Mediation analysis in longitudinal studies in the presence of measurement error and missing data

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

John Mbaziira Ssenkusu

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosopy

Advised by David M. Vock, PhD

May, 2018

Acknowledgments

I am grateful to David M. Vock, my adviser, for his training during graduate school and guidance throughout my dissertation. Thank you to James S. Hodges, a committee member and a mentor in collaborative research. I am grateful to my committee members, Julian Wolfson and Sarah Cusick, for reviewing my thesis and providing feedback that greatly improved my dissertation.

To illustrate the perfomance of statistical methods on real data, we used severe malaria data and HIV-LIVE (Longitudinal Interrelationships of Viruses and Ethanol) data. Collection and analysis of the severe malaria and cognition data were supported by the National Institute of Neurologic Disorders and Stroke (grant: R01NS055349). The malaria studies in Chapter 4 were supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (grant: P01 HD059454 and 1R01HD086124-01). I also acknowledge NIAAA grants R01-AA13216, U24AA020778, U24AA020779 and NIDA grant R25DA013582 that support the data repository for HIV-LIVE data. Thanks to Daowen Zhang for providing the software for the seminonparametric (SNP) approach.

Thank you to the Fogarty International Centre for funding my graduate school through the D43 training grant (grant: D43 NS078280). Thank you to Molly McCoy, the grant administrator, and to Shailey Prasad and Debra Olson, who were site (University of Minnesota) principal investigators for the training

grant. Finally, I am grateful to Chandy C. John, a mentor and the principal investigator of the D43 training grant.

Dedication

To my wife Susan Ssenkusu Hope, children Mariana Talikoma and Joseph Balungi.

Abstract

Mediation analysis hypothesizes that an exposure causes a mediator and in turn the mediator causes the outcome, so mediation is inherently longitudinal. Unfortunately, potential mediators may be measured with error and regression estimators obtained by ignoring measurement error can be severely biased. This can induce bias in the estimation of causal direct and indirect effects. In Chapter 2, using regression calibration, we show how to adjust for measurement error in longitudinal studies with repeated measurements of the mediator, and evaluate the effect of ignoring measurement error on direct and indirect effects. Rather than assuming normality for the random effects in the linear mixed effects calibration model, we correct for measurement error in the mediator allowing flexibility in the distribution of subject-specific random effects. On the other hand, longitudinal studies face challenges of missing data resulting from loss to follow-up, death, or withdrawal. In mediation analysis, multiple imputation has been shown to perform well for data missing completely at random (MCAR) and missing at random (MAR) in cross-sectional studies, but it is unclear how it performs in longitudinal studies under misspecification of the imputation model, specifically, where the misspecification ignores clustering by subject. In Chapter 3, we examine the impact of ignoring clustering on mediated effect estimates under MCAR and MAR mechanisms with varying degrees of missingness. In Chapter 4, using data from a randomized controlled trial, we examine the mediation effects on child neurodevelopment of intermittent preventive malaria treatment in pregnant women. Chapter 5 concludes and discusses future work.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Many disciplines are interested in mediation, in which a third variable (mediator) M partly describes the effect of an exposure X on an outcome Y (Figure 1.1). In a cross-sectional study, assume X_i is binary, and M_i and Y_i are con-

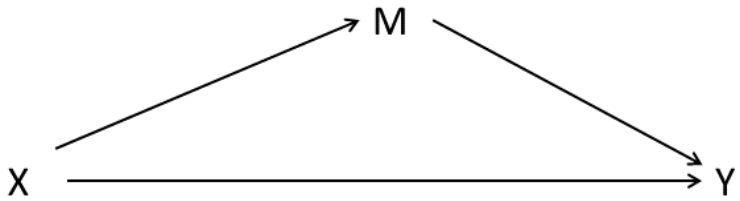

Figure 1.1 Mediation directed acyclic graph (DAG).

tinuous variables for subject i. Direct and indirect (mediated) effects can be parameterized using the linear models

$$\mathbb{E}(M_i|X_i) = \zeta_0 + \zeta_x X_i, \tag{1.1}$$

$$\mathbb{E}(Y_i|X_i, M_i) = \gamma_0 + \gamma_x X_i + \gamma_m M_i. \tag{1.2}$$

In the absence of an exposure-mediator interaction in model (1.2), the direct effect is given by γ_x and the indirect effect by $\zeta_x \gamma_m$. The direct effect measures the exposure effect on the outcome after adjusting for the mediator while the indirect effect measures the change in the outcome as a result of the exposure's effect on the mediator.

Unfortunately, potential mediators may be measured with error and regression estimators obtained by ignoring measurement error can be severely biased (Valeri et al., 2014). This can induce bias in the estimators of causal direct and indirect effects. Regression calibration is one method widely used to correct for measurement error (Brown and Fuller, 1990; Carroll et al., 2006) and may be useful when both the mediator and outcome are repeatedly measured. However, when the outcome is repeatedly measured, as is usually the case in longitudinal studies, this method generally requires its user to assume a distribution for the random effects and inferences may be sensitive to misspecification of this distribution.

In addition, longitudinal studies are prone to missing data due to missing visits, withdrawals, lost to follow up or death. When the mediator or outcome are missing, estimators can be biased and the missing data can also lead to loss of information, decreased statistical power and weakened generalizability of findings (Dong and Peng, 2013). Biased parameter estimators can induce bias in the estimation of direct and indirect effects. In mediation analysis, the mediator is both an outcome and a covariate, so any proposed method needs to consider the mediator in both capacities. Multiple imputation (MI), a highly recommended method for dealing with missing data, can be implemented in current commercial software. However, all such software ignore correlation among observations on the same unit (Mistler, 2013), which is characteristic of longitudinal studies.

In Chapter 2, we describe the potential outcome framework for estimating direct and indirect effects when the mediator is measured with error, and propose a method to adjust for measurement error in a repeatedly measured mediator. In Chapter 3, we examine the impact of ignoring clustering in MI for longitudinal data on estimation of direct and indirect effects.

The methodological development in Chapters 2 and 3 is motivated by a prospective cohort study that examined the effects of cerebral malaria (CM), a severe form of malaria, on neurodevelopment among chidren < 5 years. Malaria is a disease caused mainly by *Plasmodium falciparum* or *Plasmodium vivax* parasites. These parasites are spread from person to person through bites of an infected female Anopheles mosquito. Globally, most malaria cases (90%) and deaths (91%) occur in sub-Saharan Africa (WHO, 2017) and the groups at most risk of malaria include infants, children under 5 years of age, and pregnant women (Brabin, 1983; Rogerson et al., 2007). Pregnant women are more susceptible to malaria infection than non-pregnant women (Brabin, 1983; Rogerson et al., 2007). Malaria infection in pregnancy can lead to placental malaria (PM) (Rogerson et al., 2007) and PM has been associated with low birth weight and preterm birth (Desai et al., 2007; Kapisi et al., 2017; Katz et al., 2013; Rogerson et al., 2007, 2003). Maternal immune response and placental changes (Muehlenbachs et al., 2006) that occur with malaria illness in pregnancy may interrupt maternal to fetus transfer of nutrients, and this could adversely affect the developing fetal brain, and lead to deficits in child neurodevelopment (ND). To protect pregnant women against malaria and its effects on the unborn children, the World Health Organization (WHO) recommends intermittent preventive treatment in pregnancy (IPTp) with sulfadoxine-pyrimethamine (SP) (WHO, 2017). However, some studies in East Africa have shown a waning effect of SP (Bigira et al., 2014; Iriemenam et al., 2012; Ndyomugyenyi et al., 2011) and proposed IPTp with dihydroartenisinin-piperaquine (DP) (Nankabirwa et al., 2016, 2014; Tarning et al., 2008). We hypothesize that the protective effect of IPTp with DP causes placental changes that reduce effects of malaria in pregnancy such as placental malaria, preterm birth, low birth weight, and low levels of hemoglobin levels, thus improving ND in the offspring. Therefore, in Chapter 4, we analyze data from a randomized controlled trial to estimate the mediated effects of

IPTp on child ND through maternal malaria, placental malaria, adverse birth outcomes (preterm birth, low birth weight, and small for gestation age), cord blood hemoglobin or child malaria.

Chapter 2

Mediation analysis for longitudinal data using regression calibration when the mediator is measured with error

2.1 Introduction

In many disciplines, it is important to understand how the effect of an exposure on an outcome is mediated through other variables. Unfortunately, potential mediators may be measured with error. The influence of measurement error in a continuous mediator on the total effects (direct and indirect effects) in the context of generalized linear models has been studied previously (Valeri et al., 2014). Since measurement error in the exposure variable (X), mediator (M) and outcome (Y) have different effects on estimates of the paths $X \rightarrow Y$, $X \rightarrow M$, and $M \rightarrow Y$ (Cole and Maxwell, 2003), regression estimators obtained by ignoring measurement error can be severely biased and thus induce bias in the estimation of causal and indirect effects (Valeri et al., 2014). However, much of the previous work concerning measurement error in mediation

analysis has assumed that the mediator and outcome are collected at a single time point.

A motivating example in this paper is a study among children of the effect of cerebral malaria on cognitive development mediated by hemoglobin levels (Cusick et al., 2016). Cerebral malaria (exposure) affects both hemoglobin levels (mediator) and cognitive ability (outcome), and hemoglobin levels affect cognitive ability in children (Boivin et al., 2016; John et al., 2008). In this longitudinal study, hemoglobin and cognitive ability were repeatedly measured over one year. To account for correlations among observations when there are repeated measurements, a routine framework for longitudinal data analysis is the linear mixed effects model (LMM). In mediation analysis, LMMs have been found to perform well relative to structural equation models (SEMs), a commonly used technique with continuous data (Blood et al., 2010), and are robust to violations of the normality assumption for the errors (Blood and Cheng, 2012). Even in the context of non-linear models, non-linear mixed models (NLMMs) have been found to perform sufficiently well in the analysis of mediated longitudinal binary data with respect to bias, coverage probability, and power (Blood et al., 2010). Though LMMs can be applied in mediation analysis for longitudinal data, the effect of a continuous mediator repeatedly measured with error on the direct and indirect effects is not well understood.

Regression calibration is one method widely used to correct for measurement error (Brown and Fuller, 1990; Carroll and Stefanski, 1990) and parametric models are often assumed for the calibration model. However, parametric assumptions about the distribution of random effects may be misleading in some cases. Some studies have documented that fixed effect estimates can be sensitive to misspecification of the random effects distribution (Agresti et al., 2004; Heagerty and Kurland, 2001; Heckman and Singer, 1984), and this can impact inference. Several methods have been proposed to allow flexibility in

the distribution of random effects including using a non-parametric maximum likelihood approach that makes no distributional assumptions though the estimated distribution of random effects is discrete (Aitkin, 1999) and requires a significant computational burden to fit (Ghidey et al., 2004); a non-parametric Bayesian approach which detects clusters with unusual results and avoids problems caused by masking in traditional parametric approaches (Burr and Doss, 2005; Ohlssen et al., 2007); a smooth non-parametric maximum likelihood approach involving a mixture of Gaussians which entails some computational burden to fit (Magder and Zeger, 1996); a mixture of normals via the EM algorithm (Verbeke and Lesaffre, 1996); and the skewed versions of the normal and t-distributions, which have been recommended to assess the robustness of conclusions (Lee and Thompson, 2008).

Blood and her collaborators focused on using LMMs to estimate direct and indirect effects either when a binary exposure, mediator and outcome are all time-varying (Blood and Cheng, 2011) or when a binary exposure and mediator are only measured at baseline but the outcome is time-varying (Blood et al., 2010), but did not consider that the mediator may be measured with error. In this Chapter, we focus on estimating direct and indirect effects using LMM when a binary exposure is only measured at baseline with an error prone time-varying mediator and a time-varying outcome. In Section 2.3, we present a potential outcome framework for estimating the direct and time dependent indirect effect in longitudinal studies when the mediator is repeatedly measured without error. Prior to estimating direct and indirect effects, we propose a method to adjust for measurement error in a repeatedly measured mediator using regression calibration. Rather than assuming normality for the random effects in the calibration model, we correct for measurement error in the mediator allowing flexibility in the distribution of random effects. We assume the random effects density belongs to a class of smooth densities (Gallant and Nychka, 1987) and use a seminonparametric (SNP) linear mixed

model described by Zhang and Davidian (2001) and Chen et al. (2002) briefly outlined in Section 2.5. We use a simulation study to compare the performance of our method to an ad hoc approach that ignores possible measurement error on estimators of the direct and indirect effects. We apply the method to a study in which hemoglobin level mediates the effect of cerebral malaria on child cognition and an HIV study where ART adherence mediates the effect of heavy alcohol use and HIV progression, and conclude this Chapter with a discussion in Section 2.8.

2.2 Methods

2.3 Mediation analysis with linear mixed models without measurement error in the mediator

Let Y_{ij}, X_{ij}, G_i, and C_{ij} denote the observed outcome, exposure, baseline covariate, and mediator-outcome confounder respectively for subject i at time $j = 1, \ldots, s$. Let M_{ij} be the mediator measured with error while M^*_{ij} is the mediator without error for subject i at time j. Since a subject's exposure status does not change in studies that motivate this work, X_{ij} is constant within subject i from time 0 through s, though in other studies it can be time-varying. We therefore subsequently drop j on X_{ij} to reflect a non time-varying exposure, X_i for subject i.

We present a directed acyclic graph (DAG) for two time points in Figure 2.1 to illustrate assumptions about mediators measured with error relative to those without error, and relationships among other variables. Bind et al. (2015) considered a time-varying exposure that directly affected the mediator and outcome at only the same time point. In this setup (Figure 2.1), the non

time-varying exposure (X_i) affects the mediator and outcome at baseline and subsequent time points. Using the potential outcome framework for longitudi-

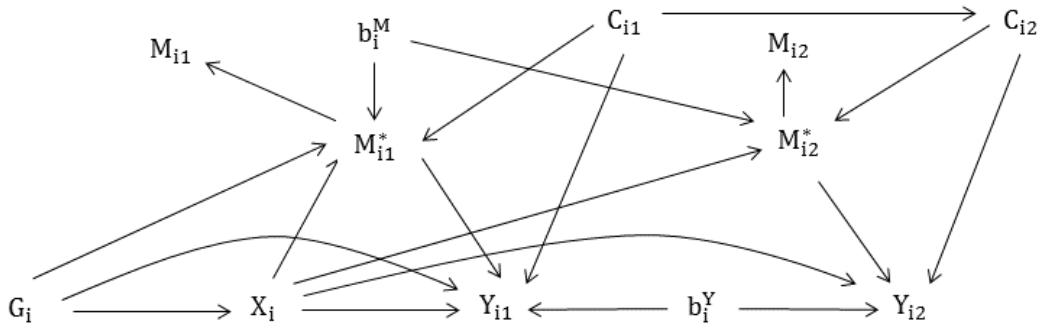

Figure 2.1 Directed acyclic graph (DAG) for time points $j = 1, 2$. M_{i1} and M_{i2} are mediators measured with error while M_{i1}^* and M_{i2}^* are without error.

nal data described by van der Laan and Petersen (2004), VanderWeele (2010) and Bind et al. (2015), let $Y_{ij}(x_i, m_{ij}^*)$ denote the counter factual outcome for subject i at time j when X_i was set to x_i and M_{ij}^* was set to m_{ij}^* and let $M_{ij}^*(x_i)$ be the counter factual mediator for subject i at time j when their exposure status was set to x_i. For consistency, we assume that $M_{ij}^*(X_i) = M_{ij}^*$, and $Y_{ij}(X_i, M_{ij}^*) = Y_{ij}$ (Cole and Frangakis, 2009; VanderWeele, 2009). Bind et al. achieved identifiability of mediation effects in a longitudinal design by conditioning on the random effects (Bind et al., 2015). A difference between the current study and Bind et al. (2015)'s study is that in the latter the mediator is measured without error. The following identification assumptions are needed to estimate mediation effects

$$
\begin{aligned}
&(1)\ Y_{ij}(x_i, m_{ij}^*) \perp\!\!\!\perp X_i | G_i = g_i, \\
&(2)\ M_{ij}^*(x_i) \perp\!\!\!\perp X_i | G_i = g_i, \\
&(3)\ Y_{ij}(x_i, m_{ij}^*) \perp\!\!\!\perp M_{ij}^* | X_i = x_i, G_i = g_i, C_{ij} = c_{ij}, \\
&(4)\ Y_{ij}(x_i, m_{ij}^*) \perp\!\!\!\perp M_{ij}^*(x_i') | G_i = g_i, C_{ij} = c_{ij}.
\end{aligned}
$$

These assumptions hold if there is no unmeasured exposure-outcome confounding, no unmeasured exposure-mediator confounding, no unmeasured mediator-outcome confounding, and no unmeasured mediator-outcome confounder af-

fected by the exposure, for assumptions (1)−(4) respectively. Assumptions (1) and (2) are automatically satisfied if the exposure is randomized. Assuming no time-varying confounding with respect to M_{ij}^* and time-varying covariates (VanderWeele and Tchetgen Tchetgen, 2017), we then define the average direct effect at time j as

$$ADE_j = \mathbb{E}\big[\mathbb{E}\big\{Y_j(1, M_j^*(x)) - Y_j(0, M_j^*(x)) \mid G_i = g_i, C_{ij} = c_{ij}, t_j\big\}\big]$$

and the average causal mediation effect

$$ACME_j = \mathbb{E}\big[\mathbb{E}\big\{Y_j(x, M_j^*(1)) - Y_j(x, M_j^*(0)) \mid G_i = g_i, C_{ij} = c_{ij}, t_j\big\}\big]$$

for some x. With an exposure-time interaction in the mediator model, we note that the mediation effects can change over the course of the study.

