear mixed model (GLMM) was used to examine the effect of breastfeeding on calcium supplementation postpartum taking into account inherent correlations among repeated measurements, and the breastfeeding status was included as a time-dependent variable. Random intercept model without covariates (Model I) was run initially to test random intercept effect u_i, and then any breastfeeding status at the different time points (denoted as $Breastfeeding_{ij}$) and an indicator variable of measurement times (denoted as $Time_{ij}$) were added into the above Model I to be a Model II. Furthermore, subject level socio-demographic covariates (included in vector $\mathbf{X}_i$) such as annual household income, maternal age and maternal education were then added into and adjusted for in the Model II to formulate a Model III. Finally, obstetric characteristics (also included in vector $\mathbf{X}_i$) such as parity, gravidity, infant gender, infant birth weight and infant gestational week, together with calcium supplementation during pregnancy, were further adjusted for in the Model III to become the final Model IV.

Model I: $\text{logit}[P(Y_{ij}=1|\, u_i)]=\beta_0 + u_i$

Model II: $\text{logit}[P(Y_{ij}=1|u_i)]=\beta_0 + \beta_1 Time_{ij} + \beta_2 Breastfeeding_{ij} + u_i$

Model III & Model IV:

$$\text{logit}[P(Y_{ij}=1|\mathbf{X}_i, u_i)] = \beta_0 + \beta_1 Time_{ij} + \beta_2 Breastfeeding_{ij} + \mathbf{X}_i^T \boldsymbol{\beta} + u_i$$

The above regression analysis was carried out for data set extracted from each cohort study separately, and the results of Model II and final Model IV were reported. In addition, a pooled effect size was calculated using a fixed-effect model given that the heterogeneity between the two studies was tested being statistically nonsignificant. All statistical analyses were performed by using SAS 9.4 (SAS Institute Inc., Cary, NC, USA).

Table 4-2 Variables included in the generalised linear mixed model (GLMM)

	Variables description	Value type	Time-dependent or not?
Outcome variable	Maternal calcium supplementation postpartum	Binary (yes or no)	Yes
Primary variable of interest	Any breastfeeding status	Binary (yes or no)	Yes
Measurement times	Indicator variable	Ordinal	Yes
Socio-demographic covariates	Household annual income, maternal age & maternal education	Categorical	No
Obstetric covariates	Parity, gravidity, infant gender, infant birth weight & gestational week	Categorical	No
Other adjusted covariates	Calcium supplementation during pregnancy	Binary (yes or no)	No

4.3.3 Results

Characteristics of participants

For each cohort, mothers' baseline socio-demographic status, obstetric characteristics and calcium supplementation during pregnancy are presented in Table 4-3. In the

Jiangyou study, 695 mothers were interviewed at baseline, and 648 and 620 mothers remained in the study at 1 month and 3 months postpartum, respectively. Any breastfeeding rate dropped slightly from 93.5% at discharge to 91.1% at 1 month postpartum then continuously to 83.7% at 3 months postpartum. In the other cohort conducted in Chengdu, 845 mothers were interviewed at baseline and 760 mothers were followed up until six months postpartum. Any breastfeeding rate declined from 93.0% at discharge to 87.9% at 1 month postpartum then substantially to 73.4% at 3 months postpartum.

Table 4-3 Characteristics of participants at baseline by breastfeeding status

Variable	Cohort in Jiangyou (*n*=695)		Cohort in Chengdu (*n*=845)	
	BF	Non-BF	BF	Non-BF
Number of participants	650 (93.5)	45 (6.5)	786 (93.0)	59 (7.0)
Household annual income (Chinese yuan)				
<2000	186 (31.0)	9 (23.1)	1 (0.2)	0 (0.0)
2000–5000	309 (51.4)	23 (59.0)	155 (23.5)	12 (24.0)
>5000	106 (17.6)	7 (17.9)	503 (76.3)	38 (76.0)
Maternal age (years)				
<25	373 (57.4)	26 (57.8)	156 (19.9)	5 (8.5)
25-29	163 (25.1)	13 (28.9)	372 (47.3)	28 (47.5)
>29	114 (17.5)	6 (13.3)	258 (32.8)	26 (44.0)
Maternal education				
Secondary school or lower	355 (54.6)	25 (55.6)	90 (11.5)	11 (18.6)
Senior school	215 (33.1)	18 (40.0)	165 (21.0)	11 (18.6)
University or higher	80 (12.3)	2 (4.4)	531 (67.5)	37 (62.8)
Parity				
Primiparous	518 (79.7)	37 (82.2)	700 (89.1)	51 (86.4)
Multiparous	132 (20.3)	8 (17.8)	86 (10.9)	8 (13.6)
Gravidity				
Primigravida	249 (38.3)	18 (40.0)	430 (54.7)	26 (44.1)
Multigravida	401 (61.7)	27 (60.0)	356 (45.3)	33 (55.9)
Infant gender				
Male	328 (50.5)	26 (57.8)	412 (52.4)	34 (57.6)
Female	322 (49.5)	19 (42.2)	374 (47.6)	25 (42.4)
Infant birth weight (g)				
<2500	10 (1.5)	2 (4.4)	13 (1.7)	0 (0.0)
≥2500	640 (98.5)	43 (95.6)	773 (98.3)	59 (100.0)
Infant gestational week				
<37	8 (1.2)	3 (6.8)	9 (1.2)	2 (3.4)
≥37	640 (98.8)	41 (93.2)	777 (98.8)	57 (96.6)
Calcium supplementation during pregnancy				
Yes	410 (63.1)	25 (55.6)	627 (79.8)	47 (79.7)
No	240 (36.9)	20 (44.4)	159 (20.2)	12 (20.3)

Calcium supplementation status postpartum

Overall, among mothers in the Jiangyou cohort, an inverted U shape of calcium supplementation rates at three different postpartum periods was observed, which corresponded to 13.4%, 19.4% and 17.7%, respectively. While in the Chengdu cohort, a constant decline trend was recorded with 22.5%, 22.2% and 12.0% reported at the three postpartum periods. When considering separately for breastfeeding and non-breastfeeding groups, as shown in Figures 4-3 and Figure 4-4, the calcium supplementation rate in the breastfeeding group was statistically significantly higher than that in the non-breastfeeding group for all the different postpartum periods, except between discharge and 1 month in the Jiangyou cohort (p=0.36). In the Jiangyou cohort, calcium supplementation rates ranged from 13.7% to 21.2% for breastfeeding mothers, and ranged from 1.7% to 8.9% for non-breastfeeding mothers. In the Chengdu cohort, calcium supplementation rates reduced from around 23% in the first 3 months postpartum to 14.5% between 3 months and 6 months in breastfeeding mothers, and ranged from 5.0% to 14.1% in non-breastfeeding mothers.

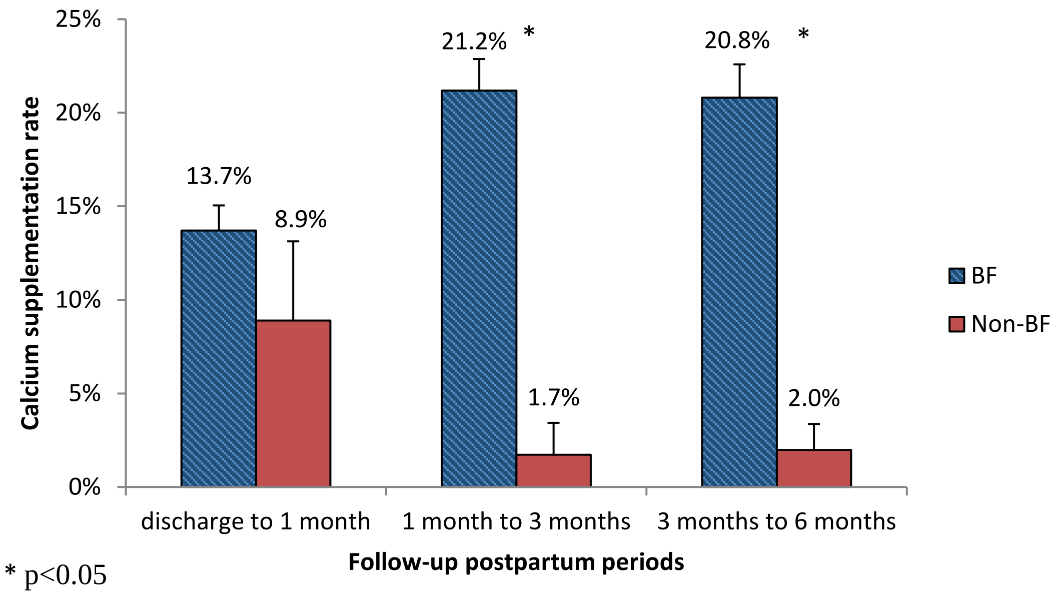


Figure 4-3 Maternal calcium supplementation postpartum in Jiangyou

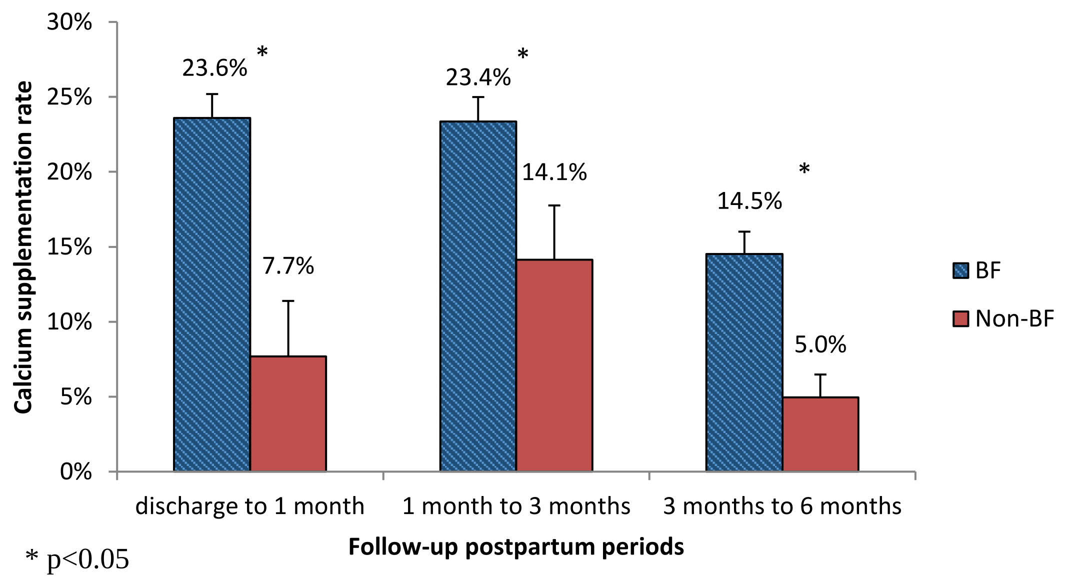


Figure 4-4 Maternal calcium supplementation postpartum in Chengdu

Association between breastfeeding and calcium supplementation postpartum

In Model I (without any covariates) for both cohorts, subject-specific random effect was found to be statistically significant. Hence, both the primary variables of interest (i.e., breastfeeding status) and the indicator variable of measurement times were subsequently added into the Model I for examining the association between breastfeeding and calcium supplementation postpartum. As shown in Table 4-4, the likelihood of calcium supplementation in breastfeeding mothers were 5.85 times (95% confidence interval (CI) (2.50, 13.72)) and 2.88 times (95% CI (1.50, 5.54)) higher of that in non-breastfeeding mothers in Jiangyou and Chengdu, respectively. After adjusting for socio-demographic and obstetric factors as well as calcium supplementation during pregnancy, the odds ratios (ORs) with and its 95% CI had changed appreciably to 6.95 and (2.68, 18.04) in the Jiangyou study, and 3.03 and (1.52, 6.02) in the Chengdu study, respectively. The heterogeneity between these two studies was not significant (I^2=0.479, p=0.17) statistically, therefore a fixed effect model was used to pool the ORs of the two studies. The pooled analysis of these two cohort studies revealed that calcium supplementation postpartum was significantly positively associated with breastfeeding with an adjusted OR=4.02 with a 95% CI of (2.30, 7.03).

Table 4-4 Association between calcium supplementation postpartum and breastfeeding status

Variable	Model II	Model IV
	Crude ORs (95% CI)	Adjusted ORs (95% CI)
Jiangyou Cohort		
Measurement times *		
At discharge (ref)	1	1
1 month	1.72 (1.24, 2.38)	1.90 (1.33, 2.70)
3 months	1.57 (1.12, 2.20)	1.69 (1.18, 2.44)
Breastfeeding status *		
Non-breastfeeding (ref)	1	1
Any breastfeeding	5.85 (2.50, 13.72)	6.95 (2.68, 18.04)
Chengdu Cohort		
Measurement times *		
At discharge (ref)	1	1
1 month	1.02 (0.74, 1.43)	1.02 (0.72, 1.45)
3 months	0.31 (0.21, 0.46)	0.30 (0.20, 0.45)
Breastfeeding status *		
Non-breastfeeding (ref)	1	1
Any breastfeeding	2.88 (1.50, 5.54)	3.03 (1.52, 6.02)
Pooled effect size of two studies		
Non-breastfeeding (ref)	-	1
Any breastfeeding	-	4.02 (2.30, 7.03)

Crude ORs (obtained from Model II): Model included breastfeeding status and the indicator variable of measurement times; Adjusted ORs (obtained from the final Model IV): Model adjusted for socio-demographic variables (household annual income, maternal age and maternal education); obstetric factors (parity, gravidity, infant gender, infant birth weight and infant gestational week); and calcium supplementation during pregnancy; * p<0.05; ref: reference category.

4.3.4 Discussion

To our knowledge, the present study is the first population-based study that determines the longitudinal trend of calcium supplementation by Chinese women from discharge to 6 months postpartum and the effect of breastfeeding on calcium supplementation. A relatively low level of calcium supplementation (less than 23%) was observed throughout the postpartum period in either breastfeeding mothers or non-breastfeeding mothers. The pooled effect size after adjusting for socio-demographic variables (household annual income, maternal age and maternal education); obstetric factors (parity, gravidity, infant gender, infant birth weight and infant gestational week); and calcium supplementation during pregnancy reveals that mothers who breastfed their babies were 4.02 times more likely to take calcium supplements compared to their non-breastfeeding counterparts during postpartum. The present result is consistent with previous findings that breastfeeding mothers consumed more calcium than non-breastfeeding counterparts (Chan et al., 2005; Laskey et al., 1998; Lopez et al., 1996). One reason leading to the higher calcium supplementation in breastfeeding mothers may be the general belief that adequate calcium intake is beneficial to breast milk production, and mothers' special attention to infants' calcium intake under the context of wide shortage of calcium intake for Chinese women, in spite of recent evidence demonstrating that calcium supplementation in lactation has no significant effect on increasing calcium content in breast milk (Jarjou et al., 2006; Kalkwarf et al., 1997; Prentice et al., 1995). The other reason might be mothers' perception of the beneficial effect of calcium supplementation on maternal bone loss during lactating. Some studies found little benefits of calcium supplementation on maternal bone loss during lactating (Cross et al., 1995; Kalkwarf et al., 1997; Zhang et al., 2016), whereas other studies carried out in the U.S. and Brazil suggested that higher calcium intake during early lactation could minimize the bone loss for the mothers who had daily calcium intake less than 500 mg (O'Brien et al., 2012; O'Brien et al., 2006). Further investigation on the factors contributing to the difference of calcium supplementation between breastfeeding mothers and non-breastfeeding mothers as well as the effect of calcium supplementation on reducing maternal bone loss during lactation or enhancing maternal skeleton remodelling and remineralisation after weaning of breastfeeding is recommended.

Given the habitually lower calcium dietary intake and relatively high lactose intolerance rate in the general Chinese population (De Vrese et al., 2001; He et al., 2007; Shin et al., 2006), calcium supplementation plays an important role on bone health, especially for exclusive breastfeeding women who provide around 300 mg of calcium per day to their babies via breast milk which accompany maternal bone calcium turnover (Kovacs, 2016).

This study had several strengths. The existing studies on the association between breastfeeding and calcium supplementation postpartum were mainly conducted in clinical settings with relatively small sample size. Their findings were mainly derived from statistical analysis such as Mann-Whitney U-test (Chan et al, 2005) and repeated-measures analysis of variance (Laskey et al, 1998 and Lopez et al, 1996) without controlling for confounders nor accounting for inherent correlation between repeated measurements. In this population-based study, we utilised data from two cohort studies to investigate the longitudinal trends of calcium supplementation at three different postpartum time points (i.e., 1 month, 3 months and 6 months postpartum) and conducted random effect regression modelling accounting for inherent dependency between the repeated measurements. Moreover, since the breastfeeding status was measured longitudinally as well in two cohorts, it was treated as a time-dependent variable in the analysis to account for possible feedback effects between the breastfeeding status and calcium supplementation at different times. In addition, our pooled analysis based on the two individual studies yielded the combined effect size with a larger sample size and higher statistical power. Furthermore, calcium supplementation during pregnancy was adjusted in the modelling to control for the consequent effect of calcium intake during pregnancy on calcium supplementation during lactation.

A caveat of this study was that both cohort studies were carried out in Sichuan Province, which may limit the results being able to generalise to other regions of China. Sichuan Basin has special geographic characteristics, where the number of cloudy or rainy days is substantially larger than that in other regions in China, which may lead to a relatively lower level of vitamin D synthesis and calcium deficiency consequently (Wang et al., 2010). However, to the best of our knowledge, no data were available currently on calcium supplementation during postpartum in other regions of China for comparison purpose.

In conclusion, calcium supplementation during postpartum in Sichuan is variable at different times during the postpartum period but is always at a relatively low level (less than 23%). Although breastfeeding has a substantive effect on calcium supplementation postpartum, dietary supplementation intervention programs and health education targeting different subgroups (e.g., breastfeeding mothers and bottle-feeding mothers) should be promoted in Chinese women, given a wide shortage of dietary calcium intake and calcium supplementation currently during postpartum.

4.4 References

Breslow, N.E., & Clayton, D.G. (1993). Approximate Inference in Generalized Linear Mixed Models. *Journal of the American Statistical Association, 88*(421), 9-25.

Chan, G.M., McMurry, M., Westover, K., Engelbert-Fenton, K., & Thomas, M.R. (1987). Effects of increased dietary calcium intake upon the calcium and bone mineral status of lactating adolescent and adult women. *Am J Clin Nutr, 46*(2), 319-323.

Chan, S.M., Nelson, E.A., Leung, S.S., & Cheng, J.C. (2005). Bone mineral density and calcium metabolism of Hong Kong Chinese postpartum women--a 1-y longitudinal study. *Eur J Clin Nutr, 59*(7), 868-876.

Chen, H., Wang, P., Han, Y., Ma, J., Troy, F.A., 2nd, & Wang, B. (2012). Evaluation of dietary intake of lactating women in China and its potential impact on the health of mothers and infants. *BMC Womens Health, 12*, 18.

Chen, Y.M., Teucher, B., Tang, X.Y., Dainty, J.R., Lee, K.K., Woo, J.L., et al. (2007). Calcium absorption in postmenopausal Chinese women: a randomized crossover intervention study. *Br J Nutr, 97*(1), 160-166.

Cross, N.A., Hillman, L.S., Allen, S.H., & Krause, G.F. (1995). Changes in bone mineral density and markers of bone remodeling during lactation and postweaning in women consuming high amounts of calcium. *Journal of Bone & Mineral Research, 10*(9), 1312-1320.

De Vrese, M., Stegelmann, A., Richter, B., Fenselau, S., Laue, C., & Schrezenmeir, J. (2001). Probiotics--compensation for lactase insufficiency. *Am J Clin Nutr, 73*(2 Suppl), 421S-429S.

Deng, J., Huang, Y.M., & Lin, Q. (2008). Main sources of dietary calcium in perimenopausal women in Changsha. *Zhong Nan Da Xue Xue Bao Yi Xue Ban*, *33*(9), 875-879.

Diggle, P., Heagerty, P., Liang, K.-Y., & Zeger, S. (2002). *Analysis of longitudinal data* (2nd ed.). Oxford: Oxford University Press.

Fitzmaurice, G.M., Laird, N.M., & Ware, J.H. (2011). *Applied Longitudinal Analysis* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Hansen, L.P. (1982). Large Sample Properties of Generalized-Method of Moments Estimators. *Econometrica, 50*(4), 1029-1054.

Hansen, L.P. (2010). Generalized method of moments estimation. In S. N. Durlauf & L. E. Blume (Eds.), *Macroeconometrics and Time Series Analysis* (pp. 105-118). London: Palgrave Macmillan UK.

He, Y., Zhai, F., Wang, Z., & Hu, Y. (2007). Status of dietary calcium intake of Chinese residents. *Wei Sheng Yan Jiu, 36*(5), 600-602.

Hedeker, D., & Gibbons, R.D. (2006). *Longitudinal data analysis*. New York John Wiley & Sons.

Hu, F.B., Goldberg, J., Hedeker, D., Flay, B.R., & Pentz, M.A. (1998). Comparison of population-averaged and subject-specific approaches for analyzing repeated binary outcomes. *Am J Epidemiol*, *147*(7), 694-703.

Huang, Z. (2014). *The dietary status of 269 lactating women and the content of minerals in breast milk.* MPH Central South University, Changsha.

Institute of Medicine. (2011). *Dietary Reference Intakes for Calcium and Vitamin D*. Washington DC: National Academies Press.

Jarjou, L.M., Prentice, A., Sawo, Y., Laskey, M.A., Bennett, J., Goldberg, G.R., et al. (2006). Randomized, placebo-controlled, calcium supplementation study in pregnant Gambian women: effects on breast-milk calcium concentrations and infant birth weight, growth, and bone mineral accretion in the first year of life. *Am J Clin Nutr, 83*(3), 657-666.

Kalkwarf, H.J., Specker, B.L., Bianchi, D.C., Ranz, J., & Ho, M. (1997). The effect of calcium supplementation on bone density during lactation and after weaning. *N Engl J Med*, *337*(8), 523-528.

Kalkwarf, H.J., Specker, B.L., & Ho, M. (1999). Effects of calcium supplementation on calcium homeostasis and bone turnover in lactating women. *J Clin Endocrinol Metab, 84*(2), 464-470.

Kong, X., Fei, J., Zhai, Y., Feng, Y., & Li, J. (2016). Dietary survey of lactating mothers during the puerperal state. *China Medical Herald*, *13*(14), 49-68.

Kovacs, C.S. (2011). Calcium and bone metabolism disorders during pregnancy and lactation. *Endocrinol Metab Clin North Am*, *40*(4), 795-826.

Kovacs, C.S. (2016). Maternal Mineral and Bone Metabolism During Pregnancy, Lactation, and Post-Weaning Recovery. *Physiol Rev, 96*(2), 449-547.

Kovacs, C.S., & Ralston, S.H. (2015). Presentation and management of osteoporosis presenting in association with pregnancy or lactation. *Osteoporosis International, 26*(9), 2223-2241.

Lai, J., Yin, S., Ma, G., Piao, J., & Yang, X. (2007). The nutrition and health survery of pregnant women in China. *Acta Nutrimenta Sinica, 29*(1), 4-8.

Lai, T.L., & Small, D. (2007). Marginal regression analysis of longitudinal data with time-dependent covariates: a generalized method-of-moments approach. *Journal of the Royal Statistical Society Series B-Statistical Methodology, 69*, 79-99.

Laskey, M.A., Prentice, A., Hanratty, L.A., Jarjou, L.M., Dibba, B., Beavan, S.R., et al. (1998). Bone changes after 3 mo of lactation: influence of calcium intake, breast-milk output, and vitamin D-receptor genotype. *Am J Clin Nutr, 67*(4), 685-692.

Lee, Y., Nelder, J.A., & Pawitan, Y. (2006). *Generalized linear models with random effects : unified analysis via h-likelihood* Boca Raton, FL: Chapman & Hall/CRC.

Li, Y., Liao, D., Miao, L., & Qian, X. (2004). Analysis on dietary investigation of pregnant and lactating women *Chinese Primary Health Care, 18*(2), 27-28.

Liang, K.Y., & Zeger, S.L. (1986). Longitudinal Data-Analysis Using Generalized Linear-Models. *Biometrika, 73*(1), 13-22.

Liu, X., Zhao, X., & Xu, L. (2004). Food sources of calcium and iron in the diet of Beijing elderly. *Wei Sheng Yan Jiu, 33*(3), 336-338.

Lopez, J.M., Gonzalez, G., Reyes, V., Campino, C., & Diaz, S. (1996). Bone turnover and density in healthy women during breastfeeding and after weaning. *Osteoporosis International, 6*(2), 153-159.

Ma, G., Li, Y., Jin, Y., Zhai, F., Kok, F.J., & Yang, X. (2007). Phytate intake and molar ratios of phytate to zinc, iron and calcium in the diets of people in China. *Eur J Clin Nutr, 61*(3), 368-374.

McCulloch, C.E., & Searle, S.R. (2001). *Generalized, linear, and mixed models* New York: New York : John Wiley & Sons.

National Health and Family Planning Commission of the People's Republic of China. (2014). Maternal and child health basic knowledge and skills (trial version) Retrieved from http://www.nhfpc.gov.cn/fys/kpxc/201405/0065ee2071204f0c9b52f7cdcbb392f2.shtml

O'Brien, K.O., Donangelo, C.M., Ritchie, L.D., Gildengorin, G., Abrams, S., & King, J.C. (2012). Serum 1,25-dihydroxyvitamin D and calcium intake affect rates of bone calcium deposition during pregnancy and the early postpartum period. *Am J Clin Nutr, 96*(1), 64-72.

O'Brien, K.O., Donangelo, C.M., Zapata, C.L., Abrams, S.A., Spencer, E.M., & King, J.C. (2006). Bone calcium turnover during pregnancy and lactation in women with low calcium diets is associated with calcium intake and circulating insulin-like growth factor 1 concentrations. *Am J Clin Nutr, 83*(2), 317-323.

Olausson, H., Goldberg, G.R., Laskey, M.A., Schoenmakers, I., Jarjou, L.M., & Prentice, A. (2012). Calcium economy in human pregnancy and lactation. *Nutr Res Rev, 25*(1), 40-67.

Pepe, M.S., & Anderson, G.L. (1994). A Cautionary Note on Inference for Marginal Regression-Models with Longitudinal Data and General Correlated Response Data. *Communications in Statistics-Simulation and Computation, 23*(4), 939-951.

Pinheiro, J.C., & Bates, D.M. (1995). Approximations to the Log-Likelihood Function in the Nonlinear Mixed-Effects Model. *Journal of Computational and Graphical Statistics, 4*(1), 12-35.

Prentice, A. (2000). Calcium in pregnancy and lactation. *Annu Rev Nutr, 20*, 249-272.

Prentice, A., Jarjou, L.M., Cole, T.J., Stirling, D.M., Dibba, B., & Fairweather-Tait, S. (1995). Calcium requirements of lactating Gambian mothers: effects of a calcium supplement on breast-milk calcium concentration, maternal bone mineral content, and urinary calcium excretion. *Am J Clin Nutr, 62*(1), 58-67.

Qiu, L., Zhao, Y., Binns, C.W., Lee, A.H., & Xie, X. (2008). A cohort study of infant feeding practices in city, suburban and rural areas in Zhejiang Province, PR China. *Int Breastfeed J, 3*, 4.

Scott, J.A., Landers, M.C., Hughes, R.M., & Binns, C.W. (2001). Factors associated with breastfeeding at discharge and duration of breastfeeding. *J Paediatr Child Health, 37*(3), 254-261.

Shin, A., Li, H., Shu, X.O., Yang, G., Gao, Y.T., & Zheng, W. (2006). Dietary intake of calcium, fiber and other micronutrients in relation to colorectal cancer risk: Results from the Shanghai Women's Health Study. *Int J Cancer, 119*(12), 2938-2942.

Thomas, M., & Weisman, S.M. (2006). Calcium supplementation during pregnancy and lactation: effects on the mother and the fetus. *Am J Obstet Gynecol, 194*(4), 937-945.

Trotter, M., & Hixon, B.B. (1974). Sequential changes in weight, density, and percentage ash weight of human skeletons from an early fetal period through old age. *Anat Rec, 179*(1), 1-18.

Wang, J., Yang, F., Mao, M., Liu, D.H., Yang, H.M., & Yang, S.F. (2010). High prevalence of vitamin D and calcium deficiency among pregnant women and their newborns in Chengdu, China. *World J Pediatr, 6*(3), 265-267.

Wang, Y., & Li, S. (2008). Worldwide trends in dairy production and consumption and calcium intake: is promoting consumption of dairy products a sustainable solution for inadequate calcium intake? *Food Nutr Bull, 29*(3), 172-185.

Wang, Y.G., Yan, Y.S., Xu, J.J., Du, R.F., Flatz, S.D., Kuhnau, W., et al. (1984). Prevalence of primary adult lactose malabsorption in three populations of northern China. *Hum Genet, 67*(1), 103-106.

Wolfinger, R., & O'Connell, M. (1993). Generalized Linear Mixed Models - a Pseudo-Likelihood Approach. *Journal of Statistical Computation and Simulation, 48*(3-4), 233-243.

World Health Organization. (2008). Indicators for assessing infant and young child feeding practices: part 1: definitions: conclusions of a consensus meeting held 6-8 November 2007 in Washington DC, USA.

Xu, F., Liu, X., Binns, C.W., Xiao, C., Wu, J., & Lee, A.H. (2006). A decade of change in breastfeeding in China's far north-west. *Int Breastfeed J, 1*, 22.

Yang, T., Zhang, Y., Ma, D., Li, W., Yang, X., & Wang, P. (2014). Survey on the nutrients intake of lactating women in three cities of China. *Acta Nutrimenta Sinica, 36*(1), 84-86.

Yoneyama, K., & Ikeda, J. (2004). The effects of increased dietary calcium intake on bone mineral density in long-term lactating women, and recovery of bone loss caused by long-term lactation with low calcium diet. *Nihon Koshu Eisei Zasshi, 51*(12), 1008-1017.

Zeger, S.L., & Liang, K.Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics, 42*(1), 121-130.

Zeger, S.L., & Liang, K.Y. (1992). An overview of methods for the analysis of longitudinal data. *Stat Med, 11*(14-15), 1825-1839.

Zhang, Z.Q., Chen, Y.M., Wang, R.Q., Huang, Z.W., Yang, X.G., & Su, Y.X. (2016). The effects of different levels of calcium supplementation on the bone mineral status of postpartum lactating Chinese women: a 12-month randomised, double-blinded, controlled trial. *Br J Nutr, 115*(1), 24-31.

Chapter 5: Modelling time-to-event breastfeeding data with time-dependent covariates

The content of this chapter is covered by a published paper "Zhao, J., Zhao, Y., Lee, A. H., & Binns, C. W. (2017). A Time-varying Covariate Approach for Survival Analysis of Paediatric Outcomes. Paediatric and perinatal epidemiology, 31(6), 598-602.". DOI: https://doi.org/10.1111/ppe.12410. (see Appendix F.2)

The statement of primary contribution of the first author and the permission to include the publication in this thesis can be found in the Appendix A. The permission to reproduce the material from the publisher can be found in the Appendix F.1.

5.1 Survival analysis

5.1.1 Time-to-event data and censoring

Survival analysis, also known as time-to-event data analysis, failure time analysis or historical event analysis, refers to a class of statistical methods for analysing time-to-event data. Time-to-event, or survival time, is defined as the length of time from a time origin to the occurrence (or failure) of an event of interest (Kalbfleisch & Prentice, 2002; Klein & Moeschberger, 2005; Lee & Wang, 2003). Time-to-event data arise commonly in longitudinally designed studies, in which participants are usually followed-up and time-to-event is thus measured from the entry of study to the occurrence (or failure) of an event of interest, such as survival time of cancer patients after treatment, time to relapse after smoking cessation and exclusive breastfeeding duration.

One distinguishing feature of time-to-event data or survival data is the concept of censoring. Censoring often occurs when the time-to-event cannot be observed for all participants due to loss to follow-up, participants having a different event or early study termination, etc. For example, in a study on the survival time of cancer patients after specific treatment, participants are followed up for a planned period. If a participant drops out from the follow-up, we will not have the information on the true survival time of that participant but knowing that the unobserved survival time will be longer than the time observed in the study. This form of censoring is generally known as right censoring. There are various types of censoring: left censoring, right censoring and interval censoring, among which right censoring is the most commonly encountered type of censoring (Klein & Moeschberger, 2005).

5.1.2 Statistical models used in survival analysis

Statistical modelling strategy for time-to-event data mainly includes three categories: nonparametric, semi-parametric and parametric approach. The classification of these approaches is based on the assumption of the baseline hazard function.

For the nonparametric survival approach, there is no parametric assumption about the form of the baseline hazard distribution. The Kaplan-Meier estimator is a typical nonparametric technique for estimating the survival function. The Nelson-Aalen estimator can be used to estimate the cumulative hazard rate function (Lee & Wang, 2003). Wilcoxon test, Cox-Mantel test, logrank test and Kruskal-Wallis test are commonly used to compare survival distributions between groups (Lee & Wang, 2003). Nonparametric estimation is useful and efficient when the baseline hazard distribution is unknown or hard to specify, however, it is difficult to incorporate covariates as predictors in the nonparametric approach.

In the circumstances, where the primary research interest is on the association between covariates/ predictors and time-to-event outcomes, semi-parametric or parametric regression models are widely used. The major difference between these two types of models is that there is no specification of the baseline hazard function required in semi-parametric regression models. Cox proportional hazards model is the most popular semi-parametric model to quantify the effects of predictors on the time-to-event with a proportional hazards assumption (i.e., a constant hazard ratio over time). In general, a Cox proportional hazards model is given below

$$\lambda(t) = \lambda_0(t)\exp(\mathbf{X}^T\boldsymbol{\beta}),$$

where $\lambda(t)$ refers to the hazard of the event occurring at the time t, $\lambda_0(t)$ is an unspecified baseline hazard function, $\mathbf{X}$ is a vector of covariates and $\boldsymbol{\beta}$ indicates a vector of regression coefficients. Considering a single exposure, the hazard ratio of any two individuals from two distinct groups, say, X_1 and X_1', stratified by two values of the specific exposure, is given as

$$\frac{\lambda(t \mid X_1)}{\lambda(t \mid X_1')} = \frac{\lambda_0(t)\exp(X_1\beta_1)}{\lambda_0(t)\exp(X_1'\beta_1)} = \frac{\exp(X_1\beta_1)}{\exp(X_1'\beta_1)} = \exp[(X_1 - X_1')\beta_1].$$

Thus, hazard ratios in Cox proportional hazards models are approximately constant over time although the hazard rates of subjects in different groups change over time.

In a parametric approach, time-to-event data are assumed to follow a theoretical distribution, and commonly used distributions include exponential, Weibull, log-normal and log-logistic distribution, etc. Parametric survival models can be further classified into proportional hazards (PH) model and accelerated failure time (AFT) model depending on whether the assumption of the proportional hazards satisfies or not. Analogous to Cox proportional hazards model, hazard ratios obtained from parametric proportional hazards model can be used to interpret the hazard difference between two groups. AFT models quantify the effect of covariates on the logarithm of the survival time, which can be interpreted as the effect of a covariate to accelerate or decelerate the survival time scale by a constant rate. Thus, the regression coefficients of AFT models can be more intuitively interpreted with respect to an expected change in median survival time.

5.2 Time-dependent bias

As time-to-event data typically arise from longitudinally designed studies, where observations are repeatedly measured over the follow-up course, time-varying covariates are commonly present in survival analysis. Typical examples of time-dependent variables are infant's height and weight, which increase as the infant grows. Consequently, the associations between these time-varying variables and time-to-event outcomes will change over time, suggesting a necessity to account for time-varying effects in the statistical analysis.

Furthermore, if the time-varying nature of exposures is not taken into account, a “time-dependent bias” can be introduced into survival analysis (Austin & Platt, 2010; Levesque et al., 2010; van Walraven et al., 2004), potentially altering the conclusions of the study (Austin et al., 2006; Levesque et al., 2010; van Walraven et al., 2004). In epidemiology, time-dependent bias is also known as immortal time bias. Immortal time refers to a period of cohort follow-up or observation time, during which, by design, the event of interest cannot occur (Jones & Fowler, 2016; Levesque et al., 2010; Rothman et al., 2008; Suissa, 2008). Due to immortal time (also known as guaranteed survival time), misclassification of exposure and non-exposure groups or treatment and non-treatment groups in statistical analysis can result in biased estimation of exposure/ treatment effect (Austin et al., 2006; Cho et al., 2017; Levesque et al., 2010).

Time-dependent bias occurs when time-varying covariates are treated as time-fixed in survival analysis such as Cox proportional hazards model without considering the time-varying nature. The problem of time-dependent bias was first debated and discussed in the famous Stanford heart transplant study, in which the misclassification of immortal time as exposed to heart transplantation was rectified using a time-varying survival modelling strategy (Mantel & Byar, 1974). Since then, time-dependent bias has been discussed in several areas of epidemiology (Wolkewitz et al., 2012), such as pharmacoepidemiology (Suissa, 2008; Zhou et al., 2005), cancer studies (Soderberg-Naucler et al., 2014), infectious disease (Glesby & Hoover, 1996), chronic obstructive pulmonary disease (Suissa, 2003) and even studies on survival in Oscar winners (Redelmeier & Singh, 2001; Sylvestre et al., 2006; Wolkewitz et al., 2010).

In paediatric studies, including breastfeeding studies, where time-varying exposures are measured, a conventional survival model ignoring the time-varying nature in covariates is still widely used, which may lead to incorrect inferences. Examples of such studies include identifying factors associated with breastfeeding duration (Scott et al., 2006), assessing the relationship between vaccination and childhood diabetes (Rousseau et al., 2016), and determining neonatal factors affecting the development of childhood epilepsy (Whitehead et al., 2006).

As shown in Figure 5-1, we illustrate two different scenarios on occurrence times of an exposure of interest in a breastfeeding study setting, where the risk set is defined as the set of individuals at risk at some specific time-to-event point (Klein et al., 2013). In Scenario 1, the exposure occurs after the start of the follow-up and before the occurrence of an outcome event (such as cessation of breastfeeding). If a time-fixed analysis is performed and the exposure is treated as a static baseline variable, the subject will be assigned to the exposed group. However, the time between initiation of the follow-up and the occurrence of exposure is ‘immortal’ to the occurrence of the event. Thus the subject within this time window should not be classified into the exposed group.

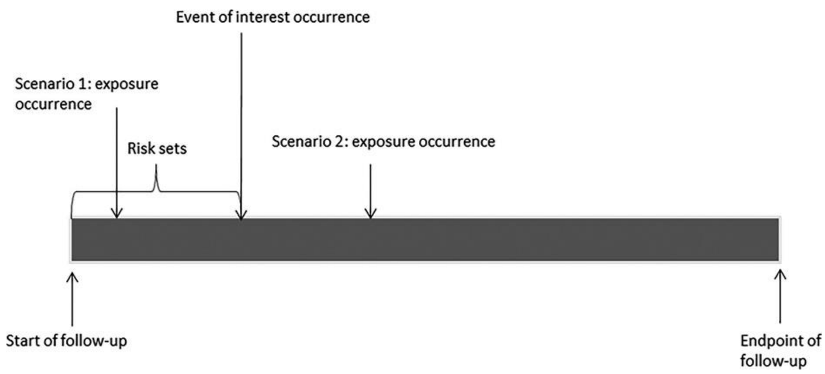


Figure 5-1 Occurrence of time-varying exposure: (1) before the event of interest occurrence (2) after the event of interest occurrence.

Scenario 2 considers an exposure is occurring after the occurrence of the event. If a time-fixed analysis is performed, the subject associated with this exposure will be classified into the exposed group. According to survival analysis theory, an exposure can contribute to the occurrence of the event only if it occurs within the time window between the onset of the follow-up and the occurrence of the event. Since the exposure happens after the occurrence of the event in Scenario 2, it cannot make any contribution to the occurrence of the event, indicating that the inclusion of this subject into the exposed group for calculating the hazard function is inappropriate. In summary, both scenarios illustrate how a time-fixed analysis could potentially lead to biased and misleading estimates of effect sizes.

5.3 Modelling time-to-event data with time-varying covariates

To account for time-dependent bias, a time-varying covariate approach can be implemented in survival analysis. Considering a semi-parametric model, the time-varying covariate approach is essentially an extended Cox regression analysis (Therneau & Grambsch, 2000). Suppose (T_i, δ_i), $i = 1, ..., n$, where T_i is the observed time-to-event outcome for the ith subject, δ_i is the censoring indicator for the event of interest. Let $\mathbf{Z}(t)$ denote the vector of time-varying covariates and $\mathbf{X}$ the vector of other time-invariant covariates. The Cox regression model incorporating the time-varying covariates for modelling the hazard of the event occurring at the time t, $\lambda(t)$, is given by:

$$\lambda(t) = \lambda_0(t)\exp\{\mathbf{Z}(t)^T\boldsymbol{\beta} + \mathbf{X}^T\boldsymbol{\beta}_0\}$$

where $\lambda_0(t)$ is an unspecified baseline hazard function for reference subjects; $\boldsymbol{\beta}$ is the vector of effect sizes of the time-varying covariates and its values represent weighted averages of log-transformed hazard ratios; $\boldsymbol{\beta}_0$ is the vector of effect magnitudes of the time-invariant covariates.

In the time-varying covariate approach, there are two different computational methods to handle the time-varying covariates, namely, the programming statement and the counting process. Both methods yield the same results although they request different format structures for survival data. In the programming statement, survival data are required to be in a standard layout. In other words, survival data are displayed as survival time and censoring indicator. The principle of the programming statement method is to create a time-varying indicator variable corresponding to a time-varying covariate in the computer memory (not in the physical dataset), and the value of this time-varying indicator changes over time according to the value of the time-varying covariate to reflect the time-varying nature in the data. The time-varying indicator is finally incorporated in the statistical parameter estimations. The programming statements can be normally realised with the help of programming loop or condition statements in statistical software, such as SAS and R. In the counting process, on the other hand, survival data are manipulated into a (start, stop] format structure. As an example shown in Figure 5-2, survival data of breastfeeding duration at 12 months and a time-varying variable presented in a standard layout (original variables: foodcomtime (time-varying variable), anybftime12 (survival time) and censor12 (censoring indicator)) are converted to a counting process form (new variables: start, foodcom, stop and event). Unlike the programming statement method, the new time-varying covariate foodcom is created in the physical dataset and incorporated into the statistical computation directly.

	ID	foodcomtime	anybftime12	censor12	start	foodcom	stop	event
1	1	165.0	1	1	0	0	1	0
2	2	75.0	188	1	0	0	75	0
3	2	75.0	188	1	75	1	188	0
4	3	90.0	182	0	0	0	90	0
5	3	90.0	182	0	90	1	182	1
6	4	NA	3	1	0	0	3	0
7	5	120.0	356	0	0	0	120	0
8	5	120.0	356	0	120	1	356	1
9	6	NA	184	1	0	0	184	0
10	7	120.0	106	0	0	0	106	1

Figure 5-2 An example of survival data in both forms: standard layout and counting process

5.4 Illustration by modelling breastfeeding data

To illustrate time-dependent bias introduced by using time-fixed variables in the conventional survival analysis, the time-varying covariate approach with the programming statements was applied to analyse breastfeeding duration data from Sichuan Province, China (Tang et al., 2015). This longitudinal study (n=695) was conducted between 2010 and 2011 with a 12-month follow-up after childbirth. A face-to-face interview was carried out at discharge, and subsequent telephone interviews were conducted at 1, 3, and 6 months postpartum to solicit details of infant feeding practices and breastfeeding problems experienced by mothers. At 12 months postpartum, only information of breastfeeding status was collected through a brief telephone interview. The breastfeeding duration was measured in days, which refers to the time length between breastfeeding initiation and breastfeeding cessation. A total of 665 women who initiated breastfeeding after delivery were included in the analysis. As shown in Figure 5-3, both scenarios in Figure 5-1 occurred for the exposures ‘solid foods introduction’ and ‘maternal return to work’.

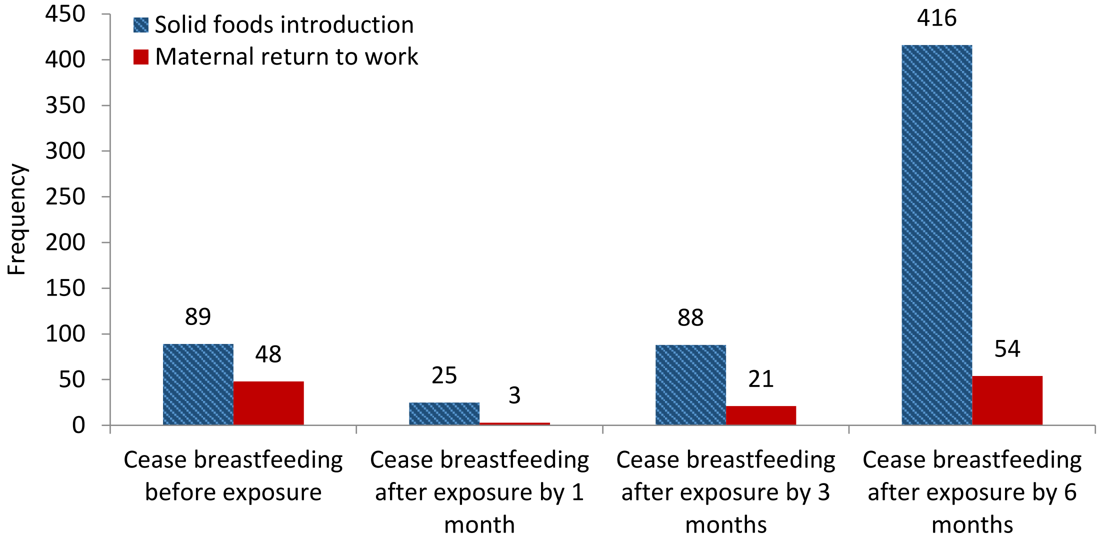


Figure 5-3 Frequency distributions of women exposed to solid foods introduction and maternal return to work within different time windows.

Two Cox regression models, in which the outcome variable was breastfeeding duration, were fitted to the data. Three significant baseline covariates (maternal age, place of delivery, and intended breastfeeding duration) identified and published previously by the original study were included in the two Cox regression models (Tang et al., 2015), but ‘solid foods introduction’ and ‘maternal return to work’ were treated as time-fixed baseline variables in Model 1, and time-varying variables in Model 2, respectively.

In Model 1, ‘solid foods introduction’ was coded as ‘0’ if no solid foods were introduced over the entire 6-month follow-up period, and ‘1’ if solid foods were introduced at any time during the follow-up. Similarly, ‘maternal return to work’ was coded as ‘0’ or ‘1’ depending on whether a mother returned to work during the follow-up.

For time-varying variables, data on ‘solid foods introduction’ was collected at 1, 3, and 6 months postpartum to measure the time mothers introduced solid foods into feeding since the last interview. In Model 2, ‘solid foods introduction’ was coded as ‘0’ from the start of follow-up and coded as ‘1’ at the first instance when solid food was introduced, on the condition that mothers were breastfeeding their babies at all times since initiation. Likewise, ‘maternal return to work’ after birth was coded as a binary variable based on the information collected at three time points (namely, 1, 3, and 6 months postpartum), with values updated at each interview time point. Statistical analysis was accomplished using SAS 9.4 (SAS Institute Inc., Cary, NC, USA) and the programming statements are presented in the Appendix F.3.

Table 5-1 summarises the results from both models. According to Model 1, ‘solid foods introduction’ by 6 months postpartum appeared to be ‘protective’ against breastfeeding cessation by 12 months, whereas ‘maternal return to work’ by 6 months was not associated with the breastfeeding duration. On the contrary, Model 2 indicated that both time-varying variables increased the hazard of discontinuing breastfeeding for the mothers by 12 months postpartum, adjusting for other time-invariant covariates. Moreover, from the perspective of goodness-of-fit, Model 2 (Akaike Information Criterion (AIC)=4527.27) performed better than Model 1 (AIC=4932.00).

Table 5-1 Comparison of effect size estimates between the time-fixed and the time-varying analyses

	Hazard ratio (95% confidence interval)	
	***Model 1 (Time-fixed analysis)**	***Model 2 (Time-varying analysis)**
Solid foods introduction by 6 months postpartum	0.61 (0.50, 0.75)	1.50 (1.17, 1.93)
Maternal return to work by 6 months postpartum	0.99 (0.73, 1.36)	1.45 (1.06, 2.00)

*Model was adjusted for time-fixed covariates: maternal age, place of delivery, and intended breastfeeding duration.

Survival curves plotted in R (R Core Development Team, 2017) from the time-fixed approach and the time-varying approach are shown in Figure 5-4, Figure 5-5, Figure 5-6 and Figure 5-7. Note that in Figure 5-6 and Figure 5-7, although both curves cross each other at some time point, the hazard ratios of these two covariates are statistically significant as (i) the proportional hazards assumption cannot hold for this extended Cox model with time-varying covariates since the hazard ratio is dependent on time; (ii) the estimated hazard ratio for each covariate is a weighted average of hazard ratios over time.

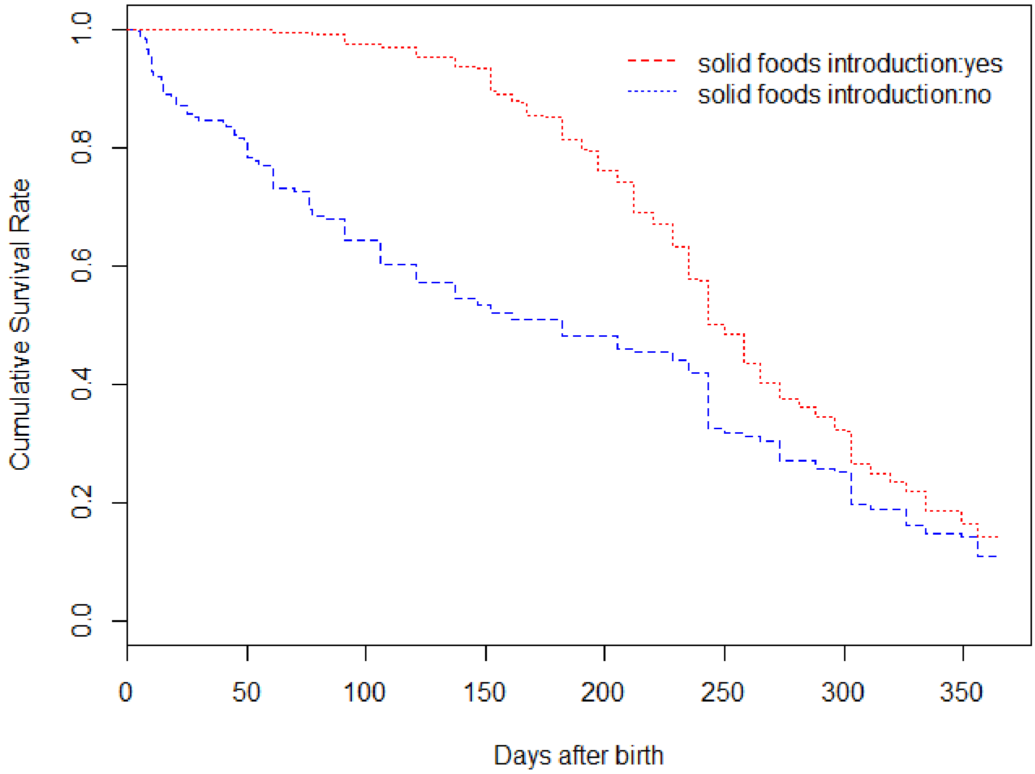


Figure 5-4 Survival curve comparing ‘yes’ versus ‘no’ for ‘solid foods introduction’ by 6 months postpartum based on the time-fixed analysis.

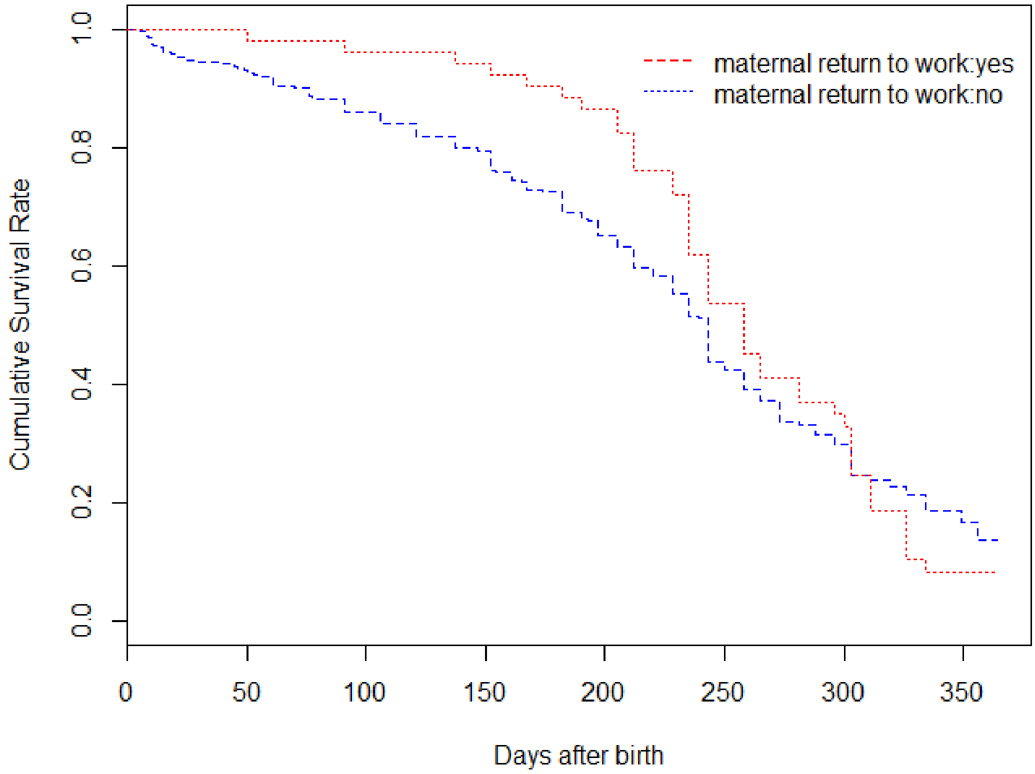


Figure 5-5 Survival curve comparing ‘yes’ versus ‘no’ for ‘maternal return to work’ by 6 months postpartum based on the time-fixed analysis.

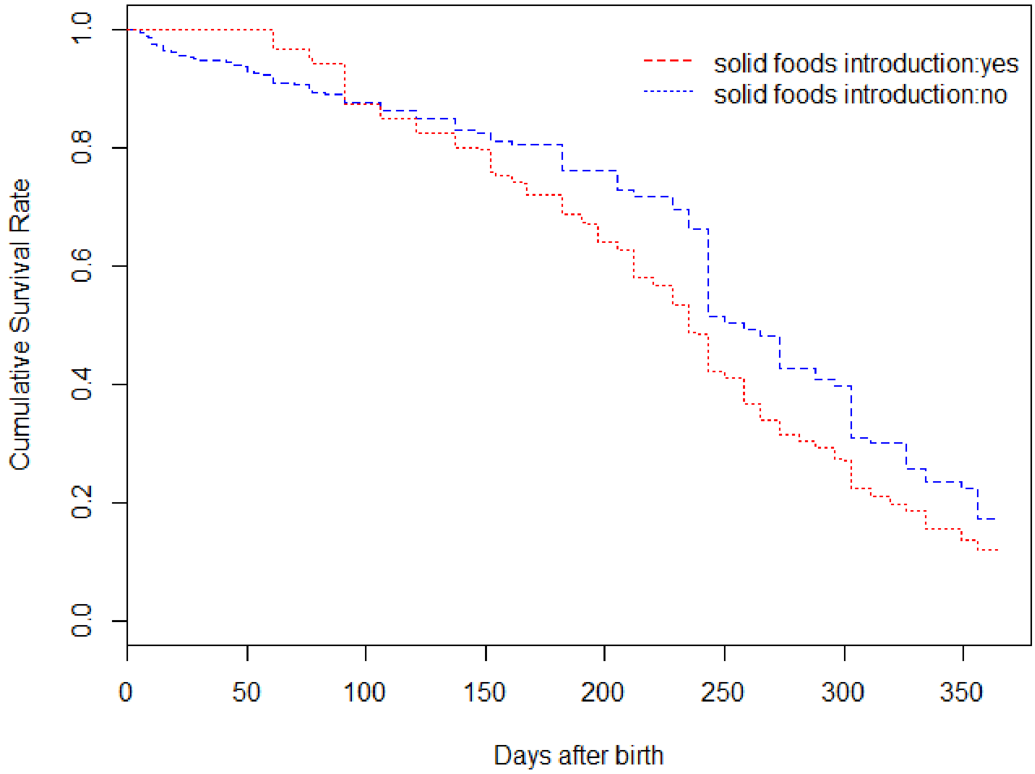


Figure 5-6 Survival curve comparing ‘yes’ versus ‘no’ for ‘solid foods introduction’ by 6 months postpartum based on the time-varying covariate approach.

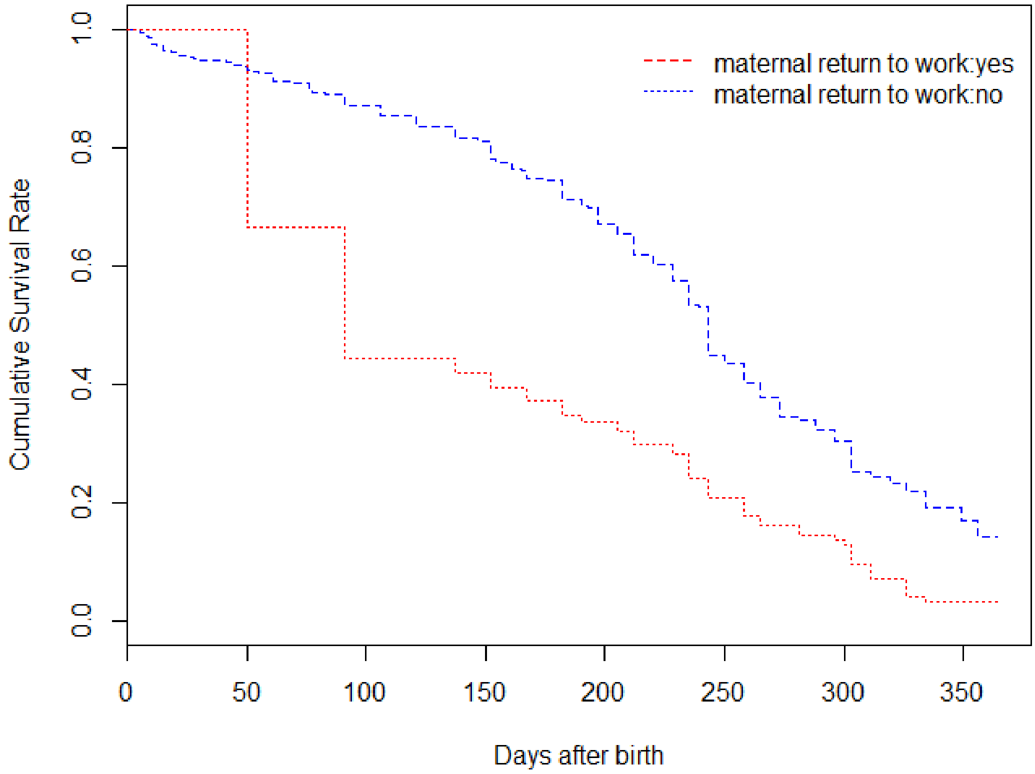


Figure 5-7 Survival curve comparing 'yes' versus 'no' for 'maternal return to work' by 6 months postpartum based on the time-varying covariate approach.

5.5 Discussion

In this chapter, we have illustrated two scenarios of time-dependent bias introduced by using time-fixed analysis of time-varying covariates in survival analysis. A solution, an extended Cox regression, was provided to avoid time-dependent bias. We used data from a breastfeeding study as an example to illustrate different results yielded from the time-fixed analysis, where the time-varying nature in covariates was ignored, and the time-varying covariate approach, where the time-varying covariates were accounted for. The time-varying modelling approach generally yields unbiased estimates, whereas the time-fixed analysis may produce biased estimates with a smaller magnitude. In our example, for both time-varying covariates, 'solid foods introduction' and 'maternal return to work', effect sizes were underestimated in the time-fixed analysis due to time-dependent bias, which was consistent with the conclusion in the literature (Austin et al., 2006; Beyersmann et al., 2008; Jones & Fowler, 2016; Suissa, 2007).

Cox model incorporating time-varying covariates is a useful approach to avoid time-dependent bias in survival analysis. In such a model, the risk sets only include a subject as survival and unexposed until the onset of the exposure. If a subject reaches the event of interest ahead of the exposure, the exposure status of this subject is treated as unexposed at all times. Computationally the estimation procedure of Cox modelling with time-varying covariates can be implemented by either the programming statement or the counting process method, and both methods result in the same estimates. Compared to the counting process method, the programming statement method is more flexible to incorporate time-varying covariates with the help of programming statements and more efficient and convenient to include a bunch of time-varying covariates simultaneously in the modelling. The counting process method, however, is disadvantaged by its time consuming and practical complexity to include a number of time-varying covariates. However, a drawback of the programming statement method is that it needs longer time to complete the modelling when the dataset is large as this method occupies computer memories to create and analyse corresponding time-varying indicator variables, which may prolong the modelling procedure, particularly when there is a large number of time-varying covariates.

Since an increasing number of breastfeeding researchers use longitudinal designs to collect data and analyse their data by using survival analysis techniques, time-dependent bias becomes an emerging concern for the validity and credibility of the study results. To avoid time-dependent bias and hence ensure unbiased estimates in the statistical analysis of breastfeeding studies, the time-varying covariate approach is recommended for survival analysis of time-to-event data in the presence of time-varying exposures.

5.6 References

Austin, P.C., Mamdani, M.M., van Walraven, C., & Tu, J.V. (2006). Quantifying the impact of survivor treatment bias in observational studies. *J Eval Clin Pract, 12*(6), 601-612.

Austin, P.C., & Platt, R.W. (2010). Survivor treatment bias, treatment selection bias, and propensity scores in observational research. *Journal of Clinical Epidemiology 63*(2), 136-138.

Beyersmann, J., Gastmeier, P., Wolkewitz, M., & Schumacher, M. (2008). An easy mathematical proof showed that time-dependent bias inevitably leads to biased effect estimation. *Journal of Clinical Epidemiology, 61*(12), 1216-1221.

Cho, I.S., Chae, Y.R., Kim, J.H., Yoo, H.R., Jang, S.Y., Kim, G.R., et al. (2017). Statistical methods for elimination of guarantee-time bias in cohort studies: a simulation study. *BMC Med Res Methodol, 17*(1), 126.

Glesby, M.J., & Hoover, D.R. (1996). Survivor treatment selection bias in observational studies: examples from the AIDS literature. *Ann Intern Med, 124*(11), 999-1005.

Jones, M., & Fowler, R. (2016). Immortal time bias in observational studies of time-to-event outcomes. *J Crit Care, 36*, 195-199.

Kalbfleisch, J.D., & Prentice, R.L. (2002). *The statistical analysis of failure time data* (2nd ed.). Hoboken, N.J.: Wiley.

Klein, J.P., & Moeschberger, M.L. (2005). *Survival Analysis: Techniques for Censored and Truncated Data*: Springer New York.

Klein, J.P., van Houwelingen, H.C., Ibrahim, J.G., & Scheike, T.H. (2013). *Handbook of Survival Analysis*: CRC Press.

Lee, E.T., & Wang, J. (2003). *Statistical methods for survival data analysis* (3rd ed.). New York: John Wiley & Sons.

Levesque, L.E., Hanley, J.A., Kezouh, A., & Suissa, S. (2010). Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ, 340*, b5087.

Mantel, N., & Byar, D.P. (1974). Evaluation of Response-Time Data Involving Transient States - Illustration Using Heart-Transplant Data. *Journal of the American Statistical Association, 69*(345), 81-86.

Redelmeier, D.A., & Singh, S.M. (2001). Survival in Academy Award-winning actors and actresses. *Ann Intern Med, 134*(10), 955-962.

Rothman, K.J., Greenland, S., & Lash, T.L. (2008). *Modern epidemiology*. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins.

Rousseau, M.C., El-Zein, M., Conus, F., Legault, L., & Parent, M.E. (2016). Bacillus Calmette-Guerin (BCG) Vaccination in Infancy and Risk of Childhood Diabetes. *Paediatr Perinat Epidemiol, 30*(2), 141-148.

Scott, J.A., Binns, C.W., Oddy, W.H., & Graham, K.I. (2006). Predictors of breastfeeding duration: evidence from a cohort study. *Pediatrics, 117*(4), e646-655.

Soderberg-Naucler, C., Peredo, I., Rahbar, A., Hansson, F., Nordlund, A., & Stragliotto, G. (2014). Use of Cox regression with treatment status as a time-dependent covariate to re-analyze survival benefit excludes immortal time bias effect in patients with glioblastoma who received prolonged adjuvant treatment with valganciclovir. *Int J Cancer, 135*(1), 248-249.

Suissa, S. (2003). Effectiveness of inhaled corticosteroids in chronic obstructive pulmonary disease: immortal time bias in observational studies. *Am J Respir Crit Care Med, 168*(1), 49-53.

Suissa, S. (2007). Immortal time bias in observational studies of drug effects. *Pharmacoepidemiol Drug Saf, 16*(3), 241-249.

Suissa, S. (2008). Immortal time bias in pharmaco-epidemiology. *American Journal of Epidemiology, 167*(4), 492-499.

Sylvestre, M.P., Huszti, E., & Hanley, J.A. (2006). Do Oscar winners live longer than less successful peers? A reanalysis of the evidence. *Annals of Internal Medicine, 145*(5), 361-363.

Tang, L., Lee, A.H., & Binns, C.W. (2015). Factors associated with breastfeeding duration: a prospective cohort study in Sichuan Province, China. *World Journal of Pediatrics, 11*(3), 232-238.

Therneau, T.M., & Grambsch, P.M. (2000). *Modeling Survival Data: Extending the Cox Model*. New York: Springer Science & Business Media.

van Walraven, C., Davis, D., Forster, A.J., & Wells, G.A. (2004). Time-dependent bias was common in survival analyses published in leading clinical journals. *Journal of Clinical Epidemiology, 57*(7), 672-682.

Whitehead, E., Dodds, L., Joseph, K.S., Gordon, K.E., Wood, E., Allen, A.C., et al. (2006). Relation of pregnancy and neonatal factors to subsequent development of childhood epilepsy: a population-based cohort study. *Pediatrics, 117*(4), 1298-1306.

Wolkewitz, M., Allignol, A., Harbarth, S., de Angelis, G., Schumacher, M., & Beyersmann, J. (2012). Time-dependent study entries and exposures in cohort studies can easily be sources of different and avoidable types of bias. *J Clin Epidemiol, 65*(11), 1171-1180.

Wolkewitz, M., Allignol, A., Schumacher, M., & Beyersmann, J. (2010). Two Pitfalls in Survival Analyses of Time-Dependent Exposure: A Case Study in a Cohort of Oscar Nominees. *American Statistician, 64*(3), 205-211.

Zhou, Z., Rahme, E., Abrahamowicz, M., & Pilote, L. (2005). Survival bias associated with time-to-treatment initiation in drug effectiveness evaluation: a comparison of methods. *Am J Epidemiol, 162*(10), 1016-1023.

Chapter 6: A two-part mixed-effects model for analysing clustered time-to-event breastfeeding data with clumping at zero

Part of this chapter is covered by a published paper "Zhao, J., Zhao, Y., Khanal, V., Lee, A. H., & Binns, C. W. (2018) Application of a Frailty Modelling Approach to Correlated Breastfeeding Duration Data. Nursing Research". DOI: https://doi.org/10.1097/NNR.0000000000000311. ((see Appendix G.2))

The statement of primary contribution of the first author and the permission to include the publication in this thesis can be found in the Appendix A. The permission to reproduce the material from the publisher can be found in the Appendix G.1.

6.1 Modelling clustered time-to-event breastfeeding data

6.1.1 Clustered time-to-event data

In longitudinally designed studies, clustered time-to-event data are commonly encountered when recruitments are conducted from subjects who are nested within same groups (clustering units). As subjects from the same clustering unit are likely to have outcomes correlated with one another, clustered time-to-event data are often correlated. This within-group correlation is generally introduced by unmeasured or unmeasurable cluster characteristics (e.g., community culture, hospital facility policies or health service resources), and needs to be taken into account in analysis, otherwise, the corresponding statistical inferences may be subject to underestimated standard error, leading to overstatement of statistical significance and an increased type I error, hence possibly resulting in incorrect conclusions (Campbell & Grimshaw, 1998; Goldstein, 2011; Raudenbush & Bryk, 2002; Zyzanski et al., 2004).

To address the correlation within groups for clustered time-to-event outcomes, a shared frailty survival model has been developed (Duchateau & Janssen, 2008; Hougaard, 2000). The frailty model specifies the correlation within groups by assuming that time-to-event outcomes within each group (clustering unit) share a common group-specific random effect. Conditional on the common random effect, time-to-event outcomes are assumed to be independent within each group. In survival analysis, an exponentially transformed common random effect term is generally named as a frailty term because it acts as a multiplicative factor on the hazard function. Thus, subjects possessing a larger frailty term within a group are more “frail” so that they have shorter event times.

Statistically, the shared frailty model can be semi-parametric (i.e., Cox frailty model) or parametric (i.e., parametric frailty models) depending on whether the distribution of baseline hazard function can be specified in advance. A general form of a proportional hazards (PH) frailty model can be given as below,

$$\lambda_{ij}(t \mid \mathbf{X}_{ij}, u_i) = \lambda_0(t)\exp\{\mathbf{X}_{ij}{}^T\boldsymbol{\beta} + u_i\}$$

where $\lambda_{ij}(t \mid \mathbf{X}_{ij}, u_i)$ is the conditional hazard function at time t for the j th subject nested in the i th clustering unit with a set of covariates $\mathbf{X}_{ij}$; u_i indicates the random effect of clustering unit i, the distributions of the frailty term (exponentially transformed u_i) can be gamma, log-normal, inverse Gaussian and compound Poisson, among which gamma and log-normal distributions are the two most commonly used distributions due to their computational simplicity (Duchateau & Janssen, 2008; Hougaard, 2000); $\lambda_0(t)$ denotes the baseline hazard function. If $\lambda_0(t)$ is unspecified, the above model is a Cox frailty model, otherwise, it is a parametric PH frailty model with a baseline hazard distribution that can be specified in advance (Duchateau & Janssen, 2008).

If the proportional hazards (PH) assumption does not hold, then an accelerated failure time (AFT) frailty model can be an alternative approach to handle the clustered time-to-event data. The AFT frailty model can be written as

$$\log T_{ij} = \theta + \mathbf{X}_{ij}{}^T\boldsymbol{\beta} + u_i + \rho\varepsilon_{ij}$$

where θ, ρ are unknown location and scale parameters, ε_{ij} has an identical distribution of survival time (or failure time) T_{ij}, and u_i denotes the random effect of clustering unit i. It is remarked that when the baseline hazard distribution of the time-to-event is Weibull or exponential, the PH parametric frailty model and the AFT frailty model coincide (Klein & Moeschberger, 2005).

6.1.2 Modelling clustered breastfeeding duration data using shared frailty survival models

In many of breastfeeding studies, it is commonly encountered that subjects are clustered (nested) in groups (clustering units), such as communities, hospitals, or health service centres. As a consequence, unobserved or unobservable heterogeneity will be introduced by the characteristics of the clustering units. Unfortunately, in the analysis of such time-to-event breastfeeding data, researchers often ignore these inherent correlations by assuming observations to be independent of each other, and conventional survival analysis techniques (e.g., Cox proportional hazards regression model) are commonly used for modelling correlated breastfeeding duration (i.e., correlated time to breastfeeding cessation). We examined the Medline database for breastfeeding duration study papers published from January 1997 to December 2017. In the past two decades, of the total 28 breastfeeding duration study articles where an inherent correlation was potentially present due to clustered structure, none of them adjusted for the inherent correlation (or the heterogeneity between clustering units) in their analyses.

Given the limited application of the shared frailty models in modelling clustered breastfeeding duration data, we applied the shared frailty model to fit clustered breastfeeding duration data from a maternal cohort in Nepal as an illustrative example to make breastfeeding researchers aware of the underlying issues. A manuscript based on this part of study has been published in Nursing Research journal, which is attached in the Appendix G.2.

6.2 Data with excess zeros and two-part models

Data with excess zeros are often encountered in medical and public health studies. This type of data are usually characterised as positively skewed and zero-inflated or clumped at zero. Failure to account for excess zeros may lead to biased statistical inference. Two-part models are commonly used for analysing data with excess zeros. Two-part models have been developed since the 1970s and widely applied in a variety of areas such as rainfall data (Cole & Sherriff, 1972; Garbutt et al., 1981; Stern, 1980), health care data (Cooper et al., 2003; Duan et al., 1983; Manning et al., 1981; Pohlmeier & Ulrich, 1995), financial data (Brown et al., 2015; J.S. Ramalho & da Silva, 2009) and physical activity data (Lee, Xiang, et al., 2010; Lee, Zhao, et al., 2010). In the two-part model framework, outcomes consisting of positive outcomes and excess zeros are usually modelled by two parts: 1) a model for predicting the probability of zero observations occurrence, and 2) a model for the continuous distribution of positive outcomes. Due to diverse types of outcomes in the second part, in the following section, we review two-part models for semicontinuous data, count data and survival data, respectively.

6.2.1 Two-part models for semicontinuous data

Semicontinuous data can be characterised as mixture outcomes of zeros and positive values with a continuous distribution. Aitchison (1955) initially proposed the idea to estimate the moments of positive random variables having a discrete probability mass at the origin by defining two processes: 1) a non-zero probability for the variable with a zero value, and 2) a conditional distribution for the positive values of the variable. Since 1980s, two-part models have been increasingly developed and applied in health econometrics. Manning et al. (1981) and Duan et al. (1983) modelled medical care expenditure using two-part models, where the probability of expenditure (use or not) was modelled in the first part and conditional on the expenditure incurred, the intensity of expenditure was modelled in the second part. As such, the general form of two-part models for semicontinuous data can be written as

The first part (binary part): $\text{logit}\{\Pr(Y_i > 0 \mid \mathbf{X}_i)\} = \mathbf{X}_i^T \boldsymbol{\beta}_1$

The second part (continuous part): $E(Y_i \mid Y_i > 0, \mathbf{X}_i) = \mathbf{X}_i^T \boldsymbol{\beta}_2$

where Y_i denotes the outcome value of the ith subject, $\mathbf{X}_i$ is the vector of covariates, $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are namely regression coefficient vectors.

In longitudinal settings or clustered settings, two-part models can also be used to handle excess zeros problem in semicontinuous data with accounting for random effects. Olsen and Schafer (2001) first extended conventional two-part regression approach to longitudinal semicontinuous data settings by incorporating random effects into both the binary and the continuous parts. In their proposed model, the random effects from these two parts were assumed jointly normally distributed and possibly correlated. Tooze et al. (2002) proposed a similar two-part model with correlated random effects using quasi-Newton optimisation of the likelihood approximated by adaptive Gaussian quadrature technique. The two-part model with correlated random effects can be specified as

The first part (binary part): $\mathrm{logit}\{\Pr(Y_{ij} > 0 \mid \mathbf{X}_{ij}, u_i)\} = \mathbf{X}_{ij}{}^T \boldsymbol{\beta}_1 + u_i$

The second part (continuous part): $E(Y_{ij} \mid Y_{ij} > 0, \mathbf{X}_{ij}, v_i) = \mathbf{X}_{ij}{}^T \boldsymbol{\beta}_2 + v_i$

Random effects u_i from the binary part and v_i the continuous part are assumed jointly normally distributed and possibly correlated,

$$\begin{pmatrix} u_i \\ v_i \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \rho\sigma_u\sigma_v \\ \rho\sigma_u\sigma_v & \sigma_v^2 \end{bmatrix} \right)$$

6.2.2 Two-part models for count data with excess zeros

Excess zeros are commonly encountered in count data (Cameron & Trivedi, 2013). A two-part model called hurdle model was proposed by Mullahy (1986) for modelling count data with excess zeros. The first part of the hurdle model is a binary model for modelling binary response: zero or positive, and the second part is a conditional model based on the positive response in the first part. Thus, the second part modelling is based on a zero-truncated Poisson distribution. The probability mass function of the two-part hurdle model can be written as

$$P(Y_i = 0) = p_i,$$

$$P(Y_i = j) = \frac{(1-p_i)}{1-e^{-\lambda_i}} \frac{e^{-\lambda_i} \lambda_i^j}{j!}, \; j = 1, 2, \ldots,$$

where Y_i denotes the outcome value of the i th subject, p_i is the probability of zeros for the i th subject, j is the observed count, λ_i is the expected Poisson count for the

i th subject. In hurdle models, zeros are generated by a binary process in the first part.

In contrast to the hurdle model, Lambert (1992) introduced zero-inflated Poisson (ZIP) regression models to account for overdispersion arising from excess zero counts. In Lambert's ZIP model, the data were treated as a mixture of zeros and outcomes with Poisson distribution. Response outcome Y_i for subject i was assumed independent and

$$Y_i \sim \begin{cases} 0 & \text{with probability } p_i \\ \text{Poisson}(\lambda_i) & \text{with probability } 1 - p_i \end{cases}$$

Thus, the resulting distribution has

$$P(Y_i = 0) = p_i + (1 - p_i)e^{-\lambda_i},$$

$$P(Y_i = j) = (1 - p_i)\frac{e^{-\lambda_i}\lambda_i^j}{j!}, \ j = 1, 2, \ldots,$$

where Y_i denotes the outcome value of the i th subject, p_i is the probability of structural zeros for the i th subject, j is the observed count, λ_i is the expected Poisson count for the i th subject. Consequently, the ZIP model has two components regarding the zero generating process. One is governed by a binary distribution that generates structural zeros, and the other is governed by a Poisson distribution that generates counts including zeros.

In longitudinal settings or clustered settings, random effects have been incorporated in hurdle models or ZIP models to account for within-subject correlations or within-cluster correlations. Yau and Lee (2001) proposed a hurdle model with uncorrelated random effects to analyse longitudinal count data with extra zeros. Min and Agresti (2005) improved the hurdle model by fitting a correlated random effects model. Hall (2000) extended Lambert's ZIP model by incorporating random effects in the Poisson process to account for within-subject correlation and between-subject heterogeneity. However, random effects in the binary part were not considered in Hall's model. A ZIP model with random effects in both processes were discussed by

Min and Agresti (2005). Random effects introduced to account for inter-hospital variations and the dependency of clustered length of hospital stay were discussed in a zero-inflated negative binomial (ZINB) model proposed by Yau et al. (2003).

6.2.3 Two-part models for survival data with a surviving fraction

In the context of survival data, two-part models mainly have been applied to analyse time-to-event outcomes with long-term survivors or a cured fraction of patients. In conventional survival analysis, it is typically assumed that all of the study participants will eventually experience the event of interest if the follow-up is sufficiently long. However, a substantial proportion of long-term survivors or cured patients are often reported in a variety of studies (Chen, 2013; De Angelis et al., 1999; Sposto, 2002; Tong et al., 2012). Mixture cure models were first proposed by Boag (1949) and later by Berkson and Gage (1952) to analyse survival data when a proportion of patients are cured. Farewell (1986) applied a mixture model composed of a logistic model and a Weibull model to data from a breast cancer clinical trial. The logistic regression was used for modelling the cured probability, and the Weibull model was used for fitting the survival function of the uncured patients. In the literature, mixture cure models have been extensively discussed by many other researchers, including Goldman (1984), Greenhouse and Wolfe (1984), Gray and Tsiatis (1989), Kuk and Chen (1992), Sposto et al. (1992), Maller and Zhou (1992, 1994, 1995), Maller and Zhou (1996), Sy and Taylor (2000), Peng and Dear (2000), and Cooner et al. (2007). In mixture cure models, the survival function $S(t)$ for the entire population can be written as

$$S(t) = \pi + (1-\pi)S^*(t),$$

where π refers to the fraction of population who are cured or long-term survivors and $S^*(t)$ denotes the survival function for the non-cured group.

In the context of longitudinal survival data or clustered survival data, Kim and Jhun (2008) imposed a common normal frailty effect into a mixture cure model to characterise the association between the cure fraction and survival model. Seppa et al. (2010) used a pair of random effects to capture the regional variation in the cure fraction and the survival model under a Bayesian framework. Peng and Taylor

(2011) proposed several estimation methods including Gaussian quadrature, rejection sampling, and importance sampling to obtain the maximum likelihood estimates of the mixture cure model with random effects and applied the model to clustered survival data from a multi-centre tonsil cancer study. Other recent extensive work on mixture cure models for analysing clustered survival data with a cure fraction can be found in studies by Xiang et al. (2011), Chen et al. (2013), Niu and Peng (2014), Gallardo et al. (2016) and Yi et al. (2018).

6.3 Clustered time-to-event breastfeeding data with clumping at zero

In longitudinal epidemiological studies consisting of a baseline stage and a follow-up stage, observations at the baseline stage may contain a notable proportion of negative responses (i.e., “No” or “False” to a research question regarding a baseline event of interest). In the follow-up stage, only subjects with positive responses (i.e., “Yes” or “True” to the same research question regarding the same baseline event of interest) at baseline are continuously followed up to measure the time (duration) up to the occurrence of a failure event of interest, as a consequence the time to the failure event occurrence for those unfollowed-up subjects is denoted as zero.

A motivating example is the data arising from longitudinal breastfeeding studies, which normally have a baseline stage for measuring the prevalence of breastfeeding at discharge and a followed-up stage for measuring the duration of breastfeeding. Breastfeeding, especially exclusive breastfeeding (EBF), is beneficial to both infants and mothers (Gartner et al., 2005; Jordan et al., 2010; Kull et al., 2004; Palmer et al., 2014). The World Health Organization has recommended EBF for at least 6 months (World Health Organization, 2018). However, the prevalence of EBF at hospital discharge varies globally and remains low in many countries, for example, 75.6% in Australia (Scott et al., 2006), 50.3% in China (Qiu et al., 2009), and 68.6% in Spain (Vila-Candel et al., 2017). In other words, there exist a high proportion of non-exclusively breastfed infants, corresponding to those “no” responses to the status of EBF, at discharge (baseline). Their EBF duration (i.e., time to cessation of EBF) would be noted as zero.

These two-stage studies produce two processes of outcome data, namely, a binary outcome to describe the prevalence of the event of interest at baseline and a time-to-event outcome measuring the duration up to the failure event occurring in the follow-up. Statistical analyses, including logistic regression and survival analysis, for identifying factors associated with the prevalence of the baseline event, and with the time to the failure event occurrence, are widely performed in the literature. It is known that standard survival analysis considers only positive time-to-event outcomes so that subjects with a negative response at baseline are excluded from the risk set for estimating the survival probability (morbidity or mortality in some instances). Consequently, the information from these subjects, which may be important to the failure event of interest, is abandoned in the analysis. Given the two-stage data structure, in the literature, a framework of two-part model has been introduced to analyse outcomes with a two-component structure such as semi-continuous data, count data with excess zeros and survival data with a surviving fraction. However, to our best knowledge, its application to time-to-event data with clumping at zero is limited to the lapse of insurance (Brockett et al., 2008).