In this longitudinal context, to estimate the time-varying average causal mediation effect (ACME) or indirect effect, two regression models would be fitted if we could observe the mediator without error. The first regression model, referred to as the outcome model, is the regression of the outcome Y_{ij} on the exposure X_i, mediator M_{ij}^*, exposure-mediator interaction, covariate G_i, and confounder C_{ij}.

$$\begin{aligned}\mathbb{E}(Y_{ij}|X_i, M_{ij}^*, G_i, C_{ij}, b_i^Y, t_j) &= \gamma_0 + \gamma_x X_i + \gamma_m M_{ij}^* + \gamma_{xm} X_i M_{ij}^* + \gamma_{t_2} t_2 + \ldots \\ &\quad + \gamma_{t_s} t_s + \gamma_g G_i + \gamma_c C_{ij} + b_i^Y \qquad (2.1)\end{aligned}$$

where t_j is the time indicator such that $t_j = I(t = j)$, b_i^Y is the subject specific random effect in the outcome model and is assumed to be normally distributed, i.e., $b_i^Y \sim \text{Normal}\{0, \sigma_{b^Y}^2\}$. The effect of the exposure X_i on the outcome Y_{ij} estimated by $\gamma_x + \gamma_{xm} m_{ij}^*$ is the average direct effect (ADE) or direct effect (Baron and Kenny, 1986; Valeri et al., 2014).

The second regression model, referred to as the mediator model, is the regression of the mediator M_{ij} on the exposure X_i, with an exposure-time interaction and covariates G_i.

$$M_{ij} = \zeta_0 + \zeta_x X_i + \zeta_{t'_2} t_2 + \cdots + \zeta_{t'_s} t_s + \zeta_{t_2} X_i t_2 + \cdots + \zeta_{t_s} X_i t_s + \zeta_g G_i + b_i^M + \epsilon_{ij}^M \quad (2.2)$$

where b_i^M is a subject specific random effect. We typically assume that $b_i^M \sim \text{Normal}\{0, \sigma^2_{b^M}\}$, and the short term biological variability and measurement error in the mediator $\epsilon_{ij}^M \sim \text{Normal}\{0, \sigma^2_{\epsilon^M}\}$ independently of b_i^M. Random effects b_i^Y and b_i^M are assumed to be independent.

Using the product method (Baron and Kenny, 1986) and with some algebra, the time dependent $ACME_j$ or time dependent indirect effect is then estimated as (also see appendix A.1 for details);

$$\begin{aligned} ACME_j &= \mathbb{E}\big\{Y_j(x', M_j^*(x)) - Y_j(x', M_j^*(\tilde{x})) \mid G_i = g_i, C_{ij} = c_{ij}, t_j\big\} \\ &= \gamma_m\big\{\zeta_x + \zeta_{t_j}\big\}(x - \tilde{x}) + \gamma_{xm} x'\big\{\zeta_x + \zeta_{t_j}\big\}(x - \tilde{x}) \end{aligned}$$

where $x = 1$, $\tilde{x} = 0$ and x' can be 0 or 1. Since the exposure coefficient in the mediator model is fixed within subject, we do not add the covariance of the coefficients as proposed by previous studies (Bauer et al., 2006). This has also been documented in the context of linear regression (VanderWeele, 2016).

In the absence of exposure-mediator interaction in the outcome model, the $ACME_j$ reduces to $\gamma_m\zeta_x + \gamma_m\zeta_{t_j}$, and in the absence of interaction in both models (2.1) and (2.2), the ACME reduces to $\gamma_m\zeta_x$, which is not time-dependent. Mediation effects in longitudinal studies have also been defined where there are multiple time-varying mediators, random slopes, interactions between multiple mediators, and when the random effects are correlated (Bind et al., 2015). The sampling distribution of the mediated effect or the ACME is not always

well-approximated by a normal distribution (MacKinnon et al., 2004), thus confidence limits based on the normal distribution for the mediated effect are often inaccurate (MacKinnon et al., 2002, 1995). Therefore, resampling methods and methods based on the sampling distribution of $\gamma_m\zeta_x$ have been recommended for better ACME coverage (MacKinnon et al., 2004).

2.4 Regression calibration when the mediator is measured with error

When the mediator is measured with error, we do not get to observe the true mediator M_{ij}^* (without error), but instead observe $M_{ij} = M_{ij}^* + \epsilon_{ij}^M$, where ϵ_{ij}^M is the measurement error. However, because we collect data longitudinally, we can use regression calibration to obtain consistent estimators for linear mixed model coefficients (Buonaccorsi et al., 2000; Carroll et al., 2006) without access to a separate validation data set. Assuming the classical measurement error model holds, then equation (2.2) gives the regression calibration model. The random effect associated with the mediator calibration model b_i^M can be assumed to be normally distributed or assumed to take a seminonparametric form described in Section 2.5. Allowing flexibility in the distribution of random effects, the calibrated mediator becomes

$$\begin{aligned}\widehat{M_{ij}^*} = \mathbb{E}(M_{ij}|X_i, t_j, G_i, \hat{b}_i^M) &= \hat{\zeta}_0 + \hat{\zeta}_x X_i + \hat{\zeta}_{t_2'} t_2 + \cdots + \hat{\zeta}_{t_s'} t_s + \hat{\zeta}_{t_2} X_i t_2 + \ldots \\ &+ \hat{\zeta}_{t_s} X_i t_s + \hat{\zeta}_g G_i + \hat{b}_i^M \qquad (2.3)\end{aligned}$$

where $\hat{b}_i^M$ is the estimated best linear unbiased predictor (BLUP). We then use $\widehat{M_{ij}^*}$ from model 2.3 in model 2.1 in the place of the unobserved M_{ij}^* and proceed to estimate the ADE and ACME as described in section 2.3. Regression calibration (RC) in linear mixed models for longitudinal studies has been documented by Buonaccorsi et al. (2000). They show that under the classical measurement error model, the mediator model (2.2), and the outcome model

(2.1), RC estimators are equivalent to pseudo-maximum likelihood estimators under normality if the random effects design matrices in models (2.1) and (2.2) are the same. Though Wang et al. (1999) show that RC estimates in generalized linear mixed measurement error models may result in inconsistent estimators, Buonaccorsi et al. (2000) show that RC estimators are consistent and that RC is highly efficient for estimation of fixed effects in linear mixed models.

2.5 Seminonparametric linear mixed model

Of course, the consistency of regression parameter estimators in the outcome model depends on making the correct assumption about the random effects in the regression calibration model. The seminonparametric (SNP) linear mixed model has been described by Zhang and Davidian (2001). In this framework, the subject specific random effects are represented as

$$b_i^M = \mu + RZ_i \tag{2.4}$$

where μ is a $(q \times 1)$ vector of parameters, R is a $(q \times q)$ lower-triangular square matrix and Z_i is a $(q \times 1)$ random vector. Usually, Z_i is assumed to be standard multivariate normal, so that $b_i^M \sim N_q(\mu, RR^T)$. Instead, Z_i and hence b_i^M is assumed to belong to a class of smooth densities (Gallant and Nychka, 1987) without unusual behavior such as kinks, jumps or oscillations that would be unrepresentative of the expected subject-to-subject heterogeneity. However, these densities are sufficiently differentiable and can be skewed, multimodal, or fat- or thin-tailed relative to the normal. Zhang and Davidian (2001) represent the density of Z_i by a standard SNP density

$$h_K(z) = P_K^2(z)\varphi(z) = \left\{ \sum_{|\lambda| \leq K} a_\lambda z^\lambda \right\}^2 \varphi(z)$$

where $\lambda = (\lambda_1, \cdots, \lambda_q)$ is a vector of nonnegative integers, q is the dimension of the random effects, z^λ is the monomial $z^\lambda = z_1^{\lambda_1} \cdots z_q^{\lambda_q}$ of order $|\lambda| = \sum_{k=1}^{q} \lambda_k$, $\varphi(z)$ is a q-dimensional standard normal density, K is the order of the polynomial acting as a tuning parameter controlling the degree of flexibility of the resulting density $h_K(z)$, and the coefficients a_λ of $P_K(z)$ are chosen such that $\int h_K(z)dz = 1$. They ensure $\int h_K(z)dz = 1$ by imposing $E\{P_K^2(U)\} = 1$, where $U \sim N_q(0, I)$. When $K = 0$, $P_K(z) \equiv 1$ and model (2.2) reduces to the usual linear mixed model with $b_i^M \sim N_q(\mu, RR^T)$. This normal density modified by a squared polynomial (i.e. $k = 2$) can approximate complicated shapes such as skewed or multimodal distributions. For the parameter space $\boldsymbol{\zeta} = \{\zeta_0, \zeta_x, \ldots, \zeta_g\}$ in model (2.2), an empirical Bayes approach is then used for inference on individual b_i^M by finding the maximizer $\hat{Z}_i$ of the posterior density $f(Z_i|M_i; \boldsymbol{\zeta}) \propto f(M_i|Z_i; \boldsymbol{\zeta}) P_K^2(Z_i)\varphi(Z_i)$ with $\boldsymbol{\zeta} = \hat{\zeta}$ and calculating $\hat{b}_i = \hat{\mu} + \hat{R}\hat{Z}_i$ (Zhang and Davidian, 2001).

2.6 Data application examples

We apply regression calibration to two data sets and estimate both the ADE and time-varying ACME with their respective confidence intervals. First, we assume the random effects in the mediator model are normally distributed, and second, allow flexibility in the distribution of random effects using the SNP approach. We also present results when no regression calibration is used, in other words, when the mediator is measured with error (Table 2.1) and this is not handled by the analysis.

2.6.1 Severe malaria and cognition data

We consider a longitudinal study conducted in Uganda, one objective of which was to elucidate the contributions of iron deficiency and malaria to neurocognitive impairment. In this prospective cohort study, children under 5 years of age with cerebral malaria (CM) and community control (CC) children were

tested for cognitive development at baseline, 6 months and 12 months using the Mullen scales of early learning that includes five subscales (gross motor, visual reception, fine motor, expressive language, and receptive language). A Mullen composite was generated by summing up four subscales excluding gross motor. Cerebral malaria, a form of severe malaria, may indirectly affect levels of hemoglobin (hgb) in children while directly affecting their neuro-cognitive development. Hemoglobin, presumed to be measured with error, mediates the association between severe malaria and neuro-cognition (Boivin et al., 2016). In this setting, we have a mediator (hemoglobin) measured at baseline, 6 months and 12 months. Covariates included study group (CM=1, CC=0), any child education (Educ), and age. The calibration model was given by

$$\begin{aligned} Hgb_{ij} &= \zeta_0 + \zeta_1 Group_i + \zeta_2 I(t=6) + \zeta_3 I(t=12) + \zeta_4 Group_i * I(t=6) \\ &+ \zeta_5 Group_i * I(t=12) + \zeta_6 Educ_i + b_{0i}^M + \epsilon_{ij}^M \end{aligned}$$

where Hgb_{ij} is the hemoglobin measurement for child i at time $j = 0, 6, 12$ months, t is the time indicator, and the ϵ_{ij}^M are assumed to have a normal distribution N(0, $\sigma^2_{\epsilon_M}$). The calibrated hemoglobin is then estimated, first assuming b_{0i}^M, the subject-specific random intercept, is normally distributed and secondly assuming it belongs to a class of SNP densities (Gallant and Nychka, 1987; Zhang and Davidian, 2001).

To compute the average causal mediation effect (ACME) and average direct effect (ADE), the outcome model was

$$\begin{aligned} Y_{ij} &= \gamma_0 + \gamma_1 Group_i + \gamma_2 Hgb_{ij}^* + \gamma_3 I(t=6) + \gamma_4 I(t=12) + \gamma_5 Age_{ij} \\ &+ \gamma_6 Educ_i + b_{0i}^Y + \epsilon_{ij}^Y \end{aligned}$$

where Y_{ij} is the Mullen composite score measurement for child i at time j, Hbg_{ij}^* is the true hemoglobin value for child i at time t, b_{0i}^Y is the random in-

tercept and ϵ_{ij}^{Y} is the random error. The interaction between study group and hemoglobin was not significant, and thus excluded from the outcome model. Since Hbg_{ij}^{*} is unknown, we use the regression calibration estimate $\widehat{Hbg}_{ij}^{*}$ described earlier. Identification assumptions were examined prior to estimating the mediation effects (appendix A.2.1). In the presence of interaction between study group and time in the mediator model, the ACME was given by $\gamma_2\zeta_1$ at baseline, $\gamma_2(\zeta_1 + \zeta_4)$ at 6 months, and $\gamma_2(\zeta_1 + \zeta_5)$ at 12 months. For the SNP calibration model, $K = 1$ was selected using AIC. Table 2.1 presents results for the ACME and ADE and their respective 95% confidence intervals for the malaria-cognition data.

Table 2.1 ACME and ADE estimates for the malaria-cognition data when the mediator is uncalibrated or calibrated assuming normality or SNP for the random effects.

Time point	Calibration	Average causal mediation effect (ACME)		
		Estimate	LP	UP
Baseline	Normal	-1.407	-5.694	2.684
	SNP	-2.099	-4.639	0.265
	Uncalibrated	-2.257	-5.302	0.794
6 months	Normal	-0.012	-0.337	0.162
	SNP	-0.018	-0.258	0.206
	Uncalibrated	-0.019	-0.339	0.183
12 months	Normal	0.027	-0.158	0.287
	SNP	0.040	-0.122	0.253
	Uncalibrated	0.043	-0.139	0.318
		Average direct effect (ADE)		
		Estimate	LP	UP
	Normal	-9.178	-14.540	-4.356
	SNP	-9.853	-16.441	-4.009
	Uncalibrated	-8.801	-14.293	-3.851

LP, Lower 0.025 bootstrap percentile; UP, Upper 0.975 bootstrap percentile; SNP, Seminonparametric approach. Number of bootstrap data sets for percentile confidence intervals = 2000.

Results from Table 2.1 suggest that mediation effects (ACME and ADE) following regression calibration assuming normality for the individual specific random effects gives results not very different from when there is flexibility in the distribution of the same random effects. For example, the ACME at baseline assuming normality for the random effects was -1.407 (95% CI -5.694, 2.684) compared to -2.099 (95% CI -4.639, 0.265) for the SNP approach. No calibration on the other hand gave ACME estimates (baseline ACME = -2.257, 95% CI -5.302, 0.794) closer to those for SNP calibration than for normal calibration, although this was not the case for ADE. However, examining the 95% CI for ACME at baseline, 6 months and 12 months for normal, SNP and uncalibrated mediators, they are all similar.

2.6.2 HIV-LIVE (Longitudinal Interrelationships of Viruses and Ethanol) data

We consider a prospective cohort study in the United States of 400 HIV-infected patients with a history of alcohol problems. This study investigated the relationship between alcohol and HIV disease progression and related factors. Heavy alcohol consumption can affect a patient's ability to adhere to medication, and this in turn can worsen HIV disease progression. Previous analyses have demonstrated (1) a significant effect of heavy alcohol consumption on CD4 cell count (Samet et al., 2007), (2) an association between any alcohol use and worse ART adherence (Samet et al., 2004) and (3) a reduction in the effect of heavy alcohol consumption on CD4 cell count after controlling for medication adherence among patients on antiretroviral therapy (ART) (Samet et al., 2007), suggesting that medication adherence could mediate the effect of heavy alcohol consumption on the CD4 cell count. Data was collected every 6 months from baseline to 42 months. Three-day adherence to ART was determined using the AIDS Clinical Trials Group Questionnaire for Adherence to Antiretroviral Medications and heavy alcohol use termed 'at-risk of drink-

ing', was evaluated according to the National Institute of Alcohol Abuse and Alcoholism (NIAAA) guidelines. Other variables included CD4 cell count, age, literacy score, homeless in last 6 months (yes/no), and HIV Quality of life scale (HIV_QOL). CD4 cell counts were skewed and so a square root transformation was performed before analysis. Three quarters (75%) of the study participants were male, mean age was 42.5 years (Standard deviation = 7.4) and 125 (31%) were 'at-risk of drinking' at baseline (Alc). The mediator calibration model was

$$\begin{aligned} Adherence_{ij} &= \beta_0 + \beta_1 Alc_i + \beta_2 Age\_BL_i + \beta_3 HIV\_QOL_{ij} + \beta_4 Literacy_i \\ &+ \beta_5 I(Time = 2) + \cdots + \beta_{11} I(Time = 8) \\ &+ \beta_{12} I(Time = 2) * Alc + \cdots + \beta_{18} I(Time = 8) * Alc \\ &+ b_{0i}^M + \epsilon_{ij}^M. \end{aligned}$$

For SNP calibration, different K values were evaluated. K = 2 gave the smallest Akaike's Information Criterion (AIC) and thus was fixed at that value. The outcome model was

$$\begin{aligned} sqrt\_CD4_{ij} &= \theta_0 + \theta_1 Alc_i + \theta_2 \widehat{Adherence}_{ij}^{*} + \theta_3 Age\_BL_i + \theta_4 HIV\_QOL_{ij} \\ &+ \theta_5 Homeless_{ij} + \theta_6 I(Time = 2) + \cdots + \theta_{12} I(Time = 8) \\ &+ b_{0i}^Y + \epsilon_{ij}^Y \end{aligned}$$

where $\widehat{Adherence}_{ij}^{*}$ is the calibrated adherence. The interaction between at risk of drinking and adherence was not significant. Appendix A.2.2 presents an examination of identification assumptions. The average direct effect of heavy alcohol consumption after controlling for adherence and other covariates were -1.901 (95% CI -3.340, -0.309) assuming normal random effects for the calibration model, -1.903 (95% CI -3.326, -0.318) assuming the SNP approach, and -2.058 (95% CI -3.477, -0.568) for the uncalibrated mediator. Since the mediator model included an interaction between heavy alcohol use and time,

the ACME at baseline was given by $\theta_2\beta_1$, and at time $j = 2, \ldots, 8$ was given by $\theta_2\beta_1 + \theta_2\beta_{1j}$, taking on different values at different time points. ACME estimates with their corresponding bootstrap confidence intervals are displayed in Figure 1. We see similar ACME estimates for normal calibration and SNP calibration at all 8 time points. When no calibration is conducted, the ACME estimates at all time points are close to zero and the confidence intervals are much smaller.