In addition, in epidemiological settings, subjects are often clustered within hospitals, communities, or health service centres. It is expected that observations exhibit intra-cluster correlation due to similar socio-economic, environmental or health conditions for individuals in the same cluster. To account for the heterogeneity between clustering units, an adjustment for the underlying correlation structure via random effects in a two-part model becomes necessary.

Furthermore, in some cases it is reasonable to assume the binary data and time-to-event data generated from the two processes are related to each other, indicating an inherent correlation between the two stages. In breastfeeding studies, for example, mothers in one community which has a higher prevalence of EBF at hospital discharge are more likely to have longer EBF durations compared to those in other communities. Ignoring such correlation may introduce bias in statistical inferences.

Therefore, we proposed a two-part mixed-effects (fixed and random effects) modelling approach to analyse clustered time-to-event data with clumping at zero, with application to EBF data as an illustrative example. Correlated random effects accounting for possible correlation between baseline and follow-up stages were taken into account in this approach.

6.4 Modelling clustered time-to-event breastfeeding data with clumping at zero using a two-part mixed-effects model

6.4.1 Model specification

We consider that longitudinal observations can be partitioned structurally into two parts. The first part is a baseline part (Part 1), where subjects' statuses regarding a 'baseline event' are observed. Examples of positive responses to the 'baseline event' include "EBF at discharge" or "hold a health insurance at baseline". The second part is a follow-up part (Part 2), where positive time to a 'failure event' of interest, conditional on the positive baseline response in the first part, is observed. Examples of the 'failure event' could be "EBF cessation" or "health insurance lapse".

Suppose the observations from a longitudinal epidemiological study are given as below:

$$(y_{ij}, \delta_{ij}),\ \ j = 1, \cdots, n_i, i = 1, \cdots, n,\ \ \mathrm{N} = \sum_{i=1}^{n} n_i,$$

where y_{ij} is the observed time to event for the jth subject within the ith cluster, and δ_{ij} is the censoring indicator for the event. For each subject $j = 1, 2, \ldots, n_i$ within a cluster $i = 1, 2, \ldots, n$, let a binary response variable $Z_{ij} = 1$ corresponding to the positive response to the baseline event and $Z_{ij} = 0$ otherwise in Part 1. The likelihood function is then given by

$$L(\boldsymbol{\gamma}, \boldsymbol{\beta}) = \prod_{(i,j):Z_{ij}=1} p_{ij} f(y_{ij}, \xi_{ij}, \boldsymbol{\beta}) \prod_{(i,j):Z_{ij}=0} (1 - p_{ij}) = [\prod_{(i,j):Z_{ij}=1} p_{ij} \prod_{(i,j):Z_{ij}=0} (1 - p_{ij})][\prod_{(i,j):Z_{ij}=1} f(y_{ij}, \xi_{ij}, \boldsymbol{\beta})],$$

where $p_{ij} = p(\eta_{ij}, \boldsymbol{\gamma}) = p(Z_{ij} = 1 | \eta_{ij})$ is the probability of the baseline event occurrence for the jth subject within the ith cluster for given covariates x_{ij} via $\eta_{ij} = x_{ij}' \boldsymbol{\gamma} + u_i$, with u_i being the random effects for adjusting the subject-level correlation at baseline and the parameter vector $\boldsymbol{\gamma}$ represents the effects of x_{ij} on the binary outcome; $f(y_{ij}, \xi_{ij}, \boldsymbol{\beta})$ denotes the probability density of the observed time to the failure event y_{ij} for covariates w_{ij} via $\xi_{ij} = w_{ij}'\boldsymbol{\beta} + v_i$, with v_i being the random effects for adjusting the subject-level correlation in the follow-up and $\boldsymbol{\beta}$ is parameter vector associated with w_{ij}. The likelihood function $L(\boldsymbol{\gamma}, \boldsymbol{\beta})$ can be factorised into two parts as

$$L_1(\boldsymbol{\gamma}) = \prod_{(i,j):Z_{ij}=1} p_{ij} \prod_{(i,j):Z_{ij}=0} (1 - p_{ij}) \text{ and } L_2(\boldsymbol{\beta}) = \prod_{(i,j):Z_{ij}=1} f(y_{ij}, \xi_{ij}, \boldsymbol{\beta}).$$

Here, this two-part mixed-effects model aims to: (i) determine factors associated with the prevalence of the baseline event of interest; (ii) assess associations between exposures and time-to-event outcomes; (iii) capture/ account for the possible correlation between the prevalence part and the time-to-event part.

Part 1: logistic mixed-effects regression model

For the first part $L_1(\boldsymbol{\gamma})$, a logistic mixed-effects regression model can be applied to achieve the first objective via

$$\log(p_{ij} / (1 - p_{ij})) = \eta_{ij} = x_{ij}'\boldsymbol{\gamma} + u_i$$

Part 2: parametric frailty model

For the second part $L_2(\boldsymbol{\beta})$, either a conditional semi-parametric Cox proportional hazards model with random effects (i.e., Cox frailty model) or a class of conditional parametric survival models incorporating random effects (i.e., parametric frailty models) can be used. However, compared to semi-parametric proportional hazards model (i.e., Cox proportional hazards model), parametric survival models provide more efficient and informative estimations when the baseline hazard function could be specified in advance (Duchateau & Janssen, 2008; Nardi & Schemper, 2003). Thus we focus on parametric frailty models in Part 2 modelling. The resulting log-likelihood function is given by

$$l_2(\boldsymbol{\beta}) = \sum_{i=1}^{n}\sum_{j=1}^{n_i}\{(1-\delta_{ij})\log S(y_{ij};\xi_{ij},\boldsymbol{\beta}) + \delta_{ij}\log f(y_{ij};\xi_{ij},\boldsymbol{\beta})\}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n_i}\{\log S(y_{ij};\xi_{ij},\boldsymbol{\beta}) + \delta_{ij}\log h(y_{ij};\xi_{ij},\boldsymbol{\beta})\},$$

where $S(y_{ij};\xi_{ij},\boldsymbol{\beta})$ is the survival function, and $h(y_{ij};\xi_{ij},\boldsymbol{\beta})$ is the hazard function, $\xi_{ij} = w_{ij}'\boldsymbol{\beta} + v_i$. The hazard function can be one of the various forms depending on different distributions of the time-to-event outcome y_{ij}, including Weibull, exponential, log-logistic and log-normal. In practice, gamma and lognormal are commonly used distributions for the frailty term (exponential transformation of the random effects) in the above frailty model (Duchateau & Janssen, 2008).

Assumption on random effects u_i and v_i

For the relationship between random effects u_i and v_i in Part 1 and Part 2, the study considers the following two assumptions:

1) If u_i and v_i are assumed to be independent, following a normal distribution $N(0,\sigma_u^2)$ and $N(0,\sigma_v^2)$, respectively, the approach is an independent two-part mixed-effects modelling.

2) If u_i and v_i are assumed to be correlated, following a bivariate normal distribution with a variance-covariance matrix $\begin{pmatrix} u_i \\ v_i \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \rho\sigma_u\sigma_v \\ \rho\sigma_u\sigma_v & \sigma_v^2 \end{bmatrix}\right)$, the approach is a correlated two-part mixed-effects modelling.

6.4.2 Parameter estimation

When the random effects u_i and v_i are independent and the likelihood function $L(\boldsymbol{\gamma},\boldsymbol{\beta})$ then allows for different sets of covariates in both parts, it is computationally feasible and fully efficient to fit the two regression models separately. When the random effects u_i and v_i are assumed to be correlated, the Laplace approximation

(Olsen & Schafer, 2001) or adaptive Gaussian quadrature (Liu et al., 2010; Tooze et al., 2002), which has been utilised to handle correlated two-part random effects for modelling semi-continuous data, can be adapted. Under either assumption, model fitting and parameters estimation can be conveniently implemented by using the adaptive Gaussian quadrature technique available in PROC NLMIXED of SAS (SAS Institute Inc., Cary, NC, USA). The SAS codes used in our illustrative example can be found in the Appendix G.3.

6.5 An illustrative example

6.5.1 Breastfeeding data

A community-based maternal cohort study was conducted in Nepal between January and October 2014 to determine factors affecting the breastfeeding duration, involving 27 randomly selected communities (15 village development committees and 12 wards of two municipalities). Details of the study design, setting and sampling had been reported previously (Khanal et al., 2015). Briefly, a total of 735 mothers were recruited and interviewed shortly after giving birth, and those mothers who confirmed they exclusively breastfed their infants were followed up for measuring the EBF duration by six months postpartum. Information about the EBF duration (i.e., time to EBF cessation) was collected for mothers who were exclusively breastfeeding their babies at baseline. The final sample (N=649) excluded 86 mothers who delivered at home.

6.5.2 Two-part modelling procedure

Study variables

The status of EBF at baseline was coded as a binary outcome with ‘1’ for EBF and ‘0’ for non-EBF and used as the outcome variable in Part 1 logistic mixed-effects regression model. For those mothers who exclusively breastfed their infants at the baseline interview, the EBF duration (days) up to 6 months postpartum obtained in the follow-up was used as the outcome variable in Part 2 parametric frailty model. Three covariates, namely, birth mode (caesarean section vs natural birth (reference group)), grandmother feeding preference (breastfeeding vs other feeding (reference group)), and mother-child bonding (yes vs no (reference group)), were included in both two parts of the model.

Identifying an appropriate baseline hazard function

For Part 2 parametric frailty model, an appropriate form of the baseline hazard function needs to be specified. With no other suggestive information available from the data, we fitted the data with three commonly used event time distributions, namely, Weibull, exponential and log-logistic, without any covariates, then compared their Akaike information criterion (AIC) values. The one with the smallest AIC value was chosen as the baseline hazard function.

Assessing correlation between two stages of random effects

The breastfeeding data were fitted with a logistic mixed-effects model, and a parametric accelerated failure time (AFT) frailty model initially assuming independent random effect u_i and v_i. Empirical Bayes estimates of the random effects were plotted against each other to examine the degree of correlation between the two random effects. The Pearson's correlation coefficient between the estimated u_i and v_i ($i = 1, 2, \ldots, 27$) was calculated as well.

Fitting data with a two-part mixed-effects model incorporating correlated random intercepts

The adaptive Gaussian quadrature technique was used for parameter estimation via the SAS NLMIXED procedure. Maybe due to data scaling issues (Kiernan et al., 2012), the estimation procedure fails to converge with this dataset. We also suspected that unequal numbers of observations used in the correlated two-part modelling may cause some problems related to a non-invertible variance-covariance matrix in parameter estimation. As a computational solution, to ensure the two-part regression models having the same number of observations, each zero EBF duration observation was transformed by adding a very small positive value (e.g., 1E-6), and was then included as a censored observation in Part 2 parametric frailty modelling.

Sensitivity analyses were performed to evaluate the robustness of the results after transformation of zero observations in Part 2 parametric frailty modelling. The parameter estimation and goodness-of-fit were compared between models excluding zero observations and including transformed zero observations.

6.5.3 Results

Among those 649 mothers, 434 mothers (66.9%) exclusively breastfed their infants at baseline, leading to 33.1% ‘zero’ EBF duration observations. Conditional on those mothers who practiced EBF at baseline, only 434 mothers were followed up to six months postpartum for measuring the EBF duration. In Part 2, the Weibull distribution, which has the minimum AIC=283.97 compared with 926.02 for the exponential distribution and 354.71 for the log-logistic distribution, was identified and chosen as the most appropriate distribution for the EBF duration data. As shown in Figure 6-1, a positive correlation (r=0.55, p=0.003) was found between the random effects u_i and v_i, suggesting an inherent correlation between the occurrence of EBF for mothers after giving birth and time to EBF cessation by 6 months postpartum.

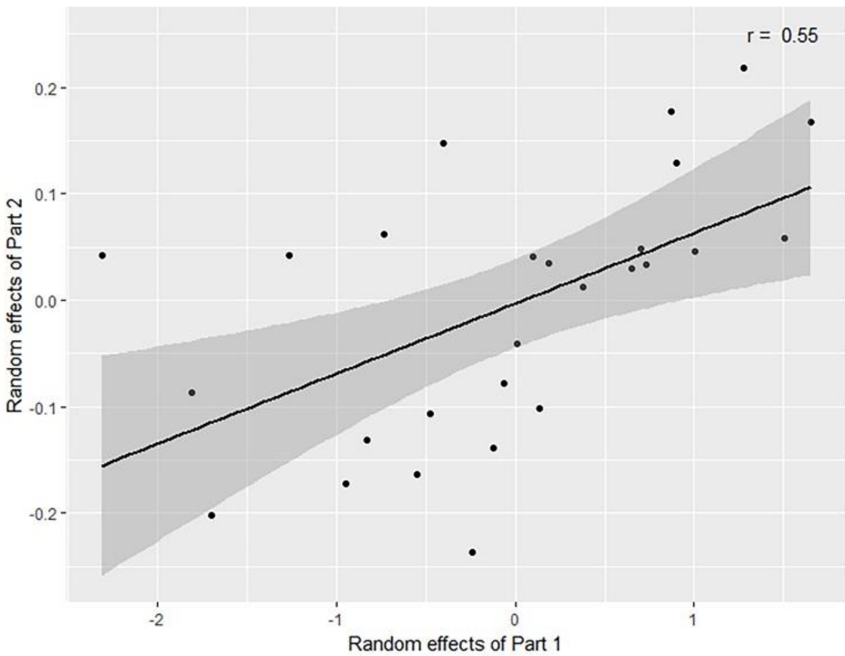


Figure 6-1 Correlation between random effects estimated from two independent models (the logistic mixed-effects model and the Weibull AFT frailty model).

A correlated two-part mixed-effects model was consequently fitted to the data while a two-part mixed-effects model assuming independent random effects was also fitted for comparison purpose. As shown in Table 6-1, a significantly positive correlation ($\rho = 0.67$, p<0.001) between the two parts was confirmed, indicating that the independent random effects assumption was inappropriate. It suggested that there was a positive correlation between the baseline prevalence of EBF and EBF duration, that is, mothers in a community which has a higher prevalence of EBF at hospital discharge intended to have longer EBF durations. There were only slight differences in the estimated regression coefficients between the independent and correlated two-part mixed-effects models; however, variance components estimated by the independent two-part model were slightly smaller than those by the correlated two-part model, which also fitted the data better with a smaller -2 log-likelihood.

Table 6-1 Parameter estimates in the two-part model with different assumptions (independent or correlated) for breastfeeding data

Parameters	Independent two-part model			Correlated two-part model		
	Estimate	SE	p	Estimate	SE	p
Part 1 logistic mixed-effects model						
Intercept	0.8361	0.4564	0.078	0.8381	0.4584	0.080
Birth mode	-1.8801	0.2777	<.0001	-1.9039	0.2785	<.0001
Grandmother feeding preference	0.4480	0.3485	0.21	0.4691	0.3472	0.19
Mother-child bonding	0.1443	0.2916	0.62	0.1407	0.2894	0.63
Part 2 Weibull AFT frailty model						
Gamma	4.9438	0.2146	<.0001	4.9465	0.2146	<.0001
Intercept	4.8690	0.0518	<.0001	4.8689	0.0510	<.0001
Birth mode	-0.0462	0.0399	0.26	-0.0491	0.0397	0.23
Grandmother feeding preference	0.0772	0.0413	0.073	0.0781	0.0410	0.068
Mother-child bonding	0.0798	0.0360	0.035	0.0736	0.0355	0.049
σ_u^2	1.4218	0.5515	0.016	1.4928	0.5702	0.015
$\rho\sigma_u\sigma_v$				0.1111	0.0474	0.027
σ_v^2	0.0182	0.0064	0.0083	0.0186	0.0065	0.0082
Correlation coefficient (ρ)				0.6668	0.1513	0.0002
-2 log-likelihood (both parts)	4277.8			4269.2		

The sensitivity analyses showed that the transformation of zero EBF durations in Part 2 time-to-event analysis did not affect the robustness of parameter estimation and model fit. Parameter estimates and the goodness-of-fit were exactly equivalent between the model without zero EBF durations and the model including transformed zero EBF durations as censored observations.

6.6 Discussion

A two-part mixed-effects modelling approach was proposed for analysing clustered time-to-event data with clumping at zero. This approach takes into account the correlation within each clustering unit by incorporating random effects in each part. This approach further takes into account the possible correlation between the two random effects, that is, the correlation between the baseline prevalence and the followed-up time-to-event outcomes. Compared to the conventional survival analysis, this approach makes full use of all available information at baseline and during the follow-up. In addition to breastfeeding studies, the approach can be applied to other epidemiological areas, in which clustered/ longitudinal time-to-event outcomes with clumping at zero arise.

Similar two-part mixed-effects modelling approaches have been discussed widely in the literature (Farewell et al., 2017), but most existing works have been focused on longitudinal semi-continuous data (Lee & Xiang, 2011; Olsen & Schafer, 2001; Su et al., 2009; Tooze et al., 2002), and longitudinal count data with excess zeros (Lee et al., 2006; Yau & Lee, 2001). Applications of the two-part mixed-effects model on clustered time-to-event data with clumping at zero are very sparse. To our best knowledge, our proposed two-part mixed-effects model is the first attempt to deal with such type of data.

As discussed in the work by Su et al. (2009), an incorrect assumption about the correlation between random effects in the two parts can introduce bias in the two-part modelling of semi-continuous data. The results obtained from the illustrative example in our study showed a consistent conclusion that the variance component was underestimated in Part 2 survival modelling if the model was misspecified as independent random effects when the correlation between two parts existed.

Clustered time-to-event data are often encountered in longitudinal epidemiological studies. The proposed correlated two-part mixed-effects modelling approach takes into account the correlations possibly presenting in a two-level hierarchical data structure, i.e., subjects nested within different clustering units, via random effect terms u_i and v_i in the model. The method can be extended to handle higher level hierarchical or a multilevel data structure, by specifying a corresponding variance-covariance structure for depicting correlations between random effects.

6.7 References

Aitchison, J. (1955). On the distribution of a positive random variable having a discrete probability mass at the origin. *Journal of the american statistical association, 50*(271), 901-908.

Berkson, J., & Gage, R.P. (1952). Survival Curve for Cancer Patients Following Treatment. *Journal of the American Statistical Association, 47*(259), 501-515.

Boag, J.W. (1949). Maximum Likelihood Estimates of the Proportion of Patients Cured by Cancer Therapy. *Journal of the Royal Statistical Society. Series B (Methodological), 11*(1), 15-53.

Brockett, P.L., Golden, L.L., Guillen, M., Nielsen, J.P., Parner, J., & Perez-Marin, A.M. (2008). Survival Analysis of a Household Portfolio of Insurance Policies: How Much Time Do You Have to Stop Total Customer Defection? *The Journal of Risk and Insurance, 75*(3), 713-737.

Brown, S., Ghosh, P., Su, L., & Taylor, K. (2015). Modelling household finances: A Bayesian approach to a multivariate two-part model. *Journal of Empirical Finance, 33*, 190-207.

Cameron, A.C., & Trivedi, P.K. (2013). *Regression analysis of count data* (Vol. 53): Cambridge university press.

Campbell, M.K., & Grimshaw, J.M. (1998). Cluster randomised trials: time for improvement. The implications of adopting a cluster design are still largely being ignored. *BMJ, 317*(7167), 1171-1172.

Chen, C.M., Lu, T.F.C., & Hsu, C.M. (2013). Association estimation for clustered failure time data with a cure fraction. *Computational Statistics & Data Analysis, 57*(1), 210-222.

Chen, T.T. (2013). Statistical issues and challenges in immuno-oncology. *J Immunother Cancer, 1*, 18.

Cole, J.A., & Sherriff, J.D.F. (1972). Some single- and multi-site models of rainfall within discrete time increments. *Journal of Hydrology, 17*(1), 97-113.

Cooner, F., Banerjee, S., Carlin, B.P., & Sinha, D. (2007). Flexible cure rate modeling under latent activation schemes. *Journal of the American Statistical Association, 102*(478), 560-572.

Cooper, N.J., Sutton, A.J., Mugford, M., & Abrams, K.R. (2003). Use of Bayesian Markov Chain Monte Carlo methods to model cost-of-illness data. *Medical Decision Making, 23*(1), 38-53.

De Angelis, R., Capocaccia, R., Hakulinen, T., Soderman, B., & Verdecchia, A. (1999). Mixture models for cancer survival analysis: application to population-based data with covariates. *Stat Med, 18*(4), 441-454.

Duan, N., Manning, W.G., Morris, C.N., & Newhouse, J.P. (1983). A Comparison of Alternative Models for the Demand for Medical Care. *Journal of Business & Economic Statistics, 1*(2), 115-126.

Duchateau, L., & Janssen, P. (2008). *The Frailty Model*. New York: Springer.

Farewell, V.T. (1986). Mixture-Models in Survival Analysis - Are They Worth the Risk. *Canadian Journal of Statistics-Revue Canadienne De Statistique, 14*(3), 257-262.

Farewell, V.T., Long, D.L., Tom, B.D.M., Yiu, S., & Su, L. (2017). Two-Part and Related Regression Models for Longitudinal Data. *Annual Review of Statistics and Its Application, Vol 4, 4*, 283-315.

Gallardo, D.I., Bolfarine, H., & Pedroso-de-Lima, A.C. (2016). Destructive weighted Poisson cure rate models with bivariate random effects: Classical and Bayesian approaches. *Computational Statistics & Data Analysis, 98*, 31-45.

Garbutt, D.J., Stern, R.D., Dennett, M.D., & Elston, J. (1981). A comparison of the rainfall climate of eleven places in West Africa using a two-part model for daily rainfall. *Archives for meteorology, geophysics, and bioclimatology, Series B, 29*(1), 137-155.

Gartner, L.M., Morton, J., Lawrence, R.A., Naylor, A.J., O'Hare, D., Schanler, R.J., et al. (2005). Breastfeeding and the use of human milk. *Pediatrics, 115*(2), 496-506.

Goldman, A.I. (1984). Survivorship Analysis When Cure Is a Possibility - a Monte-Carlo Study. *Statistics in Medicine, 3*(2), 153-163.

Goldstein, H. (2011). *Multilevel Statistical Models* (4th ed.). Chichester, UK: Wiley.

Gray, R.J., & Tsiatis, A.A. (1989). A Linear Rank Test for Use When the Main Interest Is in Differences in Cure Rates. *Biometrics, 45*(3), 899-904.

Greenhouse, J.B., & Wolfe, R.A. (1984). A Competing Risks Derivation of a Mixture Model for the Analysis of Survival-Data. *Communications in Statistics-Theory and Methods, 13*(25), 3133-3154.

Hall, D.B. (2000). Zero-inflated Poisson and binomial regression with random effects: a case study. *Biometrics, 56*(4), 1030-1039.

Hougaard, P. (2000). *Analysis of multivariate survival data*. New York: Springer.

J.S. Ramalho, J., & da Silva, J.V. (2009). A two-part fractional regression model for the financial leverage decisions of micro, small, medium and large firms. *Quantitative Finance, 9*(5), 621-636.

Jordan, S.J., Siskind, V., A, C.G., Whiteman, D.C., & Webb, P.M. (2010). Breastfeeding and risk of epithelial ovarian cancer. *Cancer Causes Control, 21*(1), 109-116.

Khanal, V., Lee, A.H., Karkee, R., & Binns, C.W. (2015). Postpartum Breastfeeding Promotion and Duration of Exclusive Breastfeeding in Western Nepal. *Birth, 42*(4), 329-336.

Kiernan, K., Tao, J., & Gibbs, P. (2012). *Tips and strategies for mixed modeling with SAS/STAT® procedures.* Paper presented at the SAS Global Forum.

Kim, Y.J., & Jhun, M. (2008). Cure rate model with interval censored data. *Stat Med, 27*(1), 3-14.

Klein, J.P., & Moeschberger, M.L. (2005). *Survival Analysis: Techniques for Censored and Truncated Data*: Springer New York.

Kuk, A.Y.C., & Chen, C.H. (1992). A Mixture Model Combining Logistic-Regression with Proportional Hazards Regression. *Biometrika, 79*(3), 531-541.

Kull, I., Almqvist, C., Lilja, G., Pershagen, G., & Wickman, M. (2004). Breast-feeding reduces the risk of asthma during the first 4 years of life. *J Allergy Clin Immunol, 114*(4), 755-760.

Lambert, D. (1992). Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. *Technometrics, 34*(1), 1-14.

Lee, A.H., Wang, K., Scott, J.A., Yau, K.K., & McLachlan, G.J. (2006). Multi-level zero-inflated poisson regression modelling of correlated count data with excess zeros. *Stat Methods Med Res, 15*(1), 47-61.

Lee, A.H., & Xiang, L. (2011). Mixture analysis of heterogeneous physical activity outcomes. *Ann Epidemiol, 21*(10), 780-786.

Lee, A.H., Xiang, L., & Hirayama, F. (2010). Modeling physical activity outcomes: a two-part generalized-estimating-equations approach. *Epidemiology, 21*(5), 626-630.

Lee, A.H., Zhao, Y., Yau, K.K., & Xiang, L. (2010). How to analyze longitudinal multilevel physical activity data with many zeros? *Prev Med, 51*(6), 476-481.

Liu, L., Strawderman, R.L., Cowen, M.E., & Shih, Y.C. (2010). A flexible two-part random effects model for correlated medical costs. *J Health Econ, 29*(1), 110-123.

Maller, R.A., & Zhou, S. (1992). Estimating the Proportion of Immunes in a Censored Sample. *Biometrika, 79*(4), 731-739.

Maller, R.A., & Zhou, S. (1994). Testing for Sufficient Follow-up and Outliers in Survival-Data. *Journal of the American Statistical Association, 89*(428), 1499-1506.

Maller, R.A., & Zhou, S. (1995). Testing for the presence of immune or cured individuals in censored survival data. *Biometrics, 51*(4), 1197-1205.

Maller, R.A., & Zhou, X. (1996). *Survival Analysis with Long-Term Survivors*. Chichester, U.K. : Wiley.

Manning, W.G., Morris, C.N., Newhouse, J.P., Orr, L.L., Duan, N., Keeler, E.B., et al. (1981). A two-part model of the demand for medical care: preliminary results from the health insurance study. *Health, economics, and health economics*, 103-123.

Min, Y.Y., & Agresti, A. (2005). Random effect models for repeated measures of zero-inflated count data. *Statistical Modelling, 5*(1), 1-19.

Mullahy, J. (1986). Specification and Testing of Some Modified Count Data Models. *Journal of Econometrics, 33*(3), 341-365.

Nardi, A., & Schemper, M. (2003). Comparing Cox and parametric models in clinical studies. *Stat Med, 22*(23), 3597-3610.

Niu, Y., & Peng, Y.W. (2014). Marginal regression analysis of clustered failure time data with a cure fraction. *Journal of Multivariate Analysis, 123*, 129-142.

Olsen, M.K., & Schafer, J.L. (2001). A two-part random-effects model for semicontinuous longitudinal data. *Journal of the American Statistical Association, 96*(454), 730-745.

Palmer, J.R., Viscidi, E., Troester, M.A., Hong, C.C., Schedin, P., Bethea, T.N., et al. (2014). Parity, lactation, and breast cancer subtypes in African American women: results from the AMBER Consortium. *J Natl Cancer Inst, 106*(10)

Peng, Y., & Dear, K.B. (2000). A nonparametric mixture model for cure rate estimation. *Biometrics, 56*(1), 237-243.

Peng, Y., & Taylor, J.M. (2011). Mixture cure model with random effects for the analysis of a multi-center tonsil cancer study. *Stat Med, 30*(3), 211-223.

Pohlmeier, W., & Ulrich, V. (1995). An Econometric Model of the Two-Part Decisionmaking Process in the Demand for Health Care. *The Journal of Human Resources, 30*(2), 339-361.

Qiu, L., Zhao, Y., Binns, C.W., Lee, A.H., & Xie, X. (2009). Initiation of breastfeeding and prevalence of exclusive breastfeeding at hospital discharge in urban, suburban and rural areas of Zhejiang China. *Int Breastfeed J, 4*, 1.

Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods* (2nd ed.). Thousand Oaks: SAGE Publications.

Scott, J.A., Binns, C.W., Oddy, W.H., & Graham, K.I. (2006). Predictors of breastfeeding duration: evidence from a cohort study. *Pediatrics, 117*(4), e646-655.

Seppa, K., Hakulinen, T., Kim, H.J., & Laara, E. (2010). Cure fraction model with random effects for regional variation in cancer survival. *Stat Med, 29*(27), 2781-2793.

Sposto, R. (2002). Cure model analysis in cancer: an application to data from the Children's Cancer Group. *Stat Med, 21*(2), 293-312.

Sposto, R., Sather, H.N., & Baker, S.A. (1992). A Comparison of Tests of the Difference in the Proportion of Patients Who Are Cured. *Biometrics, 48*(1), 87-99.

Stern, R.D. (1980). Analysis of daily rainfall at Samaru, Nigeria, using a simple two-part model. *Archiv für Meteorologie, Geophysik und Bioklimatologie, Serie B, 28*(1), 123-135.

Su, L., Tom, B.D., & Farewell, V.T. (2009). Bias in 2-part mixed models for longitudinal semicontinuous data. *Biostatistics, 10*(2), 374-389.

Sy, J.P., & Taylor, J.M. (2000). Estimation in a Cox proportional hazards cure model. *Biometrics, 56*(1), 227-236.

Tong, E.N.C., Mues, C., & Thomas, L.C. (2012). Mixture cure models in credit scoring: If and when borrowers default. *European Journal of Operational Research, 218*(1), 132-139.

Tooze, J.A., Grunwald, G.K., & Jones, R.H. (2002). Analysis of repeated measures data with clumping at zero. *Stat Methods Med Res, 11*(4), 341-355.

Vila-Candel, R., Duke, K., Soriano-Vidal, F.J., & Castro-Sanchez, E. (2017). Effect of Early Skin-to-Skin Mother-Infant Contact in the Maintenance of Exclusive Breastfeeding. *J Hum Lact*, 890334416676469.

World Health Organization. (2018). Breastfeeding.Retrieved from http://www.who.int/topics/breastfeeding/en/

Xiang, L.M., Ma, X.M., & Yau, K.K.W. (2011). Mixture cure model with random effects for clustered interval-censored survival data. *Statistics in Medicine, 30*(9), 995-1006.

Yau, K.K., & Lee, A.H. (2001). Zero-inflated Poisson regression with random effects to evaluate an occupational injury prevention programme. *Stat Med, 20*(19), 2907-2920.

Yau, K.K.W., Wang, K., & Lee, A.H. (2003). Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros. *Biometrical Journal, 45*(4), 437-452.

Yi, N., Lixin, S., Yufeng, L., & Yingwei, P. (2018). Modeling clustered long-term survivors using marginal mixture cure model. *Biometrical Journal*, *60*(4), 780-796.

Zyzanski, S.J., Flocke, S.A., & Dickinson, L.M. (2004). On the nature and analysis of clustered data. *Ann Fam Med*, *2*(3), 199-200.

Chapter 7: Conclusion and discussion

7.1 Summary

The work presented in this thesis is intended to shed light on some unsolved common methodological problems/ issues in statistical modelling of diverse types of breastfeeding data.

In line with its specific objectives, this thesis started with two systematic review and meta-analysis studies (as presented in Chapter 2 and Chapter 3) to examine the associations between two pertinent factors (namely caesarean section and maternal education) and breastfeeding practices in China using aggregated data from publications (original studies if applicable).

Next, different statistical models and approaches for modelling breastfeeding individual participant data sourced from longitudinally designed studies were proposed and presented in Chapter 4, Chapter 5 and Chapter 6. Some key epidemiological or statistical methodological problems were addressed, including time-dependent exposures modelling in longitudinal breastfeeding data analysis, time-dependent bias in modelling time-to-event breastfeeding data, and the issue of heterogeneity between clustering units and excess zeros in modelling clustered time-to-event breastfeeding data. Furthermore, this thesis also presented the applications of these statistical models and approaches on modelling empirical breastfeeding data collected from different studies as illustrative examples.

From a methodological perspective, we presented an updated statistical analysis framework for modelling breastfeeding data from various sources.

- In Chapter 2 and Chapter 3, we demonstrated the use of meta-analysis techniques to synthesise evidence from aggregated data. Cutting-edge meta-analysis techniques including fixed-effects/ random-effects methods, meta-regression, subgroup analysis, sensitivity analysis, publication bias test and visual inspection techniques (Galbraith plot and funnel plot) were used for modelling aggregated breastfeeding data.

- The method outlined in Chapter 4 addressed a crucial modelling issue of time-dependent exposures in longitudinal breastfeeding data analysis. We incorporated time-dependent exposures into the generalised linear mixed-effects model (GLMM). The time-dependent nature of the exposures and the inherent correlation within subjects are simultaneously accounted for in this subject-specific conditional model. Effect sizes of the time-dependent exposures estimated from the GLMM are averaged over all subjects with justification of subject-specific random-effects. Additionally, comparisons between repeated responses can be drawn from the modelling results. The GLMM incorporated with time-dependent exposures and random effects was applied to longitudinal breastfeeding data from two prospective cohort studies in China.
- In Chapter 5, we discussed time-dependent bias introduced by ignoring the time-varying nature of exposures and misclassification of time-varying exposures in the time-fixed analysis of time-to-event breastfeeding data. Biased estimation of exposure/ treatment effect can result from statistical analysis when time-dependent bias is not accounted for. We illustrated two different scenarios on occurrence times of time-varying exposures in a breastfeeding study setting followed by an illustrative example of modelling time-to-event breastfeeding data using a time-varying covariate approach via an extended Cox model. Two different computational methods in the time-varying covariate approach, namely, the programming statement and the counting process, were discussed and compared in this chapter. In the presence of time-varying exposures, the time-varying covariate approach generally yields unbiased estimates, whereas the time-fixed analysis may produce biased estimates with a smaller magnitude.
- In Chapter 6, we discussed the heterogeneity (between clustering units) issue in the statistical analysis of clustered time-to-event data, and both semi-parametric and parametric shared frailty models were applied to clustered time-to-event breastfeeding data from a community-based maternal cohort study in Nepal. Furthermore, when simultaneously dealing with the heterogeneity issue and another problem of excess zero observations, a novel two-part mixed-effects model for analysing clustered time-to-event data with clumping at zero was formulated. Two-part model framework has been

extensively applied to semi-continuous data (Duan et al., 1983; Manning et al., 1981; Olsen & Schafer, 2001; Tooze et al., 2002), zero-inflated count data (Cameron & Trivedi, 2013; Lambert, 1992; Min & Agresti, 2005; Yau & Lee, 2001; Yau et al., 2003), but sparsely in time-to-event data with clumping at zero (Brockett et al., 2008), especially clustered time-to-event data with clumping at zero. In our proposed two-part mixed-effects model, we structurally partitioned the longitudinal observations into two parts: a baseline binary process and a time-to-event process conditional on the positive responses in the first part. Both parts can be linked with one another via a possible correlation at the clustering unit level. Thus, this approach makes full use of all available information at baseline and during the follow-up.

From a public health implication perspective, this thesis provided some new insights and evidence for future relevant breastfeeding promotion and health policy decision making within infant nutrition or wider area.

- The results of the first systematic review and meta-analysis on the association between caesarean section and breastfeeding practices confirmed that the likelihood of breastfeeding was significantly lower after caesarean section compared with natural birth in China. Given the notably high rate of caesarean section in China, health policy and measures to improve breastfeeding outcomes should target the reduction of caesarean rate and health intervention after caesarean delivery in China.
- The second systematic review and meta-analysis on the association between maternal education and breastfeeding revealed a negative association between maternal education and breastfeeding practices, indicating that higher maternal education is associated with lower breastfeeding prevalence in China. This is a different finding compared to those conclusions drawn from studies in most Western countries. This finding suggests that relevant breastfeeding promotion programs may need to target special groups including better-educated mothers. Enhancing community acceptability of breastfeeding integrated with the supportive role of midwives is also an important aspect.

- The results from the study on the association between calcium supplementation postpartum and breastfeeding (Chapter 4) suggested that calcium supplementation postpartum was at a relatively low level for Chinese women, and breastfeeding mothers were more likely to take calcium supplements compared to their non-breastfeeding counterparts during the postpartum period. It is well known that calcium supplementation has significant impacts on bone health, especially for women who exclusively breastfeed their babies. Given the habitually lower calcium dietary intake, relatively high lactose intolerance rate and current wide shortage of dietary calcium intake and calcium supplementation in the general Chinese population, the findings of Chapter 4 suggest dietary supplementation intervention programs and health education should be promoted in Chinese women with special emphasis on different subgroups, such as breastfeeding mothers and bottle-feeding mothers.

7.2 Recommendations for future research directions

- Based on the synthesised evidence from the two systematic review and meta-analysis, future interventions focusing on promoting and improving breastfeeding practice and taking priority attentions to special groups, including mothers with caesarean delivery or attaining higher education, need to be pursued in China.
- Marginal approach for modelling longitudinal data with time-dependent covariates

  In this thesis, we addressed the issue of time-dependent exposures in statistical modelling longitudinal breastfeeding data by using a conditional approach via a GLMM, which estimates coefficients with subject-specific interpretations. When population-averaged interpretations of estimated coefficients are needed, another commonly used longitudinal data analysis approach, marginal model (including GEE and GMM) (Hu et al., 1998; Lai & Small, 2007; Lalonde et al., 2014; Pepe & Anderson, 1994; Zeger & Liang, 1986), can be utilised in the similar context.
- Simulation studies based on different scenarios evaluating the statistical property of candidate methods such as landmark analysis and time-dependent

Cox regression to account for time-dependent bias are recommended in future research.

- Two-part model development

  We proposed a two-part mixed-effects model to address the issue of augmented zeros in statistical modelling clustered time-to-event data in this thesis. The adaptive Gaussian quadrature technique was used to handle correlated random-effects between the two parts. However, in the illustrative example dataset, the estimation process of the two-part mixed-effects model was difficult to get converged. A Bayesian approach based on Markov Chain Monte Carlo methods may be an alternative approach for the statistical computation, although much longer computing time is anticipated (Cooper et al., 2007; Ghosh & Albert, 2009; Neelon et al., 2011; Zhang et al., 2006). Furthermore, comprehensive statistical simulation experiments to evaluate the properties and sensitivity of the proposed two-part mixed-effects model under different scenarios and conditions are warranted in further investigations.

7.3 Conclusion

This thesis contributes to the body of knowledge on statistical modelling of breastfeeding data. A series of updated and rigorous statistical models or approaches have been developed or utilised to address some key methodological problems in modelling breastfeeding data. Therefore, these models or approaches will be useful and beneficial to other researchers working in the field of infant nutrition and similar areas.

7.4 References

Brockett, P.L., Golden, L.L., Guillen, M., Nielsen, J.P., Parner, J., & Perez-Marin, A.M. (2008). Survival Analysis of a Household Portfolio of Insurance Policies: How Much Time Do You Have to Stop Total Customer Defection? *The Journal of Risk and Insurance, 75*(3), 713-737.

Cameron, A.C., & Trivedi, P.K. (2013). *Regression analysis of count data* (Vol. 53): Cambridge university press.

Cooper, N.J., Lambert, P.C., Abrams, K.R., & Sutton, A.J. (2007). Predicting costs over time using Bayesian Markov chain Monte Carlo methods: an application to early inflammatory polyarthritis. *Health Econ, 16*(1), 37-56.

Duan, N., Manning, W.G., Morris, C.N., & Newhouse, J.P. (1983). A Comparison of Alternative Models for the Demand for Medical Care. *Journal of Business & Economic Statistics, 1*(2), 115-126.

Ghosh, P., & Albert, P.S. (2009). A Bayesian analysis for longitudinal semicontinuous data with an application to an acupuncture clinical trial. *Computational Statistics & Data Analysis, 53*(3), 699-706.

Hu, F.B., Goldberg, J., Hedeker, D., Flay, B.R., & Pentz, M.A. (1998). Comparison of population-averaged and subject-specific approaches for analyzing repeated binary outcomes. *Am J Epidemiol, 147*(7), 694-703.

Lai, T.L., & Small, D. (2007). Marginal regression analysis of longitudinal data with time-dependent covariates: a generalized method-of-moments approach. *Journal of the Royal Statistical Society Series B-Statistical Methodology, 69*, 79-99.

Lalonde, T.L., Wilson, J.R., & Yin, J. (2014). GMM logistic regression models for longitudinal data with time-dependent covariates and extended classifications. *Stat Med, 33*(27), 4756-4769.

Lambert, D. (1992). Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. *Technometrics, 34*(1), 1-14.

Manning, W.G., Morris, C.N., Newhouse, J.P., Orr, L.L., Duan, N., Keeler, E.B., et al. (1981). A two-part model of the demand for medical care: preliminary results from the health insurance study. *Health, economics, and health economics*, 103-123.

Min, Y.Y., & Agresti, A. (2005). Random effect models for repeated measures of zero-inflated count data. *Statistical Modelling, 5*(1), 1-19.

Neelon, B., O'Malley, A.J., & Normand, S.L.T. (2011). A Bayesian Two-Part Latent Class Model for Longitudinal Medical Expenditure Data: Assessing the Impact of Mental Health and Substance Abuse Parity. *Biometrics, 67*(1), 280-289.

Olsen, M.K., & Schafer, J.L. (2001). A two-part random-effects model for semicontinuous longitudinal data. *Journal of the American Statistical Association, 96*(454), 730-745.

Pepe, M.S., & Anderson, G.L. (1994). A Cautionary Note on Inference for Marginal Regression-Models with Longitudinal Data and General Correlated Response Data. *Communications in Statistics-Simulation and Computation, 23*(4), 939-951.