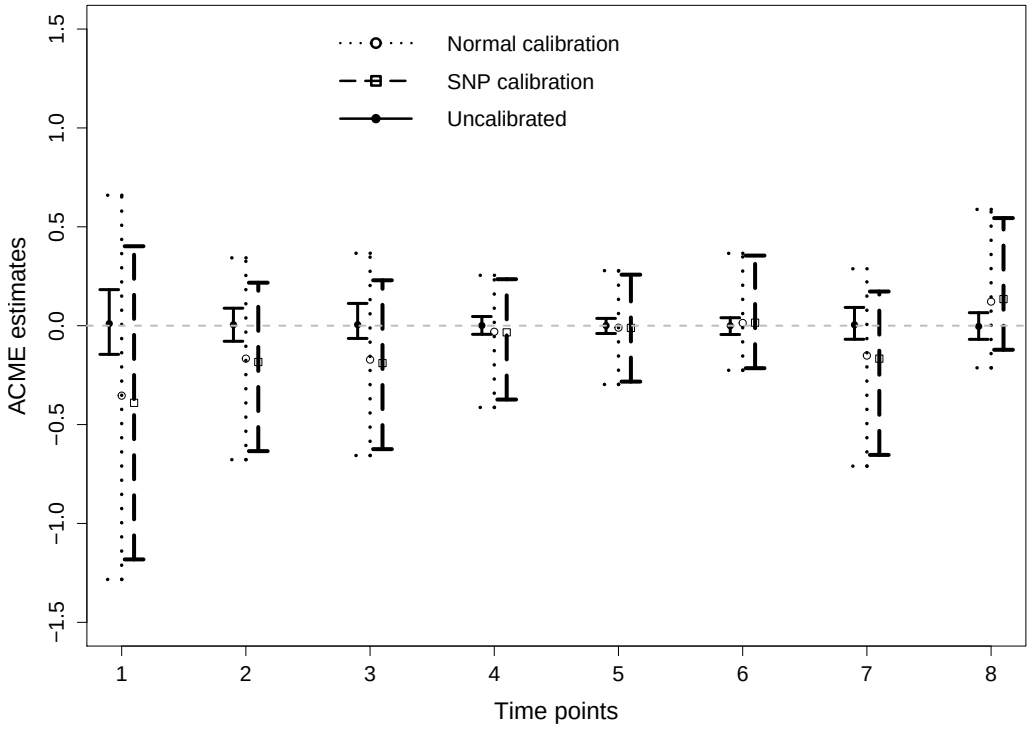

Figure 2.2 ACME estimates with respective bootstrap confidence intervals for HIV-LIVE data by time points either assuming normality or SNP approach for the random effects in the mediator calibration model.

2.7 Simulation study

We conducted a simulation study to investigate how measurement error in the mediator affects ADE and ACME. The residual errors in the mediator and the outcome models were independent. A sample of $i = 1, \cdots, m = 600$

subjects, each with $j = 1, \cdots, 6$ observations, was generated. The mediator data-generating model was

$$M_{ij} = \zeta_0 + \zeta_1 X_i + \zeta_2 I(j = 2) + \cdots + \zeta_6 I(j = 6) + b_{0i}^M + \epsilon_{M_{ij}}$$

where x_i is the exposure indicator (50% exposed and the rest non-exposed), I is an indicator function allowing differing values of the continuous mediator and outcome within individual at different time points, $\zeta_0 = 4$, $\zeta_1 = -1.7$, $\zeta_2 = 3.3$, $\zeta_3 = 3.8$, $\zeta_4 = 4.3$, $\zeta_5 = 4.8$, $\zeta_6 = 5.3$; $\epsilon_{M_{ij}}$ and the subject specific random intercept b_{0i}^M are independent. To examine the benefit of allowing flexibility in the distribution of random effects in the mediator model versus assuming normality, b_{0i}^M were simulated under three distributions scenarios, taking on a

i. Normal distribution, $N(0, 4)$

ii. Chi-square distribution, $\chi^2(4)$.

iii. Mixture of normals distribution, $0.65N(0, 1) + 0.35N(4, 1)$

To obtain the calibrated mediator, separate models were fit where b_{0i}^M was first assumed normally distributed, and secondly, assumed to come from a class of smooth densities (SNP). For the SNP approach, K was chosen using the Akaike's Information Criteria (AIC) after fitting three models with $K = 0, 1,$ or 2. The K from a model with the least AIC was chosen. The outcome data-generating model was

$$Y_{ij} = \gamma_0 + \gamma_1 X_i + \gamma_2 M_{ij}^* + \gamma_3 I(j = 2) + \cdots + \gamma_7 I(j = 6) + b_{0i}^Y + \epsilon_{Y_{ij}}$$

where x_i, and I are as described before, $\gamma_0 = -1.0$, $\gamma_1 = -0.6$, $\gamma_2 = 0.8$, $\gamma_3 = 0.15$, $\gamma_4 = 0.25$, $\gamma_5 = 0.4$, $\gamma_6 = 0.55$, $\gamma_7 = 0.7$; $\epsilon_{Y_{ij}}$ and $\epsilon_{M_{ij}}$ were from a multivariate normal distribution, with means zero, $Var(\epsilon_{Y_{ij}}) = 1$, $Var(\epsilon_{M_{ij}}) = 1.5$, and $Cov(\epsilon_{Y_{ij}}, \epsilon_{M_{ij}}) = 0$.

2000 data sets were generated under each distribution scenario for M_{ij}. The calibrated mediator $\widehat{M}^{*}_{ij}$ was computed for each data set, first, assuming b^{M}_{0i} are normally distributed, and second, assuming they come from a class of smooth densities. The computed ADE and ACME under each assumption for b^{M}_{0i} was compared to true estimates (ACME = -1.360, ADE = -0.60) (Table 2.2). We computed the mean square error (MSE) between the predicted random effects and the simulated random effects under both the normality assumption and the SNP approach (see Table 2.3). All simulations were conducted in SAS version 9.4.

Without mediator calibration, both ACME and ADE estimators were biased for all mediator random effects distributions used to generate the data. In particular the ACME was over estimated while the ADE was under estimated. Under these simulation assumptions, we note that the ACME was biased towards the null while the ADE was biased away from the null. However, results improved substantially using regression calibration. ACME and ADE estimators assuming a normal calibration model were approximately unbiased irrespective of the random effects distribution used to generate the data. This was also the case when the SNP approach was used for the calibration model. Assuming normality for the calibration model, irrespective of random effects distribution, gave similar results (both ACME and ADE estimates) to allowing flexibility in the random effects distribution using the SNP approach. Further simulations comparing the mean square error (MSE) for the true random effects and the estimated random effects showed that the random effects mean square errors when normality was assumed for the calibration model were very similar to those when flexibility of the random effects distribution was assumed using the SNP approach. On the other hand, without calibration, the ACME estimator at baseline was away from the null for the severe malaria and cognition data whereas the ADE was towards the null compared to normal and SNP calibration results. This suggests that no calibration for a mediator mea-

sured with error could bias mediation effects in any direction of the null. This simulation included six time points; we also conducted simulations with only three time points and results were similar (appendix tables A.1 and A.2).

Table 2.2 Simulation results: average causal mediation effect (ACME) and average direct effect (ADE) estimates either assuming normality or SNP approach for the random effects in the mediator calibration model.

Distribution of RE	Average causal mediation effect (ACME)								
	Normal calibration model			SNP calibration model			Using mediator with error		
	Estimate (-1.360)	Bias	MSE	Estimate (-1.360)	Bias	MSE	Estimate (-1.360)	Bias	MSE
Normal	-1.364	-0.004	0.020	-1.364	-0.004	0.020	-0.207	1.153	1.330
$\chi^2(4)$	-1.362	-0.002	0.037	-1.362	-0.002	0.037	-0.269	1.091	1.194
Mixture of Normals	-1.358	0.002	0.023	-1.360	-0.000	0.023	-0.220	1.140	1.301
	Average direct effect (ADE)								
	Estimate (-0.600)	Bias	MSE	Estimate (-0.600)	Bias	MSE	Estimate (-0.600)	Bias	MSE
Normal	-0.600	0.000	0.011	-0.600	0.000	0.011	-1.757	-1.157	1.358
$\chi^2(4)$	-0.600	0.000	0.010	-0.600	0.000	0.010	-1.694	-1.094	1.227
Mixture of Normals	-0.601	-0.001	0.010	-0.599	0.001	0.010	-1.739	-1.139	1.320

RE, Random effects; MSE, Mean square error; SNP, Seminonparametric approach. Number of Monte-Carlo data sets = 2000.

Table 2.3 Simulation results: mean square errors and standard deviation comparing true random effects distribution and estimated random effects distribution assuming normality or using the SNP approach.

Distribution of Random effects	Normal		χ^2_4		Mixture of Normals	
	MSE	SD	MSE	SD	MSE	SD
Normal calibration	0.248	0.019	0.268	0.030	0.252	0.021
SNP calibration	0.250	0.024	0.261	0.033	0.242	0.019

RE, Random effects; MSE, Mean square error; SD, Standard deviation. The mediator and outcome are independent. Number of Monte-Carlo data sets = 2000.

2.8 Discussion

We have studied how measurement error in the mediator affects estimators of the direct and indirect effects. Rather than assuming normality for the random effects in the mediator calibration model, we allowed flexibility in the distribution of random effects and evaluated its impact on direct and indirect

effect estimators.

Estimation of the average causal mediation effect depends on estimates of the effect of the exposure on the mediator and the effect of the mediator on the outcome. Thus, any bias in these two estimators is carried forward to estimation of the ACME. Measurement error in the mediator which is not accounted for in the analysis, leads to biased parameter estimators in the outcome model and biases estimators of the causal effect. However, when the mediator is collected longitudinally, as is done in many studies, regression calibration may be used to obtain improved estimators of the causal effect without the need for an external validation dataset. A key finding in our study is that misspecifying the random effect distribution in the mediator calibration model had no impact on the ACME estimates for the cases considered. That is, when the random effects distribution deviates from normality such as chi-square (skewed) or mixture of normals, our results suggest that assuming normality for the mediator calibration still gives approximately unbiased estimators of the ACME and ADE, an indication of robustness. We used the seminonparametric approach (SNP) to allow flexibility in the distribution of random effects; other methods could be used.

Dividing the calibration mediator in model (2.3) into three components, i.e., fixed effects, intercept and random effects, potential biases during regression calibration are likely to emerge from three sources: (1) estimation of the fixed effects, (2) estimation of the intercept, and (3) estimation of the random effects. Though previous studies show mixed results on how misspecification of the random effects distribution affects the estimation of fixed effects (Agresti et al., 2004), several studies have documented that fixed effect estimators are approximately unbiased (McCulloch and Neuhaus, 2011; Neuhaus et al., 1992). The continuous mediator in our setting is a within-cluster covariate, which means that it varies within cluster but has a constant average between clus-

ters (McCulloch and Neuhaus, 2011). Studies have shown little impact of random effects distribution misspecification on the estimation of the effect of a within-cluster covariate (McCulloch and Neuhaus, 2011; Neuhaus et al., 1992; Zhang and Davidian, 2001). It follows, therefore, that the effect of the calibrated mediator on the outcome will be unbiased whether we allow flexibility in the distribution of random effects in the calibration model or make the common normality assumption. The estimation of the intercept also affects the calibrated mediator. Neuhaus et al. (1992) showed that estimates of the intercept for nonlinear models may be biased when the random effects distribution is far from normal, however, this is not the case for linear mixed models (McCulloch and Neuhaus, 2011). Our simulation results are consistent with findings by McCulloch and Neuhaus (2011), i.e. that assuming normality for the random effect distribution for the mediator model has little or no effect on the estimation of the intercept.

Another source of bias is the prediction of the random effects. This study compared predicted random effects for the mediator model assuming normality versus predicted random effects assuming the SNP approach. The SNP approach had similar mean square error compared to assuming normality for the different random effects distribution assumptions. These results are consistent with other studies that found no change (Magder and Zeger, 1996) or not very large differences (Agresti et al., 2004) in the estimation of random effects when the assumed random effects distribution was not normal yet normality was assumed. Agresti et al. (2004) studied effects of assuming normality for the random effects for binary response data and their approach did not perform well when the true distribution was a two-point mixture with a large variance component, which is unrealistically extreme. In the study by Magder and Zeger (1996), they found that allowing flexibility in the distribution of random effects using the nonparametric maximum likelihood estimate (NPMLE) and the smooth nonparametric maximum likelihood estimate (SNPMLE) when

the true distribution was Gaussian, skewed or discrete, yielded similar mean square errors for the random effects compared to assuming normality. The current study results bolster the notion that assuming normality for the distribution of random effects when they are non-normal has little or no impact on predicted random effects.

Our simulations included only random intercept models. McCulloch and Neuhaus (2011) points out that when the mean of the random effects distribution depends on a covariate, a relationship between the covariate and the random effect distribution is introduced, creating a serious bias in estimating the relationship between the covariate and the outcome. We did not examine this scenario, and the potential effects of this relationship on the estimation of ACME and ADE need to be investigated further.

A strong assumption in our study was that the residual error in the mediator and outcome models were uncorrelated. In appendix A.4, we show that as the mediator-outcome correlation increases, with an increase in within subject variability, the ACME and ADE estimates are more biased. More research towards bias correction in estimates is needed, especially when mediator and outcome are correlated.

We have documented a method to calibrate an error prone repeatedly measured mediator in mediation analysis for longitudinal data. In the presence of measurement error in the mediator, regression calibration improves mediation effect estimates compared to no calibration at all.

Chapter 3

Missing data and mediation analysis in longitudinal studies

3.1 Introduction

Longitudinal studies, even those that are well designed and executed, face challenges of missing data that often result from loss to follow-up, death, or withdrawal. Missing data can lead to biased estimators, loss of information, decreased statistical power and weakened generalizability of study findings (Dong and Peng, 2013). Patterns of missing data are classified as missing completely at random (MCAR), where the probability of missingness does not depend on observed or unobserved data; missing at random (MAR), where the probability of missingness depends on the observed data but not on the unobserved data; and missing not-at-random (MNAR), where the probability of missingness depends on the unobserved data (Little and Rubin, 1987; Rubin, 1976). Methods proposed to deal with missing data include complete case analysis, pairwise deletion, mean substitution, which are older methods that are highly discouraged, and popular contemporary methods such as multiple imputation (MI) and maximum likelihood (ML) methods (Graham, 2009). MI and ML methods are often used because they preserve data characteristics, result in parameter estimators that are unbiased (Graham, 2009; Schafer and

Graham, 2002) when data is MCAR or MAR, and can be implemented in commercially available software such as PROC MI in SAS or MI IMPUTE in STATA. MI is preferred to ML methods in some situations because it creates complete data sets (which allow use of statistical methods that work only with complete data (Schafer, 2001; Wu and Jia, 2013)), allows for missing values in the predictors, separates the imputation phase from the analysis phase (allowing variables that explain nonresponse, but are not of interest in analysis, to be used in the imputation phase (Schafer, 2001)), and allows the use of different models for analysis and imputation.

Mediation analysis hypothesizes that the exposure causes the mediator and in turn the mediator causes the outcome (MacKinnon, 2008; MacKinnon et al., 2007), so that mediation is inherently longitudinal. In mediation analysis, MI has been shown to perform well for cross-sectional data for both MCAR and MAR data with and without auxiliary variables (Wu and Jia, 2013; Zhang et al., 2015), but, it is unclear how MI performs in longitudinal studies when researchers are interested in delayed mediated effects. An example is a study of children under 5 years (Cusick et al., 2016) that examined the effect of cerebral malaria on cognitive development, mediated by hemoglobin level. Researchers may be interested in the mediated effects of (i) baseline hemoglobin, and (ii) change in hemoglobin level from baseline to 6 months or to 12 months, on the effect of cerebral malaria on cognitive development at 12 months or on the change in cognitive development from baseline to 12 months.

MI in current commercial software typically ignores correlation among observations on the same unit or cluster (Mistler, 2013), which is characteristic of longitudinal studies. In cluster randomized trials (CRTs), ignoring correlation among observations has been found to bias cluster means towards the grand mean and to underestimate between-cluster variance (Taljaard et al., 2008; Zhou et al., 2014). Using MI in these commercial software packages with-

out modification can bias fixed effects estimators in CRTs (Andridge, 2011; Black et al., 2011). To our knowledge, no one has examined the impact of ignoring clustering in the imputation model on mediated effects in longitudinal studies. To preserve the hierarchical structure of data during imputation in CRTs, Graham (2009) proposed to include a cluster-specific dummy variable as a fixed effect in the imputation model. However, CRTs have a few large clusters, while longitudinal studies, where study participants are treated as clusters, have moderate to large number of clusters. Moreover, including cluster dummy variables makes imputed values biased towards the cluster means and this inflates differences between clusters (Zhou et al., 2014). Previously, studies that compared ignoring versus considering clustering in imputation models have been in CRTs (Andridge, 2011; Black et al., 2011; Taljaard et al., 2008; Yucel and Demirtas, 2010) and not in longitudinal studies. Furthermore, these studies have examined the effect of ignoring clustering on fixed effects without mediation. In the context of mediation analysis, the mediator is both an outcome and a covariate and, therefore, any method proposed to address missingness in the mediator needs to consider the mediator in both capacities.

In this Chapter, we examine the impact of ignoring clustering on mediated effects estimates under MCAR and MAR mechanisms in longitudinal studies with varying degrees of missingness. We describe a Markov chain Monte Carlo (MCMC) approach for imputing missing longitudinal data considering clustering (Schafer, 2001), followed by two imputation approaches that ignore clustering. We apply the methods to the severe malaria and cognition data (Cusick et al., 2016), conduct a simulation study, present results from the three imputation approaches and results when no imputation is conducted, and conclude the Chapter with a discussion.

3.2 Methods

3.2.1 Mediation analysis

Consider an outcome Y_{ij}, potential mediator M_{ij}, treatment or exposure T_i, and a vector of potential confounders U_i, $i = 1, \cdots, n$, $j = 1, \cdots, s$ for subject i at time j. Suppose we are interested in whether the effect of the exposure T_i on the outcome Y_{is} is mediated through M_{ij}. The total effect of the treatment on the outcome can be decomposed into the average causal mediation effect (ACME), the effect of treatment on the outcome only through the treatment effect on the mediator and the average direct effect (ADE), the effect of treatment on the outcome adjusting for the mediator. To compute the ACME and ADE in a simple mediation analysis, we fit two models (i) $E(Y_{is}|T_i = t, M_{ij} = m, U_i = u) = \alpha_0 + \alpha_1 T_i + \alpha_2 M_{ij} + U_i^T \alpha_u^T$, and (ii) $E(M_{ij}|T_i = t, U_i = u) = \beta_0 + \beta_1 T_i + U_i^T \beta_u^T$. Under the assumptions of no unmeasured confounding of the exposure-mediator relationship, exposure-outcome relationship, mediator-outcome relationship, and no mediator-outcome confounders associated with the exposure (Bind et al., 2015; Valeri et al., 2014; van der Laan and Petersen, 2004; VanderWeele, 2010), the ACME is given by $\alpha_2\beta_1$ and the ADE by α_1 (Baron and Kenny, 1986) so that the total effect is $\alpha_1 + \alpha_2\beta_1$.

3.2.2 Multiple imputation (MI) in mediation analysis

3.2.2.1 Considering clustering using an MCMC approach

In longitudinal data, subjects are repeatedly measured over time and measurements for subject i are clustered within subject, and are correlated. Multiple imputation methods for correlated data in longitudinal studies have been developed by Schafer (1997) and Liu et al. (2000). These methods feature a Markov chain Monte Carlo (MCMC) algorithm that uses a Gibbs sampler for a multivariate linear mixed effects model for incomplete data (Liu et al.,

2000; Schafer, 1997; Schafer and Yucel, 2002). To do imputation, we assume a multivariate extension of the linear mixed effects model

$$R_i = X_i\lambda + Z_i b_i + \varepsilon_i \tag{3.1}$$

where R_i is an $n_i \times r$ matrix of responses with missing values, $X_i(n_i \times p)$ and $Z_i(n_i \times q)$ are covariate matrices, $\lambda(p \times r)$ is a matrix of regression coefficients, rows of the residual matrix are independently distributed as $\varepsilon_i \sim N(0, \Sigma)$, and $vec(b_i) \sim N(0, \Psi)$ is a vector of random coefficients, where b_i is $q \times r$ and vec denotes the vectorization of a matrix by stacking its columns. Variables with missing values are included in R_i regardless of whether they are responses or covariates. When there is missingness in the outcome Y_{ij} and mediator M_{ij}, responses for participant i can be arranged in an $s \times r$ matrix with a column for each variable as (in this case $r = 2$)

$$R_i = \begin{bmatrix} y_{i1} & m_{i1} \\ y_{i2} & m_{i2} \\ \vdots & \vdots \\ y_{is} & m_{is} \end{bmatrix}.$$

The rows indicate measurements at different time points. In model (3.1), the covariates are the same for both the outcome and mediator. This poses no problem since the objective is to impute missing responses while preserving relations within the data (Schafer, 2001).

Priors for Ψ and Σ are typically chosen to be weak to limit their influence on results. Schafer (2001) and Schafer and Yucel (2002) recommended independent Wishart priors $\Sigma^{-1} \sim W(\nu_1, \Lambda_1)$ and $\Psi^{-1} \sim W(\nu_2, \Lambda_2)$, where $W(\nu, \Lambda)$ denotes a Wishart variate with $\nu > 0$ degrees of freedom and mean $\nu\Lambda > 0$. These priors allow an unstructured Ψ and in practice hyperparameters are chosen such that $\nu_2 = qr$ where r is the number of variables with missing data, q

is the dimension of the random effects and $\Lambda_2^{-1} = \nu_2\hat{\Psi}$. The value of ν_1 is set to r and $\Lambda_1^{-1} = \nu_1\hat{\Sigma}$. The values for $\hat{\Psi}$ and $\hat{\Sigma}$ can be obtained from the data using maximum likelihood estimation (Schafer and Yucel, 2002; Yucel, 2015). The prior for λ is usually an improper uniform density over $\mathbb{R}^{pr}$ (Schafer, 1997; Yucel and Demirtas, 2010). The MCMC algorithm is run until convergence. After convergence, the algorithm draws from the posterior distribution of the parameters and then imputes missing data values conditional on the drawn parameter values. These methods are implemented in PAN, an open source R package that accounts for clustering in MI for longitudinal data (Zhao and Schafer, 2016). PAN has been described previously (Schafer, 2001; Schafer and Yucel, 2002).

3.2.2.2 Ignoring clustering using multivariate imputation by chained equations (MICE)

Multivariate imputation by chained equations (MICE), also refered to as fully conditional specification (Azur et al., 2011; Van Buuren, 2007) or sequential regression multiple imputation (Azur et al., 2011; Raghunathan et al., 2001) has been used in a variety of fields to multiply impute missing data (Buuren and Groothuis-Oudshoorn, 2011). MICE imputes data on a variable-by-variable basis with the flexibility to specify a different model for each variable. Suppose Y_{ij} and M_{ij} are partially observed and X_{ij} is a set of fully observed covariates. Data for each subject are stacked together in columns so that subjects have multiple rows and each row is a subject's measurement at one time, j. Since this method ignores clustering, subscripts i and j are subsequently dropped from Y_{ij}, M_{ij} and X_{ij}. Let Y^{obs} and M^{obs} denote observed values and Y^{mis} and M^{mis} denote missing values for Y and M respectively. Initially, missing values in Y and M are filled in by simple random sampling with replacement from observed values (White et al., 2011). Using only observed values for Y for all subjects, Y^{obs} is regressed on M^{obs} and covariates X, producing a set of maximum likelihood estimates $\hat{\theta}_y$ of θ_y. A random draw is taken from the posterior

predictive distribution of θ_y, to produce a new set of coefficients θ_y^*. To generate sufficient variability in the imputed values, θ_y^* is drawn from a multivariate normal distribution with mean $\hat{\theta}_y$ and the estimated covariance of $\hat{\theta}_y$ with an additional random draw for the residual variance. Using a draw of θ_y^*, predicted values for Y are generated for all cases, Y^{obs} and Y^{mis}. For each Y^{mis}, a set of five Y^{obs} whose predicted values are nearest to the predicted value for the case with a missing value are identified. From these nearest neighbors, one value is randomly chosen and Y^{mis} is replaced by its corresponding observed value. Then M^{obs} is regressed on Y (including the imputed values) and covariates X, producing another set of coefficients θ_m. θ_m^* is drawn from the posterior predictive distribution of θ_m as was done for Y and used to generate predicted values for all M^{obs} and M^{mis}. For each case with M^{mis}, five cases with M^{obs} are identified with predicted values nearest to the predicted value for the case with a missing value. From the five nearest neighbors, one value is randomly chosen and its corresponding observed value substituted for M^{mis}. This process is repeated g cycles until regression parameters become stable (Bouhlila and Sellaouti, 2013; White et al., 2011) to produce a single imputed data set. To stabilize the distribution of regression parameters, 5-10 cycles have been found to yield satisfactory performance with moderate amounts of missing data (Brand, 1999; Van Buuren et al., 2006).

3.2.2.3 Linear model (LM) method

This method can be considered when the analysis model is a generalized linear model rather than a generalized linear mixed model that is consistent with longitudinal data. It emulates situations when researchers may be interested in the mediation effect of a change in the mediator from baseline to the end of the study on the effect of an exposure on the outcome at the end of the study. This method is similar to ignoring clustering in that it does not use a random effect. However, only variables at time points needed for analysis are used in the imputation phase, contrary to using all the data. The LM method uses

MICE with a non time-varying exposure, outcome, mediator and covariates to generate imputed data sets.

3.2.3 Pooling results

For each imputation method, several complete data sets are generated with missing values imputed, say k complete data sets. The k data sets are then used to fit k models and fixed effect estimates from these models are combined using Rubin's rules (Rubin, 1987), also described by White et al. (2011). Briefly, assume that k data sets are imputed and the estimate of interest is $\hat{\beta}$ with estimated variance $\hat{V}$. If $\hat{\beta}_l$ and $\hat{V}_l$ are fixed effect estimates and variance respectively from the l^{th} $(l = 1, 2, \ldots, k)$ imputed data set, then the combined estimate is $\hat{\boldsymbol{\beta}} = (1/k)\sum_{l=1}^{k}\hat{\beta}_l$ and the total variance is $\text{Var}(\hat{\boldsymbol{\beta}}) = W + [1 + (1/k)]B$, where W and B are the within-imputation and between-imputation variances given by $W = (1/k)\sum_{l=1}^{k}\hat{V}_l$ and $B = [1/(k-1)]\sum_{l=1}^{k}(\hat{\beta}_l - \hat{\boldsymbol{\beta}})^2$ respectively. Some studies have suggested that 3-5 imputed data sets are adequate (Rubin, 1987; Schafer, 1999; Schafer and Olsen, 1998), though some have suggested up to 10 data sets (Schafer, 1999) or even more (Graham et al., 2007).

3.3 Simulation study

We simulated longitudinal data for the mediator M and the outcome Y from linear mixed models. Y_{ij} and M_{ij} were generated for subjects $i = 1, 2, \ldots, n = 300$ at time points $j = 1, 2, \ldots, m = 6$ such that

$$\begin{aligned}
M_{ij} &= \zeta_0 + \zeta_2 T_i + \zeta_3 I(j=2) + \cdots + \zeta_7 I(j=6) + \zeta_8 X_1 + \zeta_9 X_2 \\
&+ \zeta_{10} T_i I(j=2) + \cdots + \zeta_{14} T_i I(j=6) + b_{0i}^{M} + \varepsilon_{M_{ij}} \\
Y_{ij} &= \gamma_0 + \gamma_2 T_i + \gamma_3 (M_{ij} - M_{i1}) + \gamma_4 I(j=2) + \cdots + \gamma_8 I(j=6) \\
&+ \gamma_9 X_1 + \gamma_{10} X_2 + b_{0i}^{Y} + \varepsilon_{Y_{ij}}
\end{aligned}$$

where T_i is the treatment for subject i, I is the indicator function, X_1 is a continuous covariate such that $X_1 \sim N(10, 1.2)$, and X_2 is a binary covariate taking each possible value (0 or 1) with probability 0.5. The vectors $\boldsymbol{\zeta}$ and $\boldsymbol{\gamma}$ are fixed effects for the mediator and outcome data generating models respectively. We set $\boldsymbol{\zeta} = (8, -1.7, 3.3, 3.8, 4.3, 4.8, 5.3, -0.3, 1.9, -0.2, -0.6,$
$-0.5, -0.7, -0.9)$ and $\boldsymbol{\gamma} = (-1, -1.5, 2.4, 1.2, 1.25, 1.45, 1.6, 1.75, 0.22, -1.55)$. The random intercepts and errors were simulated from a multivariate normal distribution so that the intra-class correlation for both outcome and mediator is 0.8.

$$\begin{pmatrix} \varepsilon_{Y_{ij}} \\ \varepsilon_{M_{ij}} \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right), \quad \begin{pmatrix} b_{0i}^{Y} \\ b_{0i}^{M} \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \right).$$

Complete data was first generated and then outcomes at different time points were set to missing according to a MCAR or MAR mechanism. Observations at the first time point were always observed and thus non-missing. The missingness was monotone for MAR, i.e., if the mediator or outcome is missing at time $j < 6$, then it is also missing at subsequent time points from $j + 1$ to time point 6. Proportions of missingness were 10%, 20%, 30% and 40%. Since the mediator is both an outcome and a covariate, the MAR mechanism was mediator dependent in half of the data set and the other half, outcome dependent. The MAR mechanism was determined by the magnitude of the drop in the outcome or mediator measure from time point to time point, however, values for the outcome or mediator from which this magnitude can be measured are retained in the data set. In a longitudinal study with withdrawal, loss to follow-up, or death, we assume that when the outcome is missing at time j, we are unable to measure the mediator at the same time j and vice versa. Thus, when the outcome was missing, the mediator was set to missing too and vice versa. To generate data consistent with the MCAR mechanism, data was made missing randomly in the complete data set from time point 2 through 6, so missingness does not depend on either the observed or unobserved data.

Multiple imputation for missing values was done, first, ignoring clustering using multivariate imputation by chained equations (MICE) (Buuren and Groothuis-Oudshoorn, 2011), second, using the analysis model (3.2) with imputation by MICE, referred to as the linear model (LM) method, and third, considering clustering in the data using PAN R package as described in the methods section. The linear model (LM) method is a natural approach and is congenial in the sense of Meng (1994) because the imputation model is similar to the analysis model. For both ignoring clustering and the LM method, the number of cycles was fixed at $g = 15$. The analysis models were

$$\begin{aligned} M_{i6} - M_{i1} &= \beta_0 + \beta_1 T_i + \beta_2 X_1 + \beta_3 X_2 + \epsilon_{i6}^M \\ Y_{i6} &= \alpha_0 + \alpha_1 T_i + \alpha_2 (M_{i6} - M_{i1}) + \alpha_3 X_1 + \alpha_4 X_2 + \epsilon_{i6}^Y \end{aligned} \quad (3.2)$$

where time point $j = 6$ is the end of the study. The true ACME was $\alpha_2\beta_1 = -2.16$ and the true ADE was $\alpha_1 = -1.50$. Missing values for M_{ij} were imputed first, and then the difference $M_{i6} - M_{i1}$ computed using observed and imputed values. For each proportion of missingness, 2000 Monte-Carlo data sets were simulated. For each data set, $k = 10$ imputations were conducted, and analyzed using the linear model (3.2) with results pooled using Rubin's method (Rubin, 1987). For each imputed data set, the ACME and ADE were estimated, after which results were pooled. ACME and ADE estimates for all methods including no imputation were then compared on bias, mean square error, and the coverage probability of 95% Wald-type confidence intervals. ACME estimates were also compared on the Sobel standard error (S-SE) (Sobel, 1982) and ADE estimates on the standard error. The ACME coverage probability was computed as the percentage of times the true parameter value was covered by the 95% confidence interval constructed using the S-SE.

3.4 Results

3.4.1 Simulation results

Simulation results in Table 3.1 indicate that when data are MCAR, for all proportions of missingness, ignoring clustering in the imputation model leads to biased ACME estimators and that bias increases as the proportion of missing data increases. Coverage probabilities are below the nominal 95% and decrease as the proportion of missing data increases. Since the Monte-Carlo standard deviation (MC-SD) was closer to the mean Sobel standard error (MS-SE), an estimator for the ACME standard error, though lower than for complete data, lower coverage probabilities are a consequence of bias in the ACME estimator, which shifts confidence intervals away from the true ACME. On the other hand, no imputation, imputing using the LM, and considering clustering gave unbiased estimators and a coverage probability close to 95% for all proportions of MCAR. Considering clustering had the lowest MSEs compared to no imputation, ignoring clustering and the LM method. We also compared the methods according to the MS-SE. For all proportions of missing data, the MS-SEs when clustering was ignored were lower than when LM was used or when clustering was considered in the imputation model. When 10% of data was MCAR, the MSEs for LM and considering clustering were close to that of complete data. However, for 20%-40% proportions of missingness, considering clustering consistently gave lower MSEs compared to the LM method and this difference in MSEs increased as the proportions of missing data increased.

When data were MAR, ACME estimators were unbiased at up to 30% proportions of missing data when clustering was considered in the imputation model (Table 3.1). ACME estimators for all considered proportions of missing data for no imputation and ignoring clustering in the imputation model were biased and bias in the estimator increased as the proportion of missing data

increased. In fact, ignoring clustering had the most biased ACME estimators and as the proportion of missing data increased, the ACME estimator became more biased towards the null. Moreover, the bias was greater when data was MAR than when it was MCAR. The ACME estimator for imputation using the LM method was biased for MAR data, but the bias was lower compared to no imputation and ignoring clustering. The MSE was highest when clustering was ignored compared with the other methods for all proportions of missing data, and increased as the proportion of missing data increased. Considering clustering had the lowest MSE for all proportions of missing data. MS-SEs ignoring clustering were five times higher than those for other methods including when no imputation was done. Coverage rates for no imputation, considering clustering and the LM method were close to the nominal 95% at all proportions of missing data but not when clustering was ignored.

Table 3.1 Simulation results: average causal mediation effect (ACME) estimates when no imputation is conducted and after multiple imputation using different approaches in longitudinal data for differing proportions of missing data.

Prop.	Method	Missing completely at random (MCAR)						Missing at random (MAR)					
		Estimate	Bias	MSE	MC-SD	MS-SE	CP	Estimate	Bias	MSE	MC-SD	MS-SE	CP
Complete		-2.16	0.00	0.17	0.41	0.40	0.95	-2.16	0.00	0.17	0.41	0.40	0.95
10%	No imputation	-2.18	-0.02	0.18	0.42	0.42	0.96	-2.26	-0.10	0.19	0.43	0.42	0.94
	IC	-1.49	0.67	0.57	0.35	0.38	0.58	-1.24	0.92	0.98	0.37	0.34	0.26
	LM	-2.16	0.00	0.18	0.42	0.42	0.96	-2.24	-0.08	0.19	0.43	0.42	0.94
	CC	-2.16	0.00	0.17	0.41	0.41	0.96	-2.13	0.03	0.17	0.42	0.41	0.95
20%	No imputation	-2.15	0.01	0.20	0.45	0.45	0.95	-2.28	-0.12	0.22	0.45	0.45	0.94
	IC	-1.07	1.09	1.26	0.28	0.34	0.11	-0.81	1.35	1.92	0.30	0.30	0.02
	LM	-2.11	0.05	0.20	0.45	0.45	0.95	-2.25	-0.09	0.21	0.45	0.45	0.94
	CC	-2.12	0.04	0.18	0.42	0.42	0.95	-2.07	0.09	0.19	0.43	0.43	0.94
30%	No imputation	-2.18	-0.02	0.23	0.48	0.48	0.95	-2.35	-0.19	0.28	0.49	0.49	0.94
	IC	-0.83	1.33	1.82	0.25	0.31	0.01	-0.58	1.58	2.56	0.26	0.27	0.00
	LM	-2.13	0.03	0.23	0.48	0.47	0.94	-2.29	-0.13	0.27	0.50	0.48	0.93
	CC	-2.13	0.03	0.20	0.44	0.44	0.94	-2.08	0.08	0.21	0.45	0.44	0.94
40%	No imputation	-2.16	0.00	0.28	0.53	0.52	0.94	-2.36	-0.20	0.33	0.54	0.54	0.94
	IC	-0.64	1.52	2.35	0.22	0.28	0.00	-0.42	1.74	3.09	0.23	0.25	0.00
	LM	-2.07	0.09	0.29	0.53	0.51	0.92	-2.26	-0.10	0.31	0.54	0.52	0.93
	CC	-2.08	0.08	0.23	0.48	0.46	0.93	-2.04	0.12	0.25	0.49	0.47	0.93

CC, Considering clustering; CP, Coverage probability; IC, Ignoring clustering; LM, Linear model; MC-SD, Monte-Carlo standard deviation; MSE, Mean square error; MS-SE, Mean Sobel standard error; Prop, Proportion of missing data. True ACME = -2.16, sample size = 300, and number of Monte-Carlo data sets = 2000.