Tooze, J.A., Grunwald, G.K., & Jones, R.H. (2002). Analysis of repeated measures data with clumping at zero. *Stat Methods Med Res, 11*(4), 341-355.

Yau, K.K., & Lee, A.H. (2001). Zero-inflated Poisson regression with random effects to evaluate an occupational injury prevention programme. *Stat Med*, *20*(19), 2907-2920.

Yau, K.K.W., Wang, K., & Lee, A.H. (2003). Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros. *Biometrical Journal*, *45*(4), 437-452.

Zeger, S.L., & Liang, K.Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, *42*(1), 121-130.

Zhang, M., Strawderman, R.L., Cowen, M.E., & T Wells, M. (2006). Bayesian inference for a two-part hierarchical model: An application to profiling providers in managed health care. *Journal of the American Statistical Association*, *101*(475), 934-945.

Appendix A Permission and statements of contribution from co-authors

This thesis contains a series of published work. The bibliographical details of the work are outlined below.

Chapter 2

Zhao, J., Zhao, Y., Du, M., Binns, C. W., & Lee, A. H. (2017). Does Caesarean Section Affect Breastfeeding Practices in China? A Systematic Review and Meta-Analysis. Maternal and child health journal, 21(11), 2008-2024. DOI: https://doi.org/10.1007/s10995-017-2369-x".

Chapter 3

Zhao, J., Zhao, Y., Du, M., Binns, C. W., & Lee, A. H. (2017). Maternal education and breastfeeding practices in China: A systematic review and meta-analysis. Midwifery, 50, 62-71. DOI: https://doi.org/10.1016/j.midw.2017.03.011

Chapter 4

Zhao, J., Zhao, Y., Binns, C. W., & Lee, A. H. (2016). Increased Calcium Supplementation Postpartum Is Associated with Breastfeeding among Chinese Mothers: Finding from Two Prospective Cohort Studies. Nutrients, 8(10), 622.". DOI: https://doi.org/10.3390/nu8100622

Chapter 5

Zhao, J., Zhao, Y., Lee, A. H., & Binns, C. W. (2017). A Time-varying Covariate Approach for Survival Analysis of Paediatric Outcomes. Paediatric and perinatal epidemiology, 31(6), 598-602.". DOI: https://doi.org/10.1111/ppe.12410

Chapter 6

Zhao, J., Zhao, Y., Khanal, V., Lee, A. H., & Binns, C. W. (2018) Application of a Frailty Modelling Approach to Correlated Breastfeeding Duration Data. Nursing Research, DOI: https://doi.org/10.1097/NNR.0000000000000311

Appendix A

Co-author permission and agreement

By signing below the co-author gives permission to include the aforementioned publications in this thesis and endorses that Jian Zhao is the primary contributor to the conception, design, execution, processing, analysis, and manuscript writing of the aforementioned publications.

Yun Zhao Signature: ___________________ Date: 13/08/2018

Andy H Lee Signature: ___________________ Date: 13/08/2018

Colin W Binns Signature: _______________ Date: 13/08/2018

Mengran Du Signature: __________________ Date: 13/08/2018

Appendix A

Vishnu Khanal Signature: ________________ Date: 13/08/2018

Appendix B Human research ethics approval letter

MEMORANDUM

Curtin University

To:	Prof Colin Binns
	School of Public Health
CC:	Mr. Jian Zhao
From:	Dr Catherine Gangell, Manager Research Integrity
Subject	Ethics approval
	Approval number: RDHS-101-15
Date:	08-Jun-15

Office of Research and Development
Human Research Ethics Office

TELEPHONE 9266 2784
FACSIMILE 9266 3793
EMAIL hrec@curtin.edu.au

Thank you for your application submitted to the Human Research Ethics Office for the project: 5987

Modelling breastfeeding data in China

Your application has been approved through the low risk ethics approvals process at Curtin University.

Please note the following conditions of approval:

1. Approval is granted for a period of four years from **08-Jun-15** to **08-Jun-19**
2. Research must be conducted as stated in the approved protocol.
3. Any amendments to the approved protocol must be approved by the Ethics Office.
4. An annual progress report must be submitted to the Ethics Office annually, on the anniversary of approval.
5. All adverse events must be reported to the Ethics Office.
6. A completion report must be submitted to the Ethics Office on completion of the project.
7. Data must be stored in accordance with WAUSDA and Curtin University policy.
8. The Ethics Office may conduct a randomly identified audit of a proportion of research projects approved by the HREC.

Should you have any queries about the consideration of your project please contact the Ethics Support Officer for your faculty, or the Ethics Office at hrec@curtin.edu.au or on 9266 2784. All human research ethics forms and guidelines are available on the ethics website.

Yours sincerely

Dr Catherine Gangell
Manager, Research Integrity

Appendix C for Chapter 2

C.1 Permission from publisher-Springer Nature

**SPRINGER NATURE LICENSE
TERMS AND CONDITIONS**

May 06, 2018

This Agreement between Mr. Jian Zhao ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4343140765220
License date	May 06, 2018
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Maternal and Child Health Journal
Licensed Content Title	Does Caesarean Section Affect Breastfeeding Practices in China? A Systematic Review and Meta-Analysis
Licensed Content Author	Jian Zhao, Yun Zhao, Mengran Du et al
Licensed Content Date	Jan 1, 2017
Licensed Content Volume	21
Licensed Content Issue	11
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	electronic
Portion	full article/chapter
Will you be translating?	no
Circulation/distribution	<501
Author of this Springer Nature content	yes
Title	Statistical modelling of breastfeeding data
Instructor name	Dr Yun Zhao
Institution name	Curtin University
Expected presentation date	Jul 2018
Requestor Location	Mr. Jian Zhao 2/82 Reynolds Rd Mount Pleasant Perth, WA 6153 Australia Attn: Mr. Jian Zhao
Billing Type	Invoice
Billing Address	Mr. Jian Zhao 2/82 Reynolds Rd Mount Pleasant Perth, Australia 6153 Attn: Mr. Jian Zhao
Total	0.00 AUD
Terms and Conditions	

Springer Nature Terms and Conditions for RightsLink Permissions

Appendix C

Springer Customer Service Centre GmbH (the Licensor) hereby grants you a non-exclusive, world-wide licence to reproduce the material and for the purpose and requirements specified in the attached copy of your order form, and for no other use, subject to the conditions below:

1. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of another entity (as credited in the published version).

   If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.

2. Where **print only** permission has been granted for a fee, separate permission must be obtained for any additional electronic re-use.

3. Permission granted **free of charge** for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.

4. A licence for 'post on a website' is valid for 12 months from the licence date. This licence does not cover use of full text articles on websites.

5. Where **'reuse in a dissertation/thesis'** has been selected the following terms apply: Print rights for up to 100 copies, electronic rights for use only on a personal website or institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/).

6. Permission granted for books and journals is granted for the lifetime of the first edition and does not apply to second and subsequent editions (except where the first edition permission was granted free of charge or for signatories to the STM Permissions Guidelines http://www.stm-assoc.org/copyright-legal-affairs/permissions/permissions-guidelines/), and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence.

7. Rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to Journalpermissions@springernature.com/bookpermissions@springernature.com for these rights.

8. The Licensor's permission must be acknowledged next to the licensed material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract, and must be hyperlinked to the journal/book's homepage. Our required acknowledgement format is in the Appendix below.

9. Use of the material for incidental promotional use, minor editing privileges (this does not include cropping, adapting, omitting material or any other changes that affect the meaning, intention or moral rights of the author) and copies for the disabled are permitted under this licence.

10. Minor adaptations of single figures (changes of format, colour and style) do not require the Licensor's approval. However, the adaptation should be credited as shown in Appendix below.

Appendix — Acknowledgements:

For Journal Content:
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

For **Advance Online Publication papers:**
Reprinted by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj.[JOURNAL ACRONYM].)

For Adaptations/Translations:
Adapted/Translated by permission from [**the Licensor**]: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication)

Note: For any republication from the British Journal of Cancer, the following credit line style applies:

Reprinted/adapted/translated by permission from [**the Licensor**]: on behalf of Cancer Research UK: : [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT]** (year of publication)

For **Advance Online Publication** papers:
Reprinted by permission from The [**the Licensor**]: on behalf of Cancer Research UK: [**Journal Publisher** (e.g. Nature/Springer/Palgrave)] [**JOURNAL NAME**] [**REFERENCE CITATION** (Article name, Author(s) Name), [**COPYRIGHT**] (year of publication), advance online publication, day month year (doi: 10.1038/sj. [JOURNAL ACRONYM])

For Book content:
Reprinted/adapted by permission from [**the Licensor**]: [**Book Publisher** (e.g. Palgrave Macmillan, Springer etc) [**Book Title**] by [**Book author**(s)] [**COPYRIGHT**] (year of publication)

Other Conditions:

Version 1.0

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

C.2 Publication 1

Appendix C

Matern Child Health J (2017) 21:2008–2024
DOI 10.1007/s10995-017-2369-x

REVIEW PAPER

Does Caesarean Section Affect Breastfeeding Practices in China? A Systematic Review and Meta-Analysis

Jian Zhao[1] · Yun Zhao[1] · Mengran Du[2] · Colin W. Binns[1] · Andy H. Lee[1]

Published online: 10 October 2017

Abstract *Objectives* To ascertain the association between caesarean delivery and breastfeeding practices in China. *Methods* We conducted a systematic review and meta-analysis following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and Meta-analysis Of Observational Studies in Epidemiology (MOOSE) guidelines. Electronic databases of CNKI, Medline, EMBASE, CINAHL, ProQuest and Science Direct were searched and screened to identify relevant articles from January 1990 to June 2015. Both fixed and random effect meta-analysis techniques were used to estimate the pooled effect size between caesarean delivery and breastfeeding outcomes at different time points. Sensitivity analysis and publication bias test were also conducted. *Results* Forty six studies were eligible for the qualitative synthesis of systematic review; among them, 27 studies were included for the meta-analysis. At the early postpartum period, the odds of exclusive breastfeeding after caesarean section was 47% (pooled OR 0.53, 95% CI 0.41, 0.68) lower than that after vaginal delivery. At 4 months postpartum, the odds of breastfeeding was similarly lower (pooled OR 0.61, 95% CI 0.53, 0.71) for caesarean mothers. Substantial heterogeneity among studies was detected for both breastfeeding outcomes. Subgroup analyses stratified by study design, time points of breastfeeding outcomes and definitions of breastfeeding all confirmed the negative association between caesarean section and breastfeeding prevalence. *Conclusions* In China, breastfeeding practices were affected adversely by caesarean delivery. Therefore, health policy to improve breastfeeding outcomes should take this into consideration.

✉ Yun Zhao
y.zhao@curtin.edu.au

1 School of Public Health, Curtin University, Kent Street, Bentley, GPO Box U 1987, Perth, WA 6845, Australia

2 West China School of Public Health, Sichuan University, No.17 Section 3 Renmin South Road, Chengdu, China

Keywords Breastfeeding · Caesarean section · Systematic review · Meta-analysis · China

Significance

Caesarean delivery has been identified as one of important factors that adversely affect breastfeeding practices, but the findings remain inconsistent. A systematic review and meta-analysis of worldwide literature confirmed the negative relationship between caesarean section and breastfeeding practices, yet systematic review and meta-analysis of the studies conducted in China published in Chinese or English have never been done. This study is the first study systematically examining the association between caesarean section and breastfeeding practices in China in such a situation that China has been cited as having one of the highest rates of caesarean section globally.

Introduction

Breastfeeding has always been the physiological norm (Berry and Gribble 2008). Non-breastfed or early weaned infants are more likely to get adverse health outcomes as well as their mothers (Eidelman et al. 2012, Ip et al. 2007). However, the rate of breastfeeding, especially exclusive breastfeeding, still remains below the optimal level worldwide (Guo et al. 2013; Merewood et al. 2005; Ryan et al. 2002). Factors affecting breastfeeding have been studied extensively (Harley et al. 2007; Pearce et al. 2012; Scott

Appendix C

et al. 1999; Senarath et al. 2010). Caesarean delivery has been identified as one of important factors negatively associated with breastfeeding practices; however, the findings remain inconsistent. For example, some studies reported the negative associations between caesarean delivery and breastfeeding outcomes such as breastfeeding initiation, breastfeeding rates at the first week and at 4 months postpartum (Guo et al. 2013; Kohlhuber et al. 2008; Patel et al. 2015), while other studies reported no association between caesarean delivery and breastfeeding rates at 42 days postpartum, and breastfeeding prevalence within 6 months postpartum (Joshi et al. 2014; Patel et al. 2015). A systematic review and meta-analysis on the topic of association between caesarean section and breastfeeding was conducted in 2012 (Prior et al. 2012), in which studies carried out in China and published in the English language were retrieved using the PubMed database. The review concluded that pre-labour caesarean section was associated with early breastfeeding negatively and no statistically significant association was found with the breastfeeding rate at 6 months postpartum (Prior et al. 2012). However, the majority of breastfeeding studies undertaken in China were published in Chinese and hence excluded from this review (Prior et al. 2012). Some degree of bias may have resulted from this omission.

Understanding the relationship between caesarean section and breastfeeding practices accurately in more depth is important to country such as China, where the rate of caesarean section is high at 46.2% in 2007–2008 and 56.1% in 2011, which are nearly double the global rate (Feng et al. 2012; Lumbiganon et al. 2010; Tang et al. 2006). An updated systematic review and meta-analysis based on both English and Chinese literature would reduce such bias and improve the knowledge of the relationship between caesarean delivery and breastfeeding practices in China. Therefore, the aim of the present study is to examine the association between caesarean delivery and breastfeeding practices including 'exclusive breastfeeding during the early postpartum period' and 'breastfeeding at 4 months postpartum'. Specifically, we test the hypothesis that caesarean delivery reduces breastfeeding intensity and prevalence.

Methods

Search Strategy

A systematic electronic search of both Chinese and English language articles on breastfeeding and method of delivery was conducted using the Chinese database China National Knowledge Infrastructure (CNKI) as well as Medline, EMBASE, CINAHL, ProQuest, and Science Direct from January 1990 to June 2015. A two-stage search strategy was adopted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and Meta-analysis Of Observational Studies in Epidemiology (MOOSE) guidelines (Moher et al. 2009; Stroup et al. 2000).

Stage 1: The following Medical Subject Headings (MeSH) terms and key words "breast feeding", "milk human", "breastfeeding duration", "breastfeeding cessation", "human lactation", "infant feed*", "breastfed", "risk factor*", "protective factor*", "determinant*", "socioeconomic factor*", "China", "mainland China", "Chinese" were used.

Stage 2: The following MeSH terms and key words "caesarean delivery", "cesarean delivery", "caesarean section", "cesarean section" and "c-section" were further added to the search process.

Corresponding Chinese terms and key words were used to search in the database of CNKI.

Literature Screening and Selection Criteria

At the initial screening stage, abstracts of eligible publications were retrieved by two independent reviewers (JZ and MRD). Relevant citations were identified after screening the abstracts and their full-texts were then obtained and evaluated. Discrepancies on relevancy were resolved through consensus or referred to a third investigator (YZ) when necessary. Articles were included if they met the following criteria: (i) published in peer-reviewed journals or theses/dissertations; (ii) observational study design; (iii) reported the association between caesarean delivery and breastfeeding quantitatively; (iv) effect size could be obtained directly or calculated from raw tabulated data. The exclusion criteria were as follows: (i) studies that did not specify sample size; (ii) studies that did not report or define time points of breastfeeding outcomes; (iii) studies that reported inappropriate statistical result (statistical error).

Data Extraction

The following information was extracted from each eligible study for qualitative and quantitative synthesis: publication year, name of the first author, study design, location of study, sample size, breastfeeding outcomes (including definitions of breastfeeding, types of breastfeeding, time points of measurements), other factors associated with the breastfeeding outcome, and raw tabulated data or effect size (odds ratios) reported via univariate analysis or multivariate analysis. In cases where the relevant results or raw data were missing, the authors were contacted by email.

Odds ratios (ORs), either crude or adjusted, estimated from logistic regression analysis of 'exclusive breastfeeding during the early postpartum period' and 'breastfeeding at 4 months postpartum' for caesarean section versus vaginal delivery and their corresponding 95% confidence intervals (CIs), were extracted from the eligible studies.

Appendix C

The definitions of breastfeeding used in the data extraction followed the World Health Organization (WHO) definitions (World Health Organization 2003, 2008):

Exclusive breastfeeding: Breastfeeding while giving no other food or liquid, not even water, with the exception of drops or syrups consisting of vitamins, mineral supplements or medicines.

Full breastfeeding: Exclusive breastfeeding or predominant breastfeeding (or almost exclusive breastfeeding). Breastmilk is the only source of milk given to the infant regardless of supplementation with other fluids such as water and orange juice.

Any breastfeeding: The child has received breastmilk (direct from the breast or expressed) with or without other drink, formula or other infant food.

Quality Assessment

To assess methodological quality of the selected studies, we developed a checklist based on the criteria proposed in Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) (von Elm et al. 2007) and the previous checklist proposed by Tooth et al. (2005). The possible score of our checklist ranges from 0 to 18, with scores above 14, between 11 and 14, and below 11 indicating high, medium and low quality, respectively (see Appendix).

Statistical Analysis

The odds ratios of 'exclusive breastfeeding during the early postpartum period' (defined as initiation of breastfeeding or exclusive breastfeeding before discharge or exclusive breastfeeding at 42 days post birth) and 'breastfeeding (including exclusive breastfeeding, full breastfeeding and any breastfeeding) at 4 months postpartum', for caesarean section versus vaginal delivery, were the primary outcomes of interest. A meta-analysis was performed to determine the pooled effect size of caesarean delivery on these two breastfeeding outcomes separately. Based on the raw data extraction from the selected studies, the natural logarithmic transformed ORs were used in the meta-analysis.

Fixed-effect [inverse variance (I-V) method of fixed-effect model] meta-analysis was performed initially and heterogeneity across studies was assessed by the I-square statistic (Higgins and Thompson 2002; Higgins et al. 2003). A random-effect model [DerSimonian and Laird (D + L) method of random-effect model] was further utilised and presented when the heterogeneity was confirmed statistically significant (DerSimonian and Laird 1986). A meta-regression was then performed to investigate potential sources contributing to the heterogeneity. Subgroup analysis was undertaken to assess the magnitude of effect on the outcome variables of interest under different stratifications.

To test the dependence of effect size reported in each study, sensitivity analysis using the jackknife approach was performed to assess the robustness of the results (Miller 1974). Such sensitivity analysis was repeated multiple times with one study removed per cycle.

In order to ascertain publication bias and small sample size bias among the studies, Begg's funnel plot/test and Egger's test were applied (Egger et al. 1997). All calculations and statistical analyses were performed using the Stata package version 14.1 (StataCorp LP, College Station, USA). A p value less than 0.05 was considered as statistically significant.

Results

Systematic Review

As shown in Fig. 1, a total of 1061 records were identified from both English and Chinese databases, and five more records were obtained from other sources. After removal of duplicates, 1054 articles were screened by manually reading titles and abstracts. As a result 197 articles were deemed eligible for full-text review. After formal review, 151 of them were excluded (86 qualitative studies, 54 not suitable for analysis, 4 full-text Chinese publications incomplete, 1 sample size not specified, 3 with statistical errors, 3 duplicated the same study). One author was contacted to obtain additional data to calculate the rate of exclusive breastfeeding initiation (Xu 2008). Finally, 46 articles (38 published in Chinese and 8 published in English, respectively) were included for the qualitative synthesis of systematic review (Table 1), among which 13 were suitable for meta-analysis of the association between caesarean delivery and 'exclusive breastfeeding during the early postpartum period', and 14 were appropriate for meta-analysis of the effect of caesarean delivery on 'breastfeeding at 4 months postpartum' (Chen et al. 2010; Fang et al. 1996; Gan et al. 2007; Guo et al. 2013; He et al. 1994; Huang and Lu 2010; Jiang 2000; Jiang and Li 2008; Kang et al. 2013; Leng 2014; Li 2014; Liu 2008; Liu and Xing 1998; Liu et al. 2014, 2012; Liu and Shao 2009; Ma et al. 2009; Qin and Hua 2013; Qiu 2008; Ruan et al. 2012; Tang 2013, 2014; Tian et al. 2008; Wang et al. 2006, 2005, 2009, 1995, 2013; Wang 2010; Wei and Li 2009; Xu 2008; Xu and Yu 2009; Xue et al. 2012; Yang and Feng 2014; Ye 2008; Yin et al. 2012; Yu 2013; Zhang and Wang 2000; Zhang and Shi 2013; Zhang and Wu 1999; Zhang et al. 2006, 2013; Zhang 2012; Zheng et al. 2008; Zhu et al. 2013, 2014). Among these 46 articles 20 of them were assessed to be high scientific quality, 24 medium quality

Appendix C

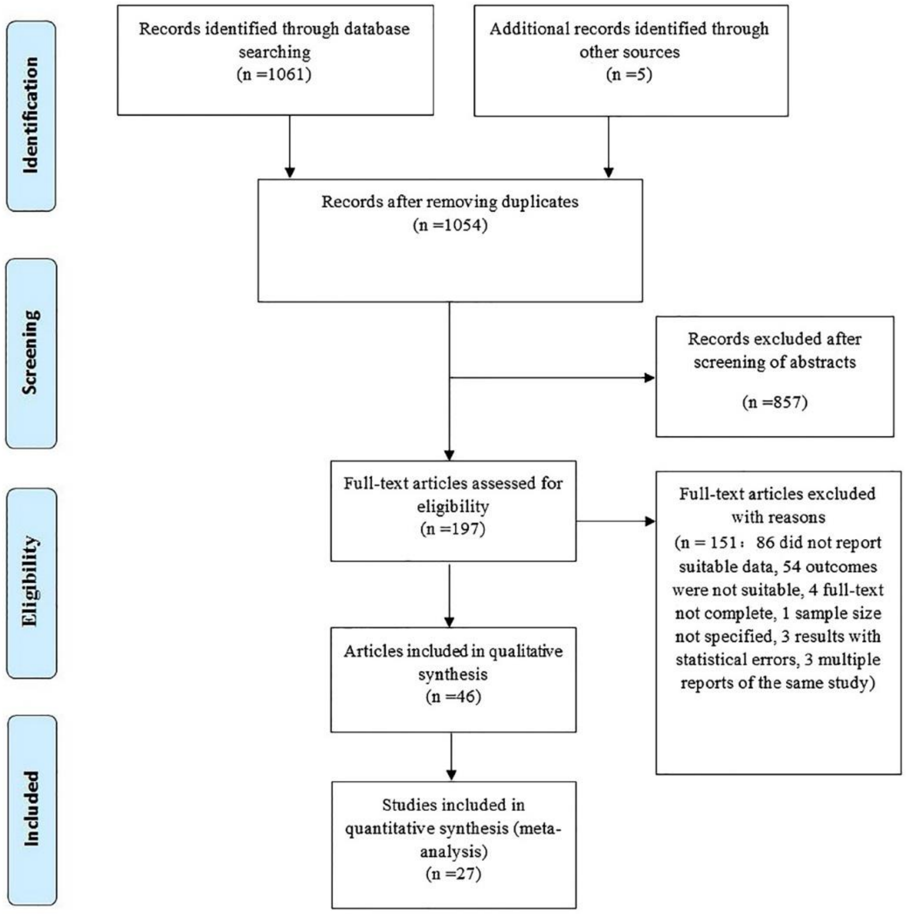


Fig. 1 PRISMA flow chart of the systematic review process

and 2 low quality, according to the methodological quality checklist scores in Appendix.

Effect of Caesarean Delivery on 'Exclusive Breastfeeding During The Early Postpartum Period'

The random-effect model meta-analysis of the 13 studies (441,044 subjects: 27,152 in the caesarean delivery group and 413,892 in the vaginal delivery group) showed that the odds of 'exclusive breastfeeding during the early postpartum period' was 47% (pooled OR 0.53, 95% CI 0.41, 0.68) lower in the caesarean delivery group than the vaginal delivery group (Fig. 2), with a significant heterogeneity in effect sizes evident across the studies ($I^2 = 90.6\%$, $p < 0.001$).

The stepwise meta-regression analysis revealed no clear main sources for heterogeneity, which may be due to the small number of studies included in this meta-analysis

Appendix C

Table 1 Characteristics of studies assessing the association between caesarean delivery and breastfeeding in China

Published year	First author	Study design	Location	Sample size	Other factors associated with breastfeeding outcomes	Breastfeeding outcome measured	Definition	Score
1994	HE,H.L	Cross-sectional	Fuzhou	216	Initiation time, milk bottle using, sleep, alcohol drinking	Breastfeeding at 1 month postpartum	Non-WHO	13
1995	WANG,S	Cross-sectional	Beijing	439	BFHI, maternal age, infant gender, gestational reaction	FB at 4 months postpartum	Reclassified with WHO definition	13
1996	FANG,L.L	Cross-sectional	Beijing	60	Initiation time	EBF at 42 days/4 months postpartum	WHO	13
1998	LIU,J	Cross-sectional	Inner Mongolia	374	Maternal health condition, complication, nutrition condition, appetite after delivery, maternal education, initiation time	Breastfeeding 0–90 days postpartum	Non-WHO	11
1999	ZHANG,S.J	Prospective cohort	Hebei	207	Knowledge of breastfeeding, confidence of breastfeeding, initiation time, breast milk substitute	Breastfeeding at 1 month postpartum	Non-WHO	15
2000	ZHANG,G.D	Prospective cohort	Chongqing	627	Postpartum hemorrhage	Breastfeeding 28 days/4 months postpartum	Non-WHO	11
2000	JIANG,G.F	Cross-sectional	Liuhe	736	Maternal education, rooming-in	EBF within 6 months postpartum	WHO	15
2005	WANG,C.X	Retrospective cohort	Jinan	853	Maternal education, health education, milk powder promotion, initiation time	EBF at 4 months postpartum	Non-WHO	16
2006	WANG,B.S	Prospective cohort	Shanghai	602	NA	FB at 1,6,12 months postpartum	Reclassified with WHO definition	13
2006	ZHANG,W.K	Prospective cohort	Beijing	802	Maternal age, early touch time, perception of breastfeeding during gestation, living space	FB 2–5 days, 42 days postpartum	Reclassified with WHO definition	16
2007	GAN,W.L	Cross-sectional	Chongqing	375	Maternal age, maternal education, occupation, maternal mood, postpartum home visit, monthly income	EBF at 4 months postpartum	Non-WHO	12
2008	XU,F.L	Prospective cohort	Xinjiang	1064	Giving breastmilk as the first feed, feeding on demand, maternal perception of the breastfeeding information received, minority ethnic group, giving birth in spring or summer, medical staff not recommending formula to parents, prelacteal feeds of water or formula	AF initiation/EBF initiation	WHO	18

Appendix C

Table 1 (continued)

Published year	First author	Study design	Location	Sample size	Other factors associated with breastfeeding outcomes	Breastfeeding outcome measured	Definition	Score
2008	LIU,F.L	Retrospective cohort	Xinxiang	288 392 415 376	NA	AF at 4 months postpartum in year 2000, 2002, 2004 and 2006 separately	Non-WHO	13
2008	ZHENG,K.Y	Cross-sectional	Hangzhou	628	Neonatal disease, initiation time, early feeding	AF before discharge	Reclassified with WHO definition	16
2008	TIAN,J.Z	Cross-sectional	Zhejiang	253	Neonate disease, early sucking, initiation time, breastfeeding confidence, maternal education	FB before discharge	Reclassified with WHO definition	14
2008	YE,C.E	Cross-sectional	Ninghai	931	Maternal education, health education	EBF at 4 months postpartum	Non-WHO	16
2008	JIANG,Z.H	Retrospective cohort	Herbin	310	Initiation time, maternal education	EBF at 42 days postpartum	Non-WHO	14
2008	QIU,L.Q	Prospective cohort	Zhejiang	917	Living in the suburb or rural areas, maternal age, mother decides to breastfeed until after birth, prelacteal feeding	EBF initiation/at discharge	WHO	18
2009	LIU,X.Q	Cross-sectional	Beijing	123	Health education, initiation time	Breastfeeding at 42 days postpartum	Non-WHO	11
2009	XU,T	Retrospective cohort	Shenyang	1025	Maternal age, family income	FB at 6 months postpartum	Reclassified with WHO definition	17
2009	WEI,X.J	Retrospective cohort	Zhengzhou	501 445 458 503 541 399 340 223 347 284 250 246	NA	FB 42 days postpartum in year 1996–2007 separately	Reclassified with WHO definition	16
2009	WANG,R	Cross-sectional	Kunming	1328	NA	EBF initiation, at 4 months postpartum	Non-WHO	15

Appendix C

Matern Child Health J (2017) 21:2008–2024

Table 1 (continued)

Published year	First author	Study design	Location	Sample size	Other factors associated with breastfeeding outcomes	Breastfeeding outcome measured	Definition	Score
2009	MA,X	Prospective cohort	Shaanxi	605	Initiation time, perception of breastfeeding, supplementary feeding, confidence of breastfeeding, milk bottle using	EBF before discharge and 1 month postpartum	Non-WHO	13
2010	CHEN,Y.F	Cross-sectional	Wuhan	445	Prenatal preparation, initiation time, right feeding method, feeding confidence, maternal nutrition, feeding setting	Breastfeeding within 4–6 months postpartum	Non-WHO	13
2010	WANG,H.Z	Cross-sectional	Changli	1296	Health education, maternal education	EBF at 6 months postpartum	WHO	13
2010	HUANG,Q.J	Cross-sectional	Shanghai	350	Early feeding, milk bottle using	EBF before discharge	WHO	13
2012	RUAN,M.J	Cross-sectional	Beijing	103	Maternal education, fixed term job, maternal age	EBF at 4 months postpartum	Non-WHO	13
2012	ZHANG,Y.M	Prospective cohort	Chengdu	268	NA	EBF at discharge	Non-WHO	15
2012	XUE,F	Cross-sectional	Changshu	126	Area (rural or urban)	EBF at 4 months postpartum	Non-WHO	13
2012	YIN,X.G	Prospective cohort	Hefei	2522	Maternal education, family income, preterm birth	AF at 2 months, 4 months postpartum	Reclassified with WHO definition	18
2012	LIU,X	Retrospective cohort	27 sites	431,704	NA	EBF before discharge	WHO	17
2013	ZHANG,Q.J	Cross-sectional	Zhengzhou	612	NA	EBF 24 hours before discharge	Non-WHO	8
2013	ZHANG,Y.X	Cross-sectional	Danyang	3057	First child breastfeeding duration, perception of breastmilk amount, psychosocial factors	FB before discharge and after discharge	WHO	15
2013	KANG,Y	Cross-sectional	Chongqing	939	Infant gender, maternal education, monthly income, birth weight, duration of maternity leave, perception of breastmilk amount, prelacteal feeding	Breastfeeding at 6 months	Non-WHO	14
2013	WANG,Z	Cross-sectional	Zhejiang	528	Infant age, infant gender, early feeding, perception of breast-milk amount	EBF within 6 months postpartum	WHO	16
2013	GUO,S.F	Cross-sectional	26 counties	2293	Maternal antenatal clinic visit, infant age	EBF initiation /within 6 months postpartum	WHO	14
2013	ZHU,P	Prospective cohort	Hefei	1602	Preterm birth, breastfeeding frequency on Day 1, life events in the third trimester, onset of lactation	AF at 2 month postpartum	Reclassified with WHO definition	14
2013	QIN,L.L	Retrospective cohort	Suzhou	1212	Maternal occupation, maternal age, birth region, breastfeeding professional instruction	EBF at 4 months postpartum	WHO	17

Table 1 (continued)

Published year	First author	Study design	Location	Sample size	Other factors associated with breastfeeding outcomes	Breastfeeding outcome measured	Definition	Score
2013	TANG,L	Prospective cohort	Jiangyou	693	Father's attitude towards breastfeeding, early breastfeeding initiation	EBF initiation within 1 hour post birth/AF at discharge/FB at discharge	WHO	18
2013	YU,C	Prospective cohort	Chengdu	845	Maternal occupation, paternal education, intention of going back to work, first feeding, mothers' friends breastfeed their babies, paternal occupation, staff encouragement, father's attitude, maternal grandmother's breastfeeding history	AF within 15 days postpartum/FB within 15 days postpartum	WHO	18
2014	TANG,Z.J	Cross-sectional	Guangzhou	315	NA	BF at 6 months postpartum	Reclassified with WHO definition	13
2014	LIU,L.F	Cross-sectional	Lishui	675	Family monthly income, early initiation, health education, neonatal disease	EBF within 6 months	WHO	16
2014	LENG,X.L	Cross-sectional	Shenzhen	1200	Maternal age, initiation time	FB within 6 months postpartum	Reclassified with WHO definition	13
2014	YANG,Y.L	Cross-sectional	Wuhan	513	Health education, maternal education, prenatal risk factors	EBF within 6 months postpartum	Non-WHO	8
2014	LI,R	Cross-sectional	Shenzhen	840	Initiation time, nipples condition	EBF before discharge	Non-WHO	13
2014	ZHU,X	Cross-sectional	3 cities	151	Occupation status, initiation time, frequency of lactation per day	EBF at 6 months postpartum	Non-WHO	14

BFHI Baby friendly hospital initiative, *FB* full breastfeeding, *EBF* exclusive breastfeeding *AF* any breastfeeding

Appendix C

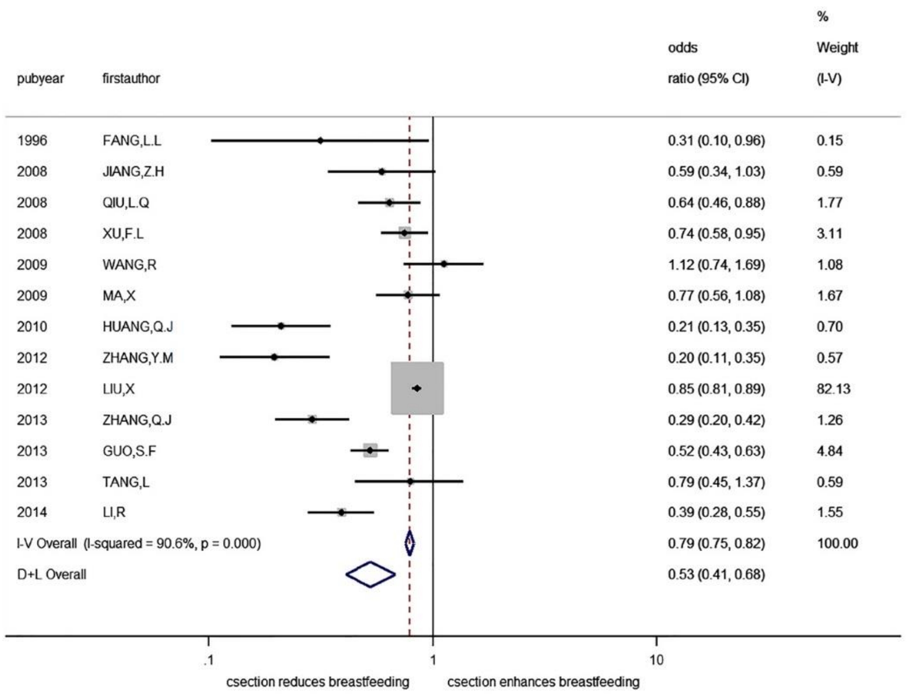


Fig. 2 Forest plot showing the fixed-effect and random-effect meta-analysis for the exclusive breastfeeding during the early postpartum period

(Higgins and Green 2011). Exploratory subgroup analyses were further conducted with regard to study design (cross-sectional, prospective cohort and retrospective cohort), exclusive breastfeeding time points (initiation, before discharge and at 42 days after birth) and definitions of breastfeeding (WHO and non-WHO).

Study Design

Of the 13 studies, six adopted a cross-sectional design, two used retrospective design and five were prospective cohort studies. Considerable heterogeneity remained present in both cross-sectional and prospective cohort subgroups ($I^2 = 86.1\%$, $p < 0.001$ and $I^2 = 79.9\%$, $p = 0.001$, respectively) whereas no significant heterogeneity was observed in the retrospective cohort subgroup ($I^2 = 38.5\%$, $p = 0.202$). Meta-analysis stratified by the study design (Fig. 3) showed that the adverse effect of caesarean delivery remained on the exclusive breastfeeding prevalence during the early postpartum period for the three study designs (cross-sectional: pooled OR 0.42, 95% CI 0.27, 0.64; retrospective: pooled OR 0.85, 95% CI 0.81, 0.89; prospective: pooled OR 0.59, 95% CI 0.41, 0.85).

Exclusive Breastfeeding Time Points

Figure 4 presents the subgroup analysis stratified by time points (initiation, before discharge and at 42 days post birth) of exclusive breastfeeding outcomes measured. Again, the negative association between caesarean delivery and the exclusive breastfeeding prevalence during the early postpartum period remained persistent (initiation: pooled OR 0.71, 95% CI 0.55, 0.91; before discharge: pooled OR 0.39, 95% CI 0.23, 0.67; at 42 days post birth:

Appendix C

Matern Child Health J (2017) 21:2008–2024 2017

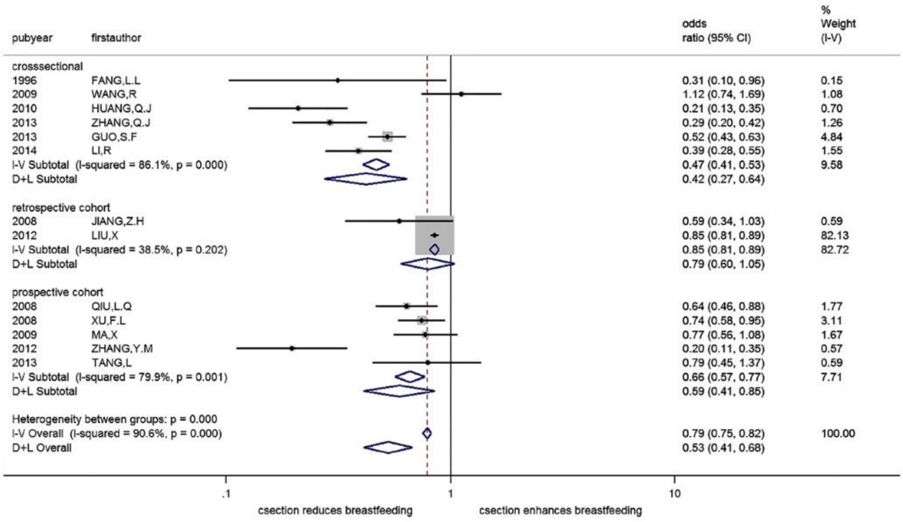


Fig. 3 Subgroup analysis for exclusive breastfeeding during the early postpartum period stratified by study design

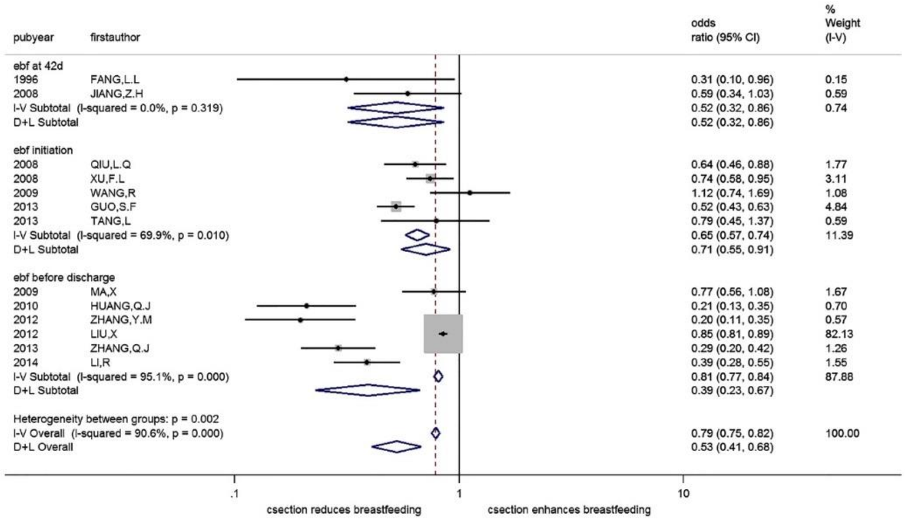


Fig. 4 Subgroup analysis for exclusive breastfeeding during the early postpartum period stratified by time points of breastfeeding outcomes measured

Springer

Appendix C

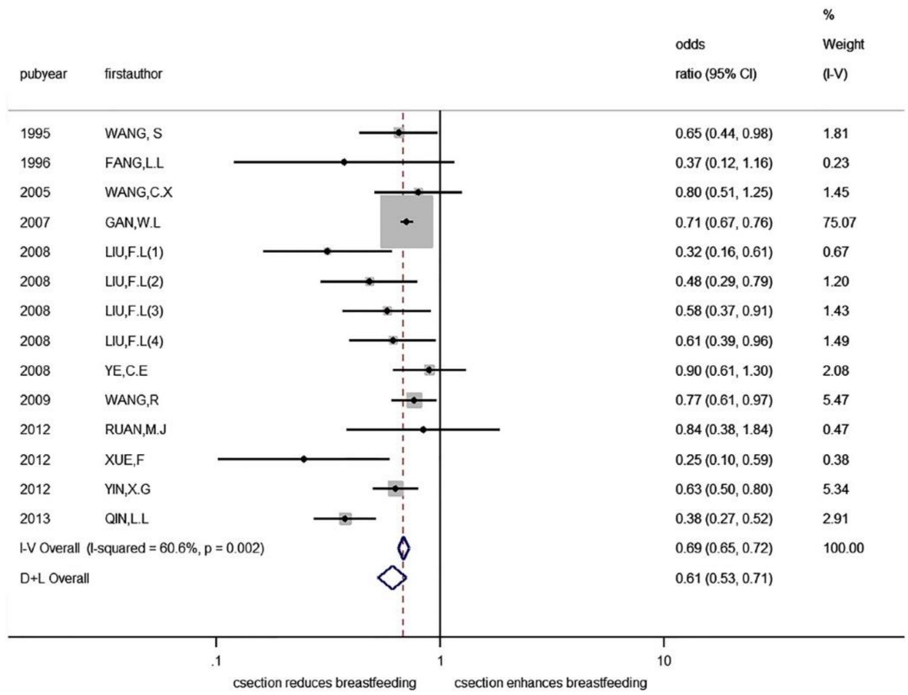


Fig. 5 Forest plot showing the fixed-effect and random-effect meta-analysis for breastfeeding at 4 months postpartum

pooled OR 0.52, 95% CI 0.32, 0.86). Except the third subgroup ($I^2 = 0.0\%$, $p = 0.319$), significant heterogeneity was evident for the first two subgroups (initiation: $I^2 = 69.9\%$, $p = 0.010$ and before discharge: $I^2 = 95.1\%$, $p < 0.001$, respectively).