When data is MCAR, average direct effect (ADE) estimators with no imputation, considering clustering in the imputation models and the LM method

were approximately unbiased but not for ignoring clustering (Table 3.2). Considering clustering gave the lowest MSE at all proportions of missing data, in fact, similar to when data is complete. On the other hand, when clustering was ignored, MSEs were higher than other methods, over eight times the MSE for complete data. Coverage probabilities were close to 95% at all proportions of missing data for all methods except for ignoring clustering.

When data were MAR, considering clustering performed better than all other methods in terms of bias, MSE and coverage rates (Table 3.2). ADE estimators for no imputation, ignoring clustering, and the LM method were biased at all proportions of missing data. In fact, ignoring clustering in the imputation model had the most biased ADE estimators and the bias increased as the proportion of missing data increased. MSEs for considering clustering were similar to when data was complete, and increased slightly as the proportion of missing data increased. Coverage probabilities for all methods except considering clustering were below the nominal 95% at all proportions of missing data. As the proportion of data MAR increased, coverage probabilities for ignoring clustering decreased.

In this simulation, we first imputed missing data for both the outcome and the mediator and then computed $M_{i6} - M_{i1}$ using both observed and imputed data. We also conducted a simulation where $M_{ij} - M_{i1}$ were first calculated and then missing values for the outcome Y_{ij} and $M_{ij} - M_{i1}$ imputed, after which ACME and ADE were estimated. Results were similar for ACME although the bias was lower for ignoring clustering than under the former approach. ADE estimators for all methods (no imputation, ignoring clustering, LM method, and considering clustering) were approximately unbiased when data was MCAR. However, except considering clustering, ADE estimators for all other methods were biased when data was MAR (appendix B).

Table 3.2 Simulation results: average direct effect (ADE) estimates when no imputation is conducted and after multiple imputation using different approaches in longitudinal data for differing proportions of missing data.

Prop.	Method	Missing completely at random (MCAR)						Missing at random (MAR)					
		Estimate	Bias	MSE	MC-SD	Avg-SE	CP	Estimate	Bias	MSE	MC-SD	Avg-SE	CP
Complete		-1.49	0.01	0.07	0.27	0.27	0.95	-1.49	0.01	0.07	0.27	0.27	0.95
10%	No imputation	-1.51	-0.01	0.08	0.28	0.29	0.95	-1.37	0.13	0.10	0.28	0.28	0.92
	IC	-2.20	-0.70	0.62	0.37	0.43	0.66	-2.54	-1.04	1.23	0.38	0.43	0.30
	LM	-1.51	-0.01	0.08	0.28	0.29	0.95	-1.37	0.13	0.10	0.28	0.28	0.92
	CC	-1.52	-0.02	0.08	0.27	0.28	0.95	-1.54	-0.04	0.08	0.27	0.28	0.95
20%	No imputation	-1.51	-0.01	0.09	0.31	0.31	0.94	-1.29	0.21	0.13	0.30	0.29	0.89
	IC	-2.59	-1.09	1.41	0.46	0.49	0.39	-2.86	-1.36	2.04	0.44	0.48	0.18
	LM	-1.51	-0.01	0.10	0.31	0.31	0.94	-1.29	0.21	0.13	0.30	0.30	0.88
	CC	-1.54	-0.04	0.08	0.29	0.28	0.94	-1.56	-0.06	0.09	0.29	0.29	0.94
30%	No imputation	-1.48	0.02	0.10	0.32	0.33	0.95	-1.21	0.29	0.18	0.31	0.32	0.85
	IC	-2.86	-1.36	2.08	0.49	0.54	0.27	-3.03	-1.53	2.59	0.48	0.53	0.16
	LM	-1.48	0.02	0.11	0.33	0.33	0.95	-1.21	0.29	0.18	0.32	0.32	0.85
	CC	-1.54	-0.04	0.08	0.28	0.29	0.95	-1.57	-0.07	0.09	0.29	0.30	0.95
40%	No imputation	-1.49	0.01	0.13	0.36	0.35	0.95	-1.17	0.33	0.23	0.35	0.35	0.84
	IC	-3.05	-1.55	2.72	0.56	0.59	0.24	-3.13	-1.63	2.99	0.56	0.59	0.20
	LM	-1.49	0.01	0.13	0.36	0.36	0.94	-1.17	0.33	0.23	0.35	0.35	0.83
	CC	-1.57	-0.07	0.09	0.29	0.30	0.96	-1.59	-0.09	0.10	0.30	0.31	0.94

CC, Considering clustering; CP, Coverage probability; IC, Ignoring clustering; LM, Linear model; MC-SD, Monte-Carlo standard deviation; MSE, Mean square error; Avg-SE, Average standard error; Prop, Proportion of missing data. True ADE = -1.50, sample size = 300 and number of Monte-Carlo data sets = 2000.

3.4.2 Data application example: Severe malaria and cognition data

In this longitudinal study, 70 children with cerebral malaria (CM) and 83 community control (CC) children 18 months to 5 years of age were enrolled to study the effects of iron deficiency and severe malaria on neurocognitive development. All children with CM were iron deficient while only 35 out of 83 CC children were iron deficient at baseline. Children with severe malarial anemia (SMA) were also enrolled in the study, but excluded from this analysis because their inclusion was restricted on the mediator (hemoglobin levels ≤ 5 mg/dl). Community control children were recruited from the nuclear family, neighbourhood, or extended family of children with CM or SMA and were within 1 year in age of a recently enrolled CM or SMA child but not matched to CM or SMA children. Cognitive ability was measured using the Mullen Scales of Early Learning (Mullen et al., 1995). Sub-scale scores for visual reception, fine motor, expressive language, and receptive language were summed

to give the Mullen composite score, a measure of overall cognitive ability. Measurements of the outcome (cognitive ability) and mediator (hemoglobin) were made at baseline (0 months), 6 months and 12 months. Variables predictive of missingness for hemoglobin and cognitive ability included socioeconomic status, baseline weight-for-height, mother's education level, and whether the child had any education yet. One CC child who was missing values for two predictive variables for missingness (socioeconomic status and any child education) was excluded from the analysis.

CM has been associated with long-term neurocognitive impairment (Bangirana et al., 2014; Boivin et al., 2007; John et al., 2008) and hemoglobin level has been suggested to mediate the effect of malaria on neurocognition (Boivin et al., 2016). In this study, data can be missing for both the outcome and mediator at all time points. Percentages of missing data by variable are displayed in Table 3.3.

Table 3.3 Percent (%) missing at each time point for model variables (N = 152).

Variable	Baseline (0 months)	6 months	12 months
Cognitive ability†	1	11	20
Hemoglobin	1	12	7
Age	1	2	3

† Cognitive ability, the outcome, is a composite score that is a sum of 4 scales, visual reception, fine motor, expressive language and receptive language.

We sought to examine the mediation effect of sustained levels of hemoglobin early in follow-up (measured by the average of hemoglobin levels at baseline and 6 months) on the effect of CM on changes in neurocognition from baseline to 12 months. First, we analyzed the data using linear regression in R without any imputation for missing data. Second, we imputed imputed missing data values ignoring clustering using MICE (Buuren and Groothuis-Oudshoorn, 2011). Third, we imputed missing values assuming a linear imputation model for average hemoglobin and change in neurocognition and lastly, we imputed

missing values considering the clustering in the data. Following imputation, missing values for average hemoglobin and change in neurocognition were calculated using observed and imputed values. Ten data sets were imputed for each imputation method and pooled results are displayed in Table 3.4. For all four methods except for ignoring clustering (which found borderline significance, p=0.093), the estimated ACMEs were significant. For example, for considering clustering, the ACME = 5.14 (S-SE = 2.56, p = 0.045) captures the effect on changes in neurocognition as a result of CM, but only through CM's effect on average hemoglobin. ACME estimates varied across methods. No imputation assumes data is MCAR, which may not be the case. The LM method does not use all the data during imputation and ignoring clustering assumes all observations within subject are independent, yet they are correlated. With non-significant direct effects for all methods, these results suggest that the effect of CM on changes in neurocognition is possibly mediated by initial levels of hemoglobin. The standard error for ACME when clustering is ignored is slightly inflated compared to other methods in the imputation model. These results suggest possible mis-estimation of standard errors when clustering is ignored in the imputation model.

Table 3.4 Malaria-cognition data: average causal mediation effect (ACME) and average direct effect (ADE) estimates after multiple imputation using a linear model, considering and ignoring clustering.

Imputation method	ACME		ADE	
	Estimate (S-SE)	p value	Estimate (SE)	p value
No imputation	8.40 (2.94)	0.004	-5.63 (3.76)	0.137
Ignoring Clustering	5.35 (3.19)	0.093	-2.95 (4.27)	0.490
LM imputation	7.05 (2.68)	0.009	-4.76 (3.49)	0.173
Considering clustering	5.14 (2.56)	0.045	-3.54 (3.35)	0.291

LM, Linear model; SE, Standard error; S-SE, Sobel standard error.

3.5 Discussion

The main objective of this Chapter was to evaluate the effects on mediation effect estimates of misspecifying the imputation model by ignoring clustering when data is clustered due to repeated measures on the same subject. We found that under the scenarios considered, ACME and ADE estimators for ignoring clustering when data is MCAR or MAR are biased at all moderate rates of missing data. Considering clustering in the imputation model performed better than ignoring clustering at all rates of missing data in terms of bias, MSE, and coverage probability. The magnitude of bias in the ACME and ADE estimators when clustering was ignored increased with increasing rates of missing data, while considering clustering provided approximately unbiased ACME and ADE estimators at rates of missing data of up to 40%. ACME and ADE estimators for no imputation and the LM method were approximately unbiased when data was MCAR. To impute missing data ignoring clustering, we used the MICE R package (Buuren and Groothuis-Oudshoorn, 2011) which assumes independent observations, suitable for normal models such as linear regression. This is inconsistent with correlated observations, which are common in longitudinal studies.

Our results are consistent with previous findings for cluster randomized trials, where considering clustering gave approximately unbiased fixed effect estimates compared to ignoring clustering (Black et al., 2011), although at 10% rate of MAR, fixed effects from multiply imputed data ignoring clustering were generally unbiased and comparable in accuracy to those considering clustering. In the same study, as rates of missing data increased under a MAR mechanism, accuracy of fixed effects decreased and the within-subject variance was overestimated when clustering was ignored (Black et al., 2011). The overestimation of within-subject variance was attributed to underestimation of the between-subject variance (Black et al., 2011; Taljaard et al., 2008). Although

the imputation methods varied, two-level clustered data was generated and analyzed using a linear mixed model.

Our simulation is not congenial in the sense of Meng (1994) since the analysis model cannot be derived from the imputation model. However, the statistical community agrees that for better inference, data relationships and characteristics within and between clusters need to be preserved during the imputation phase. A difference in the current study is that data was simulated using a linear mixed model but analyzed using a linear model. In addition, data that was MAR was set to depend on both the outcome and the mediator. Compared to results for complete data, ACME standard error estimates when clustering was ignored were consistently lower for all missing data proportions. Black et al. (2011)'s results in cluster randomized trials, indicating potential bias in variance estimates when the multiple imputation model is misspecified, are consistent with the current results where ACME standard errors are underestimated. In estimating the mediation effect of $M_{i6} - M_{i1}$, these results suggest that ignoring clustering in the imputation model can cause underestimation of ACME standard errors. To compute the ACME, we compute the product of two fixed effects. When clustering is ignored in the imputation model, the bias inherent in each of the fixed effect estimates is amplified in their product. When fixed effects are of interest and there is a high rate of missingness, or when inferences on variance components are of interest, studies conducted in cluster randomized trials discourage ignoring clustering in the imputation model (Black et al., 2011).

Another method that has been used to account for clustering in the imputation model is including indicators of cluster membership as fixed effects. This drastically increases the number of fixed effects to estimate in longitudinal studies because such studies can have hundreds of individuals considered as clusters. Moreover, including dummy variables indicating cluster membership

in the context of CRTs has been found to lead to severe over-estimation of variance of group means, and that the over-estimation is severe for small cluster sizes and small intra-class correlation (Andridge, 2011). A hybrid strategy where half of the time one imputes from a model with cluster dummy variables as fixed effects and the other half multiply imputing from normal model imputation has also been tested. Using the MCMC approach to allow for clustering still performed better than this hybrid strategy (Zhou et al., 2014).

In this study, we imputed a continuous outcome and mediator. A limitation in using PAN to account for clustering in the imputation model for longitudinal data is that it uses a multivariate linear mixed effects model. When categorical data is missing and treated as multivariate normal in the PAN model, it is unclear what kind of biases this can introduce. Horton et al. (2003) point out that rounding off multivariate normal imputed values for dummy variables to make them plausible can cause more bias than using the originally imputed implausible value. On the other hand, one could use probability mean matching (pmm), available in MICE, though the bias associated with 'pmm' for categorical variables in clustered data needs more investigation.

We have demonstrated that ignoring clustering in the imputation model when data is MCAR or MAR provides biased ACME and ADE estimates in longitudinal studies. To get unbiased ACME and ADE estimates and good coverage probabilities at moderate rates of missing data, considering clustering under the linear mixed effects model in the imputation model is recommended.

Chapter 4

Mediation effects of malaria prevention in pregnant women on child neurodevelopment: a randomized controlled trial

4.1 Introduction

Children under five years are particularly vulnerable to malaria illness, infection and death. In 2015, out of the estimated 303,000 malaria deaths globally in children under 5 years, 96% were from the African region (WHO, 2015). The effects of malaria on childhood development may start *in utero*. Particularly, pregnant women are more at risk of *Plasmodium falciparum* malaria infection than non-pregnant women (Brabin, 1983; Rogerson et al., 2007). Malaria infection in pregnancy leads to placental malaria (PM) (Rogerson et al., 2007) and PM, seen in 26-28% (Guyatt and Snow, 2004; Steketee et al., 2001) of all pregnant women, has been associated with low levels of hemoglobin, low birth weight and pre-term birth (Conroy et al., 2013; Desai et al., 2007; Kapisi et al., 2017; Katz et al., 2013; Rogerson et al., 2007, 2003). In malaria endemic areas,

malaria in pregnancy accounts for 8-36% of low birth weight babies (Steketee et al., 2001). Moreover, malaria is an important cause of maternal anemia (Desai et al., 2007), and anemia is an independent risk factor for low birth weight and fetal growth restriction (Hendrix and Berghella, 2008). Animal models suggest that maternal infection can lead to impairment in neurologic function in the offspring, even in the absence of fetal infection (Ozawa et al., 2006; Shi et al., 2005; Urakubo et al., 2001). Maternal systemic and placental changes such as placental inflammation (Muehlenbachs et al., 2006) and PM that occur with malaria illness in pregnancy can adversely affect the developing fetal brain, and fetal brain injury affects long term childhood neurodevelopment (ND). The inflammatory response against malaria triggered in pregnant mothers reduces the amount of key nutrients in maternal circulation and thus available for transfer to the fetus, leading to functional deficiencies of several key micronutrients, notably iron and zinc, that are essential to brain development in the first 1000 days of life (Nelson et al., 2001; Okoko et al., 2003). Thus, preventing malaria during pregnancy is fundamental.

In areas in Africa with moderate to high malaria transmission, the World Health Organization (WHO) recommends intermittent preventive treatment in pregnancy (IPTp) with sulfadoxine-pyrimethamine (SP), as part of antenatal care services (WHO, 2016). However, some studies in infants and pregnant women found an increased risk of anemia among those who received intermittent preventive treatment (IPT) with SP (Bigira et al., 2014; Gosling et al., 2009; Harrington et al., 2011). In addition, one study that compared the effectiveness between IPTp with SP and insecticide treated nets (ITNs) found no differences in low birth weight (LBW), maternal anemia, placental parasitemia, and peripheral parasitemia (Ndyomugyenyi et al., 2011). WHO also recommends that all pregnant women receive iron and folic acid supplementation as a part of routine antenatal care (WHO, 2016). However, folic acid can interfere with the efficacy of SP prophylaxis because SP inhibits dihydro-

folate reductase and, therefore, dihydroartenisinin-piperaquine (DP) prophylaxis, presumably unaffected by folic acid, could be an alternative option. IPT with DP in children has shown promising results in protecting them against clinical malaria, parasitemia, and anemia (Nankabirwa et al., 2014), though the effects of its use in pregnancy on the children's ND is not well understood.

In a high malaria transmission setting, IPTp with DP has shown superiority in lowering the burden of malaria among pregnant women compared to IPTp with SP, with higher doses of DP providing more protection (Kakuru et al., 2016). Malaria infection that spreads to the placenta (Rogerson et al., 2007), coupled with maternal anemia, can cause preterm birth and intrauterine growth retardation (IUGR) (Conroy et al., 2013; Lagerberg, 2008); individually or in combinations these effects can affect a child's ND. Asymptomatic or severe forms of malaria in children have been associated with deficits in neurodevelopment (Al Serouri et al., 2000; Boivin et al., 2007; John et al., 2008; Nankabirwa et al., 2013) while children with lower hemoglobin have been associated with poorer test scores (Al Serouri et al., 2000; Fernando et al., 2003). Although one study found that anemia mediates the effect of repeated episodes of malaria illness on cognitive development in young children (Boivin et al., 2016), the role of cord blood hemoglobin in mediating the effect of maternal asymptomatic or symptomatic malaria during pregnancy on their child's ND is unclear. In this study, we expect that the protective effect of IPTp causes placental changes that reduce effects of malaria in pregnancy such as placental malaria, preterm birth, low birth weight, and low levels of cord blood hemoglobin levels, thus improving ND in the offspring.

Formally, we hypothesize that maternal malaria infection or placental malaria, adverse birth outcomes measured by preterm birth, LBW or small for gestation age (SGA), cord blood hemoglobin levels, and child malaria (referred to as mediators) mediate the effect of IPTp on children's ND. The purpose of

this study, therefore, was to estimate indirect (mediation) effects of IPTp in pregnant women on child ND.