Definitions of Breastfeeding

Figure 6 in Appendix presents the result of subgroup analysis stratified by the definitions of breastfeeding (WHO and non-WHO). The adverse effect of caesarean delivery on the exclusive breastfeeding prevalence during the early postpartum period was confirmed significant for both subgroups (WHO: pooled OR 0.58, 95% CI 0.43, 0.77; non-WHO: pooled OR 0.48, 95% CI 0.29, 0.78). Significant heterogeneity was detected as well (WHO: $I^2 = 89.3\%$, $p < 0.001$; non-WHO: $I^2 = 88.2\%$, $p < 0.001$).

Effect of Caesarean Delivery on 'Breastfeeding at 4 Months Postpartum'

Fourteen studies were included for this meta-analysis, among which one study (Liu 2008) had four sub-studies carried out independently in four different periods and consequently these sub-studies were regarded as separate studies in the meta-analysis. Figure 5 shows that the prevalence of breastfeeding at 4 months postpartum was significantly lower after caesarean section (pooled OR = 0.61, 95% CI 0.53, 0.71). Significant heterogeneity ($I^2 = 60.6\%$, $p = 0.002$) was found.

Among the 14 studies, four used the WHO definitions. For the subgroup of studies using WHO definitions, the odds of breastfeeding at 4 months postpartum reduced by 48% (pooled OR 0.52, 95% CI 0.38, 0.72) when compared with that using non-WHO definitions (pooled OR 0.66, 95% CI 0.56, 0.77), as indicated in Figure 7 in Appendix.

Appendix C

Heterogeneity was found marginally significant in both subgroups (WHO: $I^2 = 62.0\%$, $p = 0.048$; non-WHO: $I^2 = 47.0\%$, $p = 0.049$).

Sensitivity Analysis

The results of sensitivity analysis for both breastfeeding outcomes showed that the pooled ORs remained significant when one study was omitted at a time during recalculation of the pooled ORs, suggesting the pooled ORs were not substantially influenced by any individual study, so that the results of meta-analysis were robust.

Publication Bias

The distribution of 13 studies involved in the meta-analysis of 'exclusive breastfeeding during the early postpartum period' showed in the funnel plot (Fig. 8 in Appendix) was considered asymmetric, however the Begg's test was not significant ($p = 0.502$), while some evidence of publication bias or small sample size effect was detected by the Egger's test ($p = 0.007$).

For the meta-analysis of 'breastfeeding at 4 months postpartum', the funnel plot (Fig. 9 in Appendix) appeared symmetric ($p = 0.049$ for Begg's test and $p = 0.059$ for Egger's test), suggesting little publication bias or small sample size effect was present.

Discussion

The present systematic review incorporated 46 studies to assess the association between caesarean delivery and breastfeeding practices in China. The meta-analysis comprised of 13 studies (441,044 subjects) on 'exclusive breastfeeding during the early postpartum period' and 14 studies (8771 subjects) on 'breastfeeding at 4 months postpartum' to estimate the pooled relationship between caesarean delivery and breastfeeding outcomes. Based on our findings, caesarean delivery was found to have an adverse effect on the breastfeeding outcomes. More specifically, compared with vaginal birth, the likelihood of mothers exclusively breastfeeding their babies during the early postpartum period was reduced by 47% (pooled OR 0.53) and 39% (pooled OR 0.61) reduction was found for the odds of breastfeeding at 4 months postpartum after caesarean section. Our finding of 'exclusive breastfeeding during the early postpartum period' (pooled OR 0.53; 95% CI 0.41, 0.68) appears to be consistent with a previous review (pooled OR 0.57, 95% CI 0.50, 0.64) (Prior et al. 2012). However, our finding of significant impact of caesarean section on 'breastfeeding at 4 months postpartum' (pooled OR 0.61, 95% CI 0.53, 0.71) is different from the non-significant result on any breastfeeding at 6 months postpartum in that same review (pooled OR 0.95, 95% CI 0.89, 1.01) (Prior et al. 2012). The current knowledge of surgery consequences suggests that postsurgical pain, haemorrhage, infections as well as some hormone issues (like prolactin level postpartum) affect breastfeeding, however these effects lessen in strength over time, indicating the impact of caesarean delivery on breastfeeding may attenuate over time (Chapman and Perez-Escamilla 1999; Liu et al. 2012; Marcus et al. 2015; Mkontwana and Novikova 2015; Prior et al. 2012; Wang et al. 2006). Previous studies showed that caesarean section remained an important barrier to early initiation of breastfeeding as well as to the implementation of hospital practices such as delayed skin-to-skin contact between mother-infant pairs (Bramson et al. 2010; Rowe-Murray and Fisher 2002). Given the negative association between caesarean section and breastfeeding outcomes revealed in the present analysis, it suggests that early initiation of breastfeeding may play a role of mediator variable between caesarean section and premature cessation of exclusive breastfeeding or any breastfeeding.

It is well known that China has a high caesarean section rate where nearly half of the babies born were delivered by caesarean section in 2010 (Hellerstein et al. 2015). The reasons for such a high rate mainly include three important factors: the obstetric care system in China (in hospital births, urbanisation, highly covered New Co-operative Medical Scheme that reduces patient costs and increases revenues for doctors and hospitals for caesarean than for vaginal deliveries, and high volume of deliveries), health care provider factors (insufficient nurses/midwives, some doctors do recommend caesarean section to avoid possible lawsuits in view of the medical malpractice environment), and cultural aspects of patient preference (demand for perfect baby, fear of pain, increasing numbers of macrosomia, increasing pregnancies in older women and delivery date choosing because of luck and belief) (Hellerstein et al. 2015; Long et al. 2012; Mi and Liu 2014). Although the association between caesarean section on maternal request and breastfeeding has been investigated in China (Liu et al. 2012), there is a gap in research on the different impacts of caesarean section on breastfeeding outcomes with respect to medical indication and maternal request (Liu et al. 2015; Tang et al. 2006).

The findings of this study suggest that practices or interventions after caesarean delivery are of benefits to improve breastfeeding behaviours. The practice of maternal-infant skin to skin contact after birth has been increasingly considered as an efficient way to promote

Appendix C

breastfeeding status postpartum especially breastfeeding initiation (Moore et al. 2016). The feasibility of an intervention of skin to skin contact after caesarean delivery in the operating room to improve breastfeeding as well as maternal satisfaction and pain perception outcomes was described and evaluated in previous studies (Hung and Berg 2011; Sundin and Mazac 2015). However, to our best knowledge, there is no intervention described to improve breastfeeding practice following caesarean section available in the literature. Future intervention studies to promote breastfeeding after caesarean delivery are warranted.

A major strength of this study was the extensive searches conducted in both Chinese and English literature to ensure all relevant articles have been included to reduce reporting bias. In addition, several studies without clarification of valid time-point for the breastfeeding outcomes had been excluded to enhance the quality of our evaluation. However, publication bias was still detected in the meta-analyses for early postpartum exclusive breastfeeding using the Egger's test (Higgins and Green 2011). Six months postpartum is the recommended period for exclusive breastfeeding by the WHO and it would be more comparable with other studies if the impact of caesarean section on 'breastfeeding at 6 months postpartum' could be examined in our study, however most of available reports in the present systematic review comprised the breastfeeding outcomes only for within 6 months postpartum. In view of the small number of studies addressing breastfeeding outcomes at 6 months postpartum, the corresponding quantitative synthesis was not performed. Another limitation concerns that the pooled weighted effect size was estimated based on both crude and adjusted ORs due to the limited number of eligible studies available. Ideally, extraction of data and analysis should be performed for adjusted ORs, as consequence the impact of caesarean section on breastfeeding practice could be assessed under the controlling for possible confounding effects. Since most of all studies retrieved were cross-sectional in this systematic review, a causal relationship between caesarean section and breastfeeding rates could not be concluded.

In conclusion, the present study confirmed that the likelihood of breastfeeding, including 'exclusive breastfeeding during the early postpartum period' and 'breastfeeding at 4 months postpartum', was significantly lower after caesarean section in China. Therefore, health policy and measures to improve breastfeeding outcomes should target the reduction of caesarean rate and health intervention after caesarean delivery in China.

Acknowledgements We acknowledge the authors who kindly provided additional data for this research. Jian Zhao would like to thank China Scholarship Council for financial support. Part of the work was presented as a poster in the 21st World Congress of Epidemiology, International Epidemiological Association (IEA) in Saitama, Japan on 19–22 August, 2017.

Appendix

Quality Assessment Checklist

1 Design of study	Score
Prospective	2
Retrospective	1
Cross-sectional	0
2 Loss to follow up /incompleteness of record	
<20%	2
≥20% or no information	0
3. Sample size	
≥100	2
<100	0
4. Description of period of recruitment	
Yes	2
No	0
5. Participants selection	
Representative of the general perinatal population	2
No	0
6. Exposure	
Define clearly	2
No	0
7. Outcomes	
Define clearly	2
No	0
8. Statistical analysis	
Proper statistical analysis with controlling confounders	2
Proper statistical analysis without controlling confounders	1
Otherwise	0
9. Result report	
Summarize key results with reference to study objectives	2
No	0

High: >14; Medium: 11-14; Low: <11

See Figs. 6, 7, 8, 9.

Appendix C

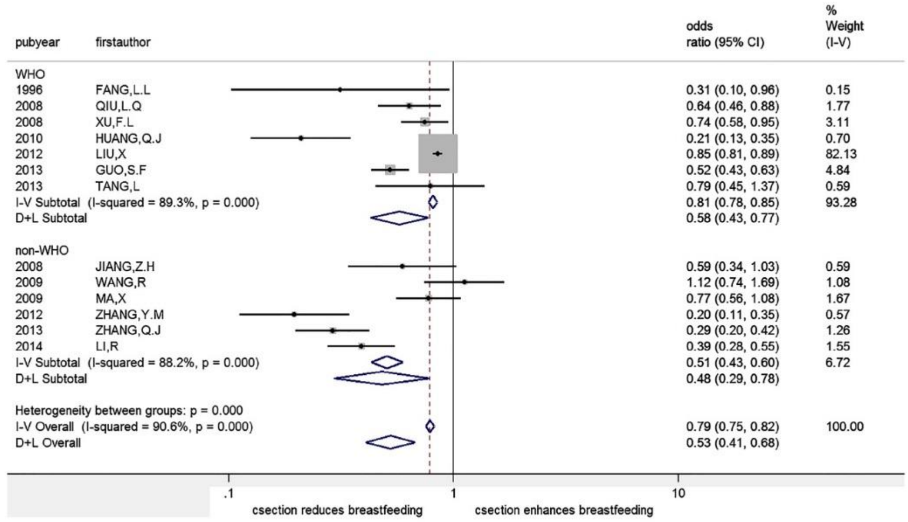


Fig. 6 Subgroup analysis for exclusive breastfeeding during the early postpartum period stratified by definitions

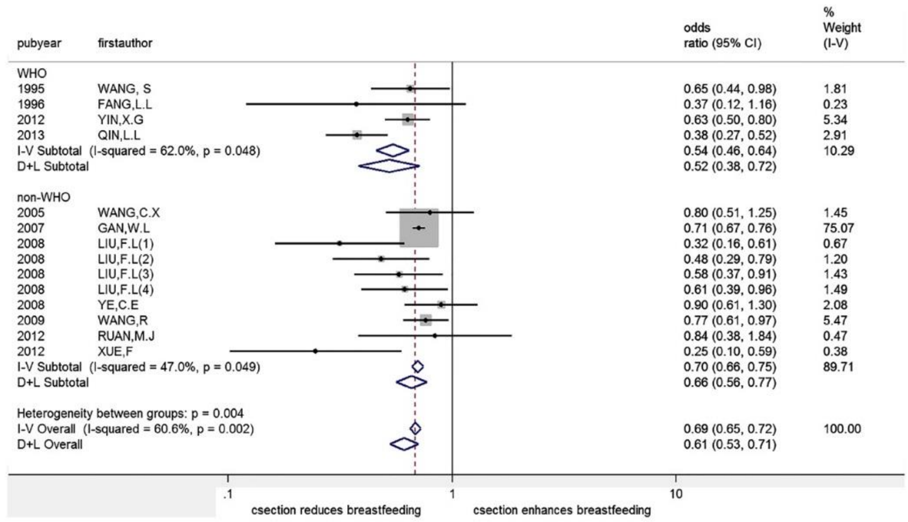


Fig. 7 Subgroup analysis for breastfeeding at 4 months postpartum stratified by definitions

Appendix C

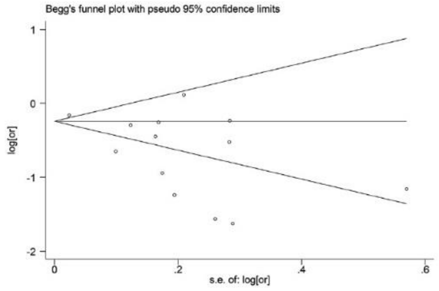


Fig. 8 Funnel plot exploring publication bias in exclusive breastfeeding during the early postpartum period

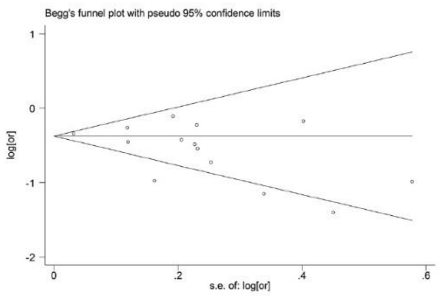


Fig. 9 Funnel plot exploring publication bias in breastfeeding at 4 months postpartum

References

Berry, N. J., & Gribble, K. D. (2008). Breast is no longer best: Promoting normal infant feeding. *Maternal & Child Nutrition, 4*(1), 74–79. doi:10.1111/j.1740-8709.2007.00100.x.

Bramson, L., Lee, J. W., Moore, E., Montgomery, S., Neish, C., Bahjri, K., et al. (2010). Effect of early skin-to-skin mother–infant contact during the first 3 h following birth on exclusive breastfeeding during the maternity hospital stay. *Journal of Human Lactation: Official Journal of International Lactation Consultant Association, 26*(2), 130–137. doi:10.1177/0890334409355779.

Chapman, D. J., & Perez-Escamilla, R. (1999). Identification of risk factors for delayed onset of lactation. *Journal of the American Dietetic Association, 99*(4), 450–454. doi:10.1016/S0002-8223(99)00109-1.

Chen, Y. F., Ouyang, Y. Q., & Wu, S. (2010). Current situation of breastfeeding of healthcare workers in Wuhan and its countermeasures. *Maternal and Child Health Care of China, 25*(27), 3921–3923. (**In Chinese**).

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials, 7*(3), 177–188.

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*(7109), 629–634.

Eidelman, A. I., Schanler, R. J., Johnston, M., Landers, S., Noble, L., Szucs, K., et al. (2012). Breastfeeding and the use of human milk. *Pediatrics, 129*(3), E827–E841. doi:10.1542/peds.2011-3552.

Fang, L. L., Zheng, X. X., & Guo, A. B. (1996). Survey on breastfeeding related factors. *Medical Journal of Chinese Civil Administration, 8*(2), 98–101. (**In Chinese**).

Feng, X. L., Xu, L., Guo, Y., & Ronsmans, C. (2012). Factors influencing rising caesarean section rates in China between 1988 and 2008. *Bulletin of the World Health Organization, 90*(1), 30–39A. doi:10.2471/BLT.11.090399.

Gan, W. L., Zhao, Y., & Wang, J. L. (2007). Breastfeeding associated factors analysis for infants 0–4 months postpartum in urban area of Chongqing. *Maternal and Child Health Care of China, 22*(20), 2805–2806. (**In Chinese**).

Guo, S., Fu, X., Scherpbier, R. W., Wang, Y., Zhou, H., Wang, X., et al. (2013). Breastfeeding rates in central and western China in 2010: Implications for child and population health. *Bulletin of the World Health Organization, 91*(5), 322–331. doi:10.2471/BLT.12.111310.

Harley, K., Stamm, N. L., & Eskenazi, B. (2007). The effect of time in the U.S. on the duration of breastfeeding in women of Mexican descent. *Maternal and Child Health Journal, 11*(2), 119–125. doi:10.1007/s10995-006-0152-5.

He, H. L., Wang, M., Zhong, Y. Y., Xu, X. X., & Xu, G. Y. (1994). Breastfeeding status 1 month postpartum and associated factors analysis. *Maternal and Child Health Care of China, 9*(4), 23–24. (**In Chinese**).

Hellerstein, S., Feldman, S., & Duan, T. (2015). China's 50% caesarean delivery rate: Is it too high? *BJOG, 122*(2), 160–164. doi:10.1111/1471-0528.12971.

Higgins, J. P., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions* (Vol. 4). Chichester: Wiley.

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539–1558. doi:10.1002/sim.1186.

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*(7414), 557–560. doi:10.1136/bmj.327.7414.557.

Huang, Q. J., & Lu, X. (2010). Investigation on breastfeeding status in baby-friendly hospital in Pudong New Area. *Maternal and Child Health Care of China, 25*(9), 1242–1244. (**In Chinese**).

Hung, K. J., & Berg, O. (2011). Early skin-to-skin after cesarean to improve breastfeeding. *MCN. The American Journal of Maternal Child Nursing, 36*(5), 318–324. doi:10.1097/NMC.0b013e3182266314.

Ip, S., Chung, M., Raman, G., Chew, P., Magula, N., DeVine, D., et al. (2007). Breastfeeding and maternal and infant health outcomes in developed countries. *Evidence Report Technology Assessment*, (153), 1–186.

Jiang, G. F. (2000). Associated factors analysis for breastfeeding. *Jiangsu Preventive Medicine, 11*(4), 49–50. (**In Chinese**).

Jiang, Z. H., & Li, S. (2008). Survey on factors influencing breastfeeding of neonates. *Maternal and Child Health Care of China, 23*(24), 3495–3496. (**In Chinese**).

Joshi, P. C., Angdembe, M. R., Das, S. K., Ahmed, S., Faruque, A. S., & Ahmed, T. (2014). Prevalence of exclusive breastfeeding and associated factors among mothers in rural Bangladesh: A cross-sectional study. *International Breastfeeding Journal, 9*, 7. doi:10.1186/1746-4358-9-7.

Kang, Y., Liang, X. H., Liu, Y. X., Chen, J., & Li, T. Y. (2013). Relative factors of breastfeeding cessation in infants before 6-month-old. *Journal of Third Military Medical University, 35*(16), 1735–1739. (**In Chinese**).

Appendix C

Kohlhuber, M., Rebhan, B., Schwegler, U., Koletzko, B., & Fromme, H. (2008). Breastfeeding rates and duration in Germany: A Bavarian cohort study. *The British Journal of Nutrition, 99*(5), 1127–1132. doi:10.1017/S0007114508864835.

Leng, X. L. (2014). Investigation of exclusive breastfeeding rate for infants under 6 months. *China Medical Engineering, 22*(7), 172–173. (**In Chinese**).

Li, R. (2014). Research of factors affecting breastfeeding in baby-area and countermeasures of nursing intervention. *Journal of Gannan Medical University, 34*(3), 453–455. (**In Chinese**).

Liu, F. L. (2008). Survey on status of breastfeeding at 4 months postpartum from 2000 to 2006 in the firtst hospital of people's Xinxiang City. *Chinese Journal of Misdiagnosis, 8*(19), 4777–4778. (**In Chinese**).

Liu, J., & Xing, E. Z. (1998). Associated factors analysis for breastfeeding. *Journal of Baotou Medical College*, 14(4), 29–30. (**In Chinese**).

Liu, L. F., Hu, J. Y., Jin, H. J., & Lei, J. (2014). Investigation and related factors analysis of breastfeeding of infants under 6 months in Lishui. *Chinese Journal of Child and Health Care, 22*(10), 1092–1095. (**In Chinese**).

Liu, X., Zhang, J., Liu, Y., Li, Y., & Li, Z. (2012). The association between cesarean delivery on maternal request and method of newborn feeding in China. *PLoS ONE, 7*(5), e37336. doi:10.1371/journal.pone.0037336.

Liu, X. H., Landon, M. B., Cheng, W. W., & Chen, Y. (2015). Cesarean delivery on maternal request in China: What are the risks and benefits? *American Journal of Obstetrics and Gynecology, 212*(6), 9. doi:10.1016/j.ajog.2015.01.043.

Liu, X. Q., & Shao, H. M. (2009). Investigation on status of postpartum breastfeeding and its influencing factors. *Occupation and Health, 25*(24), 2746–2748. (**In Chinese**).

Long, Q., Klemetti, R., Wang, Y., Tao, F., Yan, H., & Hemminki, E. (2012). High Caesarean section rate in rural China: Is it related to health insurance (New Co-operative Medical Scheme)? *Social Science & Medicine, 75*(4), 733–737. doi:10.1016/j.socscimed.2012.03.054.

Lumbiganon, P., Laopaiboon, M., Gulmezoglu, A. M., Souza, J. P., Taneepanichskul, S., Ruyan, P., et al. (2010). Method of delivery and pregnancy outcomes in Asia: The WHO global survey on maternal and perinatal health 2007-08. *Lancet, 375*(9713), 490–499. doi:10.1016/S0140-6736(09)61870-5.

Ma, X., Ma, L., Gan, L., Jia, H., & Wang, B. Z. (2009). Cohort study on the impact of different delivery methods on breastfeeding. *Chinese Journal of Child Health Care, 17*(4), 428–429. (**In Chinese**).

Marcus, H., Gerbershagen, H. J., Peelen, L. M., Aduckathil, S., Kappen, T. H., Kalkman, C. J., et al. (2015). Quality of pain treatment after caesarean section: Results of a multicentre cohort study. *European Journal of Pain, 19*(7), 929–939. doi:10.1002/ejp.619.

Merewood, A., Mehta, S. D., Chamberlain, L. B., Philipp, B. L., & Bauchner, H. (2005). Breastfeeding rates in US baby-friendly hospitals: Results of a national survey. *Pediatrics, 116*(3), 628–634. doi:10.1542/peds.2004-1636.

Mi, J., & Liu, F. (2014). Rate of caesarean section is alarming in China. *Lancet, 383*(9927), 1463–1464. doi:10.1016/S0140-6736(14)60716-9.

Miller, R. G. (1974). The jackknife-a review. *Biometrika, 61*(1), 1–15.

Mkontwana, N., & Novikova, N. (2015). Oral analgesia for relieving post-caesarean pain. *Cochrane Database of Systematic Reviews, 3*, CD010450. doi:10.1002/14651858.CD010450.pub2.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internatinal Medicine, 151*(4), 264–269.

Moore, E. R., Bergman, N., Anderson, G. C., & Medley, N. (2016). Early skin-to-skin contact for mothers and their healthy newborn infants. *Cochrane Database of Systematic Reviews, 11*, CD003519. doi:10.1002/14651858.CD003519.pub4.

Patel, A., Bucher, S., Pusdekar, Y., Esamai, F., Krebs, N. F., Goudar, S. S., et al. (2015). Rates and determinants of early initiation of breastfeeding and exclusive breast feeding at 42 days postnatal in six low and middle-income countries: A prospective cohort study. *Reproductive Health, 12* (Suppl 2), S10. doi:10.1186/1742-4755-12-S2-S10.

Pearce, A., Li, L., Abbas, J., Ferguson, B., Graham, H., Law, C., et al. (2012). Childcare use and inequalities in breastfeeding: Findings from the UK Millennium Cohort Study. *Archives of Disease in Childhood, 97*(1), 39–42. doi:10.1136/adc.2009.177337.

Prior, E., Santhakumaran, S., Gale, C., Philipps, L. H., Modi, N., & Hyde, M. J. (2012). Breastfeeding after cesarean delivery: A systematic review and meta-analysis of world literature. *The American Journal of Clinical Nutrition, 95*(5), 1113–1135. doi:10.3945/ajcn.111.030254.

Qin, L. L., & Hua, J. F. (2013). Investigation and analysis of infant feeding within four months among floating population in Suzhou. *Nursing Practice and Research, 10*(10), 138–140.

Qiu, L. Q. (2008). *Factors that influence breastfeeding initiation and duration in urban, suburban and rural areas of Zhejiang Province, Peoples Republic of China*. (Ph.D thesis). Curtin University, Perth.

Rowe-Murray, H. J., & Fisher, J. R. (2002). Baby friendly hospital practices: cesarean section is a persistent barrier to early initiation of breastfeeding. *Birth, 29*(2), 124–131.

Ruan, M. J., Ba, L., Wang, Z., & Song, H. (2012). Study on breastfeeding status and the related influencing factors in Chaoyang district of Beijing. *Maternal and Child Health Care of China, 27*(22), 3478–3480. (**In Chinese**).

Ryan, A. S., Wenjun, Z., & Acosta, A. (2002). Breastfeeding continues to increase into the new millennium. *Pediatrics, 110*(6), 1103–1109.

Scott, J. A., Aitkin, I., Binns, C. W., & Aroni, R. A. (1999). Factors associated with the duration of breastfeeding amongst women in Perth, Australia. *Acta Paediatrica, 88*(4), 416–421.

Senarath, U., Dibley, M. J., & Agho, K. E. (2010). Factors associated with nonexclusive breastfeeding in 5 east and southeast Asian countries: a multilevel analysis. *Journal of Human Lactation: Official Journal of International Lactation Consultant Association, 26*(3), 248–257. doi:10.1177/0890334409357562.

Stroup, D. F., Berlin, J. A., Morton, S. C., Olkin, I., Williamson, G. D., Rennie, D., et al. (2000). Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA, 283*(15), 2008–2012.

Sundin, C. S., & Mazac, L. B. (2015). Implementing Skin-to-Skin Care in the Operating Room After Cesarean Birth. *MCN. The American Journal of Maternal Child Nursing, 40*(4), 249–255. doi:10.1097/NMC.0000000000000142.

Tang, L. (2013). *A cohort study of infant feeding practices and maternal breastfeeding problems in a rural area of Sichuan Province, PR China*. (Ph.D thesis). Curtin University, Perth.

Tang, S., Li, X., & Wu, Z. (2006). Rising cesarean delivery rate in primiparous women in urban China: Evidence from three nationwide household health surveys. *American Journal of Obstetrics and Gynecology, 195*(6), 1527–1532. doi:10.1016/j.ajog.2006.03.044.

Tang, Z. J. (2014). Investigation on current situation and influencing factors of infant breastfeeding in Liwan District. Guangzhou City. *Chinese Journal of Women and Children Health, 5*(1), 73–76. (**In Chinese**).

Tian, J. Z., Tian, X. S., & Li, L. (2008). Breastfeeding influencing factors analysis in 183 mothers before discharge. *Zhejiang Journal of Preventive Medicine, 20*(7), 45–55. (**In Chinese**).

Appendix C

Tooth, L., Ware, R., Bain, C., Purdie, D. M., & Dobson, A. (2005). Quality of reporting of observational longitudinal research. *American Journal of Epidemiology, 161*(3), 280–288. doi:10.1093/aje/kwi042.

Victora, C. G., & Barros, A. J. D. (2000). Effect of breastfeeding on infant and child mortality due to infectious diseases in less developed countries: A pooled analysis. *Lancet, 355*(9202), 451–455. doi:10.1016/S0140-6736%2800%2982011-5.

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gotzsche, P. C., Vandenbroucke, J. P., et al. (2007). The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Medicine, 4*(10), e296. doi:10.1371/journal.pmed.0040296.

Wang, B. S., Zhou, L. F., Zhu, L. P., Gao, X. L., & Gao, E. S. (2006). Prospective observational study on the effects of caesarean section on breastfeeding. *Zhonghua Fu Chan Ke Za Zhi, 41*(4), 246–248.

Wang, C. X., Chen, H., & Liu, Y. Z. (2005). Breastfeeding failure factors analysis and countermeasures. *Journal of Community Medicine*, 3(9), 11–12. (**In Chinese**).

Wang, H. Z. (2010). Status analysis on breastfeeding in 1296 infants. *Practical Preventive Medicine, 17*(3), 519–520. (**In Chinese**).

Wang, R., Liu, J., & Liu, J. T. (2009). Effects of delivery mode on breastfeeding rate and child growth. *Maternal and Child Health Care of China, 24*(25), 3522–3523. (**In Chinese**).

Wang, S., Liang, Y. S., & Li, L. (1995). Survey on breastfeeding associated factors. *Chinese Journal of Family Planning* 20(6), 358–359. (**In Chinese**).

Wang, Z., Gao, J. K., Zhang, Y., Wang, F. H., Jin, J. M., & Zhu, B. H., et al. (2013). Present status and influencing factors analysis on breastfeeding for infants in Zhejiang Province. *Nursing and Rehabilitation Journal, 12*(7), 628–630. (**In Chinese**).

Wei, X. J., & Li, Q. (2009). Investigation on breastfeeding status of 4537 infants 42 days postpartum from 1996 to 2007. *Chinese Journal of Child Health Care, 17*(1), 106–108. (**In Chinese**).

World Health Organization. (2003). *Infant and Young Child Feeding: A Tool for Assessing National Practices, Policies and Programmes*. Geneva: World Health Organization

World Health Organization. (2008). Indicators for assessing infant and young child feeding practices: part 1: definitions: conclusions of a consensus meeting held 6–8 November 2007 in Washington DC, USA.

Xu, F. L. (2008). *The Determinants of Breastfeeding Initiation and Duration in Xinjiang, Peoples Republic of China*. (Ph.D thesis). Curtin University, Perth.

Xu, T., & Yu, X. S. (2009). Breastfeeding influencing factors analysis. *Chinese Journal of Health Statistics, 26*(4), 406–410. (**In Chinese**).

Xue, F., Gong, J. F., & Luo, Y. (2012). Study on breastfeeding influencing factors in Changshu. *Maternal and Child Health Care of China, 27*(10), 1536–1537. (**In Chinese**).

Yang, Y. L., & Feng, G. Z. (2014). Related factors of breastfeeding. *Journal of Mathematical Medicine, 27*(4), 427–428. (**In Chinese**).

Ye, C. E. (2008). Breastfeeding status analysis. *Maternal and Child Health Care of China, 23*(19), 2770. (**In Chinese**).

Yin, X. G., Gui, J. J., Wang, H. Q., Huang, X. M., Zhou, M. Y., & Zhu, P. (2012). Follow-up survey of risk factors for early cessation of breastfeeding in early infancy. *Chinese Journal of Neonatology, 27*(3), 148–152. (**In Chinese**).

Yu, C. (2013). *Breastfeeding and health outcomes in infants who receive continuing care from hospitals or community health centres in Chengdu Sichuan Province, People's Republic of China*. (Ph.D thesis). Curtin University, Perth.

Zhang, G. D., & Wang, J. L. (2000). Obstetric factors and breast-feeding. *Chongqing Medicine*, 29(2), 137–138. (**In Chinese**).

Zhang, Q. J., & Shi, Y. J. (2013). Breastfeeding influencing factors analysis for 612 women after delivery. *Maternal and Child Health Care of China, 28*(30), 5095–5096. (**In Chinese**).

Zhang, S. J., & Wu, L. X. (1999). Breastfeeding 1 month postpartum and associated factors analysis. *Journal of North China Coal Medical College*, 1(1), 61. (**In Chinese**).

Zhang, W. K., Hao, B., Chen, L. J., Zhao, G. L., Bao, Y. Q., & Zhou, M., et al. (2006). The effects of method of delivery on rate of breastfeeding. *Chinese Journal of Child Health Care, 14*(5), 461–463. (**In Chinese**).

Zhang, Y. M. (2012). Impact of different delivery modes on breastfeeding. *Seek Medical and Ask the Medicine 10*(10), 423. (**In Chinese**).

Zhang, Y. X., Dai, N., Zhang, A. H., Li, P. X., Yan, Y. H., & Li, N. (2013). Breastfeeding influencing factors analysis for 3057 women after delivery in Danyang City. *Journal of Qiqihar University of Medicine, 34*(18), 2736–2738. (**In Chinese**).

Zheng, K. Y., Zhang, X. H., Li, L., Shu, Y., & Qiu, L. Q. (2008). Factors influencing breastfeeding before discharge for delivery women in Hangzhou City. *Chinese Journal of Child Health Care, 16*(6), 716–718. (**In Chinese**).

Zhu, P., Hao, J., Jiang, X., Huang, K., & Tao, F. (2013). New insight into onset of lactation: mediating the negative effect of multiple perinatal biopsychosocial stress on breastfeeding duration. *Breastfeeding Medicine, 8*, 151–158. doi:10.1089/bfm.2012.0010.

Zhu, X., Tian, J., Chen, G., & Christensson, K. (2014). Predictors of breastfeeding exclusivity in three cities of China. *Breastfeeding Medicine, 9*(2), 103–104. doi:10.1089/bfm.2013.0053.

Appendix D for Chapter 3

D.1 Permission from publisher-Elsevier

Copyright Clearance Center RightsLink®

Home | Account Info | Help

Title: Maternal education and breastfeeding practices in China: A systematic review and meta-analysis

Author: Jian Zhao,Yun Zhao,Mengran Du,Colin W. Binns,Andy H. Lee

Publication: Midwifery

Publisher: Elsevier

Date: July 2017

Logged in as:
Jian Zhao
LOGOUT

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: https://www.elsevier.com/about/our-business/policies/copyright#Author-rights

BACK | CLOSE WINDOW

D.2 Publication 2

Appendix D

Midwifery 50 (2017) 62–71

Contents lists available at ScienceDirect

Midwifery

journal homepage: www.elsevier.com/locate/midw

Maternal education and breastfeeding practices in China: A systematic review and meta-analysis

Jian Zhao, MMed Ph.D. Candidate[a], Yun Zhao, Ph.D. Senior Lecturer[a,*], Mengran Du[b], Colin W. Binns, Ph.D. John Curtin Distinguished Emeritus Professor[a], Andy H. Lee, Ph.D. Professor[a]

[a] School of Public Health, Curtin University, Perth, Australia
[b] West China School of Public Health, Sichuan University, Chengdu, China

ARTICLE INFO

Keywords:
Breastfeeding
Maternal education
Systematic review
Midwives
China

ABSTRACT

Objective: to examine the association between maternal education and breastfeeding prevalence in China.
Methods: a systematic review and meta-analysis was conducted based on the literature of observational studies retrieved from electronic databases of CNKI, Medline, Embase, CINHAL, ProQuest and Science Direct. Maternal education was recoded into two binary categorical variables using different cut-off points. Both fixed and random effect models were used to estimate the pooled association between maternal education and breastfeeding prevalence in China. Visual inspection of Galbraith plot for heterogeneity detection, sensitivity analysis and publication bias test were performed.
Findings: a total of 31 studies were included in the systematic review, and 15 and 26 studies were suitable for meta-analysis in terms of two different cutoff points of maternal education respectively. In the group using 6-year education cut-off (Group 1), the odds of breastfeeding was 10% (pooled OR=0.90, 95% CI: 0.83, 0.97) lower in mothers who had been educated for 'more than 6 years' compared to mothers with '6 years or less' education.

In the group using 12-year education cut-off (Group 2), the odds of breastfeeding was 9% (pooled OR=0.91, 95% CI: 0.86, 0.96) lower in mothers who had 'more than 12 years' education compared to mothers who attained '12 years or less' education. There was substantial heterogeneity across the studies in both groups. Through meta-regression analysis, sample size of studies was detected contributing to the heterogeneity in Group 1; however none of study level factors were found to be a source of heterogeneity in Group 2.
Conclusion: in the Chinese culture and employment environment, mothers who have attained a higher level of education are less likely to breastfeed their babies compared to mothers with lower education levels.

Introduction

Breastfeeding is the optimal feeding method for infants as breast-milk provides all nutrients that infants need for healthy growth and development up to six months of age (Gillman et al., 2001; Schanler et al., 1999). After six months of age, the benefits continue after the introduction of complementary foods. The benefits of breastfeeding to both infant and maternal health, either short term or long term, are well documented (Bhandari et al., 2003; Howie et al., 1990; Ip et al., 2007). However, the rates of breastfeeding still remain below optimal levels in many countries (Amir and Donath, 2008; Chalmers et al., 2009; Ryan and Zhou, 2006). According to the most recent community-based, cross-sectional survey of 12 central and western provinces in China, overall any breastfeeding rate was noted as 98.3%, however, only 28.7% of infants younger than 6 months had been exclusively breastfed (Guo et al., 2013).

Because of the public health benefits of breastfeeding, many countries have initiated health promotion interventions to support and promote breastfeeding practices, especially exclusive breastfeeding, based on the factors affecting breastfeeding outcomes (Bernaix et al., 2010; Howell et al., 2014; Kramer et al., 2001; Oken et al., 2013; Pisacane et al., 2012). Among the factors affecting breastfeeding practices, maternal education has been investigated widely but the findings and conclusions are inconsistent. The effect of maternal education has often been different in studies carried out in China compared to western countries (Heck et al., 2006; Kehler et al., 2009; Ludvigsson and Ludvigsson, 2005; Scott and Binns, 1999). In the U.S., Australia and European countries, studies report that maternal education is positively associated with higher breastfeeding prevalence (Heck et al., 2006; Hornell et al., 1999; Scott and Binns, 1999), and a longer

* Corresponding author.
E-mail address: y.zhao@curtin.edu.au (Y. Zhao).

http://dx.doi.org/10.1016/j.midw.2017.03.011
Received 6 April 2016; Received in revised form 20 September 2016; Accepted 22 March 2017

Appendix D

duration compared to low-educated mothers. While some studies in China agreed with these results (Jiang, 2000; Ma et al., 2014) others have reached the opposite conclusions (Huang et al., 2012; Qiu, 2008; Ye et al., 2007).There has been no published systematic review of the impact of maternal education on breastfeeding prevalence in China. Because of the inconsistent results a systematic review and meta-analysis of breastfeeding studies carried out in China has been undertaken to explore the association between maternal education and breastfeeding prevalence.

The findings of this study have the potential to guide the provision of focused and evidence-based interventions aimed to promote breastfeeding in China.

Methods

Search strategy

Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009), a systematic electronic search was conducted to retrieve literature published in Chinese or English using the Chinese database China National Knowledge Infrastructure (CNKI) and English databases Medline, Embase, CINHAL, ProQuest and Science Direct from January 1990 to June 2015. The search strategy adopted included the following two stages:

Stage 1: The following Medical Subject Headings (MeSH) terms and key words "breast feeding", " human milk", "breastfeeding duration", "breastfeeding cessation", "human lactation", "infant feed*", "breastfed", "risk factor*", "protective factor*", "determinant*", "socioeconomic factor*", "China", "Mainland China", "Chinese" were used.

Stage 2: The following MeSH terms and key words "maternal education", "education" and "education status" were further added to the search process.

Literature screening and selection criteria

The abstracts of the identified records were screened by two independent reviewers (JZ and MRD) to check whether they were appropriate to be included in the present study. Full texts were then extracted and evaluated after the appropriateness was confirmed by the abstract screening. The inclusion criteria for the studies were: (i) published in peer-reviewed journals or theses/dissertations; (ii) observational study design; (iii) reported and evaluated the relationship between maternal education and breastfeeding prevalence quantitatively; (iv) effect size with 95% confidence intervals (CIs) could be obtained directly or calculated from the raw tabulated data. The exclusion criteria were: (i) sample size not specified; (ii) time points of breastfeeding outcomes not reported or defined; (iii) inappropriate statistical result reported (statistical error); (iv) unspecified category of maternal education. Discrepancies were discussed between JZ and MRD until a consensus could be reached or referred to a third investigator (YZ) when necessary.

Data extraction

Information, including published year, first author name, study design, location of study, sample size, breastfeeding outcomes measured, maternal education status, other factors associated with breastfeeding outcomes, results of tabulated data or effect size (odds ratio) with their corresponding 95% CIs reported by univariate analysis (crude ORs) or multivariable analysis with adjustment for confounders (adjusted ORs), was retrieved from each eligible study for qualitative and quantitative synthesis. In cases in which essential information was not provided in articles, the authors were contacted by email.

We followed the World Health Organization (WHO) definitions in the data extraction (World Health Organization, 2003, 2008):

Exclusive breastfeeding: Breastfeeding while giving no other food or liquid, not even water, with the exception of drops or syrups consisting of vitamins, mineral supplements or medicines.

Full breastfeeding: Exclusive breastfeeding or predominant breastfeeding (or almost exclusive breastfeeding). Breastmilk is the only source of milk given to the infant regardless of supplementation with other fluids such as water and orange juice.

Any breastfeeding: The child has received breastmilk (direct from the breast or expressed) with or without other drink, formula or other infant food.

Quality assessment

Each selected study had been assessed its methodological quality using a formal checklist 'the Standard Quality Assessment Criteria for Evaluating Primary Research Papers from a Variety of Fields' (Kmet et al., 2004). Two independent researchers performed the scoring process according to the each criterion of the checklist. Each item on studies was given a score (0-No, 1-Partial, 2-Yes) based on the degree to which the criteria were met. The summary score was calculated by summing the total score obtained across relevant items then divided by the total possible score. Therefore, study quality scores ranged from 0 to 1, where a higher score corresponds to higher quality. Any discrepancies between the assessors were resolved by consensus.