4.2 Methods

4.2.1 Study description

The data are from a prospective longitudinal cohort study (PROTECT) conducted in Tororo District, Uganda. Tororo is in Eastern Uganda where malaria is highly endemic throughout the year with an estimated entomological inoculation rate of 310 infectious bites per person year in 2012 (Kamya et al., 2015). This study was nested within a randomized, double-blinded, controlled trial (PROMOTE-II) (Kakuru et al., 2016) of which the objective was to evaluate promising interventions to reduce the burden of malaria among pregnant women and improve maternal child health. In the parent study (PROMOTE-II), pregnant women presenting for routine care at the Tororo District hospital antenatal clinic were approached for enrollment between June and October 2014. HIV uninfected women between 12-20 weeks of gestation, 16 years of age or older, living within 30 kilometers of the study clinic, and intending to deliver at the hospital were enrolled into the PROMOTE-II study. They were randomized to receive either three-dose SP, three-dose DP or monthly DP for IPTp. All doses of SP administered were directly observed in the clinic and the first of three doses of DP were directly observed in the clinic. The second and third doses of DP were administered at home. Children born to mothers randomized to receive three-dose SP during pregnancy received DP every three months between 2-24 months of age. Children born to mothers randomized to receive three-dose DP or monthly DP during pregnancy received either DP every three months or monthly DP between 2-24 months of age. Pregnant women who had a history of serious adverse events to SP or DP, had a chronic condition requiring frequent medication, had prior SP prevention therapy or

any other antimalarial therapy during pregnancy, or intended to move more than 30 kilometers from the study clinic were excluded from PROMOTE-II study. In the current study, children (study participants) born to mothers in the PROMOTE-II study were enrolled if they were HIV uninfected, 12 months of age, and within 30 km of the clinic. The exclusion criteria for the children included serious adverse events to study drugs, active illness at enrollment, history of head trauma or coma, cerebral palsy or other neurologic disease, known chronic illness requiring medical care, major medical abnormalities, and known development delay. All mothers received a long lasting insecticide-treated bed net (LLIN) at birth.

4.2.2 Neurodevelopment outcome and mediator measurements

Children's ND was assessed at 12 and 24 months using the Bayley scales of Infant and Toddler Development (third edition) (Bayley, 2006). The Bayley scales are easy to use to test children from ages 1 month to 42 months and are widely used to assess neuropsychological domains in children. They include tests of cognitive ability, expressive language, receptive language, fine motor and gross motor. Potential mediators of maternal IPTp included placental malaria, adverse birth outcomes, cord blood hemoglobin, maternal malaria measured monthly throughout pregnancy, and child malaria measured monthly throughout followup. Measures of placental malaria at delivery included detection of malaria parasites in placental blood by both microscopy and loop-mediated isothermal amplification (LAMP), and histopathologic evidence of placental malaria (parasites and/or pigment) from placental biopsies (Kakuru et al., 2016; Kapisi et al., 2017). A child had an adverse birth outcome if they either had a pre-term birth ($<$ 37 weeks gestation age), were small for gestation age ($<$10th percentile) or had a low birth weight ($<$2.5 kilograms according to East African fetal weight standards) (Schmiegelow et al., 2012).

4.2.3 Laboratory procedures

At enrollment, mothers were evaluated for *Plasmodium* parasitemia and thereafter were evaluated monthly throughout their pregnancy. At birth, biological specimens were collected, including placental tissue and placental blood, to determine their placental malaria status. Pregnant women were encouraged to deliver in the hospital that was adjacent to the study clinic and when they delivered at home, they were visited by the study staff at delivery or soon afterwards. Blood smears were stained with 2% Giemsa and read by trained laboratory technologists not involved in patient care. A blood smear was considered negative when the examination of 100 high-power fields did not reveal asexual parasites. All slides were read by a second reader and a third reader settled any discrepancy between readings. Placental tissues were processed for histologic evidence of placental malaria as described previously (Natureeba et al., 2014). Histological slides were read in duplicate by two trained independent readers, and the results were recorded on a standardized case-record form; any discrepant results were resolved by a third reader. Placental malaria was defined as presence of infected erythrocytes in the intervillious space (IVS) by histological examination of placental biopsies. For mothers who gave birth to twins, delivery outcomes were based on whether the outcome was present in either child or in the placenta. Children who had a temperature $> 38.0°C$ or reported a history of fever in the past 24 hours were tested for malaria. Blood was obtained by finger prick for a thick blood smear in very young children, but heel pricks were used too to substitute for finger pricks.

4.2.4 Ethical approval

Informed consent was obtained from the caregivers of all children at enrollment. The Makerere University School of Medicine Research and Ethics Committee, the Indiana University Institutional Review Board, and the Uganda National Council of Science and Technology reviewed and approved the PRO-

TECT study. PROMOTE-II, the parent study study was also approved by the University of California, San Francisco, Committee on Human Research.

4.3 Randomization

4.3.1 Allocation and blinding

Pregnant women were randomized in a ratio of 2:1:1:1:1 to one of the five treatment arms A-E in Table 4.1 respectively. A randomization list was com-

Table 4.1 Malaria intermittent preventive treatment (IPT) assignment during pregnancy and infancy.

Phase of IPT	Intermittent preventive treatment arm				
	A	B	C	D	E
During pregnacy	3 dose SP	3 dose DP	3 dose DP	Monthly DP	Monthly DP
During infancy	3 monthly DP	3 monthly DP	Monthly DP	3 monthly DP	Monthly DP

SP, Sulfadoxine-Pyrimethamine; DP, Dihydroartemisinin-Piperaquine.

puter generated by a project member not directly involved in the conduct of the study using permuted blocks of size 6 and 12. Prior to the PROMOTE-II study, a set of sequentially numbered, opaque, sealed envelopes were prepared. Each envelope was marked with a treatment allocation number and the envelope contained a piece of paper with both the treatment allocation number and the IPTp group assignment. IPTp allocation to pregnant women was done by the responsible study pharmacist. All study drugs were prepackaged by a study pharmacist and administered by a study nurse blinded to the pregnant women's IPTp group. Study personnel who conducted ND testing for the children were blinded to both the mother's IPTp and child's chemopreventive malaria prophylaxis.

4.3.2 Statistical methods

Characteristics for the pregnant women and children born to them were summarized using frequencies (percent) for categorical data and means (standard

deviations) for continuous data. Data were analyzed using the intention-to-treat approach. IPTp effects on each mediator were estimated by regressing each mediator on IPTp using linear regression for cord blood hemoglobin and logistic regression for binary mediators. Mediator effects on ND outcomes were estimated by regressing each ND outcome on each mediator by fitting linear regression models.

To estimate the mediation effects, the effect of IPTp on ND through each mediator, we fitted two models; 1) regressed each mediator on IPTp, adjusting for the child's malaria chemopreventive prophylaxis and 2) separately for each ND outcome measure at 12 and 24 months, we regressed each outcome on IPTp, adjusting for the child's chemopreventive prophylaxis, a mediator, and we included a mediator-IPTp interaction term. Logistic regression models were fitted for the mediators maternal malaria, placental malaria, adverse birth outcome, and child malaria, while linear regression models were fitted for the mediator cord blood hemoglobin and for ND outcome measures. Mediation effects were then computed using the product method (Baron and Kenny, 1986; Valeri et al., 2014) (supplementary materials C.5). Mediation effects of three-dose DP or monthly DP during pregnancy on child ND outcome measures through the mediators were computed among maternal three-doses DP and monthly DP study groups. We present 95% percentile bootstrap confidence intervals for the indirect effects. Some pregnant women with placental malaria were not diagnosed with malaria during pregnancy and vice versa. A binary mediator "any maternal malaria" (either placental or maternal malaria or both) versus neither, was created and mediation effects estimated.

Due to potential interaction between IPTp and child malaria prophylaxis in preventing child malaria, an exposure variable (mother-child prophylaxis) that included different mother-child prophylaxis combinations was created. Mediation effects of any maternal malaria (M_1) and child malaria (M_2) were esti-

mated concurrently by 1) regressing any maternal malaria or child malaria on mother-child prophylaxis and 2) separately for each ND outcome measure at 24 months, we regressed each outcome on mother-child prophylaxis, adjusting for any maternal malaria and for child malaria. We estimated the total indirect effect and mediation effects through only M_1, through only M_2, through both M_1 and M_2 (supplementary materials C.6) with their 95% percentile bootstrap confidence intervals. Malaria episodes by the 12-month followup were very few and, therefore, 12-month outcomes were excluded from this analysis. The analysis was for hypothesis generation and, therefore, no adjustment was done for multiple testing. All statistical analyses were conducted using R version 3.2.2 (R Core Team, 2015).

4.4 Results

4.4.1 Trial profile, baseline, during pregnancy and at birth characteristics

We screened 386 pregnant women, 300 were randomized to one of three malaria prophylaxes, and 289 were followed up to delivery. There were 291 live births out of 297 births and 272 children of randomized pregnant women were enrolled and tested at 12 months. Among the live births, 6 died, 3 withdrew consent, 3 moved out of the study area, 4 were unable to be located > 60 days, 2 were unable to comply with the study protocol, and 1 declined to participate in the study. Mothers of 94, 87, and 91 children had been randomized to receive three-dose SP, three-dose DP, and monthly DP respectively (Figure 4.1). Baseline and clinical characteristics at enrollment were similar across IPTp groups (Table 4.2). Among the 272 children who were enrolled at 12 months, 251 (92.3%) were followed and tested at 24 months (Figure 4.1).

Thirty three percent (31/94) of pregnant women who received three-dose SP

had at least one malaria episode during pregnancy compared to 16.1% and 11.0% among those who received three-dose DP and monthly DP respectively. Placental malaria at birth was highest among pregnant women who received three-dose SP (51.6%) compared to 36.0% and 27.8% among those who received three-dose DP and monthly DP groups respectively. There was no significant difference in adverse birth outcomes (low birth weight, small for gestation age or preterm birth) between IPTp groups (P = 0.744). Any maternal malaria (i.e. either maternal or placental malaria) was significantly higher among pregnant women who received three-dose SP compared to other IPTp groups (P = 0.002) (Table 4.2). Neurodevelopment outcomes at both 12 and 24 months did not differ across IPTp groups (Table 4.3).

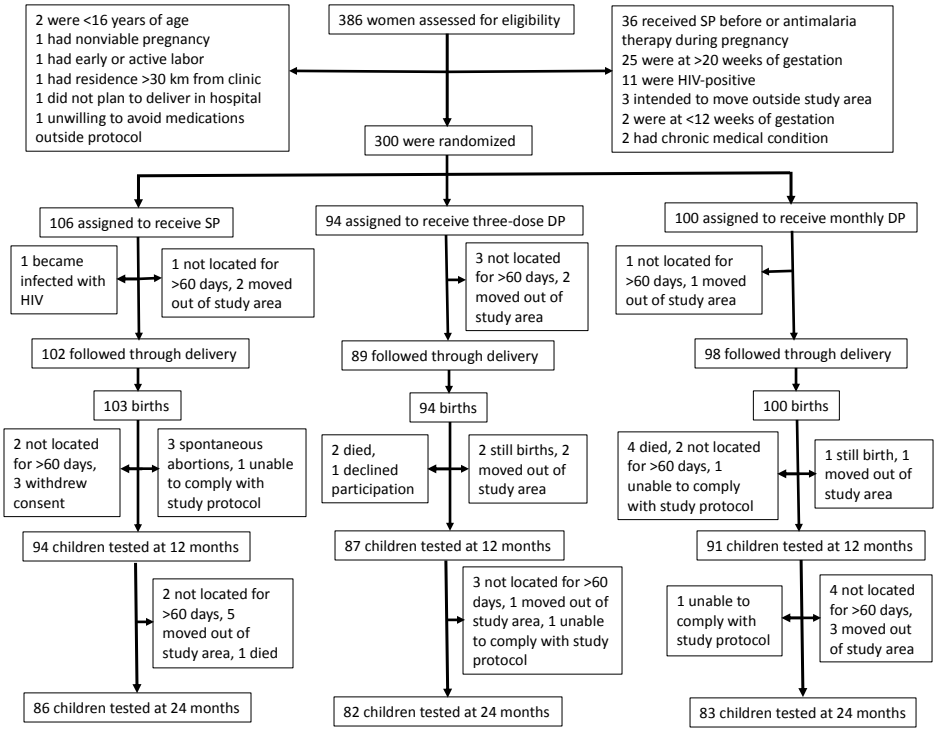

Figure 4.1 Enrollment, randomization and follow-up for study participants and their mothers.

Table 4.2 Baseline and clinical characteristics of study participants and their mothers.

Characteristics	Three-dose SP N=94	Three-dose DP N=87	Monthly DP N=91	P
At enrollment				
Child's sex, female (%)	48 (51.1)	42 (48.3)	46 (50.5)	0.925
Highest education level, no. (%)				0.348
None	6 (6.4)	3 (3.4)	1 (1.1)	
Primary	70 (74.5)	66 (75.9)	64 (70.3)	
O level	16 (17.0)	17 (19.5)	25 (27.5)	
A level and higher	2 (2.1)	1 (1.1)	1 (1.1)	
Gestational age in weeks, mean (sd)	15.16 (2.05)	15.36 (1.90)	15.30 (1.94)	0.776
Gravidity, no.(%)				0.973
1	34 (36.2)	30 (34.5)	32 (35.2)	
2	30 (31.9)	25 (28.7)	28 (30.8)	
≥3	30 (31.9)	32 (36.8)	31 (34.1)	
Household wealth index, no. (%)				0.471
Lowest tertile	34 (36.2)	28 (32.2)	32 (35.2)	
Middle tertile	28 (29.8)	35 (40.2)	26 (28.6)	
Highest tertile	32 (34.0)	24 (27.6)	33 (36.3)	
Prior malaria episodes, mean (sd)	0.09 (0.28)	0.03 (0.18)	0.11 (0.31)	0.160
Hemoglobin in g/dl at enrollment, mean (sd)	11.87 (1.47)	11.84 (1.11)	12.01 (1.41)	0.660
During pregnancy or at birth				
Total episodes of malaria, mean (sd)	0.46 (0.74)	0.17 (0.41)	0.11 (0.31)	<0.001
Malaria episodes diagnosed, any (%)	31 (33.0)	14 (16.1)	10 (11.0)	<0.001
Maternal blood microscopy, no.(%)				0.139
Negative	90 (95.7)	86 (98.9)	90 (98.9)	
Positive	4 (4.3)	1 (1.1)	0 (0.0)	
Results missing	0 (0.0)	0 (0.0)	1 (1.1)	
Any placental malaria, Yes (%)	47 (51.6)	31 (36.0)	25 (27.8)	0.004
Positive samples by LAMP, any (%)	88 (93.6)	69 (79.3)	66 (72.5)	0.001
Cord blood hemoglobin in g/dl, mean (sd)	14.29 (1.93)	13.65 (2.71)	14.51 (2.00)	0.040
Maternal anemia at birth (Hb<11 g/dl), Yes (%)	16 (17.4)	16 (18.4)	15 (16.5)	0.945
Birth weight in kg, mean (sd)	2.95 (0.45)	2.90 (0.53)	2.94 (0.38)	0.754
Preterm birth, Yes (%)	8 (8.5)	13 (14.9)	4 (4.4)	0.050
Small for gestation age (<10th percentile), Yes (%)	24 (25.5)	14 (16.1)	26 (28.6)	0.124
Adverse birth outcome, Yes (%)†	32 (34.0)	25 (28.7)	29 (31.9)	0.744
Any maternal malaria, Yes (%)‡	57 (61.3)	38 (44.2)	32 (35.6)	0.002

g/dl, grams per deciliter; Hb, Hemoglobin; kg, Kilograms; LAMP, Loop-mediated isothermal amplification; sd, Standard deviation; SP, Sulfadoxine-Pyrimethamine; DP, Dihydroartemisinin-Piperaquine.

†Composite for children who either had a low birth weight (< 2.5kg), were small for their gestation age (<10th percentile) or whose birth was preterm.

‡Includes mothers who either had maternal or placental malaria versus neither of the two.

Table 4.3 Neurodevelopment outcomes at 12 and 24 months follow up by IPTp group.

Neurodevelopment measures	Three-dose SP N=94	Three-dose DP N=87	Monthly DP N=91	P
At 12 months				
Cognitive ability, mean (sd)	10.63 (2.15)	10.74 (2.73)	11.01 (2.58)	0.560
Expressive language, mean (sd)	11.46 (1.64)	11.59 (1.90)	11.55 (1.49)	0.867
Receptive language, mean (sd)	8.40 (2.43)	8.95 (2.10)	8.36 (2.09)	0.142
Fine motor, mean (sd)	11.59 (2.52)	11.05 (2.79)	11.56 (2.81)	0.326
Gross motor, mean (sd)	11.01 (1.89)	10.61 (2.69)	11.04 (2.00)	0.347
At 24 months	N=86	N=82	N=83	
Cognitive ability, mean (sd)	6.55 (1.75)	6.56 (1.81)	6.93 (1.37)	0.243
Expressive language, mean (sd)	8.28 (1.63)	8.22 (1.63)	8.20 (1.52)	0.950
Receptive language, mean (sd)	8.03 (2.12)	7.68 (2.10)	8.19 (2.16)	0.290
Fine motor, mean (sd)	9.15 (1.48)	9.01 (1.64)	9.17 (1.91)	0.807
Gross motor, mean (sd)	9.13 (2.57)	8.71 (2.78)	9.47 (2.89)	0.205

DP, Dihydroartemisinin-Piperaquine; sd, Standard deviation; SP, Sulfadoxine-Pyrimethamine.

4.4.2 Effects of malaria prophylaxis on mediators and mediator effects on neurodevelopment outcomes

We regressed each mediator on IPTp to test their association. Compared to three-dose SP, three-dose DP was protective against maternal malaria ($P = 0.01$), placental malaria ($P = 0.04$), and child malaria until 24 months ($P = 0.04$). On the other hand, compared to three-dose SP, monthly DP was protective against maternal malaria ($P < 0.01$) and placental malaria ($P < 0.01$) (supplementary materials C.1).

We tested the effect of each mediator (maternal malaria, placental malaria, adverse birth outcomes, cord blood hemoglobin, and child malaria) on each ND outcome (cognitive ability, expressive language, receptive language, fine motor, and gross motor). At 12 months, no mediator was associated with any ND outcome measure. At 24 months, having an adverse birth outcome was associated with lower child scores in expressive language ($P = 0.01$) and fine motor ($P = 0.02$). Other associations did not reach statistical significance (supplementary materials C.2).

4.4.3 Mediation effects of maternal malaria prophylaxis on child neurodevelopment outcomes

We estimated the effect of three-dose DP or monthly DP on the five ND outcome measures mediated by each potential mediator (maternal malaria, placental malaria, adverse birth outcomes, cord blood hemoglobin, and child malaria). The IE of IPTp on a ND outcome would be the change in ND resulting from IPTp's effect on the mediator.

At 12 months, placental malaria mediated the effect of three-dose DP (compared to three-dose SP) on cognitive ability (Effect = 0.87, 95% CI 0.20, 1.54), and the effect of monthly DP on expressive language (Effect = -0.50, 95% CI -1.06, -0.05). The positive mediation effect suggests that on average, a child's cognitive ability improves by 0.87 as a result of three-dose DP's effect on placental malaria during pregnancy compared to three-dose SP, while the negative mediation effect suggests that on average, a child's expressive language declines by -0.50 as a result of monthly DP's effect on placental malaria. Maternal malaria mediated the effect of three-dose DP on expressive language (Effect = -1.28, 95% CI -1.95, -0.67). Maternal malaria, placental malaria, adverse birth outcomes, or cord blood hemoglobin did not mediate the effects of three-dose DP or monthly DP on receptive language, fine motor and gross motor (Figure 4.2).

At 24 months, maternal malaria mediated the effect of three-dose DP (compared to three-dose SP) on fine motor (Effect = 0.79, 95% CI 0.15, 1.56) and placental malaria mediated the effect of three-dose DP on expressive language (Effect = 0.68, 95% CI 0.19, 1.20) and receptive language (Effect = 0.76, 95% CI 0.19, 1.48). Adverse birth outcomes mediated the effect of three-dose DP on expressive language (Effect = 0.77, 95% CI 0.24, 1.33), receptive language (Effect = 0.89, 95% CI 0.18, 1.64), and fine motor (Effect = 0.60, 95% CI 0.10,

1.11). Cord blood hemoglobin and child malaria did not mediate the effects of three-dose DP or monthly DP on all ND outcomes at 24 months (Figure 4.3). We also estimated mediation effects of any maternal malaria, a binary variable that combined maternal malaria and placental malaria into one mediator. Any maternal malaria mediated the effect of three-dose DP on cognitive ability (Effect = 0.62, 95% CI 0.04, 1.27) at 12 months, and on expressive language (Effect = 0.39, 95% CI 0.00, 0.83) and receptive language (Effect = 0.55, 95% CI 0.07, 1.12) at 24 months (Figure 4.4). Results in Figures 4.2, 4.3, and 4.4 are also presented in tables in supplementary materials C.3, C.3, and C.4 respectively.

We concurrently estimated the mediation effects of any maternal malaria and child malaria on the effect of mother-child prophylaxis combinations on ND outcomes. At 24 months, compared to mother three-dose SP followed by child three-monthly DP, we found that child malaria mediated the effect of mother three-dose DP followed by child monthly DP on expressive language (Effect = 0.15, 95% CI 0.01, 0.31). Child malaria also mediated the effect of mother monthly DP followed by child monthly DP on expressive language (Effect = 0.12, 95% CI 0.01, 0.26). Any maternal malaria mediation effects of mother-child prophylaxis on all ND outcomes did not reach statistical significance. Mediation effects of mother-child prophylaxis through both mediators (any maternal malaria and child malaria) on all ND outcomes were all approximately zero and did not reach statistical significance (Table 4.4).

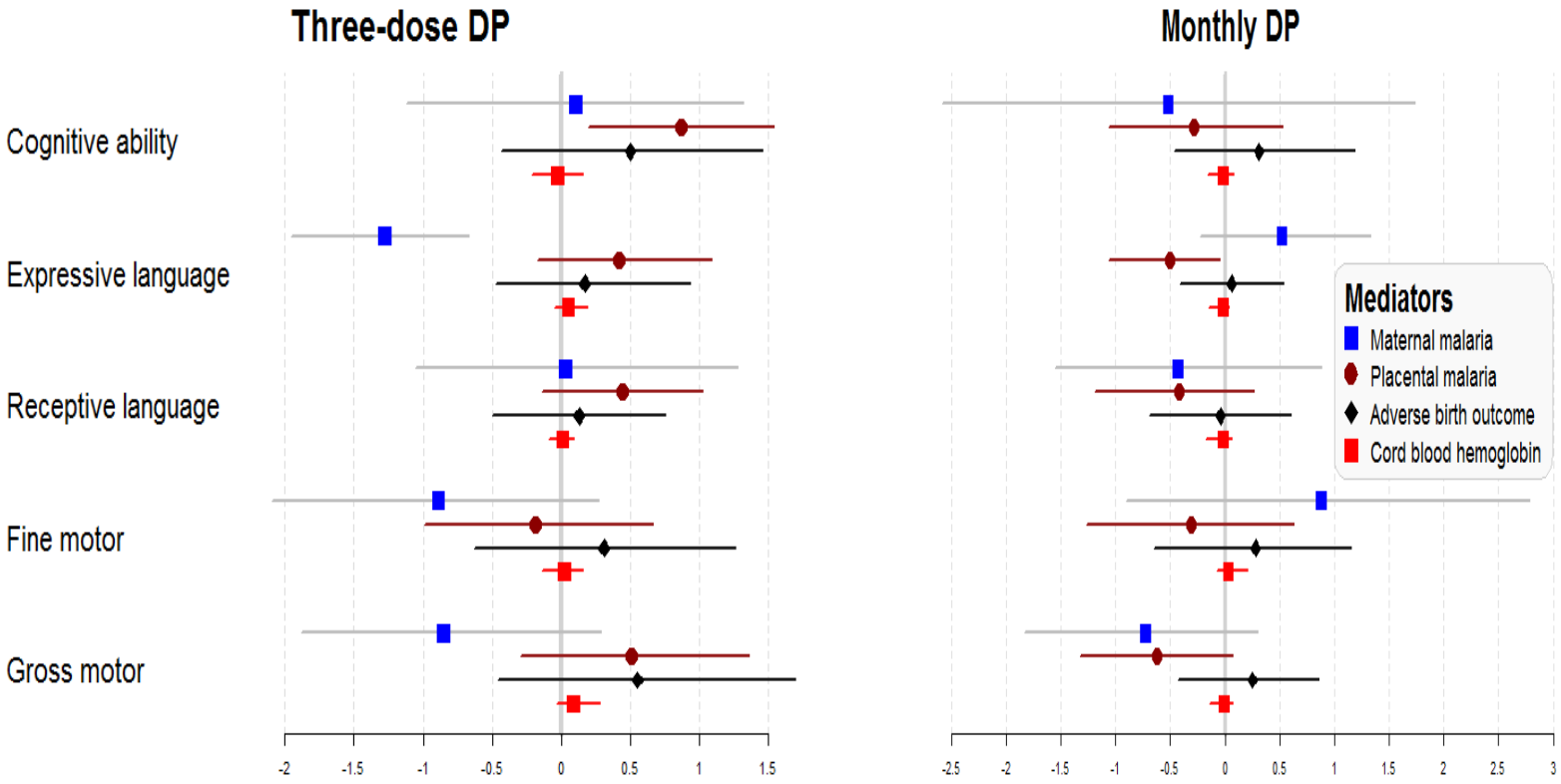

Figure 4.2 Mediator indirect effects of maternal malaria prophylaxis in pregnancy on 12-month neurodevelopment outcomes among children whose mothers received either three-dose Dihydroartenisinin-piperaquine (DP) or monthly DP groups. Indirect effects bars are 95% percentile bootstrap confidence intervals of 1000 samples.

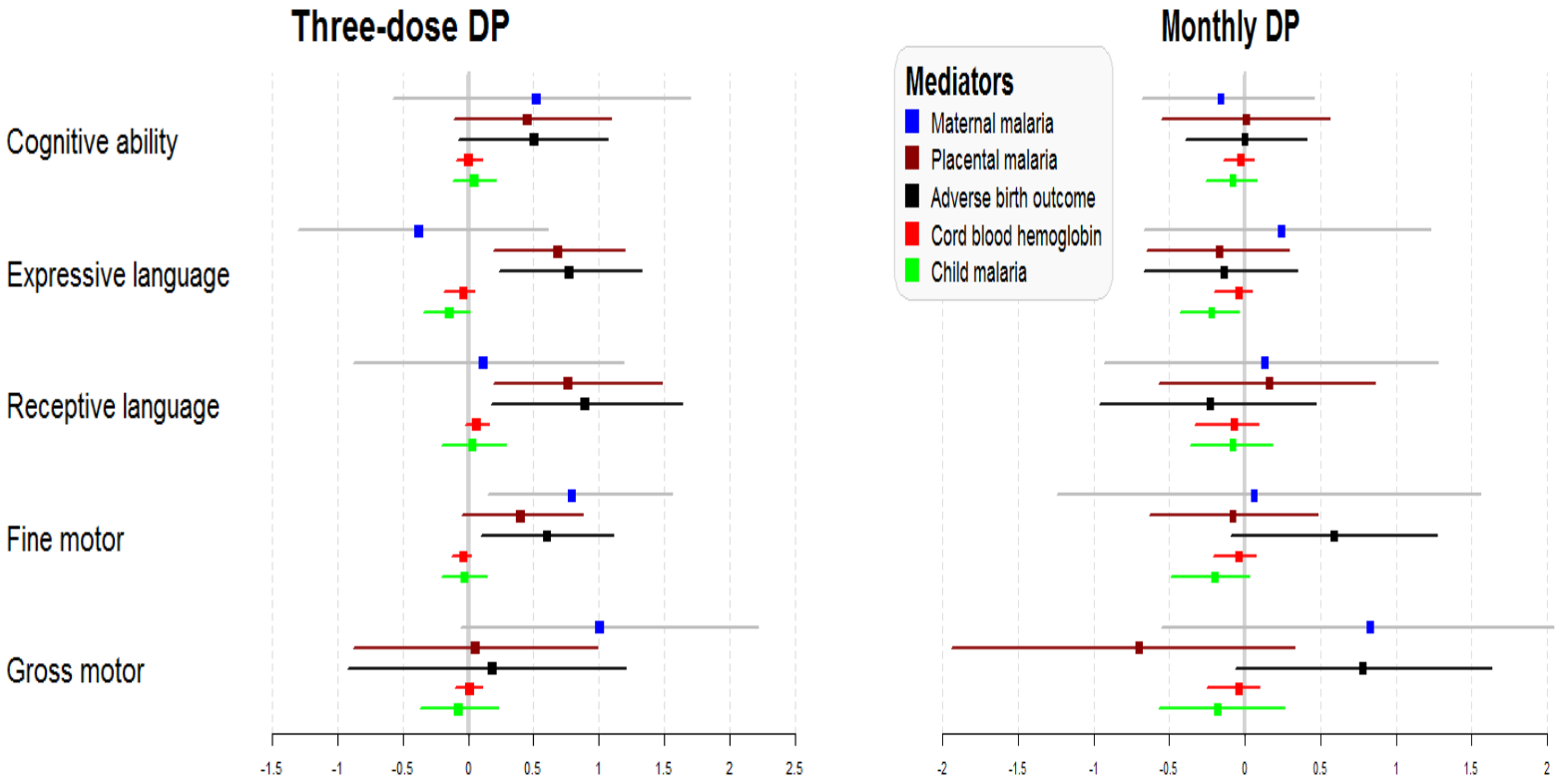

Figure 4.3 Mediator indirect effects of maternal malaria prophylaxis in pregnancy on 24-month neurodevelopment outcomes among children whose mothers received either three-dose Dihydroartenisinin-piperaquine (DP) or monthly DP groups. Indirect effects bars are 95% percentile bootstrap confidence intervals of 1000 samples.

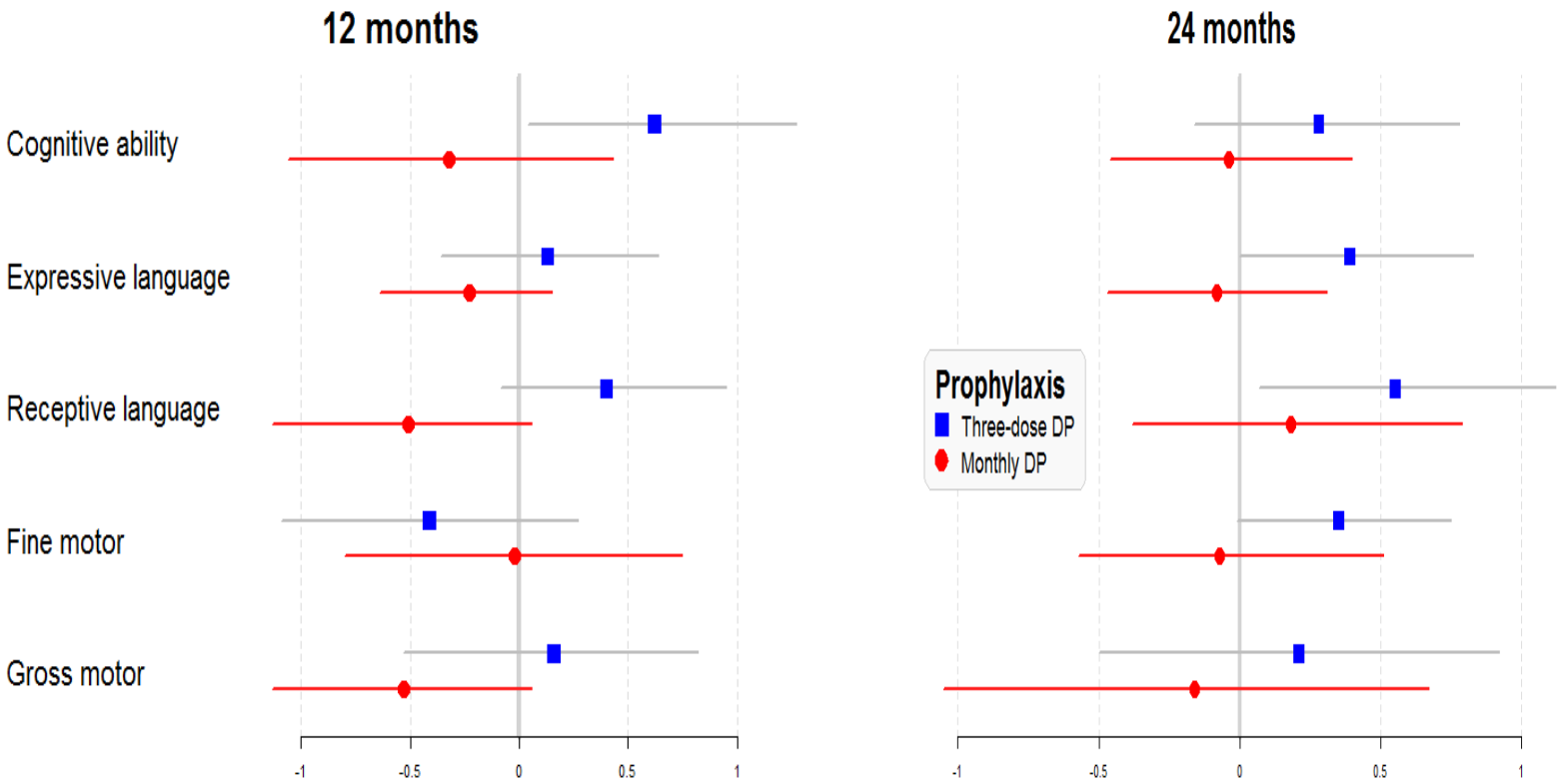

Figure 4.4 The effect of maternal malaria prophylaxis in pregnancy on 12- and 24-month neurodevelopment outcomes among children whose mothers received either three-dose dihydroartenisinin-piperaquine (DP) or monthly DP groups mediated by any malaria (maternal or placental malaria or both). Indirect effects bars are 95% percentile bootstrap confidence intervals of 1000 samples.

Table 4.4 The effect of maternal and child malaria prophylaxis combinations on child neurodevelopment outcomes at 24 months mediated by any maternal malaria (M1) and child malaria (M2).

Prophylaxis combination	Through only M1 Effect (95% CI)	Through only M2 Effect (95% CI)	Through both M1 and M2 Effect (95% CI)	Total IE Effect (95% CI)
Cognitive ability				
Mom 3DP, child 3DP	0.02 (-0.09, 0.13)	0.00 (-0.07, 0.06)	0.00 (-0.01, 0.01)	0.01 (-0.12, 0.13)
Mom 3DP, child MDP	0.01 (-0.05, 0.08)	0.03 (-0.13, 0.20)	0 (0, 0)	0.04 (-0.12, 0.02)
Mom MDP, child 3DP	0.02 (-0.11, 0.16)	-0.01 (-0.10, 0.06)	0.00 (-0.02, 0.02)	0.01 (-0.14, 0.17)
Mom MDP, child MDP	0.02 (-0.09, 0.14)	0.03 (-0.11, 0.16)	0 (0, 0)	0.04 (-0.12, 0.20)
Expressive language				
Mom 3DP, child 3DP	0.01 (-0.08, 0.12)	-0.02 (-0.14, 0.08)	0.00 (-0.02, 0.03)	-0.01 (-0.16, 0.13)
Mom 3DP, child MDP	0.00 (-0.05, 0.08)	0.15 (0.01, 0.31)*	0 (0, 0)	0.15 (0.01, 0.32)*
Mom MDP, child 3DP	0.01 (-0.10, 0.14)	-0.05 (-0.18, 0.05)	0.00 (-0.02, 0.04)	-0.03 (-0.20, 0.13)
Mom MDP, child MDP	0.01 (-0.08, 0.12)	0.12 (0.01, 0.26)*	0.00 (-0.01, 0.01)	0.13 (-0.02, 0.29)
Receptive language				
Mom 3DP, child 3DP	0.04 (-0.08, 0.20)	-0.01 (-0.09, 0.08)	0.00 (-0.02, 0.01)	0.04 (-0.10, 0.20)
Mom 3DP, child MDP	0.02 (-0.04, 0.13)	0.05 (-0.14, 0.25)	0 (0, 0)	0.07 (-0.13, 0.29)
Mom MDP, child 3DP	0.05 (-0.08, 0.24)	-0.02 (-0.13, 0.06)	0.00 (-0.02, 0.02)	0.04 (-0.13, 0.23)
Mom MDP, child MDP	0.05 (-0.08, 0.20)	0.04 (-0.11, 0.2)	0.00 (-0.01, 0.01)	0.09 (-0.11, 0.30)
Fine motor				
Mom 3DP, child 3DP	0.03 (-0.06, 0.14)	0.00 (-0.07, 0.05)	0.00 (-0.01, 0.01)	0.03 (-0.08, 0.15)
Mom 3DP, child MDP	0.02 (-0.03, 0.09)	0.01 (-0.14, 0.17)	0 (0, 0)	0.03 (-0.14, 0.20)
Mom MDP, child 3DP	0.04 (-0.08, 0.18)	0.00 (-0.09, 0.06)	0.00 (-0.01, 0.02)	0.03 (-0.10, 0.17)
Mom MDP, child MDP	0.03 (-0.06, 0.15)	0.01 (-0.11, 0.14)	0 (0, 0)	0.04 (-0.12, 0.22)
Gross motor				
Mom 3DP, child 3DP	-0.01 (-0.21, 0.17)	-0.02 (-0.21, 0.12)	0.00 (-0.02, 0.04)	-0.03 (-0.28, 0.20)
Mom 3DP, child MDP	0.00 (-0.14, 0.10)	0.18 (-0.07, 0.43)	0 (0, 0)	0.17 (-0.09, 0.45)
Mom MDP, child 3DP	-0.01 (-0.24, 0.21)	-0.06 (-0.25, 0.07)	0.00 (-0.03, 0.05)	-0.06 (-0.35, 0.19)
Mom MDP, child MDP	-0.01 (-0.21, 0.16)	0.14 (-0.05, 0.38)	0.00 (-0.01, 0.01)	0.13 (-0.13, 0.39)

Reference: Mom three-dose sulfadoxine-pyrimethamine (SP), child three-dose dihydroartenisinin-piperaquine (DP); CI, 95% percentile confidence interval from 1000 bootstrap samples; IE, Indirect effect; 3DP, Three-dose DP; MDP, Monthly DP.
*Significant i.e. 95% percentile confidence interval excludes the null (zero).