Assessment of risk of bias

Based on the Cochrane Collaboration's tool for assessing the risk of bias and GRADE guidelines of assessing study limitations (risk of bias) in observational studies (Guyatt et al., 2011; Higgins et al., 2011), we summarized key criteria to assess the risk of bias across studies included for five domains: incomplete outcome data (attrition bias), selective reporting (reporting bias), lack of adjustment for baseline characteristics, failure to develop and apply appropriate eligibility criteria and flawed measurement of both exposure and outcome. For each individual domain, we classified studies into low, unknown, and high risk of bias.

Statistical analysis

In this study, maternal education (in years) refers to the formal education mothers receive in primary schools, high schools or tertiary institutes and was recoded into two binary categorical variables. All eligible studies were classified into two groups based on different cut-off values used in the recoding of maternal education. In Group 1, maternal education was recoded into two categories 'more than 6years (> 6 years) education' versus '6 years or less (< =6 years) education', while in Group 2 the coding criteria was 'more than 12 years (> 12 years) education' versus '12 years or less (< =12 years) education', respectively. The ORs of breastfeeding (including exclusive breastfeeding, full breastfeeding and any breastfeeding within 12 months postpartum) comparing higher educational status to lower educational status were the primary statistical measures and were transformed into the logarithmic scale for meta-analysis. Considering the effect of confounding, crude ORs and adjusted ORs were pooled separately in either group.

A fixed-effect meta-analysis by the inverse variance (I-V) method was conducted to pool the OR between maternal education and breastfeeding prevalence (Higgins and Green, 2011). Visual inspection of a Galbraith plot was conducted to detect the heterogeneity due to individual studies (Bax et al., 2009). The I^2 statistic was used to assess the heterogeneity across the studies (Higgins and Thompson, 2002; Higgins et al., 2003). A random-effect meta-analysis by DerSimonian and Laird (D+L) method was applied when significant heterogeneity

Appendix D

was present (DerSimonian and Laird, 1986). A meta-regression was conducted to investigate the potential sources of heterogeneity. Subgroup analysis was carried out if some factors were detected being significant in meta-regression. Considering the difference between fixed time point prevalence of breastfeeding for cohort study and period average prevalence of breastfeeding for cross-sectional study, subgroup analysis based on the design of the study (cohort study or cross-sectional study) was performed for both Group 1 and Group 2. To test the dependence of effect size reported in each study, a sensitivity analysis using the Jack-knife method was performed to assess the robustness of the results (Miller, 1974). Begg's test and Egger's test were conducted to investigate the publication bias or small sample size bias across studies (Egger et al., 1997). All statistical analyses were performed using the Stata package version 14.1 (StataCorp LP, College Station, USA). A p value less than 0.05 was considered statistically significant.

Findings

Systematic review

As shown in Fig. 1, a total of 181 studies were identified from English and Chinese databases and 4 additional records were obtained via ProQuest Dissertations and Theses. After removal of duplicate articles, 54 articles were retrieved to assess eligibility and ultimately, 31 studies were included in the systematic review (Guo et al., 2013; Huang et al., 2012; Jiang, 2000; Jiang et al., 2013; Jiang and Li, 2008; Kang et al., 2013; Ke, 1993; Li et al., 2003; Lin et al., 1997; Liu et al., 2013; Liu and Zhou, 2013; Ma et al., 2014; Qin and Hua, 2013; Qiu, 2008; Ruan et al., 2012; Tang, 2013, 2014; Tian et al., 2008; Wang et al., 2005; Wang, 2010; Wang et al., 2013; Wu and Qiu, 2015; Xiao, 2001; Xiong et al., 2006; Xu, 2008; Yang and Feng, 2014; Ye, 2008; Ye et al., 2007; Yu, 2013; Yu and Song, 2000; Zhang et al., 1998), see Table 1. The 31 studies included 23 cross-sectional studies (Guo et al., 2013; Huang et al., 2012; Jiang, 2000; Jiang et al., 2013; Kang et al., 2013; Ke, 1993; Li et al., 2003; Lin et al., 1997; Liu et al., 2013; Liu and Zhou, 2013; Ruan et al., 2012; Tang, 2014; Tian et al., 2008; Wang, 2010; Wang et al., 2013; Wu and Qiu, 2015; Xiao, 2001; Xiong et al., 2006; Yang and Feng, 2014; Ye, 2008; Ye et al., 2007; Yu and Song, 2000; Zhang et al., 1998), 4 retrospective studies (Jiang and Li, 2008; Ma et al., 2014; Qin and Hua, 2013; Wang et al., 2005) and 4 prospective studies (Qiu, 2008; Tang, 2013; Xu, 2008; Yu, 2013). The time points of breastfeeding outcomes measured included 'before hospital discharge', 15 days, 28 days, 1 month, 42 days, 2 months, 3 months, 4 months and 6 months postpartum. The quality score of each study included was assessed and reported in Table 1 with a mean of 0.80 and standard deviation of 0.15, ranging from 0.50 to 1.00. Assessment of risk of bias across studies is shown in Table 2. Incomplete outcome data (attrition bias) were rated as "high risk of bias" in only 6 (19%) studies, selective reporting (reporting bias) was rated as "high risk of bias" in only 4 (13%) studies, while "high risk of bias" was identified in 18 (58%) studies in terms of lack of adjustment for baseline characteristics. In other two domains of risk of bias (failure to develop and apply appropriate eligibility criteria and flawed measurement of both exposure and outcome), studies were rated as either "low risk of bias" or "unknown risk of bias".

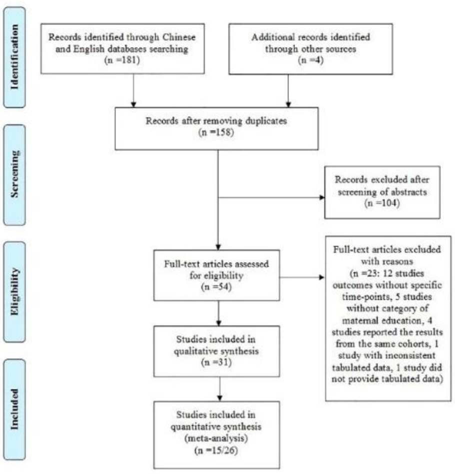


Fig. 1. Flow chart of systematic review following PRISMA.

Fifteen studies were included in Group 1 for meta-analysis (Jiang, 2000; Jiang et al., 2013; Kang et al., 2013; Ke, 1993; Liu et al., 2013; Liu and Zhou, 2013; Ma et al., 2014; Qiu, 2008; Ruan et al., 2012; Wang et al., 2005; Wang, 2010; Wu and Qiu, 2015; Xu, 2008; Ye, 2008; Yu, 2013) and 26 studies were involved in Group 2 for meta-analysis (Jiang and Li, 2008; Kang et al., 2013; Li et al., 2003; Lin et al., 1997; Liu et al., 2013; Liu and Zhou, 2013; Ma et al., 2014; Qin and Hua, 2013; Qiu, 2008; Ruan et al., 2012; Tang, 2013; Tang, 2014; Tian et al., 2008; Wang et al., 2005; Wang et al., 2013; Wu and Qiu, 2015; Xiao, 2001; Xiong et al., 2006; Xu, 2008; Yang and Feng, 2014; Ye, 2008; Yu, 2013; Yu and Song, 2000; Zhang et al., 1998), among which 2 individual studies were classified into 4 sub-studies due to the various breastfeeding outcomes measured (Tang, 2013; Yu, 2013). Note that there were 11 studies commonly shared by both groups as they provided results suitable for both cut-off values (Kang et al., 2013; Liu et al., 2013; Liu and Zhou, 2013; Ma et al., 2014; Qiu, 2008; Ruan et al., 2012; Wang et al., 2005; Wu and Qiu, 2015; Xu, 2008; Ye, 2008; Yu, 2013) and 3 studies excluded in either group as they were not suitable to be quantitatively synthesized using either cut-off values (Guo et al., 2013; Huang et al., 2012; Ye et al., 2007). Four studies reported appropriate adjusted odds ratios from multiple logistic regressions which controlled for potential confounders (Guo et al., 2013; Huang et al., 2012; Li et al., 2003; Ye et al., 2007), but only one study could be used in the quantitative synthesis based on the preselected cut-off value (Li et al., 2003). The remaining studies reported crude (unadjusted) ORs or provided tabulated data to calculate the crude ORs for the meta-analysis.

Meta-analysis for Group 1: '> 6 years' versus '< =6 years'

In Group1, where the maternal education was categorized using the cut-off of '6-year education', 15 studies (13395 subjects) were included in the meta-analysis of the association between maternal education and breastfeeding. The random-effect meta-analysis indicated that the odds of breastfeeding were 10% (pooled OR=0.90, 95% CI: 0.83, 0.97) lower in mothers who had been educated for 'more than 6 years' compared to mothers who had '6 years or less' education. There was substantial heterogeneity across the studies (I^2 =92.5%, $p < 0.001$) as shown in Fig. 2. A visual inspection of Galbraith plot (omitted for brevity) was conducted and suggested several individual studies (Jiang, 2000; Jiang et al., 2013; Liu and Zhou, 2013; Ma et al., 2014; Qiu, 2008; Wu and Qiu, 2015; Xu, 2008; Ye, 2008; Yu, 2013) were the potential sources of the heterogeneity.

A meta-regression analysis including the study level covariates such as published year, sample size, study design, definitions of breastfeeding, and quality assessment scores of studies was performed to investigate the sources of heterogeneity. None of the above covariates

Appendix D

Table 1
Characteristics of studies (n=31) assessing the association between maternal education and breastfeeding prevalence in China.

Published year	First author	Study design	Location	Sample size	Other factors associated with breastfeeding outcomes	Breastfeeding outcome measured	Definition	Score
1993	KE,F.F	Cross-sectional	Quanzhou	459	Rooming in, breastmilk initiation, income, nutrition during lactation, gender, disease, productive plan, menstruation resumed	EBF at 4mo postpartum	WHO	0.73
1997	LIN,L	Cross-sectional	Guangdong	418	Fatigue after delivery, rooming in, lack of breastfeeding knowledge and skills, income per capita, maternity leave	EBF at 4mo postpartum	Non-WHO	0.64
1998	ZHANG,H.L	Cross-sectional	Wuhan	344	Maternal occupation, rooming in, breastfeeding on demand	FB at 4mo postpartum	WHO	0.68
2000	YU,H.M	Cross-sectional	Karamay	368	Maternity leave, ethnicity	FB within 4mo postpartum	WHO	0.82
2000	JIANG,G.F	Cross-sectional	Liuhe	736	Rooming in, delivery method	EBF within 6mo postpartum	WHO	0.77
2001	XIAO,D.Q	Cross-sectional	Beijing	286	Delivery method	EBF within 28 days postpartum	WHO	0.68
2003	LI,L.B	Cross-sectional	Beijing	251	Maternal antenatal feeding plan	Breastfeeding during 6-12mo postpartum	WHO	0.95
2005	WANG,C.X	Retrospective cohort	Jinan	622	Delivery method, health education, milk powder promotion, initiation time	EBF at 4mo postpartum	Non-WHO	0.77
2006	XIONG,H.J	Cross-sectional	Beijing	146	NA	EBF at 1,2,3,4mo postpartum	WHO	0.59
2007	YE,J.L	Cross-sectional	Xinjiang	2076	Ethnicity, maternal age, boiled water drinking	EBF at 4mo postpartum	WHO	0.95
2008	XU,F.L	Prospective cohort	Xinjiang	1136	Giving breastmilk as the first feed, feeding on demand, mother feeling given enough information about breastfeeding, minority ethnic group, giving birth in spring or summer, medical staff not recommending formula to parents, prelacteal feeds of water or formula	AF initiation/ EBF initiation	WHO	1.00
2008	TIAN,J.Z	Cross-sectional	Zhejiang	253	Neonate disease, early sucking, initiation time, breastfeeding confidence, delivery method	FB before discharge	WHO	0.73
2008	YE,C.E	Cross-sectional	Ninghai	975	Delivery method, health education	EBF at 4mo postpartum	Non-WHO	0.68
2008	JIANG,Z.H	Retrospective cohort	Harbin	310	Initiation time, delivery method	EBF at 42d postpartum	Non-WHO	0.59
2008	QIU,L.Q	Prospective cohort	Zhejiang	1511	Living in the suburb and rural areas, maternal age, mother decides to breastfeed until after birth, prelacteal feeding	EBF initiation/ at discharge	WHO	1.00
2010	WANG,H.Z	Cross-sectional	Changli	1296	Health education, delivery method	EBF at 6mo postpartum	WHO	0.64
2012	RUAN,M.J	Cross-sectional	Beijing	101	Delivery method, fixed term job, maternal age	EBF at 4mo postpartum	Non-WHO	0.73
2012	HUANG,H.T	Cross-sectional	5 provinces	2522	Infant gender, ethnicity, region, premature birth, low birth mass, pregnancy disease, maternal age, history of abortion, history of induced labour, parity	EBF within 4mo postpartum	Non-WHO	0.95
2013	JIANG,Y	Cross-sectional	4 provinces	1272	Health education, ethnicity, low birth weight	EBF within 6mo postpartum	WHO	0.95
2013	LIU,L	Cross-sectional	Shanghai	210	Household income monthly, maternal occupation	EBF within 6mo postpartum	WHO	0.73
2013	KANG,Y	Cross-sectional	Chongqing	938	Infant gender, delivery method, income monthly, birth weight, duration of maternal leave, perception of breastmilk amount, prelacteal feeding	Breastfeeding at 6mo postpartum	Non-WHO	0.91
2013	WANG,Z	Cross-sectional	Zhejiang	528	Infant age, infant gender, early feeding, perception of breastmilk amount, delivery method	EBF within 6mo postpartum	WHO	0.86
2013	GUO,S.F	Cross-sectional	26 counties	2293	Maternal antenatal clinic visit, child's age, delivery method	EBF within 6mo postpartum	WHO	0.91
2013	LIU,J.H	Cross-sectional	Jintan	1280	Delivery method, gender, grandmother's education, maternal occupation, main caregiver, birth order	EBF within 6mo postpartum	Non-WHO	0.91
2013	QIN,L.L	Retrospective cohort	Suzhou	612	Maternal occupation, maternal age, birth region, breastfeeding professional instruction, delivery method	EBF at 4mo postpartum	WHO	0.91
2013	TANG,L	Prospective cohort	Jiangyou	695	Encouragement from facility staff, paternal feeding preference, time of deciding feeding method, maternal grandmother's feeding preference	AF at discharge/ FB at discharge	WHO	1.00
2013	YU,C	Prospective cohort	Chengdu	845	Maternal occupation, paternal education, intention of going back to work, first feeding, mothers' friends breastfeed their babies, paternal job, staff encouragement, father's attitude, maternal grandmother's breastfeeding history, delivery method	AF within 15 d postpartum/ FB within 15d postpartum	WHO	1.00
2014	TANG,Z.J	Cross-sectional	Guangzhou	315	Delivery method, household income per capita	BF at 6mo postpartum	WHO	0.59
2014	MA,Y.Q	Retrospective cohort	Xining & Xiaan	502	NA	EBF at 4mo postpartum	Non-WHO	0.86
2014	YANG,Y.L	Cross-sectional	Wuhan	513	Health education, delivery method, prenatal high risk factors	EBF within 6mo postpartum	Non-WHO	0.50
2015	WU,Y	Cross-sectional	Yongkang	667	Maternal age, maternal occupation, household income monthly	FB at 3mo postpartum	WHO	0.91

FB: full breastfeeding; EBF: exclusive breastfeeding; AF: any breastfeeding.
Score: quality assessment score according to the formal checklist developed by Kmet et al.

Appendix D

Table 2
Assessment of risk of bias across studies (n=31) included in the systematic review.

Published year	Study	Incomplete outcome data(Attrition bias)	Selective reporting (Reporting bias)	Lack of adjustment for baseline characteristics	Failure to develop and apply appropriate eligibility criteria	Flawed measurement of both exposure and outcome
1993	KE,F.F	+	+	–	?	+
1997	LIN,L	+	–	–	+	+
1998	ZHANG,H.L	+	+	–	?	+
2000	YU,H.M	+	–	–	+	+
2000	JIANG,G.F	+	+	–	+	+
2001	XIAO,D.Q	+	+	–	?	?
2003	LI,L.B	+	+	+	+	+
2005	WANG,C.X	+	+	–	+	+
2006	XIONG,H.J	+	–	–	+	+
2007	YE,J.L	?	+	+	+	+
2008	XU,F.L	+	+	+	+	+
2008	TIAN,J.Z	+	+	–	?	?
2008	YE,C.E	–	+	–	?	?
2008	JIANG,Z.H	+	?	–	?	?
2008	QIU,L.Q	+	+	+	+	+
2010	WANG,H.Z	+	+	–	?	+
2012	RUAN,M.J	–	+	–	+	+
2012	HUANG,H.T	+	+	+	+	+
2013	JIANG,Y	+	+	+	+	+
2013	LIU,L	–	+	–	?	?
2013	KANG,Y	+	+	+	+	+
2013	WANG,Z	–	+	+	?	?
2013	GUO,S.F	?	+	+	+	+
2013	LIU,J.H	–	?	+	+	+
2013	QIN,L.L	–	?	–	?	?
2013	TANG,L	+	+	+	+	+
2013	YU,C	+	+	+	+	+
2014	TANG,Z.J	+	?	–	?	?
2014	MA,Y.Q	+	–	–	+	+
2014	YANG,Y.L	+	?	–	?	?
2015	WU,Y	+	?	+	+	+

+Low risk of bias.
-High risk of bias.
?Unknown risk of bias.

were found contributing to the heterogeneity between studies.

Subgroup meta-analysis

Stratified by sample size. Studies in Group 1 were divided into two subgroups: one subgroup of studies with a sample size smaller than the mean of sample sizes and the other subgroup of studies with a sample size larger than the mean of sample sizes.

A subgroup meta-analysis stratified by sample size was performed as shown in Fig. 3. In the subgroup of studies with a sample size smaller than the average sample size, the pooled OR is 0.77 (95% CI:

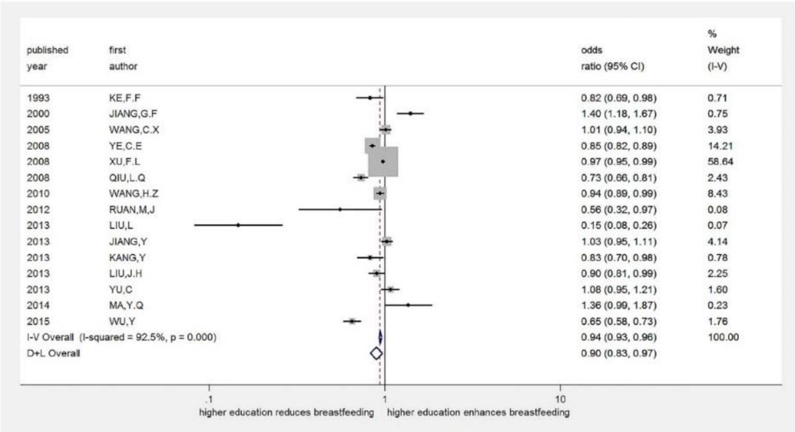


Fig. 2. Forest plot of the association in Group 1.

66

Appendix D

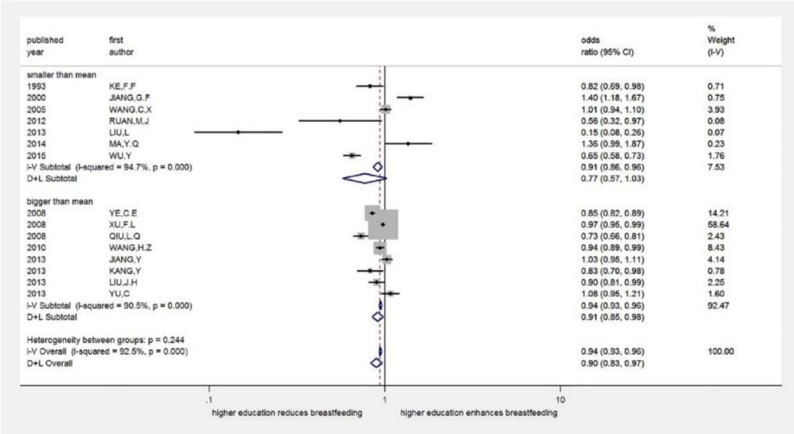


Fig. 3. Subgroup analysis stratified by sample size in Group 1.

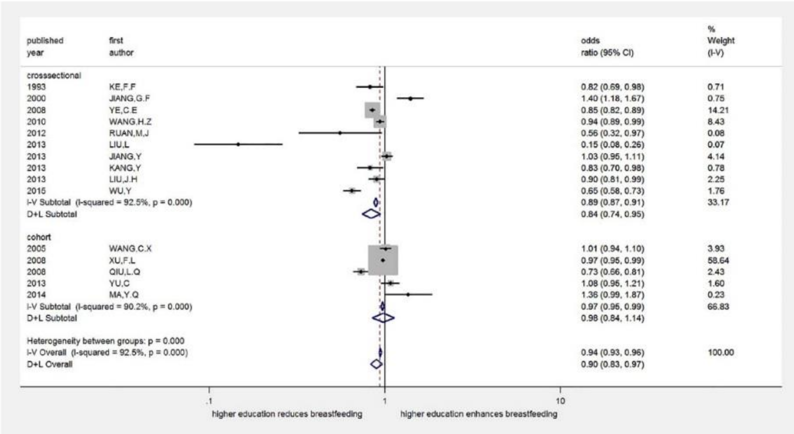


Fig. 4. Subgroup analysis stratified by study design in Group 1.

0.57, 1.03), suggesting a nonsignificant association between maternal education and breastfeeding prevalence, while in the other subgroup with a sample size larger than the average sample size, the pooled OR is 0.91 (95% CI: 0.85, 0.98), revealing a negative impact of higher maternal education on breastfeeding prevalence.

Stratified by study design. Studies in Group 1 were classified into two subgroups: cross-sectional studies and cohort studies as shown in Fig. 4. In the subgroup of cross-sectional studies, the pooled OR is 0.84 (95% CI: 0.74, 0.94) showing a significant negative impact of maternal education on breastfeeding, while in the cohort studies subgroup, the pooled OR is 0.97 (95% CI: 0.86, 1.10) without significant impact.

Meta-analysis for Group 2: '> 12 years' versus '< =12 years'

A total of 26 studies (15366 subjects) were included in Group 2 for the meta-analysis to investigate the pooled effect of maternal education on breastfeeding when the maternal education was categorized using the cut-off of '12-year education'. As Fig. 5 shows, a random-effect meta-analysis revealed that compared to the mothers who had '12 years or less' education, the odds of breastfeeding in mothers who attained 'more than 12 years' education was significantly lower by 9% (pooled OR=0.91, 95% CI: 0.86, 0.96). An evident heterogeneity was present (I^2 =85.2%, $p < 0.001$) and the corresponding Galbraith plot (omitted for brevity) identified that the individual studies (Jiang and Li, 2008; Lin et al., 1997; Liu and Zhou, 2013; Qin and Hua, 2013; Qiu, 2008; Tang, 2013, 2014; Wu and Qiu, 2015; Xiong et al., 2006; Yang

Appendix D

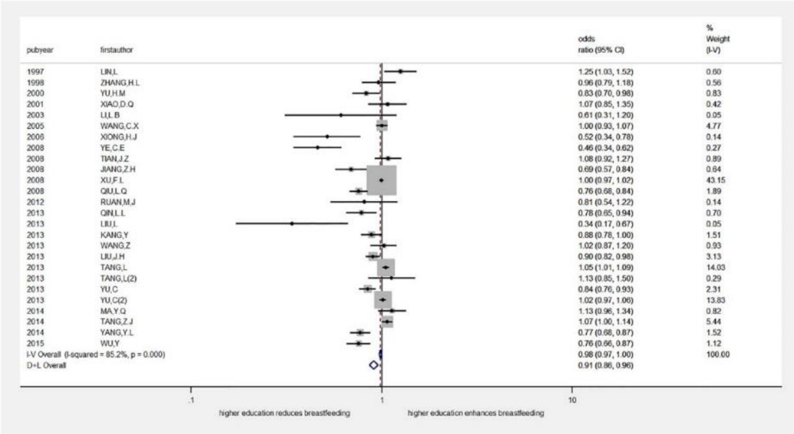


Fig. 5. Forest plot of the association in Group 2.

and Feng, 2014; Ye, 2008; Yu, 2013) were the potential source of the heterogeneity.

A meta-regression analysis was performed to investigate whether study level factors (published year, sample size, study design, definitions of breastfeeding, and quality assessment scores of studies) contribute to the heterogeneity; however, none of them were found to be a source of heterogeneity.

Subgroup analysis stratified by study design

Subgroup analysis stratified by study design (cross-sectional study and cohort study) was performed as shown in Fig. 6. The pooled OR in the subgroup of cross-sectional studies is 0.82 (95% CI: 0.73, 0.93) indicating a significant negative association between maternal education and breastfeeding, however, in the subgroup of cohort studies the pooled OR is 0.97 (95% CI: 0.92, 1.02), suggesting nonsignificant impact.

Sensitivity analysis

The sensitivity analysis omitting each study in turn indicated that the pooled effect size remained significant during each recalculation of the pooled ORs, which suggested that the results of the meta-analysis were robust.

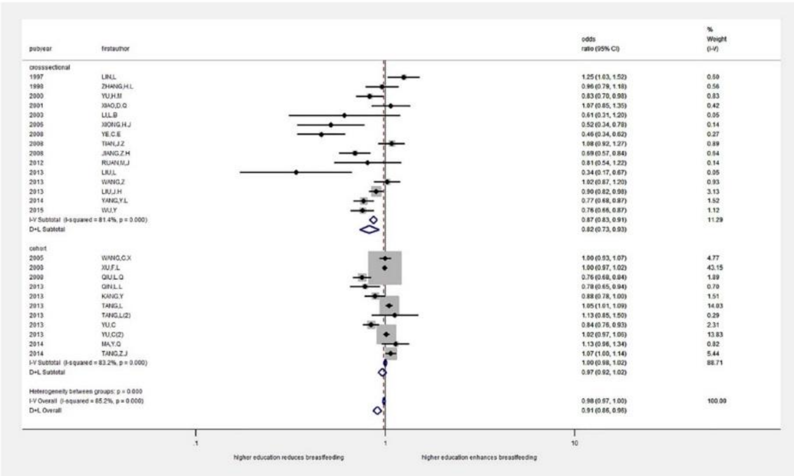


Fig. 6. Subgroup analysis stratified by study design in Group 2.

68

Appendix D

Publication bias

For the meta-analysis of 15 studies in Group1, the funnel plot showed symmetric, consistent with the Begg's test (p=0.373) and Egger's test (p=0.236), suggesting that no publication bias or small study effect was present.

The funnel plot of 26 studies in the meta-analysis of Group 2 appeared asymmetric. Although the Begg's test showed no significant publication bias or small study effect (p=0.146), the Egger's test suggested significant publication bias or small study effect (p=0.005).

Discussion

To our knowledge, this is the first systematic review and meta-analysis examining the impact of maternal education level on breastfeeding prevalence in China. The results of the present study showed that the odds of breastfeeding of mothers who had 'more than six years' or 'more than twelve years' education was reduced by 10% or 9% respectively, compared to mothers obtaining 'six years or less' or 'twelve years or less' education, respectively.

These results differ from published results of studies carried out in other countries, where a positive association between maternal education and breastfeeding prevalence (Al-Sahab et al., 2010; Barria et al., 2008; Baxter et al., 2009; Esteves et al., 2014) has frequently been reported, indicating higher maternal education is associated with higher odds of breastfeeding. By contrast the majority of studies carried out in China reported a negative or non-significant association between maternal education and breastfeeding prevalence regardless the association was assessed independently (Jiang et al., 2013; Wu and Qiu, 2015; Ye, 2008) with no adjustment of confounders or with an adjustment for potential confounders (Guo et al., 2013; Huang et al., 2012; Tang et al., 2013; Wang et al., 2013; Xu et al., 2007; Ye et al., 2007). Esteves et al. (2014) discussed the findings of one Brazilian study which also reported that higher schooling was a risk factor for delayed breastfeeding initiation within the first hour after birth in Brazil (Silveira et al., 2008). They speculated that this 'negative' association was caused by not adjusting for the potential impact of other variables on breastfeeding. However, based on our present systematic review, several studies from China concluded that the negative association between maternal education and breastfeeding remained even after adjusting potential confounders (Huang et al., 2012; Ye et al., 2007), similar to the findings of the Brazilian study. Maternal occupation and maternity leave are two commonly recognized important factors, which may potentially confound or mediate the effect of the maternal education on breastfeeding prevalence (Baxter, 2008; Jiang et al., 2012; Tan, 2011). Mothers who have higher education normally have better paid but heavy workload full time job in China. As these mothers are paid well and then have the financial resource to afford purchasing foreign imported infant formula, which is much demanding and expensive after the melamine contamination affair in 2008, for their babies rather feeding breastmilk (Gong and Jackson, 2012; Guan et al., 2009; Parry, 2008). Furthermore, high educated mothers are usually working full time in a responsible position and tend to have shorter breastfeeding duration (Kang et al., 2013; Wu and Qiu, 2015). In China, if workplaces do not provide a supportive environment, breastfeeding in public spaces is doubtful to be encouraged or even accepted and full time working mothers may choose ceasing breastfeeding. At present, female employees in China are entitled to 98 days of maternity leave and among the 98 days, 15 may be taken before giving birth, which means that mothers have to return to work around 3 months after delivery (Bohong et al., 2009). In that case, mothers who attained higher education are more likely to return to formal employment after birth compared to the mothers with lower education level.

In addition, to some extent, some incorrect traditional perceptions in China also have strong adverse influence on breastfeeding practices (Xu et al., 2009). One popular typical incorrect traditional perception about breastfeeding among mothers is that breastfeeding could impact on mothers' health or change mothers' shape (Chen and Ji, 1993). Therefore, some mothers especially mothers with higher education or higher income who pay more attentions to their own health or shape after giving birth are willing to choose infant formula to feed their babies instead of breast milk.

With the short maternity leave that is available, a busy and heavy workload, and a less supportive, even unaccepting attitude to breastfeeding in public, as well as some incorrect traditional perceptions on breastfeeding, all these together might cause a number of the more educated mothers to choose to rely on infant formula rather than breastfeeding.

Generally, it is recommended that breastfeeding women have access to midwives who are capable of providing information and support for breastfeeding (Australian Nursing and Midwifery Federation, 2015). However, in China, there is a severe shortage of midwives, with on average only around 3 midwives available for every 100,000 people, which will be exacerbated now that the one child policy has been cancelled (Wang, 2015). The mothers who have attained high education levels usually have a better paid job which allows them the possibility of access to midwives' information and support. A skilled midwives training program has been launched by The National Health and Family Planning Commission in 2014 (Wang, 2015), which may become a novel and effective breastfeeding promotion measure aiming at mothers with high educational status but low breastfeeding prevalence.

The present systematic review retrieved both English and Chinese literature on associations between maternal education and breastfeeding prevalence to reduce reporting bias. In addition, different cut-off points were used to improve the sensitivity of pooled effect size estimate.

However, there are several limitations of this study should be taken into account when considering the results of the study. Firstly, the majority of studies eligible for our systematic review only reported crude (unadjusted) odds ratios without adjustment for any confounders. Due to the limited number of studies that provided adjusted odds ratio, the crude odds ratios were also included in the meta-analysis, which might lead to some concerns about the accuracy and precision of the pooled effect size estimation due to the missing of adjustment for possible confounding effects. If future studies conducted in China could address the possible confounding effects, new investigations are recommended for uncovering the more accurate adjusted relationship between maternal education and breastfeeding. Secondly, no evidence showed that publication bias or small study effect was present in the meta-analysis for Group 1, however, Egger's test and funnel plot both suggested publication bias or small study effect were present in the analysis for Group 2.

In conclusion, the results of our study show a negative association between the maternal education and breastfeeding prevalence within 12 months postpartum. Mothers who have attained a higher level of education are less likely to breastfeed their babies compared to mothers with lower education levels in the context of the Chinese culture and employment environment. Intervention programs with a major focus on the factors related to the effect of maternal education on breastfeeding discussed in this study should be trialled to improve breastfeeding practices in China. These programs would need to promote breastfeeding in general but would give some priority to special groups, including better educated mothers. A further priority is to promote breastfeeding supportive workplaces. The role of the midwife in breastfeeding is essential in promoting and supporting breastfeeding mothers, especially help mothers to practice breastfeeding in a professional manner.

Appendix D

Therefore, enhancing community acceptability of breastfeeding integrated with the supportive role of midwives is also an important aspect. Future research should evaluate intervention studies to promote breastfeeding by mothers of all levels of education, including those who are better educated.

Acknowledgements

We acknowledge the authors who kindly provided additional data for this research. JZ would like to thank China Scholarship Council for financial support (award NO: 201406240008).

References

Al-Sahab, B., Lanes, A., Feldman, M., Tamim, H., 2010. Prevalence and predictors of 6-month exclusive breastfeeding among Canadian women: a national survey. BMC Pediatrics 10, 20.
Amir, L.H., Donath, S.M., 2008. Socioeconomic status and rates of breastfeeding in Australia: evidence from three recent national health surveys. The Medical Journal of Australia 189, 254–256.
Australian Nursing & Midwifery Federation, 2015. Promoting Breastfeeding.
Barria, R.M., Santander, G., Victoriano, T., 2008. Factors associated with exclusive breastfeeding at 3 months postpartum in Valdivia, Chile. Journal of Human Lactation 24, 439–445.
Bax, L., Ikeda, N., Fukui, N., Yaju, Y., Tsuruta, H., Moons, K.G.M., 2009. More Than Numbers: the Power of Graphs in Meta-Analysis. American Journal of Epidemiology 169, 249–255.
Baxter, J., 2008. Breastfeeding, employment and leave: an analysis of mothers in growing up in Australia. Family Matters 80, 17–26.
Baxter, J., Cooklin, A.R., Smith, J., 2009. Which mothers wean their babies prematurely from full breastfeeding? An Australian cohort study. Acta Paediatrica 98, 1274–1277.
Bernaix, L.W., Beaman, M.L., Schmidt, C.A., Harris, J.K., Miller, L.M., 2010. Success of an educational intervention on maternal/newborn nurses' breastfeeding knowledge and attitudes. Journal of Obstetric, Gynecologic, and Neonatal Nursing: JOGNN / NAACOG 39, 658–666.
Bhandari, N., Bahl, R., Mazumdar, S., et al., 2003. Effect of community-based promotion of exclusive breastfeeding on diarrhoeal illness and growth: a cluster randomised controlled trial. Lancet 361, 1418–1423.
Bohong, L., Yongying, Z., Yani, L., 2009. Reconciling Work and Family: Issues and Policies in China. International Labour Organization.
Chalmers, B., Levitt, C., Heaman, M., et al., 2009. Breastfeeding rates and hospital breastfeeding practices in Canada: a national survey of women. Birth 36, 122–132.
Chen, J., Ji, P., 1993. The process to promote breastfeeding in Beijing. Maternal and Child Health Care of China 3, 27–28, (Chinese).
DerSimonian, R., Laird, N., 1986. Meta-analysis in clinical trials. Controlled Clinical Trials 7, 177–188.
Egger, M., Davey Smith, G., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. British Medical Journal 315, 629–634.
Esteves, T.M., Daumas, R.P., Oliveira, M.I., Andrade, C.A., Leite, I.C., 2014. Factors associated to breastfeeding in the first hour of life: systematic review. Revista de Saude Publica 48, 697–708.
Gillman, M.W., Rifas-Shiman, S.L., Camargo, C.A., et al., 2001. Risk of overweight among adolescents who were breastfed as infants. JAMA-Journal of American Medical Association 285, 2461–2467.
Gong, Q., Jackson, P., 2012. Consuming Anxiety? Food, Culture & Society 15, 557–578.
Guan, N., Fan, Q., Ding, J., et al., 2009. Melamine-contaminated powdered formula and urolithiasis in young children. The New England Journal of Medicine 360, 1067–1074.
Guo, S., Fu, X., Scherpbier, R.W., et al., 2013. Breastfeeding rates in central and western China in 2010: implications for child and population health. Bulletin of the World Health Organization 91, 322–331.
Guyatt, G.H., Oxman, A.D., Vist, G., et al., 2011. GRADE guidelines: 4. Rating the quality of evidence-study limitations (risk of bias). Journal of Clinical Epidemiology 64, 407–415.
Heck, K.E., Braveman, P., Cubbin, C., Chavez, G.F., Kiely, J.L., 2006. Socioeconomic status and breastfeeding initiation among California mothers. Public Health Report 121, 51–59.
Higgins, J.P., Altman, D.G., Gotzsche, P.C., et al., 2011. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. British Medical Journal 343, d5928.
Higgins, J.P., Green, S., 2011. Cochrane Handbook for Systematic Reviews of Interventions. John Wiley & Sons.
Higgins, J.P., Thompson, S.G., 2002. Quantifying heterogeneity in a meta-analysis. Statistics in Medicine 21, 1539–1558.
Higgins, J.P., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. Measuring inconsistency in meta-analyses. British Medical Journal 327, 557–560.
Hornell, A., Aarts, C., Kylberg, E., Hofvander, Y., Gebre-Medhin, M., 1999. Breastfeeding patterns in exclusively breastfed infants: a longitudinal prospective study in Uppsala, Sweden. Acta Paediatrica 88, 203–211.
Howell, E.A., Bodnar-Deren, S., Balbierz, A., Parides, M., Bickell, N., 2014. An intervention to extend breastfeeding among black and Latina mothers after delivery. American Journal of Obstetrics and Gynecology 210, e231–e235.
Howie, P.W., Forsyth, J.S., Ogston, S.A., Clark, A., Florey, C.D., 1990. Protective effect of breast feeding against infection. British Medical Journal 300, 11–16.
Huang, H.T., Liang, H., Hu, J.W., et al., 2012. Breastfeeding and associated factors in five areas of China. Journal of Soochow University Medical Science Edition 32, 454–458, (Chinese).
Ip, S., Chung, M., Raman, G., et al., 2007. Breastfeeding and maternal and infant health outcomes in developed countries. Evidence Report/Technology Assessment, 1–186.
Jiang, G.F., 2000. Associated factors analysis for breastfeeding. Jiangsu Prev Med 11 (4): 49-50 (Chinese).
Jiang, H., Li, M., Yang, D.L., et al., 2012. Awareness, Intention, and Needs Regarding Breastfeeding: findings from First-Time Mothers in Shanghai, China. Breastfeeding Medicine 7, 526–534.
Jiang, Y., Guo, L.N., Zhang, L., Wang, C., Cui, Y., 2013. Analysis of feeding and its influence factors among rural infants in 4 provinces of Midwest China. Chinese Journal of Health Education 29, 394–397, (Chinese).
Jiang, Z.H., Li, S., 2008. Survey on factors influencing breastfeeding of neonates. Maternal and Child Health Care of China 23, 3495–3496, (Chinese).
Kang, Y., Liang, X.H., Liu, Y.X., Chen, J., Li, T.Y., 2013. Relative factors of breastfeeding cessation in infants before 6-month-old. Journal of Third Military Medical University 35, 1735–1739, (Chinese).
Ke, F.F., 1993. A survey on breastfeeding status in a rural area. Maternal and Child Health Care of China 8, 54–58, (Chinese).
Kehler, H.L., Chaput, K.H., Tough, S.C., 2009. Risk factors for cessation of breastfeeding prior to six months postpartum among a community sample of women in Calgary, Alberta. Canadian Journal of Public Health 100, 376–380.
Kmet, L.M., Lee, R.C., Cook, L.S., 2004. Standard Quality Assessment Criteria for Evaluating Primary Research Papers from a Variety of Fields. Alberta Heritage Foundation for Medical Research Edmonton.
Kramer, M.S., Chalmers, B., Hodnett, E.D., et al., 2001. Promotion of Breastfeeding Intervention Trial (PROBIT): a randomized trial in the Republic of Belarus. JAMA 285, 413–420.
Li, L., Li, S., Ali, M., Ushijima, H., 2003. Feeding practice of infants and their correlates in urban areas of Beijing, China. Pediatrics International 45, 400–406.
Lin, L., Ma, R.J., Peng, X.H., Wen, J.Y., 1997. Investigation of factors influencing exclusive breastfeeding of 4-month infants Maternal and Child Health Care of China 12. pp. 93–95 (Chinese).
Liu, J., Shi, Z., Spatz, D., Loh, R., Sun, G., Grisso, J., 2013. Social and demographic determinants for breastfeeding in a rural, suburban and city area of South East China. Contemporary Nurse: A Journal for the Australian Nursing Profession 45, 234–243, (210p).
Liu, L., Zhou, W.L., 2013. Present situation investigation of baby breastfeeding within six months in Chengjiaqiao community Shanghai Medicine 34 (12): 62-64 (Chinese).
Ludvigsson, J.F., Ludvigsson, J., 2005. Socio-economic determinants, maternal smoking and coffee consumption, and exclusive breastfeeding in 10205 children. Acta Paediatrica 94, 1310–1319.
Ma, Y.Q., Xu, D., Wang, L., Wang, Y.P., 2014. Breast feeding studies in Hui nationality infants in Xining and Xi'an areas. Chinese Journal of Child Health Care 22, 871–896, (Chinese).
Miller, R.G., 1974. The jackknife-a review. Biometrika 61, 1–15.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., Group, P., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Annals of Internal Medicine 151, 264–269, (W264).
Oken, E., Patel, R., Guthrie, L.B., et al., 2013. Effects of an intervention to promote breastfeeding on maternal adiposity and blood pressure at 11.5 y postpartum: results from the Promotion of Breastfeeding Intervention Trial, a cluster-randomized controlled trial. The American Journal of Clinical Nutrition 98, 1048–1056.
Parry, J., 2008. Contaminated infant formula sickens 6200 babies in China. British Medical Journal 337, a1738.
Pisacane, A., Continisio, P., Filosa, C., Tagliamonte, V., Continisio, G.I., 2012. Use of baby carriers to increase breastfeeding duration among term infants: the effects of an educational intervention in Italy. Acta Paediatrica 101, e434–e438.
Qin, L.L., Hua, J.F., 2013. Investigation and analysis of infant feeding within four months among floating population in Suzhou. Nursing Practices and Research 10, 138–139, (Chinese).
Qiu, L.Q., 2008. Factors that Influence Breastfeeding Initiation and Duration in Urban, Suburban and Rural Areas of Zhejiang Province, Peoples Republic of China. Curtin University, Perth.
Ruan, M.J., Ba, L., Wang, Z., Song, H., 2012. Study on breastfeeding status and the related influencing factors in Chaoyang district of Beijing. Maternal and Child Health Care of China 27, 3478–3480, (Chinese).
Ryan, A.S., Zhou, W.J., 2006. Lower breastfeeding rates persist among the special supplemental nutrition program for women, infants, and children participants, 1978–2003. Pediatrics 117, 1136–1146.
Schanler, R.J., Shulman, R.J., Lau, C., 1999. Feeding strategies for premature infants: beneficial outcomes of feeding fortified human milk versus preterm formula. Pediatrics 103, 1150–1157.
Scott, J.A., Binns, C.W., 1999. Factors associated with the initiation and duration of breastfeeding: a review of the literature. Breastfeeding Review 7, 5–16.
Silveira, R.Bd, Albernaz, E., Zuccheto, L.M., 2008. Fatores associados ao início da amamentação em uma cidade do sul do Brasil. Rev bras saúde matern infant 8, 35–43.
Tan, K.L., 2011. Factors associated with exclusive breastfeeding among infants under six

Appendix D

months of age in peninsular malaysia. International Breastfeeding Journal 6, 2, (2).

Tang, L., 2013. A Cohort Study of Infant Feeding Practices and Maternal Breastfeeding Problems in a Rural Area of Sichuan Province, PR China. Curtin University, Perth.

Tang, L., Binns, C.W., Lee, A.H., Pan, X., Chen, S., Yu, C., 2013. Low prevalence of breastfeeding initiation within the first hour of life in a rural area of Sichuan Province, China. Birth 40, 134–142.

Tang, Z.J., 2014. Investigation on current situation and influencing factors of infant breastfeeding in Liwan District, Guangzhou City. Chinese Journal of Women and Children Health 5, 73–76, (Chinese).

Tian, J.Z., Tian, X.S., Li, L., 2008. Analysis of impact factors on breastfeeding of 183 mothers before discharge. Zhejiang Prev Med 20, 45–55 (Chinese).

Wang, C.X., Chen, H., Liu, Y.Z., 2005. Breastfeeding failure factors analysis and countermeasures. Journal of Community Medicine 3, 11–12, (Chinese).

Wang, H.Z., 2010. Status analysis on breastfeeding in 1296 infants. Practical Preventive Medicine 17, 519–520, (Chinese).

Wang, X.D., 2015. Midwife shortage may grow crucial. China Daily.

Wang, Z., Gao, J.K., Zhang, Y., Wang, F.H., Jin, J.M., Zhu, B.H., et al., 2013. Present status and influencing factors analysis on breastfeeding for infants in Zhejiang Province. Nursing and Rehabilitation Journal 12, 628–630, (Chinese).

World Health Organization, 2003. Infant and Young Child Feeding: A Tool for Assessing National Practices, Policies and Programmes.

World Health Organization, 2008. Indicators for Assessing Infant and Young Child Feeding Practices: Part 1: Definitions: Conclusions of a Consensus Meeting Held 6-8 November 2007 in Washington DC, USA.

Wu Y., Qiu L.Q., 2015. An analysis on the status and influencing factors of breastfeeding among infants within six months of age Zhejiang Prev Med 27, 245–248 (Chinese).

Xiao, D.Q., 2001. The analysis of the condition on breast-feeding of newborn infants. Chinese Journal of Natural Medicine 3, 97–98, (Chinese).

Xiong, H.J., Yang, W.P., Jin, L.N., 2006. Investigation of breastfeeding status. Chinese Journal of General Practitioners 5, 747–749, (Chinese).

Xu, F., Binns, C., Yu, P., Bai, Y., 2007. Determinants of breastfeeding initiation in Xinjiang, PR China, 2003–2004. Acta Paediatrica 96, 257–260.

Xu, F., Qiu, L., Binns, C.W., Liu, X., 2009. Breastfeeding in China: a review. International Breastfeeding Journal 4, 6.

Xu, F.L., 2008. The Determinants of Breastfeeding Initiation and Duration in Xinjiang, Peoples Republic of China. Curtin University, Perth.

Yang, Y.L., Feng, G.Z., 2014. Related factors of breastfeeding. J Math Med 27, 427–428 (Chinese).

Ye, C.E., 2008. Analysis of breastfeeding current situation. Maternal and Child Health Care of China 23, 2770, (Chinese).

Ye, J.L., Cui, Y., Pan, X.P., 2007. Study of the breast feeding and the influence factors in Xinjiang rural areas. Maternal and Child Health Care of China 22, 2420–2422, (Chinese).

Yu, C., 2013. Breastfeeding and Health Outcomes in Infants Who Receive Continuing Care from Hospitals or Community Health Centres in Chengdu Sichuan Province. People's Republic of China. Curtin University, Perth.

Yu, H.M., Song, J.G., 2000. Breastfeeding among 368 infants in Karamay City, Xinjiang, and influential factors thereof Chin J Health Education 16. pp. 581-583 (Chinese).

Zhang, H.L., Wang, F., Pu, C.X., 1998. The current situation of breast-fed infants in the urban community in Wuhan and the affected factors. Medicine and Society 11, 28–30, (Chinese).

Appendix E for Chapter 4

E.1 Publication 3

Appendix E

nutrients

MDPI

Article

Increased Calcium Supplementation Postpartum Is Associated with Breastfeeding among Chinese Mothers: Finding from Two Prospective Cohort Studies

Jian Zhao, Yun Zhao *, Colin W. Binns and Andy H. Lee

School of Public Health, Curtin University, Perth 6102, Australia; jian.zhao@postgrad.curtin.edu.au (J.Z.); c.binns@curtin.edu.au (C.W.B.); andy.lee@curtin.edu.au (A.H.L.)

* Correspondence: y.zhao@curtin.edu.au; Tel.: +61-8-9266-7045

Received: 15 August 2016; Accepted: 26 September 2016; Published: 9 October 2016

Abstract: The calcium supplementation status during the postpartum period among Chinese lactating women is still unclear. The objective of this study is to utilize data from two population-based prospective cohort studies to examine the calcium supplementation status and to identify whether breastfeeding is associated with increased calcium supplementation among Chinese mothers after child birth. Information from 1540 mothers on breastfeeding and calcium supplementation measured at discharge, 1, 3, and 6 months postpartum were extracted to evaluate the association between breastfeeding and calcium supplementation postpartum. A generalized linear mixed model was applied to each study initially to account for the inherent correlation among repeated measurements, adjusting for socio-demographic, obstetric factors and calcium supplementation during pregnancy. In addition, breastfeeding status measured at different follow-up time points was treated as a time dependent variable in the longitudinal analysis. Furthermore, the effect sizes of the two cohort studies were pooled using fixed effect model. Based on the two cohort studies, the pooled likelihood of taking calcium supplementation postpartum among breastfeeding mothers was 4.02 times (95% confidence interval (2.30, 7.03)) higher than that of their non-breastfeeding counterparts. Dietary supplementation intervention programs targeting different subgroups should be promoted in Chinese women, given currently a wide shortage of dietary calcium intake and calcium supplementation postpartum.

Keywords: calcium supplementation; breastfeeding; postpartum; infant; nutrients; generalized linear mixed model; time dependent variable; pooled analysis; China

1. Introduction

The mineral accretion rate of a neonate reaches about 30–40 mg/kg per day, while calcium transfer between mothers and infants is on average 210 mg per day [1–3]. For babies who are breastfed exclusively through the first 6 months, the amount of mineral demand from the mothers is four times greater than that during 9 months of pregnancy [4]. The calcium requirement of mothers during lactation has been the subject of much discussion [5–7]. In 2011, the Institute of Medicine published the calcium dietary reference intakes by life stage, in which Estimated Average Requirement (EAR) of calcium for pregnant and lactating adult women is recommended as 800 mg [8].

Compared to western countries, the lower consumption of dairy products in China results in that most of Chinese residents have calcium intake lower than the adequate intake (AI) [9–11]. In a prospective cohort study of women's health from Shanghai, the median intake of calcium was 485 mg/day, 60% of calcium from plant sources, and only 20% from milk, which was lower than the age group specific AI (800 mg/day for 18–49 years group and 1000 mg/day for over 50 years

Appendix E

group) [11,12]. Only 6.25% of perimenopausal women reached the standard of calcium intake in Changsha [13]. The average intake of calcium of Beijing elderly was 505 mg/day, which was about one half of the recommended adequate intake for the elderly [14]. In the National Nutrition and Health Survey of 2002, fewer than 5% reached the adequate intake levels of calcium for all age groups and the prevalence of calcium supplementation during pregnancy was 41.4% [15,16]. Besides cultural preferences, the lower consumption of dairy products in China is attributed to the high rate of lactose intolerance, which is around 80% to 95% [17,18].

The Chinese National Health and Family Planning Commission recommends that pregnant women should have a dietary calcium intake of 1000 mg per day from the second trimester and increase to 1200 mg per day from the third trimester until the end of lactation [19]. However, low dietary calcium intake in lactating women has been reported in different regions of China, as shown in Table 1. This suggests that calcium supplementation for lactating women is an important public health issue to mothers in China based on the current evidence about the benefits of calcium intake during lactation on reducing maternal bone loss [20–23].

Table 1. Dietary calcium intake of lactating women in different regions of China.

Study Location	Study Design	Study Period	Average Daily Dietary Calcium Intake (Postpartum)
Guangzhou [24]	Prospective cohort	2002	786.45 mg (12 weeks)
Hunan [25]	Cross-sectional	2011–2012	426 mg
Beijing, Suzhou & Guangzhou [26]	Cross-sectional	2011–2012	401.4 mg (0–1 month) 585.3 mg (1–2 months) 591.2 mg (2–4 months) 649.0 mg (4–8 months)
Fujian [27]	Prospective cohort	2012	428 mg (2 days) 454 mg (7 days) 595 mg (30 days) 544 mg (90 days)
Shanghai [28]	Prospective cohort	2014–2015	749.3 mg (1–3 days) 781.1 mg (7–9 days) 762.3 mg (14–17 days) 768.4 mg (25–27 days) 678.5 mg (39–41 days)

The calcium supplementation status during postpartum period among Chinese lactating women is still unclear. The objective of the present study is to utilize data from two population-based prospective cohort studies to examine the calcium supplementation status and to identify whether breastfeeding is associated with increased calcium supplementation among Chinese mothers after child birth.

2. Materials and Methods

2.1. Study Participants

Two prospective cohort studies were conducted in an urban area, Chengdu (capital city) and a rural area, Jiangyou (county-level city), Sichuan Province, China between 2010 and 2012. Mothers who gave birth to a healthy singleton infant were invited to participate before discharge. These two studies used the same methodology based on same questionnaires, which had been used in Australia and China [29–31] previously, to interview all consented women face-to-face at discharge, and followed up the participants at one, three and six months postpartum by telephone interviews. The baseline interview collected detailed information on mothers and newborns, including socio-demographic, obstetric characteristics and dietary supplements during pregnancy. The follow-up interviews collected detailed information on lactation patterns and durations and dietary supplements during the postpartum period. The World Health Organization (WHO) standard definition of any breastfeeding was used in these two studies; 'Any breastfeeding' is defined as the infant has received breast milk (direct from the breast or expressed) with or without other drink, formula or other infant food [32].

2.2. Ethical Approval

The two cohort studies were approved by the Human Research Ethics Committee of Curtin University, Perth, Western Australia (approval numbers: HR169/2009 and HR168/2009, respectively). The present study was also approved by the Human Research Ethics Committee of Curtin University (approval number: RDHS-101-15). The data used in this study were de-identified.

2.3. Statistical Analysis

The outcome of the present study is maternal calcium supplementation status (yes or no) measured longitudinally during three different postpartum periods (from discharge to 1 month, from 1 month to 3 months, and from 3 months to 6 months, respectively) at three follow-up time points (namely, 1 month, 3 months and 6 months postpartum). The main variable of interest, any breastfeeding status, was measured longitudinally at three different postpartum time points (discharge, 1 month and 3 months postpartum). Descriptive statistics of mothers' socio-demographic status, obstetric characteristics, calcium supplementation during pregnancy and the three postpartum periods, and any breastfeeding status at the three postpartum time points were obtained and reported. Chi-square test was conducted to compare the calcium supplementation rates between breastfeeding group and non-breastfeeding group at the different follow-up time points. Generalized linear mixed model (GLMM) was used to examine the effect of breastfeeding on calcium supplementation postpartum taking into account inherent correlations among repeated measurements. Furthermore, the breastfeeding status was included as a time-dependent variable in the longitudinal analysis. Random intercept model without covariates (Model I) was run initially to test random intercept effect, and then any breastfeeding status at the different time points and an indicator variable of measurement times were added into the above Model I to be a Model II. Furthermore, subject level socio-demographic covariates such as household annual income, maternal age and maternal education were then added into and adjusted in the Model II to formulate a Model III. Finally, obstetric characteristics such as parity, gravidity, infant gender, infant birth weight and infant gestational week, together with calcium supplementation during pregnancy, were further adjusted in the Model III to become the final Model IV. The above regression analysis was carried out for data set extracted from each cohort study separately, and the results of Model II and final Model IV were reported. In addition, a pooled effect size was calculated using a fixed effect model given that the heterogeneity between the two studies was tested being statistically nonsignificant. All statistical analyses were performed by using SAS 9.4 (SAS Institute Inc., Cary, NC, USA).

3. Results

3.1. Characteristics of Participants

For each cohort, mothers' baseline socio-demographic status, obstetric characteristics and calcium supplementation during pregnancy are presented in Table 2. In the Jiangyou study, 695 mothers were interviewed at baseline, and 648 and 620 mothers remained in the study at 1 month and 3 months postpartum, respectively. Any breastfeeding rate dropped slightly from 93.53% at discharge to 91.05% at 1 month postpartum then continuously to 83.71% at 3 months postpartum. In the other cohort conducted in Chengdu, 845 mothers were interviewed at baseline and 760 mothers were followed up until six months postpartum. Any breastfeeding rate declined from 93.02% at discharge to 87.89% at 1 month postpartum then substantially to 73.42% at 3 months postpartum.

Appendix E

Table 2. Characteristics of participants at baseline by breastfeeding status.

Variable	Cohort in Jiangyou (n = 695)		Cohort in Chengdu (n = 845)	
	BF	**Non-BF**	**BF**	**Non-BF**
Number of participants	650 (93.5)	45 (6.5)	786 (93.0)	59 (7.0)
Household annual income (Chinese yuan)				
<2000	186 (31.0)	9 (23.1)	1 (0.2)	0 (0.0)
2000–5000	309 (51.4)	23 (59.0)	155 (23.5)	12 (24.0)
>5000	106 (17.6)	7 (17.9)	503 (76.3)	38 (76.0)
Maternal age (years)				
<25	373 (57.4)	26 (57.8)	156 (19.9)	5 (8.5)
25–29	163 (25.1)	13 (28.9)	372 (47.3)	28 (47.5)
>29	114 (17.5)	6 (13.3)	258 (32.8)	26 (44.0)
Maternal education				
Secondary school or lower	355 (54.6)	25 (55.6)	90 (11.5)	11 (18.6)
Senior school	215 (33.1)	18 (40.0)	165 (21.0)	11 (18.6)
University or higher	80 (12.3)	2 (4.4)	531 (67.5)	37 (62.8)
Parity				
Primiparous	518 (79.7)	37 (82.2)	700 (89.1)	51 (86.4)
Multiparous	132 (20.3)	8 (17.8)	86 (10.9)	8 (13.6)
Gravidity				
Primigravida	249 (38.3)	18 (40.0)	430 (54.7)	26 (44.1)
Multigravida	401 (61.7)	27 (60.0)	356 (45.3)	33 (55.9)
Infant gender				
Male	328 (50.5)	26 (57.8)	412 (52.4)	34 (57.6)
Female	322 (49.5)	19 (42.2)	374 (47.6)	25 (42.4)
Infant birth weight (g)				
<2500	10 (1.5)	2 (4.4)	13 (1.7)	0 (0.0)
≥2500	640 (98.5)	43 (95.6)	773 (98.3)	59 (100.0)
Infant gestational week				
<37	8 (1.2)	3 (6.8)	9 (1.2)	2 (3.4)
≥37	640 (98.8)	41 (93.2)	777 (98.8)	57 (96.6)
Calcium supplementation during pregnancy				
Yes	410 (63.1)	25 (55.6)	627 (79.8)	47 (79.7)
No	240 (36.9)	20 (44.4)	159 (20.2)	12 (20.3)

Data are presented as n (%); BF: any breastfeeding; Non-BF: non-breastfeeding.

3.2. Calcium Supplementation Status during Postpartum Period

Overall, among mothers in the Jiangyou cohort, an inverted U shape of calcium supplementation rates at three different postpartum periods was observed, which corresponded to 13.4%, 19.4% and 17.7%, respectively. While in the Chengdu cohort, a constant decline trend was recorded with 22.5%, 22.2% and 12.0% reported at the three postpartum periods. When considering separately for breastfeeding and non-breastfeeding groups, as shown in Figures 1 and 2, the calcium supplementation rate in the breastfeeding group was statistically significantly higher than that in the non-breastfeeding group for all the different postpartum periods, except between discharge and 1 month in the Jiangyou cohort (p = 0.36). In the Jiangyou cohort, calcium supplementation rates ranged from 13.7% to 21.2% for breastfeeding mothers, and ranged from 1.7% to 8.9% for non-breastfeeding mothers. In the Chengdu cohort, calcium supplementation rates reduced from around 23% in the first 3 months postpartum to 14.5% between 3 months and 6 months in breastfeeding mothers, and ranged from 5.0% to 14.1% in non-breastfeeding mothers.

Appendix E

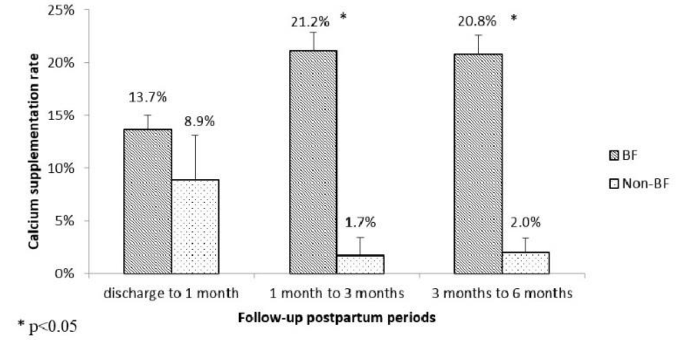


Figure 1. Calcium supplementation postpartum in Jiangyou.

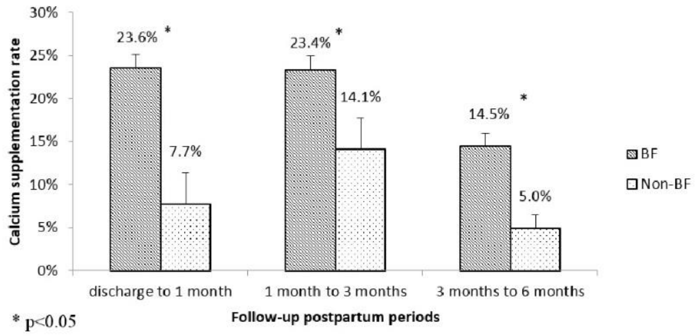


Figure 2. Calcium supplementation postpartum in Chengdu.

3.3. The Association between Breastfeeding and Calcium Supplementation Postpartum

In Model I (without any covariates) for both cohorts, subject random effect was found to be statistically significant. Hence, both the primary variables of interest (i.e., breastfeeding status and the indicator variable of measurement times) were subsequently added into the Model I for examining the association between breastfeeding and calcium supplementation postpartum. As shown in Table 3, the likelihood of calcium supplementation in breastfeeding mothers were 5.85 times (95% confidence interval (CI) (2.50, 13.72)) and 2.88 times (95% CI (1.50, 5.54)) higher of that in non-breastfeeding mothers in Jiangyou and Chengdu, respectively. After adjusting for socio-demographic and obstetric factors as well as calcium supplementation during pregnancy, the odds ratios (ORs) and its 95% CI had changed appreciably to 6.95 and (2.68, 18.04) in the Jiangyou study, and 3.03 and (1.52, 6.02) in the Chengdu study, respectively. The heterogeneity between these two studies was not significant ($I^2 = 0.479$, $p = 0.17$) statistically, therefore a fixed effect model was used to pool the ORs of the two studies. The pooled analysis of these two cohort studies revealed that calcium supplementation postpartum was significantly positively associated with breastfeeding with an adjusted OR = 4.02 with a 95% CI of (2.30, 7.03).

Appendix E

Table 3. Association between breastfeeding status and calcium supplementation postpartum.

Variable	Model II	Model IV
	Crude ORs (95% CI)	**Adjusted ORs (95% CI)**
Jiangyou Cohort		
Measurement times *		
At discharge (ref)	1	1
1 month	1.72 (1.24, 2.38)	1.90 (1.33, 2.70)
3 months	1.57 (1.12, 2.20)	1.69 (1.18, 2.44)
Breastfeeding status *		
Non-breastfeeding (ref)	1	1
Any breastfeeding	5.85 (2.50, 13.72)	6.95 (2.68, 18.04)
Chengdu Cohort		
Measurement times *		
At discharge (ref)	1	1
1 month	1.02 (0.74, 1.43)	1.02 (0.72, 1.45)
3 months	0.31 (0.21, 0.46)	0.30 (0.20, 0.45)
Breastfeeding status *		
Non-breastfeeding (ref)	1	1
Any breastfeeding	2.88 (1.50, 5.54)	3.03 (1.52, 6.02)
Pooled effect size of two studies		
Non-breastfeeding (ref)	-	1
Any breastfeeding	-	4.02 (2.30, 7.03)

Crude ORs (obtained from Model II): Model included breastfeeding status and the indicator variable of measurement times; Adjusted ORs (obtained from the final Model IV): Model adjusted for socio-demographics variables (household annual income, maternal age and maternal education); obstetric factors (parity, gravidity, infant gender, infant birth weight and infant gestational week); and calcium supplementation during pregnancy; * $p < 0.05$; ref: reference category.

4. Discussion

To our knowledge, the present study is the first population-based study that determines the longitudinal trend of calcium supplementation by Chinese women from discharge to 6 months postpartum and the effect of breastfeeding on calcium supplementation. A relatively low level of calcium supplementation (less than 23%) was observed throughout the postpartum period in either breastfeeding mothers or non-breastfeeding mothers. The pooled effect size after adjusting for socio-demographics variables (household annual income, maternal age and maternal education); obstetric factors (parity, gravidity, infant gender, infant birth weight and infant gestational week); and calcium supplementation during pregnancy reveals that mothers who breastfed their babies were 4.02 times more likely to take calcium supplements compared to their non-breastfeeding counterparts during postpartum. The present result is consistent with previous findings that breastfeeding mothers consumed more calcium than non-breastfeeding counterparts [33–35]. One reason leading to the higher calcium supplementation in breastfeeding mothers may be the general belief that adequate calcium intake is beneficial to breast milk production, and mothers' special attention to infants' calcium intake under the context of wide shortage of calcium intake for Chinese women, in spite of recent evidence demonstrating that calcium supplementation in lactation has no significant effect on increasing calcium content in breast milk [36–38]. The other reason might be mothers' perception of the beneficial effect of calcium supplementation on maternal bone loss during lactating. Some studies found little benefits of calcium supplementation on maternal bone loss during lactating [36,39,40], whereas other studies carried out in the U.S. and Brazil suggested that higher calcium intake during early lactation could minimize the bone loss for the mothers who had daily calcium intake less than 500 mg [20,21]. Further investigation on the factors contributing to difference of calcium supplementation between breastfeeding mothers and non-breastfeeding mothers as well as the effect of calcium supplementation

Appendix E

on reducing maternal bone loss during lactation or enhancing maternal skeleton remodeling and remineralization after weaning of breastfeeding is recommended.

Given the habitually lower calcium dietary intake and relatively high lactose intolerance rate in the general Chinese population [12,15,17], calcium supplementation plays an important role on bone health, especially for exclusive breastfeeding women who provide around 300 mg of calcium per day to their babies via breast milk which accompany maternal bone calcium turnover [41].

This study had several strengths. We utilized data from two cohort studies to investigate the longitudinal trends of calcium supplementation at three different postpartum time points (i.e., 1 month, 3 months and 6 months postpartum) and conducted random effect regression modelling accounting for inherent dependency between the repeated measurements. Moreover, since the breastfeeding status was measured longitudinally as well in two cohorts, it was treated as a time-dependent variable in the analysis to account for possible feedback effects between the breastfeeding status and calcium supplementation at different times. In addition, our pooled analysis based on the two individual studies yielded the combined effect size with a larger sample size and higher statistical power. Moreover, calcium supplementation during pregnancy was adjusted in the modelling to control for the consequent effect of calcium intake during pregnancy on calcium supplementation during lactation.

A caveat of this study was that both cohort studies were carried out in Sichuan Province, which may limit the results being able to generalize to other regions of China. Sichuan Basin has special geographic characteristics, where the number of cloudy or rainy days is substantially larger than that in other regions in China, which may lead to a relatively lower level of vitamin D synthesis and calcium deficiency consequently [42]. However, to the best of our knowledge, no data were available currently on calcium supplementation during postpartum in other regions of China for comparison purpose.

5. Conclusions

In conclusion, calcium supplementation during postpartum in Sichuan is variable at different times postpartum with a relatively low level (less than 23%). Although breastfeeding has a substantive effect on calcium supplementation postpartum, dietary supplementation intervention programs and health education targeting different subgroups (e.g., breastfeeding mothers and bottle feeding mothers) should be promoted in Chinese women, given currently a wide shortage of dietary calcium intake and calcium supplementation during postpartum.

Acknowledgments: We would like to thank Li Tang and Chuan Yu from West China Hospital, Sichuan University for providing the two datasets. Jian Zhao would like to acknowledge China Scholarship Council (CSC) for their financial support for his doctoral studies.

Author Contributions: J.Z., Y.Z. and C.W.B. designed the research; J.Z. performed statistical analyses and Y.Z., C.W.B. and A.H.L. provided theoretical and methodological guidance; J.Z., Y.Z. and A.H.L. wrote the manuscript. All authors read and approved the final version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Trotter, M.; Hixon, B.B. Sequential changes in weight, density, and percentage ash weight of human skeletons from an early fetal period through old age. *Anat. Rec.* **1974**, *179*, 1–18. [CrossRef] [PubMed]
2. Olausson, H.; Goldberg, G.R.; Laskey, M.A.; Schoenmakers, I.; Jarjou, L.M.; Prentice, A. Calcium economy in human pregnancy and lactation. *Nutr. Res. Rev.* **2012**, *25*, 40–67. [CrossRef] [PubMed]
3. Kalkwarf, H.J.; Specker, B.L.; Ho, M. Effects of calcium supplementation on calcium homeostasis and bone turnover in lactating women. *J. Clin. Endocrinol. Metab.* **1999**, *84*, 464–470. [CrossRef] [PubMed]
4. Kovacs, C.S.; Ralston, S.H. Presentation and management of osteoporosis presenting in association with pregnancy or lactation. *Osteoporos. Int.* **2015**, *26*, 2223–2241. [CrossRef] [PubMed]
5. Thomas, M.; Weisman, S.M. Calcium supplementation during pregnancy and lactation: Effects on the mother and the fetus. *Am. J. Obstet. Gynecol.* **2006**, *194*, 937–945. [CrossRef] [PubMed]

Appendix E

6. Prentice, A. Calcium in pregnancy and lactation. *Annu. Rev. Nutr.* **2000**, *20*, 249–272. [CrossRef] [PubMed]
7. Kovacs, C.S. Calcium and bone metabolism disorders during pregnancy and lactation. *Endocrinol. Metab. Clin. N. Am.* **2011**, *40*, 795–826. [CrossRef] [PubMed]
8. Institute of Medicine. *Dietary Reference Intakes for Calcium and Vitamin D*; National Academies Press: Washington, DC, USA, 2011.
9. Chen, Y.M.; Teucher, B.; Tang, X.Y.; Dainty, J.R.; Lee, K.K.; Woo, J.L.; Ho, S.C. Calcium absorption in postmenopausal Chinese women: A randomized crossover intervention study. *Br. J. Nutr.* **2007**, *97*, 160–166. [CrossRef] [PubMed]
10. Ma, G.; Li, Y.; Jin, Y.; Zhai, F.; Kok, F.J.; Yang, X. Phytate intake and molar ratios of phytate to zinc, iron and calcium in the diets of people in China. *Eur. J. Clin. Nutr.* **2007**, *61*, 368–374. [CrossRef] [PubMed]
11. Wang, Y.; Li, S. Worldwide trends in dairy production and consumption and calcium intake: Is promoting consumption of dairy products a sustainable solution for inadequate calcium intake? *Food Nutr. Bull.* **2008**, *29*, 172–185. [CrossRef] [PubMed]
12. Shin, A.; Li, H.; Shu, X.O.; Yang, G.; Gao, Y.T.; Zheng, W. Dietary intake of calcium, fiber and other micronutrients in relation to colorectal cancer risk: Results from the Shanghai Women's Health Study. *Int. J. Cancer* **2006**, *119*, 2938–2942. [CrossRef] [PubMed]
13. Deng, J.; Huang, Y.M.; Lin, Q. Main sources of dietary calcium in perimenopausal women in Changsha. *Zhong Nan Da Xue Xue Bao Yi Xue Ban* **2008**, *33*, 875–879. [PubMed]
14. Liu, X.; Zhao, X.; Xu, L. Food sources of calcium and iron in the diet of Beijing elderly. *Wei Sheng Yan Jiu* **2004**, *33*, 336–338. [PubMed]
15. He, Y.; Zhai, F.; Wang, Z.; Hu, Y. Status of dietary calcium intake of Chinese residents. *Wei Sheng Yan Jiu* **2007**, *36*, 600–602. [PubMed]
16. Lai, J.; Yin, S.; Ma, G.; Piao, J.; Yang, X. The nutrition and health survery of pregnant women in China. *Acta Nutr. Sin.* **2007**, *29*, 4–8.
17. De Vrese, M.; Stegelmann, A.; Richter, B.; Fenselau, S.; Laue, C.; Schrezenmeir, J. Probiotics—Compensation for lactase insufficiency. *Am. J. Clin. Nutr.* **2001**, *73*, 421S–429S. [PubMed]
18. Wang, Y.G.; Yan, Y.S.; Xu, J.J.; Du, R.F.; Flatz, S.D.; Kuhnau, W.; Flatz, G. Prevalence of primary adult lactose malabsorption in three populations of northern China. *Hum. Genet.* **1984**, *67*, 103–106. [PubMed]
19. National Health and Family Planning Commission of the People's Republic of China. Maternal and Child Health Basic Knowledge and Skills (Trial Version). Available online: http://www.nhfpc.gov.cn/fys/kpxc/201405/0065ee2071204f0c9b52f7cdcbb392f2.shtml (accessed on 5 July 2016).
20. O'Brien, K.O.; Donangelo, C.M.; Ritchie, L.D.; Gildengorin, G.; Abrams, S.; King, J.C. Serum 1,25-dihydroxyvitamin D and calcium intake affect rates of bone calcium deposition during pregnancy and the early postpartum period. *Am. J. Clin. Nutr.* **2012**, *96*, 64–72. [CrossRef] [PubMed]
21. O'Brien, K.O.; Donangelo, C.M.; Zapata, C.L.; Abrams, S.A.; Spencer, E.M.; King, J.C. Bone calcium turnover during pregnancy and lactation in women with low calcium diets is associated with calcium intake and circulating insulin-like growth factor 1 concentrations. *Am. J. Clin. Nutr.* **2006**, *83*, 317–323. [PubMed]
22. Chan, G.M.; McMurry, M.; Westover, K.; Engelbert-Fenton, K.; Thomas, M.R. Effects of increased dietary calcium intake upon the calcium and bone mineral status of lactating adolescent and adult women. *Am. J. Clin. Nutr.* **1987**, *46*, 319–323. [PubMed]
23. Yoneyama, K.; Ikeda, J. The effects of increased dietary calcium intake on bone mineral density in long-term lactating women, and recovery of bone loss caused by long-term lactation with low calcium diet. *Nihon Koshu Eisei Zasshi* **2004**, *51*, 1008–1017. [PubMed]
24. Li, Y.; Liao, D.; Miao, L.; Qian, X. Analysis on dietary investigation of pregnant and lactating women. *Chin. Prim. Health Care* **2004**, *18*, 27–28.
25. Huang, Z. The Dietary Status of 269 Lactating Women and the Content of Minerals in Breast Milk. Master's Thesis, Central South University, Changsha, China, 2014.
26. Yang, T.; Zhang, Y.; Ma, D.; Li, W.; Yang, X.; Wang, P. Survey on the nutrients intake of lactating women in three cities of China. *Acta Nutr. Sin.* **2014**, *36*, 84–86.
27. Chen, H.; Wang, P.; Han, Y.; Ma, J.; Troy, F.A., II; Wang, B. Evaluation of dietary intake of lactating women in China and its potential impact on the health of mothers and infants. *BMC Women's Health* **2012**, *12*, 18. [CrossRef] [PubMed]

Appendix E

28. Kong, X.; Fei, J.; Zhai, Y.; Feng, Y.; Li, J. Dietary survey of lactating mothers during the puerperal state. *China Med. Her.* **2016**, *13*, 49–68.
29. Qiu, L.; Zhao, Y.; Binns, C.W.; Lee, A.H.; Xie, X. A cohort study of infant feeding practices in city, suburban and rural areas in Zhejiang Province, PR China. *Int. Breastfeed. J.* **2008**, *3*, 4. [CrossRef] [PubMed]
30. Xu, F.; Liu, X.; Binns, C.W.; Xiao, C.; Wu, J.; Lee, A.H. A decade of change in breastfeeding in China's far north-west. *Int. Breastfeed. J.* **2006**, *1*, 22. [CrossRef] [PubMed]
31. Scott, J.A.; Landers, M.C.; Hughes, R.M.; Binns, C.W. Factors associated with breastfeeding at discharge and duration of breastfeeding. *J. Paediatr. Child Health* **2001**, *37*, 254–261. [CrossRef] [PubMed]
32. World Health Organization. Indicators for assessing infant and young child feeding practices: Part 1: Definitions. In Proceedings of the Conclusions of a Consensus Meeting, Washington, DC, USA, 6–8 November 2007.
33. Chan, S.M.; Nelson, E.A.; Leung, S.S.; Cheng, J.C. Bone mineral density and calcium metabolism of Hong Kong Chinese postpartum women—A 1-y longitudinal study. *Eur. J. Clin. Nutr.* **2005**, *59*, 868–876. [CrossRef] [PubMed]
34. Laskey, M.A.; Prentice, A.; Hanratty, L.A.; Jarjou, L.M.; Dibba, B.; Beavan, S.R.; Cole, T.J. Bone changes after 3 mo of lactation: Influence of calcium intake, breast-milk output, and vitamin D-receptor genotype. *Am. J. Clin. Nutr.* **1998**, *67*, 685–692. [PubMed]
35. Lopez, J.M.; Gonzalez, G.; Reyes, V.; Campino, C.; Diaz, S. Bone turnover and density in healthy women during breastfeeding and after weaning. *Osteoporos. Int.* **1996**, *6*, 153–159. [CrossRef] [PubMed]
36. Kalkwarf, H.J.; Specker, B.L.; Bianchi, D.C.; Ranz, J.; Ho, M. The effect of calcium supplementation on bone density during lactation and after weaning. *N. Engl. J. Med.* **1997**, *337*, 523–528. [CrossRef] [PubMed]
37. Prentice, A.; Jarjou, L.M.; Cole, T.J.; Stirling, D.M.; Dibba, B.; Fairweather-Tait, S. Calcium requirements of lactating Gambian mothers: Effects of a calcium supplement on breast-milk calcium concentration, maternal bone mineral content, and urinary calcium excretion. *Am. J. Clin. Nutr.* **1995**, *62*, 58–67. [PubMed]
38. Jarjou, L.M.; Prentice, A.; Sawo, Y.; Laskey, M.A.; Bennett, J.; Goldberg, G.R.; Cole, T.J. Randomized, placebo-controlled, calcium supplementation study in pregnant Gambian women: Effects on breast-milk calcium concentrations and infant birth weight, growth, and bone mineral accretion in the first year of life. *Am. J. Clin. Nutr.* **2006**, *83*, 657–666. [PubMed]
39. Zhang, Z.Q.; Chen, Y.M.; Wang, R.Q.; Huang, Z.W.; Yang, X.G.; Su, Y.X. The effects of different levels of calcium supplementation on the bone mineral status of postpartum lactating Chinese women: A 12-month randomised, double-blinded, controlled trial. *Br. J. Nutr.* **2016**, *115*, 24–31. [CrossRef] [PubMed]
40. Cross, N.A.; Hillman, L.S.; Allen, S.H.; Krause, G.F. Changes in bone mineral density and markers of bone remodeling during lactation and postweaning in women consuming high amounts of calcium. *J. Bone Miner. Res.* **1995**, *10*, 1312–1320. [CrossRef] [PubMed]
41. Kovacs, C.S. Maternal mineral and bone metabolism during pregnancy, lactation, and post-weaning recovery. *Physiol. Rev.* **2016**, *96*, 449–547. [PubMed]
42. Wang, J.; Yang, F.; Mao, M.; Liu, D.H.; Yang, H.M.; Yang, S.F. High prevalence of vitamin D and calcium deficiency among pregnant women and their newborns in Chengdu, China. *World J. Pediatr. WJP* **2010**, *6*, 265–267. [CrossRef] [PubMed]

Appendix F for Chapter 5

F.1 Permission from publisher-John Wiley and Sons

Appendix F

JOHN WILEY AND SONS LICENSE
TERMS AND CONDITIONS

May 06, 2018

This Agreement between Mr. Jian Zhao ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

License Number	4342861172680
License date	May 06, 2018
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Paediatric and Perinatal Epidemiology
Licensed Content Title	A Time-varying Covariate Approach for Survival Analysis of Paediatric Outcomes
Licensed Content Author	Colin W Binns, Andy H Lee, Yun Zhao, et al
Licensed Content Date	Sep 22, 2017
Licensed Content Volume	31
Licensed Content Issue	6
Licensed Content Pages	5
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Electronic
Portion	Full article
Will you be translating?	No
Title of your thesis / dissertation	Statistical modelling of breastfeeding data
Expected completion date	Jul 2018
Expected size (number of pages)	1
Requestor Location	Mr. Jian Zhao 2/82 Reynolds Rd Mount Pleasant Perth, WA 6153 Australia Attn: Mr. Jian Zhao
Publisher Tax ID	EU826007151
Total	0.00 AUD
Terms and Conditions	

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a"Wiley Company") or handled on behalf of a society with which a Wiley Company has exclusive publishing rights in relation to a particular work (collectively "WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your RightsLink account (these are available at any time at http://myaccount.copyright.com).

Terms and Conditions

- The materials you have requested permission to reproduce or reuse (the "Wiley Materials") are protected by copyright.

- You are hereby granted a personal, non-exclusive, non-sub licensable (on a stand-alone basis), non-transferable, worldwide, limited license to reproduce the Wiley Materials for the purpose specified in the licensing process. This license, **and any CONTENT (PDF or image file) purchased as part of your order,** is for a one-time use only and limited to any maximum distribution number specified in the license. The first instance of republication or reuse granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before the end date may be distributed thereafter). The Wiley Materials shall not be used in any other manner or for any other purpose, beyond what is granted in the license. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Wiley Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Wiley Material. Any third party content is expressly excluded from this permission.

- With respect to the Wiley Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Wiley Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Wiley Materials without the prior permission of the respective copyright owner.**For STM Signatory Publishers clearing permission under the terms of the STM Permissions Guidelines only, the terms of the license are extended to include subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and does not involve the separate exploitation of the permitted figures or extracts,** You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Wiley Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Wiley Materials on a stand-alone basis, or any of the rights granted to you hereunder to any other person.

- The Wiley Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc, the Wiley Companies, or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Wiley Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Wiley Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto

- NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS

OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

- WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

- You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

- IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

- Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

- The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

- This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

- Any fee required for this permission shall be non-refundable after thirty (30) days from receipt by the CCC.

- These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and

WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

- In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.
- WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.
- This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.
- This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

WILEY OPEN ACCESS TERMS AND CONDITIONS

Wiley Publishes Open Access Articles in fully Open Access Journals and in Subscription journals offering Online Open. Although most of the fully Open Access journals publish open access articles under the terms of the Creative Commons Attribution (CC BY) License only, the subscription journals and a few of the Open Access Journals offer a choice of Creative Commons Licenses. The license type is clearly identified on the article.

The Creative Commons Attribution License

The Creative Commons Attribution License (CC-BY) allows users to copy, distribute and transmit an article, adapt the article and make commercial use of the article. The CC-BY license permits commercial and non-

Creative Commons Attribution Non-Commercial License

The Creative Commons Attribution Non-Commercial (CC-BY-NC)License permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.(see below)

Creative Commons Attribution-Non-Commercial-NoDerivs License

The Creative Commons Attribution Non-Commercial-NoDerivs License (CC-BY-NC-ND) permits use, distribution and reproduction in any medium, provided the original work is properly cited, is not used for commercial purposes and no modifications or adaptations are made. (see below)

Use by commercial "for-profit" organizations

Use of Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee.

Further details can be found on Wiley Online Library
http://olabout.wiley.com/WileyCDA/Section/id-410895.html

Other Terms and Conditions:

v1.10 Last updated September 2015

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

F.2 Publication 4

Appendix F

Paediatric and Perinatal Epidemiology

598 doi: 10.1111/ppe.12410

Brief Report

A Time-varying Covariate Approach for Survival Analysis of Paediatric Outcomes

Jian Zhao, Yun Zhao, Andy H Lee, Colin W Binns

Department of Epidemiology and Biostatistics, School of Public Health, Curtin University, Perth, WA, Australia

Abstract

Background: Conventional survival analysis is commonly applied in the analysis of time-to-event data in paediatric studies, where the exposure variables of interest are often treated as time-fixed. However, the values of these exposure variables can vary over time and time-fixed analysis may introduce time-dependent bias.

Methods: Time-dependent bias is illustrated graphically considering two scenarios in longitudinal study settings for paediatric time-to-event outcomes. As an illustrative example, the time-varying covariate approach was applied to survival analysis of breast-feeding data (n = 695) collected in China between 2010 and 2011, with an emphasis on the effects of covariates 'solid foods introduction' and 'maternal return to work' on breast-feeding duration up to 12 months postpartum.

Results: Time-varying exposures could occur before or after the occurrence of an event of interest so that time-fixed analysis can lead to biased and imprecise parameter estimates. In the illustrative example, the reduced risk of 'solid foods introduction' (hazard ratio (HR) 0.61, 95% confidence interval (CI) 0.50, 0.75) on breast-feeding cessation and an absence of an association with 'maternal return to work' (HR 0.99, 95% CI 0.73, 1.36) from the time-fixed analysis reversed (HR 1.50, 95% CI 1.17, 1.93) and became significant (HR 1.45, 95% CI 1.06, 2.00), respectively, based on the time-varying covariate model.

Conclusions: The time-varying covariate approach is preferable for survival analysis of time-to-event data in the presence of time-varying exposures.

Keywords: *time-varying covariate, time-dependent bias, survival analysis, paediatric outcomes.*

Background

In prospective cohort study designs, survival analysis is commonly used in the analysis of time-to-event paediatric outcomes. In such designs, repeated observations are collected longitudinally throughout the follow-up period and some information may vary over time. Typical examples of time-dependent variation are infant's height and weight, which increase as the infant grows older. Consequently, the associations between these time-varying variables and paediatric outcomes will change over time, suggesting a necessity to account for time-varying effects in the statistical analysis.

If time-varying nature of exposure factors is not taken into account, a 'time-dependent bias' can be introduced into survival analysis,[1–3] potentially altering the conclusions of the study.[2–4] However, in some paediatric studies, where time-varying exposures are measured, conventional survival model based on time-fixed analysis is still widely adopted. Some example studies include identifying factors associated with breast-feeding duration,[5] assessing the relationship between vaccination and childhood diabetes,[6] and determining neonatal factors affecting the development of childhood epilepsy.[7] Time-varying exposures are often treated as time-fixed variables over the entire follow-up period, which may lead to incorrect inferences.[8,9]

The aim of this study is to illustrate the time-dependent bias introduced using time-fixed variables in the conventional survival analysis of paediatric outcomes. We illustrate this bias in an application of breast-feeding duration with solid foods introduction and maternal return to work as both time-fixed and time-varying covariates.

Correspondence:
Yun Zhao, School of Public Health, Curtin University, Perth, Australia.
E-mail: y.zhao@curtin.edu.au

Appendix F

Time-dependent bias

Figure 1 illustrates two different occurrence times of an exposure of interest in a paediatric study setting, where the risk set is defined as the set of individuals at risk at some specific time-to-event point.[10] In Scenario 1, the exposure occurs after the start of the follow-up and before the occurrence of an outcome event (such as childhood diabetes or cessation of breast feeding). If a time-fixed analysis is performed and the exposure is treated as a static baseline variable, the subject will be assigned to the exposed group. However, the time between initiation of the follow-up and the occurrence of exposure is 'immortal' to the occurrence of the event, thus the subject within the above time window should not be classified into the exposed group.

Scenario 2 considers an exposure occurring after the occurrence of the event. If a time-fixed analysis is performed, the subject associated with this exposure will be classified into the exposed group. According to survival analysis theory, an exposure can make a contribution to the occurrence of the event only if it occurs within the time window between the onset of the follow-up and the occurrence of the event. Since the exposure happens after the occurrence of the event in Scenario 2, it cannot make any contribution to the occurrence of the event, indicating that the inclusion of this subject into the exposed group for calculating the hazard function is inappropriate. In summary, both scenarios show how a time-fixed analysis could potentially lead to biased and misleading estimates of effect sizes.

Time-varying covariate approach

The time-varying covariate approach is implemented with an extended Cox regression model in survival analysis.[11] Suppose $(T_i, \delta_i), i = 1, \ldots, n$, where T_i is the observed time-to-event outcome for the ith subject, δ_i is the censoring indicator for the event of interest. Let $\mathbf{Z}(t)$ denote the vector of time-varying covariates and $\mathbf{X}$ the vector of other time-invariant covariates. The Cox regression model incorporating the time-varying covariates for modelling the hazard of the event occurring at time t, $\lambda(t)$, is given by:

$$\lambda(t) = \lambda_0(t)\exp\{\boldsymbol{\beta}'\mathbf{Z}(t) + \boldsymbol{\beta}_0'\mathbf{X}\}$$

where $\lambda_0(t)$ is an unspecified baseline hazard function for reference subjects; $\boldsymbol{\beta}$ is the vector of effect sizes of the time-varying covariates and its values represent weighted averages of log-transformed hazard ratios (HR); $\boldsymbol{\beta}_0$ is the vector of effect magnitudes of the time-invariant covariates.

Results

Illustrative example

The time-varying covariate approach was applied to analyse breast-feeding duration data from Sichuan Province, China.[12] This longitudinal study (n = 695) was conducted between 2010 and 2011 with a 12-month follow-up after childbirth. A face-to-face interview was carried out at discharge and subsequent telephone interviews were conducted at 1, 3, and 6 months postpartum to solicit details of infant feeding practices and breast-feeding problems experienced by mothers. At 12 months postpartum, only information of breast-feeding status was collected through a brief telephone interview. The breast-feeding duration was measured in days, which refers to the time length between breast-feeding initiation and breast-feeding cessation. A total of 665 women who

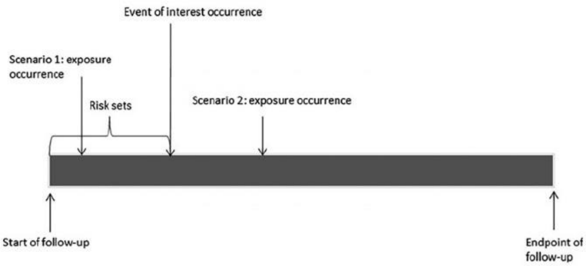


Figure 1. Occurrence of time-dependent exposure: (1) before event of interest occurrence (2) after event of interest occurrence.

Appendix F

initiated breast feeding after delivery were included in the analysis. As shown in Figure 2, both scenarios of Figure 1 occurred for the exposures 'solid foods introduction' and 'maternal return to work'.

Two Cox regression models, in which the outcome variable was breast-feeding duration, were fitted to the data. Three significant baseline covariates (maternal age, place of delivery, and intended breast-feeding duration) identified and published previously by the original study were included in the two Cox regression models,[12] but 'solid foods introduction' and 'maternal return to work' were treated as time-fixed baseline variables in Model 1 and time-varying variables in Model 2, respectively.

In Model 1, 'solid foods introduction' was coded as '0' if no solid foods were introduced over the entire 6-month follow-up period, and '1' if solid foods were introduced at any time during the follow-up. Similarly, 'maternal return to work' was coded as '0' or '1' depending on whether a mother returned to work during the follow-up.

For time-varying variables, data on 'solid foods introduction' was collected at 1, 3, and 6 months postpartum to measure the time mothers introduced solid foods into feeding since the last interview. In Model 2, 'solid foods introduction' was coded as '0' from the start of follow-up and coded as '1' at the first instance when solid food was introduced, on the condition that mothers were breast feeding their babies at all times since initiation. Likewise, 'maternal return to work' after birth was coded as a binary variable based on the information collected at three time points (namely, 1, 3, and 6 months postpartum), with values updated at each interview time point. Statistical analysis was accomplished using SAS 9.4 (SAS Institute Inc., Cary, NC, USA) and the corresponding programming statements are presented in the Appendix S1 SAS example codes.

Table 1 summarizes the results from both models. According to Model 1, 'solid foods introduction' by 6 months postpartum appeared to be 'protective' against breast-feeding cessation by 12 months, whereas 'maternal return to work' by 6 months was not associated with the breast-feeding duration. On the contrary, Model 2 indicated that both time-varying variables increased the hazard of discontinuing breast feeding for the mothers by 12 months postpartum, adjusting for other time-invariant covariates (Figures S1-S4 survival curves from the time-fixed analysis and the time-varying analysis). Moreover, from the perspective of goodness-of-fit, Model 2 (Akaike Information Criterion (AIC) = 4527.27) performed better than Model 1 (AIC = 4932.00).

Comment

Time-dependent bias has been a subject of substantial controversy and debate since the famous Stanford heart transplant study, in which the misclassification of immortal time as exposed to heart

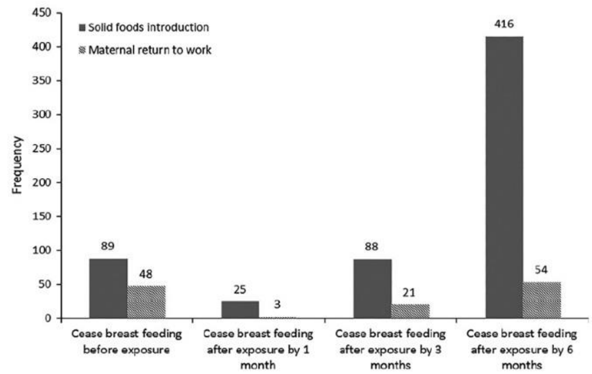


Figure 2. Frequency distributions of women exposed to solid foods introduction and maternal return to work within different time windows.

Appendix F

Table 1. Comparison of effect size estimates between the time-fixed and the time-varying analyses

	Hazard ratio (95% confidence interval)	
	Model 1 (Time-fixed analysis)	Model 2 (Time-varying analysis)
Solid foods introduction by 6 months postpartum	0.61 (0.50, 0.75)	1.50 (1.17, 1.93)
Maternal return to work by 6 months postpartum	0.99 (0.73, 1.36)	1.45 (1.06, 2.00)

Models adjusted for time-fixed covariates: maternal age, place of delivery, and intended breast-feeding duration.

transplantation was rectified using a time-varying survival modelling strategy.[13] In recent years, it has been shown that time-dependent bias would inevitably lead to biased estimation of effect size.[14] The time-varying modelling approach generally yields unbiased estimates, whereas the time-fixed analysis may produce biased estimates with a smaller magnitude.[3,14]

We specifically focus on the appropriate approach in the survival analysis context. To avoid time-dependent bias in paediatric studies, the time-varying covariate approach is recommended for survival analysis of time-to-event data in the presence of time-varying exposures.

Acknowledgements

The first author is grateful to the China Scholarship Council which supported this study.

References

1 Austin PC, Platt RW. Survivor treatment bias, treatment selection bias, and propensity scores in observational research. *Journal of Clinical Epidemiology* 2010; 63:136–138.

2 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.

3 van Walraven C, Davis D, Forster AJ, Wells GA. Time-dependent bias was common in survival analyses published in leading clinical journals. *Journal of Clinical Epidemiology* 2004; 57:672–682.

4 Austin PC, Mamdani MM, van Walraven C, Tu JV. Quantifying the impact of survivor treatment bias in observational studies. *Journal of Evaluation in Clinical Practice* 2006; 12:601–612.

5 Scott JA, Binns CW, Oddy WH, Graham KI. Predictors of breastfeeding duration: evidence from a cohort study. *Pediatrics* 2006; 117:e646–e655.

6 Rousseau MC, El-Zein M, Conus F, Legault L, Parent ME. Bacillus Calmette-Guerin (BCG) Vaccination in Infancy and Risk of Childhood Diabetes. *Paediatric and Perinatal Epidemiology* 2016; 30:141–148.

7 Whitehead E, Dodds L, Joseph KS, Gordon KE, Wood E, Allen AC, *et al.* Relation of pregnancy and neonatal factors to subsequent development of childhood epilepsy: a population-based cohort study. *Pediatrics* 2006; 117:1298–1306.

8 Wolkewitz M, Allignol A, Harbarth S, de Angelis G, Schumacher M, Beyersmann J. Time-dependent study entries and exposures in cohort studies can easily be sources of different and avoidable types of bias. *Journal of Clinical Epidemiology* 2012; 65:1171–1180.

9 Beyersmann J, Gastmeier P, Wolkewitz M, Schumacher M. An easy mathematical proof showed that time-dependent bias inevitably leads to biased effect estimation. *Journal of Clinical Epidemiology* 2008; 61:1216–1221.

10 Klein JP, vanHouwelingen HC, Ibrahim JG, Scheike TH. *Handbook of Survival Analysis*. Boca Raton: CRC Press, 2013.

11 Therneau TM, Grambsch PM. *Modeling Survival Data: Extending the Cox Model*. New York: Springer Science & Business Media, 2000.

12 Tang L, Lee AH, Binns CW. Factors associated with breastfeeding duration: a prospective cohort study in Sichuan Province, China. *World Journal of Pediatrics* 2015; 11:232–238.

13 Mantel N, Byar DP. Evaluation of response-time data involving transient states - illustration using heart-transplant data. *Journal of the American Statistical Association* 1974; 69:81–86.

14 Beyersmann J, Wolkewitz M, Schumacher M. The impact of time-dependent bias in proportional hazards modelling. *Statistics in Medicine* 2008; 27:6439–6454.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1. SAS example codes of time-varying modelling.

Appendix S2. Survival curves plotted in R package from the time-fixed approach and the time-varying approach:

Figure S1. Survival curve comparing 'yes' vs. 'no' for 'solid foods introduction' by 6 months postpartum based on the time-fixed analysis.

Figure S2. Survival curve comparing 'yes' vs. 'no' for 'maternal return to work' by 6 months postpartum based on the time-fixed analysis.

Appendix F

Figure S3. Survival curve comparing 'yes' vs. 'no' for 'solid foods introduction' by 6 months postpartum based on the time-varying covariate approach.

Figure S4. Survival curve comparing 'yes' vs. 'no' for 'maternal return to work' by 6 months postpartum based on the time-varying covariate approach.

F.3 SAS example codes of time-varying modelling

```
/*Example 1: solid foods introduction*/
proc phreg data=survival dataset;
class placeofdelivery(ref=last);
model anybftime12*censor12(1)=placeofdelivery foodintro;
 /*foodintro: time-varying covariate; placeofdelivery: time-invariant covariate*/
```

```
if foodintrotime>anybftime12 OR foodintrotime=. then foodintro=0;
else foodintro=1;
/*foodintrotime: a variable measured to indicate the time when mothers introduced
solid foods into feeding*/
run;
/*time-varying covariate only created in the computer memory for computing, not a
real variable in the survival dataset*/

/*Example 2: maternal return to work*/
proc phreg data=survival dataset;
class placeofdelivery(ref=last);
model anybftime12*censor12(1)=placeofdelivery yesreturnwork;
/*yesreturnwork: time-varying covariate; placeofdelivery: time-invariant covariate*/
array time(*)time1-time3;
/*time1-time3: 3 different follow-up time points after birth*/
array yesrework(*) yesrework1-yesrework3;
/*yesrework1-yesrework3: maternal return to work status measured at 3 follow-up
points, respectively; binary variable*/
Do j=1 to 3;
if anybftime12>time[j] AND time[j] NE . then yesreturnwork=yesrework[j];
end;
run;
```

Appendix G for Chapter 6

G.1 Permission from publisher-Wolters Kluwer Health, Inc.

Copyright Clearance Center RightsLink®

Home | Account Info | Help

Wolters Kluwer

Title: Application of a Frailty Modeling Approach to Correlated Breastfeeding Duration Data

Author: Jian Zhao, Yun Zhao, Vishnu Khanal, et al

Publication: Nursing Research

Publisher: Wolters Kluwer Health, Inc.

Date: Aug 2, 2018

Logged in as:
Jian Zhao
Account #:
3001283143
LOGOUT

License Not Required

This request is granted gratis and no formal license is required from Wolters Kluwer. Please note that modifications are not permitted. Please use the following citation format: author(s), title of article, title of journal, volume number, issue number, inclusive pages and website URL to the journal page.

BACK | CLOSE WINDOW

G.2 Publication 5

Application of a Frailty Modeling Approach to Correlated Breastfeeding Duration Data

Jian Zhao, MMed
PhD candidate
School of Public Health, Curtin University
Perth, WA, Australia

Yun Zhao, PhD
Senior Lecturer
School of Public Health, Curtin University
Perth, WA, Australia

Vishnu Khanal, PhD
Nepal Development Society
Chitwan, Nepal

Andy H. Lee, PhD
Professor
School of Public Health, Curtin University
Perth, WA, Australia

Colin W. Binns, PhD
John Curtin Distinguished Emeritus Professor
School of Public Health, Curtin University
Perth, WA, Australia

Appendix G

Jian Zhao, MMed, is PhD candidate, School of Public Health, Curtin University, Perth, WA, Australia.

Yun Zhao, PhD, is Senior Lecturer, School of Public Health, Curtin University, Perth, WA, Australia.

Vishnu Khanal, PhD, Nepal Development Society, Chitwan, Nepal.

Andy H. Lee, PhD, is Professor, School of Public Health, Curtin University, Perth, WA, Australia.

Colin W. Binns, PhD, is John Curtin Distinguished Emeritus Professor, School of Public Health, Curtin University, Perth, WA, Australia.

This study was partially funded by the China Scholarship Council (grant number: 201406240008).

The original study was approved by the Human Research Ethics Committee of Curtin University, Australia (HR 184/2013), and the Nepal Health Research Council, Nepal (773/2014). The present study utilized only de-identified data of the original study.

The authors have no conflicts of interest to report.

Corresponding author: Yun Zhao, PhD, School of Public Health, Curtin University, GPO Box U 1987, Perth, WA, Australia, 6845 (e-mail: y.zhao@curtin.edu.au).

Appendix G

Abstract

Background

Correlated breastfeeding duration data are very common in infant feeding research using cohort designs. Intracluster correlation within the same clustering group is expected and needs to be taken into account in statistical analysis, otherwise, the corresponding statistical inferences may be subject to an increased type I error.

Objectives

The aims of this study were to illustrate the necessity of adjusting for the intracluster correlation in correlated breastfeeding duration data analysis and demonstrate different frailty modeling approaches.

Methods

An introduction to shared frailty models was presented under the assumption of proportional hazards (PH). Then, two different approaches—the Cox frailty model (semiparametric approach) and the parametric frailty model (parametric approach)—were used to fit the data from a maternal cohort in Nepal as an illustrative example.

Results

For the semiparametric approach, random effects denoting the variations in the hazard of breastfeeding cessation shared by mothers living in the 27 distinct communities were estimated and graphically presented. Compared with the conventional Cox model, Cox frailty model reduced the chance of type I error occurring, providing a better model fit in the presence of correlated survival data. Among candidate parametric approaches, a Weibull PH model with a gamma frailty term was selected as an appropriate model fitting the breastfeeding data.

Discussion

Shared frailty models can be used in other research areas in the presence of correlated time-to-event data. Model selection depends on the assumption of PH, the specification of the baseline hazard function, and also the study purpose.

Key words

Breastfeeding; clustered samples; frailty model; intra-cluster correlation; time-to-event data

Appendix G

In view of the widely recognized benefits of breastfeeding for both infants and mothers (Danforth et al., 2007; Duijts, Jaddoe, Hofman, & Moll, 2010; Ip et al., 2007; Victora et al., 2016), an increasing number of studies have been undertaken to identify factors associated with breastfeeding duration. In many of these studies, subjects are clustered (i.e., nested) in groups (i.e., clustering units) such as communities, hospitals, or health service centers. It is known that subjects who are nested within the same groups are likely to have outcomes (e.g., breastfeeding duration) correlated with one another (Goldstein, 2011; Raudenbush & Bryk, 2002). This within-community correlation, or homogeneity, is often introduced by unmeasured or unmeasurable cluster characteristics (e.g., community culture, hospital facility policies, or health service resources) that need to be taken into account in analysis, otherwise, the corresponding statistical inferences may be subject to underestimated standard error, leading to overstatement of statistical significance and an increased type I error, and possibly resulting in incorrect conclusions (Campbell & Grimshaw, 1998; Goldstein, 2011; Raudenbush & Bryk, 2002; Zyzanski, Flocke, & Dickinson, 2004). However, in analyzing their data, researchers often ignore these inherent correlations by assuming observations to be independent of each other, and using conventional survival analysis techniques (e.g., Cox proportional hazards [PH] regression model) for modeling correlated breastfeeding duration (i.e., correlated time to breastfeeding cessation). We examined the Medline database for breastfeeding duration study papers published from January 1997 to December 2017. In the past two decades, of the total 28 breastfeeding duration study articles where an inherent correlation was potentially present due to clustered structure, none of them adjusted for the inherent correlation (or the heterogeneity between clustering units) in their analyses.

In the recent literature, to address the correlation (i.e., dependence) between observations within the same groups (i.e., clustering units), mixed-effects models—which are also known as multilevel models or hierarchical models (Goldstein, 2011; Raudenbush & Bryk, 2002)—have been increasingly used. In these mixed-effects models, the heterogeneity between clustering units is modeled by random effects (Cheng & Kelly, 2011; Goldstein, 2011; Raudenbush & Bryk, 2002; Shin, 2009). When the outcomes are correlated (clustered) time-to-event data, a shared frailty survival model has been developed (Duchateau & Janssen, 2008; Hougaard, 2000). In such a shared frailty survival model, a frailty term—which is an exponentially transformed random effect—is used to account for the correlation shared by the observations in one clustering unit, or the unobserved heterogeneity between clustering units. Statistically, the shared frailty model can be semiparametric (e.g., Cox frailty model) or parametric (e.g., parametric frailty models), depending on whether the distribution of baseline hazard function can be specified in advance. Nardi and Schemper (2003) have pointed out that compared with semiparametric models, parametric models are more powerful with more efficient parameter estimations. However, not only the application of the shared frailty models, both semiparametric and parametric, in modeling correlated breastfeeding duration data is limited, but also the discussions or comparisons between semiparametric and parametric frailty models are not commonly found in the current literature.

The objective of this paper was to discuss the underlying issues of the intracluster correlation in correlated breastfeeding duration data analysis and to demonstrate different frailty modeling approaches by applying the shared frailty model to fit correlated breastfeeding duration data from a maternal cohort in Nepal as an illustrative example.

Appendix G

Methods

Shared Frailty Models

Without loss of generality, in this paper we only discuss survival analysis models under the assumption of PH. As shown in Figure 1, based on the specification of baseline hazard function and the independence of observations, survival analysis models can be classified into different types of specific models. A shared frailty model, in which a frailty term such as an unobservable random effect is incorporated as a multiplicative term in the hazard function of the survival model, is an extension of a conventional survival model. The shared frailty term describes the characteristics shared by subjects clustered in one specific group. For example, mothers in a specific community share the same frailty and tend to fail (e.g., cease breastfeeding) at a similar rate. A general form of a PH frailty model can be given as below,

$$\lambda_{ij}(t \mid \mathbf{X}_{ij}, u_i) = \lambda_0(t)\exp\{\boldsymbol{\beta}'\mathbf{X}_{ij} + u_i\}, \quad (1)$$

where $\lambda_{ij}(t \mid \mathbf{X}_{ij}, u_i)$ is the conditional hazard function at time t for the j^{th} subject nested in the i^{th} clustering unit with a set of covariates $\mathbf{X}_{ij}$; u_i indicates the random effect of clustering unit i, the distributions of the frailty term (exponentially transformed u_i) can be gamma, log-normal, inverse Gaussian and compound Poisson, among which gamma and log-normal distributions are the two most commonly used distributions due to their computational simplicity (Duchateau & Janssen, 2008; Hougaard, 2000); $\lambda_0(t)$ denotes the baseline hazard function. If $\lambda_0(t)$ is unspecified, (1) is a Cox frailty model, otherwise, it is a parametric PH frailty model with a baseline hazard distribution that can be specified in advance (Duchateau & Janssen, 2008).

Empirical Study

The data used in the present study were obtained from a community-based maternal cohort study conducted in Nepal between January and October 2014. Details of the study design, sampling frame, and setting have been published previously (Khanal, Lee, Karkee, & Binns, 2015). A total of 735 mothers from 27 randomly selected communities (15 village development committees and 12 wards of two municipalities) were recruited; mothers were interviewed shortly after giving birth. Excluding 86 mothers who delivered at home, 649 mothers were included in the final sample to determine the factors associated with exclusive breastfeeding duration. The study outcome, exclusive breastfeeding duration, was defined as the time to exclusive breastfeeding cessation since breastfeeding initiation. Applying a conventional Cox proportional hazard model, Khanal and colleagues (2015) found that four covariates were associated with exclusive breastfeeding duration: mode of delivery, wealth status, two variables on breastfeeding promotion advice, and breastfeeding on demand and not to provide pacifier or teats. The original study was approved by the Human Research Ethics Committee of Curtin University, Australia (HR 184/2013), and the Nepal Health Research Council, Nepal (773/2014). The present study used only de-identified data from the original study.

Statistical Analysis

The Nepal breastfeeding duration data were fitted by Cox frailty model and parametric frailty model, respectively. A gamma frailty term was used in both modeling approaches.

Semiparametric approach. In this approach, a gamma frailty term, or exponentially transformed random effect, was incorporated into the previous published Cox proportional hazard model (Khanal et al., 2015) to accommodate the heterogeneity between communities.

A likelihood ratio test was used to compare the goodness of fit between the conventional Cox PH model and the Cox frailty model.

Parametric approach. We started this parametric approach by identifying an appropriate distribution for the baseline hazard function $\lambda_0(t)$. Three candidate baseline hazard distributions were evaluated: Weibull, exponential, and log-logistic. Their Akaike information criterion (AIC) values were calculated and compared to assess the goodness-of-fit between the three different parametric models. Then the model with the minimal AIC was fitted to the data with a gamma frailty term. The exact same four significant covariates published in the previous results (Khanal et al., 2015) were included in this parametric model. SAS version 9.4 (SAS Institute, Cary NC) and Stata version 14.2 (Stata Corp, College Station, TX) were used for statistical modeling. Specifically, the *streg* Stata package was used to fit the parametric model with a gamma frailty (Gutierrez, 2002).

Results

Semiparametric Approach

Modeling outcomes of Cox frailty model with a gamma frailty term. After incorporating the gamma-distributed frailty term in the Cox proportional hazards model, the random effect of each community was estimated. As shown in Figure 2, these random effects denote the variations in the hazard of breastfeeding cessation shared by mothers living in the 27 distinct communities. For instance, the random effect of the third community is 0.66, which could be interpreted that mothers living in this community were 1.93 (exp (0.66) = 1.93) times more likely to cease exclusive breastfeeding compared with the average hazard of ceasing exclusive breastfeeding for all communities. Mothers living in the 24th community with a random effect of -0.49 were 39% less likely to cease exclusive breastfeeding. As shown in Table 1, the heterogeneity between communities was confirmed to be statistically significant (p < .001). Two significant effect sizes revealed by the Cox PH model, wealth status, and breastfeeding on demand (p = .003 and p = .006, respectively) changed to nonsignificant (p = .293 and p =. 339, respectively) by the Cox frailty model as its frailty term explains a substantial proportion of the total variation in the exclusive breastfeeding duration (Table 1). These two underestimated p values obtained from the Cox PH model would lead to an overstatement of the effect of the two variables on the exclusive breastfeeding duration. The other two covariates behaved in a way consistent with the results of Khanal and colleague's study (2013), cesarean delivery increased and not to provide pacifier or teats to infants decreased the hazard of breastfeeding cessation. It is also noted that the corresponding p value of not to provide pacifier or teats to infants is underestimated in the Cox PH model. As shown in Table 1, standard errors are all smaller (underestimated) in the Cox PH model, leading to narrower confidence intervals. The likelihood ratio test confirmed that the Cox frailty model taking account of the heterogeneity between communities provided a better model fit to the data than the conventional Cox PH model (likelihood ratio test statistic = 70.542, df = 1, p < .001).

Parametric Approach

The outcome of identifying an appropriate baseline hazard distribution. Weibull, exponential, and log-logistic parametric hazard models were fitted to the breastfeeding duration data without covariates; AICs for the three different distributions were calculated as 2412.725, 2458.511, and 2589.088, respectively. It is obvious that the Weibull parametric model had the minimal AIC, indicating Weibull was a preferred baseline hazard distribution for the next step parametric modeling incorporating a frailty term.

Modeling outcomes of Weibull PH frailty models with a gamma frailty term. A Weibull PH model with a gamma frailty term was then fitted to the breastfeeding duration data (Table 1). Comparing results from the semi-parametric (Cox frailty model) and the parametric model (Weibull PH frailty model), one covariate, not to provide pacifier or teats, that was statistically significant in the Cox frailty model (p = .041) turned to nonsignificant in the parametric model (p = .081). Practically, the estimated hazard ratios obtained based on the Weibull PH frailty model can be interpreted similarly to that in the Cox frailty model for comparing relative risk of ceasing exclusive breastfeeding between two groups, vaginal birth versus cesarean birth. However, the goodness-of-fit of the parametric approach cannot be directly compared with the semiparametric one due to different likelihood function structures (Kalbfleisch & Prentice, 2002).

Discussion

In this study, we present a methodological framework to analyze correlated breastfeeding duration data by using data from a maternal cohort in Nepal as an example and provide some insights into using Cox and parametric PH frailty models to analyze correlated breastfeeding duration data. Due to its easy and convenient implementation by most statistical software, the conventional Cox model is the most commonly used approach in breastfeeding duration research—even in studies where observations are clustered. It has been indicated that using conventional statistical models, which often ignore the dependence between outcome observations, increases the risk of making false-positive conclusion, inflating the Type I error in analyzing clustered data (Aarts, Verhage, Veenvliet, Dolan, & van der Sluis, 2014; Zyzanski et al., 2004). Our example of Nepal clustered breastfeeding duration data showed support for Cox and parametric PH frailty models by showing that some previously published significant covariates were nonsignificant after frailty models were fitted. In addition, it has been reported that once an appropriate distribution of baseline hazard is confirmed, parametric frailty models could provide more efficient and powerful parameter estimation (Nardi & Schemper, 2003; Oakes, 1977).

In the present illustration, we assumed the assumption of PH is met; however, in some practical applications, this PH assumption may not be satisfied. Accelerated failure time (AFT) model is an alternative modeling approach when the PH assumption does not hold. Unlike PH models which estimate the effect of covariates on the hazard of an event occurring, such as the cessation of exclusive breastfeeding, AFT models quantify the effect of covariates on the logarithm of survival time (e.g., the duration of exclusive breastfeeding). The interpretation of AFT parametric frailty models is more straightforward for a study where survival time is investigated directly. Therefore, in analyzing correlated breastfeeding duration data, researchers could choose which frailty model to use according to whether the PH can be assumed, whether the baseline hazard function can be specified in advance, and whether the study aims to compare hazard or survival time between groups.

The present framework can be extended to other research when subjects are nested in the same sampling unit, such as hospitals, communities, or health service centers, when correlated time-to-event or survival data are collected, or when recurrent outcomes, such as recurrent admissions of patients with cardiovascular diseases or chronic obstructive pulmonary disease, are measured, and where random effects are subject-specific. The shared frailty survival modeling, including semiparametric or parametric, PH, or AFT, for fitting correlated survival data can be readily used by invoking some user-friendly, built-in code functions in a variety of statistical software packages, including Stata, SAS, and R.

Conclusion

To appropriately analyze correlated breastfeeding duration data or other correlated time-to-event data and hence to ensure reliable and reproducible results, a shared frailty survival model is recommended. The model fitting can be implemented by standard statistical software; the choice of which frailty model to use depends on the assumption of PH, the specification of the baseline hazard function, and the study purpose.

References

Aarts, E., Verhage, M., Veenvliet, J. V., Dolan, C. V., & van der Sluis, S. (2014). A solution to dependency: Using multilevel analysis to accommodate nested data. *Nature Neuroscience, 17,* 491-496. doi:10.1038/nn.3648

Campbell, M. K., & Grimshaw, J. M. (1998). Cluster randomised trials: Time for improvement: The implications of adopting a cluster design are still largely being ignored. *BMJ, 317,* 1171-1172.

Cheng, A.-L., & Kelly, P. J. (2011). Application of a hierarchical model incorporating intrafamily correlation and cluster effects. *Nursing Research, 60,* 208-212. doi:10.1097/NNR.0b013e31820a3dbe

Danforth, K. N., Tworoger, S. S., Hecht, J. L., Rosner, B. A., Colditz, G. A., & Hankinson, S. E. (2007). Breastfeeding and risk of ovarian cancer in two prospective cohorts. *Cancer Causes & Control, 18,* 517-523. doi:10.1007/s10552-007-0130-2

Duchateau, L., & Janssen, P. (2008). *The frailty model*. New York, NY: Springer.

Duijts, L., Jaddoe, V. W. V., Hofman, A., & Moll, H. A. (2010). Prolonged and exclusive breastfeeding reduces the risk of infectious diseases in infancy. *Pediatrics,* peds-2008. doi:10.1542/peds.2008-3256

Goldstein, H. (2011). *Multilevel statistical models* (4th ed.). Chichester, UK: Wiley.

Gutierrez, R. G. (2002). Parametric frailty and shared frailty survival models. *Stata Journal, 2,* 22-44.

Hougaard, P. (2000). *Analysis of multivariate survival data*. New York, NY: Springer.

Ip, S., Chung, M., Raman, G., Chew, P., Magula, N., DeVine, D., . . . Lau, J. (2007). Breastfeeding and maternal and infant health outcomes in developed countries. *Evidence*

Report/Technology Assessment, 153, 1-186. AHRQ Publication No. 07-E007. Rockville, MD: Agency for Healthcare Research and Quality

Kalbfleisch, J. D., & Prentice, R. L. (2002). *The statistical analysis of failure time data* (2nd ed.). Hoboken, N.J.: Wiley.

Khanal, V., Lee, A. H., Karkee, R., & Binns, C. W. (2015). Postpartum breastfeeding promotion and duration of exclusive breastfeeding in western Nepal. *Birth, 42,* 329-336. doi:10.1111/birt.12184

Nardi, A., & Schemper, M. (2003). Comparing Cox and parametric models in clinical studies. *Statistics in Medicine, 22,* 3597-3610. doi:10.1002/sim.1592

Oakes, D. (1977). The asymptotic information in censored survival data. *Biometrika, 64,* 441-448. doi:10.2307/2345319

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Shin, J. H. (2009). Application of repeated-measures analysis of variance and hierarchical linear model in nursing research. *Nursing Research, 58,* 211-217. doi:10.1097/NNR.0b013e318199b5ae

Victora, C. G., Bahl, R., Barros, A. J., França, G. V., Horton, S., Krasevec, J., . . . Rollins, N. C. (2016). Breastfeeding in the 21st century: Epidemiology, mechanisms, and lifelong effect. *Lancet, 387,* 475-490. doi:10.1016/S0140-6736(15)01024-7

Zyzanski, S. J., Flocke, S. A., & Dickinson, L. M. (2004). On the nature and analysis of clustered data. *Annals of Family Medicine, 2,* 199-200. doi:10.1370/afm.197

Appendix G

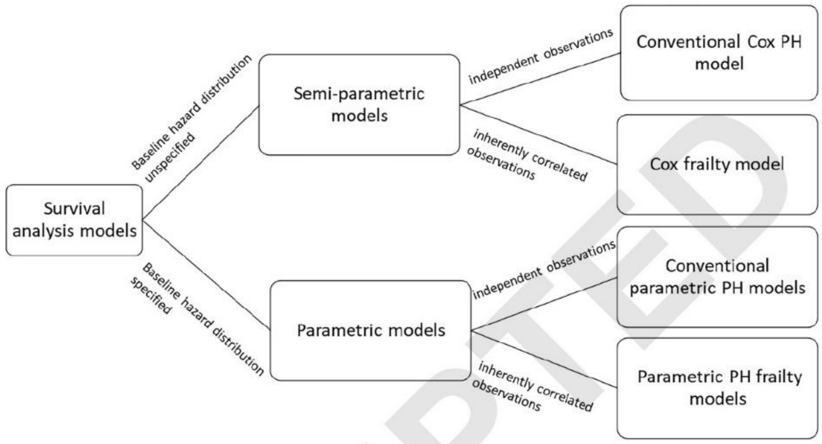
Legends for figures:

Figure 1 Survival analysis models by (a) specification of baseline hazard function and (b) independence of observations

Figure 2 Estimated random effect in each community for the breastfeeding duration data as given by the Cox frailty model

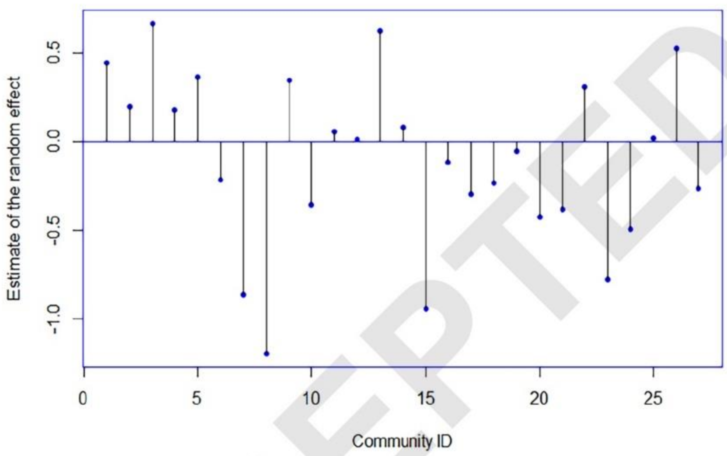
Appendix G

Figure 1



<u>Appendix G</u>

Figure 2



Appendix G

Table 1

Results of Conventional Cox PH Model, Cox Frailty Model, and Weibull PH Frailty Model

Variable	Cox PH model (previously published)			Cox frailty model			Weibull PH frailty model		
	HR [95% CI]	*SE*	*p*	*HR* [95% CI]	*SE*	*p*	*HR* [95% CI]	*SE*	*p*
Mode of delivery									
Vaginal (ref)	1.00		<.001	1.00		<.001	1.00		<.001
Cesarean	1.60 [1.27, 2.01]	.118		1.81 [1.42, 2.30]	.123		2.31 [1.82, 2.92]	.277	
Wealth status									
Poor (ref)	1.00		.003	1.00		.293	1.00		.372
Middle	1.39 [1.15, 1.69]	.097		1.16 [0.92, 1.46]	.119		1.16 [0.92, 1.46]	.137	
Rich	1.17 [0.93, 1.47]	.117		.99 [0.75, 1.32]	.145		1.02 [0.76, 1.36]	.150	
Breastfeeding promotion advice: breastfeeding on demand									
No (ref)	1.00		.006	1.00		.339	1.00		.191
Yes	.74 [0.59, 0.92]	.116		.89 [0.71, 1.13]	.120		0.86 [0.68, 1.08]	.101	
Breastfeeding promotion advice: not to provide pacifier or teats									
No (ref)	1.00		.026	1.00		.041	1.00		.081
Yes	.82 [0.68, 0.97]	.096		.80 [0.64, 0.99]	.110		0.83 [0.68, 1.02]	.087	
Variance component	-	-	-	Estimate	*SE*	*p*	Estimate	*SE*	*p*
	-	-		0.254	.086	<.001	.200	.072	<.001
-2 log likelihood	6318.306			6247.764			2206.659		

Note. CI = confidence Interval; HR = hazard ratio; ref = reference; *SE* = standard error.

G.3 SAS codes for two-part modelling

```
proc nlmixed data= tpm;
bounds gamma>0;
***Logistic component***;
eta1=betal_0+betal_1*birthmode+betal_2*grandmapref+betal_3*bonding+ranl;/*ranl is the random effect in the logistic part*/
p=exp(eta1)/(1+exp(eta1));
if ebfbase=0 then loglik=log(1-p);/*log likelihood for the first Logistic regression part*/
***Survival component***;
if ebfbase=1 then do;
eta2=betas_0+betas_1*birthmode+betas_2*grandmapref+betas_3*bonding+rans;/*rans is the random effect in the survival part*/
alpha=exp(-eta2);
```

```
loglik=log(p)-(alpha*ebfduration)**gamma+(censor=0)*(-gamma*eta2+(gamma-
1)*log(ebfduration)+log(gamma));/*log likelihood for the Weibull survival part
censor=1 indicates censored observation*/
end;
model ebfduration~general(loglik);
random ranl rans ~ normal([0,0],[exp(2*logsigl),cov_l_s,exp(2*logsigs)])
subject=CommunityCode;
estimate 'correlation coeffecient(ranl_rans_rho)' cov_l_s/(exp(logsigl)*exp(logsigs));
estimate 'variancel' exp(2*logsigl);
estimate 'variances' exp(2*logsigs);
run;
```