4.5 Discussion

In this randomized, double-blinded, controlled trial of malaria prophylaxis during pregnancy, we sought to establish indirect (mediated) effects of potential mediators of IPTp effects on a child's ND at 12 and 24 months. Potential mediators of IPTp's effect on child ND included maternal malaria, placental malaria, adverse birth outcomes, cord blood hemoglobin, and child malaria.

We found that placental malaria mediated the effect of three-dose DP on cognitive ability at 12 months, and on expressive and receptive language at 24 months. We did not find significant indirect effects of three-dose DP or

monthly DP on ND outcomes through cord blood hemoglobin or child malaria at 24 months. When we combined placental malaria and maternal malaria into one mediator, any maternal malaria, we found that any maternal malaria mediated the effect of three-dose DP on cognitive ability at 12 months and on expressive and receptive language at 24 months. In this study, pregnant women randomized to either three-dose DP or monthly DP had significantly fewer episodes of malaria during pregnancy. Malaria infection if untreated can lead to placental malaria (PM) (Rogerson et al., 2007). Placental malaria was measured at birth and its presence may not only reflect current malaria infections but also infections one month preceding delivery (Rogerson et al., 2007). In fact, there were mothers diagnosed with placental malaria at birth but who were not diagnosed with malaria throughout their pregnancy. A proposed mechanism for this is that placental malaria decreases placental blood flow, and placental inflammation resulting from malaria infection impairs the nutrition transport function of the placenta (Rogerson et al., 2007). These placental infections affecting nutrient transport in the third trimester when nutrients are in highest demand, have grave implications for fetal growth including ND (Conroy et al., 2013; Scifres and Nelson, 2009). In addition, maternal immune response, in particular proinflammatory cytokines, are thought to interfere with normal fetal brain development (Kronfol and Remick, 2000; Nawa et al., 2000). Animal models have indicated that offspring exposed to maternal infections *in utero* experience disruptions to normal neurodevelopment including gene expression in the brain and impaired neurotransmitter function (Golan et al., 2005; Meyer et al., 2006; Ozawa et al., 2006). It is possible that the activation of maternal immune response affects ND, even in the absence of fetus infection. It is therefore imperative to use efficacious malaria intermittent preventive treatment during pregnancy to prevent detrimental effects of placental malaria on the unborn child. Unexpected findings at 12 months were the negative monthly DP effect on expressive language mediated by placental malaria and a negative three-dose DP effect on expressive

language mediated by maternal malaria. However, these findings did not persist at 24 months follow up. Although substantial progress has been made to develop antimalarials that are efficacious with less side effects, older antimalarials of the quinoline class have been associated with psychiatric effects (Nevin and Croft, 2016). It is unclear if this is also the case for artemisinin-based combination therapies (ACTs) (Ramos-Martín et al., 2014), specifically affecting expressive language, but this certainly warrants further investigation.

When children were born, they were randomly assigned to receive either three-dose DP or monthly DP between 2-24 months. A study in Gambia followed children who had been randomized to malaria prophylaxis and placebo at age 3-59 months and measured their neurocognition at 14-20 years of age. Investigators found no significant differences in attention, memory, reason, knowledge and language (Jukes et al., 2006). A limitation in their study was that after the first trial, all children were offered malaria prevention with dapsone/pyremethamine, making it difficult to detect malaria prevention effects over 14 years later. In this study, significant mediated effects of IPTp with three-dose DP were observed after adjusting for the child's malaria chemopreventive prophylaxis. Mediation effects of placental malaria on the effect three-dose DP on cognitive ability, receptive language, and receptive language indicate potential early benefits of IPTp at a critical time of infant brain development. However, further studies are required to understand mediated effects and interactions between maternal and child chemopreventive malaria prophylaxis on child ND outcomes.

We also found that adverse birth outcomes mediated the effect of three-dose DP on expressive language, receptive language, and fine motor. We hypothesized that the protective effect of three-dose DP or monthly DP against malaria compared to three-dose SP could prevent adverse birth outcomes and consequently lead to better ND outcomes. Findings from the same cohort as the

current study have shown that IPTp with DP significantly reduces parasite prevalence and the risk of placental malaria (Kakuru et al., 2016). A study that used indoor residual spraying of insecticide (IRS) to prevent malaria during pregnancy found that IRS significantly reduced malaria incidence and women with IRS protection had a lower prevalence of placental parasitaemia, lower risk of LBW, preterm birth and fetal or neonatal deaths (Muhindo et al., 2016). Interventions that lower the incidence of maternal or placental malaria may lower the incidence of adverse birth outcomes. In fact, children whose mothers were randomized to three-dose DP had a lower prevalence of adverse birth outcomes compared to those randomized to three-dose SP. These findings together with another study that found an elevated risk of neurocognitive impairment among SGA infants (Shah et al., 2011), suggest a potential improvement of child ND outcomes as a result IPTp with DP's effect on adverse birth outcomes. Although studies have associated maternal and placental malaria with low birth weight and preterm birth (Desai et al., 2007; Kapisi et al., 2017; Kasumba et al., 2000; Menendez et al., 2000), malaria in pregnancy only accounts for 20% of LBW (Desai et al., 2007). Causes of preterm birth have been found to include genetic influence, stress, smoking, alcohol consumption, poor nutrition status, intrauterine infection, and non-genital tract infections (Piso et al., 2014). In addition, successful interventions against malaria have not all registered improvements in birth outcomes (Desai et al., 2016; Ndyomugyenyi et al., 2011). However, it is possible that prevention of maternal malaria or placental malaria during pregnancy using three-dose DP prophylaxis contributed to improved birth outcomes.

The efficacy of the WHO-recommended and widely-used IPTp with SP (WHO, 2017) has recently gained attention. Studies among infants found that IPT with SP had no protective efficacy against malaria and was associated with an increased risk of moderate to severe malaria (Bigira et al., 2014; Gosling et al., 2009). In pregnant women, a study that compared the effectiveness

on IPTp with SP, insecticide-treated nets (ITNs), and a combination of ITNs and IPTp with SP found no differences between interventions in low birth weight, prevalence of maternal anemia (hemoglobin < 11.0 g/dl), prevalence of peripheral parasitemia, and no difference in placental parasitemia among women who delivered at a health unit (Ndyomugyenyi et al., 2011). In the current study, findings of elevated episodes of malaria and an increased risk of placental malaria among pregnant women who received IPTp with SP are consistent with observed resistance to IPT with SP among pregnant women in western Kenya (Iriemenam et al., 2012). On the other hand, a study conducted among school children in the same region as our study found that IPTp with DP protected children against clinical malaria, parasitemia, and anemia (Nankabirwa et al., 2014). DP provides rapid killing of parasites by dihydroartemisinin and piperaquine provides prolonged protection against any remaining parasites (Nankabirwa et al., 2016). The long period of protection DP provides due to the long half-life of piperaquine (Hoglund et al., 2012; Nankabirwa et al., 2016; Tarning et al., 2008) protects against any placental malaria (Kakuru et al., 2016) and thus could prevent any delays in a child's ND. Therefore, the waning efficacy of IPT with SP irrespective of level of malaria transmission (Iriemenam et al., 2012; Ndyomugyenyi et al., 2011) in part explains these results.

We also found that cord blood hemoglobin did not mediate the effects of three-dose DP or monthly DP IPTp on child ND. Monthly IPT with DP in children improves hemoglobin levels and reduces the risk of anemia (Boivin et al., 2016; Nankabirwa et al., 2014). However, studies conducted among pregnant women and infants found that IPT with SP increased the risk of anemia in mothers and children (Bigira et al., 2014; Gosling et al., 2009; Harrington et al., 2011). Iron is essential in a child's development and deficiencies in the early stages of life could affect their ND (Lozoff et al., 1998; Nelson et al., 2001; Okoko et al., 2003). In the current study, the proportion of maternal anemia at birth

and the prevalence of cord blood hemoglobin below 11.0 g/dl was equally distributed among IPTp groups. A study in Ugandan children found anemia to mediate the effect of malaria on cognition (Boivin et al., 2016). However, after treatment cessation, no differences were seen in episodes of anemia between children on IPT with DP and IPT with SP (Boivin et al., 2016). It is possible that malaria infection effects such as maternal anemia were resolved by IPTp early during pregnancy before birth.

Maternal nutrition during pregnancy is a major determinant of birth weight, and nutritional supplementation has been shown to improve birth outcomes (Ceesay et al., 1997). In this study we did not measure maternal nutrition status throughout pregnancy. However, a strength of this study is that IPTp was randomized, and therefore we judge that non-measurement of nutritional status during pregnancy would not alter the study results.

Prevention of malaria during pregnancy remains a major public health intervention in protecting unborn children against placental malaria, which may impact children's ND. Even in the absence of peripheral parasitemia, when infection may be undetectable by blood smears, IPTp remains central to clearing placental parasitemia.

Chapter 5

Conclusion and future work

5.1 Conclusion

In longitudinal studies, understanding mechanisms by which an exposure or treatment affects an outcome in mediation analysis can be complicated by measurement error and missing data. This dissertation has explored the effects of missing data and measurement error in longitudinal studies in the context of mediation analysis.

In Chapter 2, we presented the counterfactual framework for mediation analysis in longitudinal studies when a repeatedly measured mediator is measured with error. The assumption of no unmeasured confounders is necessary to identify the direct and indirect effects. Appendices A.2.1 and A.2.2 illustrate how one can check on these assumptions in real practice. We found that when the mediator is repeatedly measured with error, regression calibration improves mediation effect estimates in comparison to no calibration. For the scenarios considered, assuming normality for the random effects in regression calibration provides approximately unbiased estimators regardless of the true distribution of the random effects.

In Chapter 3, we presented the effects of misspecification of the imputation

model in longitudinal studies, specifically ignoring clustering by subject, on mediated effects when the mediator and outcome are partly missing. We compared complete case analysis, an LM imputation method, an approach that ignores clustering by subject in the imputation model, and an MCMC approach that considers clustering, according to bias, MSE, and coverage probability when data was MCAR or MAR. We found that for moderate amounts of missing data, considering clustering in the multiple imputation model gives approximately unbiased mediation estimators unlike the other approaches.

In Chapter 4, we investigated the mediation effects of IPTp on child ND outcomes in a randomized clinical trial. We found that placental malaria mediates the effect of three-dose DP prophylaxis on child cognitive ability at 12 months and on child expressive and receptive language at 24 months. These findings suggest benefits of IPTp with DP in averting potential detrimental placental malaria effects on child ND.

5.2 Future Work

Chapter 2's simulations considered models with only a random intercept. McCulloch and Neuhaus (2011) points out that when the mean of the random effects distribution depends on a covariate, a relationship between the covariate and the random effect distribution is introduced, creating a serious bias in estimating the relationship between the covariate and the outcome. We did not examine this scenario, and the potential effects of this relationship on the estimation of ACME and ADE need to be investigated further. We assumed that the residual error in the mediator and outcome models were uncorrelated. In a smaller simulation study (appendix A.4), we observed that as the mediator-outcome correlation increases, with an increase in within subject variability, the ACME and ADE estimates are more biased. More research towards relaxing the independence assumption or developing a bias correction

approach is needed, especially when mediator and outcome are correlated.

Chapter 3 considered missing data mechanisms in a time-varying continuous mediator and outcome. The MAR mechanism depended on the magnitude of time-point to time-point reduction in either the mediator or outcome measures. Investigation of other MAR mechanisms, such as MAR mechanisms that depend on auxiliary variables, may offer improvements when auxiliary variables are included in the imputation model.

In Chapter 4, we did not find significant any maternal malaria mediation effects or mediation effects through both any maternal malaria and child malaria (Table 4.4). Splitting the exposure groups into mother-child prophylaxis combinations made group sizes small and this may have affected the power to detect mediation effects. Larger studies with sufficient power would help to estimate mediated effects of combinations of mother-child prophylaxis on the child's neurodevelopment.

References

Agresti, A., Caffo, B., and Ohman-Strickland, P. (2004). Examples in which misspecification of a random effects distribution reduces efficiency, and possible remedies. *Computational Statistics & Data Analysis*, 47(3):639–653.

Aitkin, M. (1999). Meta-analysis by random effect modelling in generalized linear models. *Statistics in Medicine*, 18(17-18):2343–2351.

Al Serouri, A. W., Grantham-McGregor, S., Greenwood, B., and Costello, A. (2000). Impact of asymptomatic malaria parasitaemia on cognitive function and school achievement of schoolchildren in the yemen republic. *Parasitology*, 121(04):337–345.

Andridge, R. R. (2011). Quantifying the impact of fixed effects modeling of clusters in multiple imputation for cluster randomized trials. *Biometrical Journal*, 53(1):57–74.

Azur, M. J., Stuart, E. A., Frangakis, C., and Leaf, P. J. (2011). Multiple imputation by chained equations: what is it and how does it work? *International Journal of Methods in Psychiatric Research*, 20(1):40–49.

Bangirana, P., Opoka, R. O., Boivin, M. J., Idro, R., Hodges, J. S., Romero, R. A., Shapiro, E., and John, C. C. (2014). Severe malarial anemia is associated with long-term neurocognitive impairment. *Clinical Infectious Diseases*.

Baron, R. M. and Kenny, D. A. (1986). The moderator–mediator variable dis-

tinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6):1173.

Bauer, D. J., Preacher, K. J., and Gil, K. M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: new procedures and recommendations. *Psychological Methods*, 11(2):142.

Bayley, N. (2006). *Bayley Scales of Infant Development and Toddler Development: Technical Manual*. The PsychCorp.

Bigira, V., Kapisi, J., Clark, T. D., Kinara, S., Mwangwa, F., Muhindo, M. K., Osterbauer, B., Aweeka, F. T., Huang, L., Achan, J., et al. (2014). Protective efficacy and safety of three antimalarial regimens for the prevention of malaria in young ugandan children: a randomized controlled trial. *PLoS Medicine*, 11(8):e1001689.

Bind, M.-A., Vanderweele, T., Coull, B., and Schwartz, J. (2015). Causal mediation analysis for longitudinal data with exogenous exposure. *Biostatistics*, 17(1):122–134.

Black, A. C., Harel, O., and Betsy McCoach, D. (2011). Missing data techniques for multilevel data: implications of model misspecification. *Journal of Applied Statistics*, 38(9):1845–1865.

Blood, E. A., Cabral, H., Heeren, T., and Cheng, D. M. (2010). Performance of mixed effects models in the analysis of mediated longitudinal data. *BMC Medical Research Methodology*, 10(1):1.

Blood, E. A. and Cheng, D. M. (2011). The use of mixed models for the analysis of mediated data with time-dependent predictors. *Journal of Environmental and Public Health*, 2011.

Blood, E. A. and Cheng, D. M. (2012). Non-linear mixed models in the analysis of mediated longitudinal data with binary outcomes. *BMC Medical Research Methodology*, 12(1):1.

Boivin, M. J., Bangirana, P., Byarugaba, J., Opoka, R. O., Idro, R., Jurek, A. M., and John, C. C. (2007). Cognitive impairment after cerebral malaria in children: a prospective study. *Pediatrics*, 119(2):e360–e366.

Boivin, M. J., Sikorskii, A., Familiar-Lopez, I., Ruiseñor-Escudero, H., Muhindo, M., Kapisi, J., Bigira, V., Bass, J. K., Opoka, R. O., Nakasujja, N., et al. (2016). Malaria illness mediated by anaemia lessens cognitive development in younger ugandan children. *Malaria Journal*, 15(1):1.

Bouhlila, D. S. and Sellaouti, F. (2013). Multiple imputation using chained equations for missing data in TIMSS: a case study. *Large-scale Assessments in Education*, 1(1):1–33.

Brabin, B. J. (1983). An analysis of malaria in pregnancy in africa. *Bulletin of the World Health Organization*, 61(6):1005.

Brand, J. (1999). *Development, Implementation and Evaluation of Multiple Imputation Strategies for the Statistical Analysis of Incomplete Data Sets*. PhD thesis, Erasmus MC: University Medical Center Rotterdam.

Brown, P. J. and Fuller, W. A. (1990). *Statistical Analysis of Measurement Error Models and Applications: Proceedings of the AMS-IMS-SIAM Joint Summer Research Conference Held June 10-16, 1989, with Support from the National Science Foundation and the US Army Research Office*, volume 112. American Mathematical Soc.

Buonaccorsi, J., Demidenko, E., and Tosteson, T. (2000). Estimation in longitudinal random effects models with measurement error. *Statistica Sinica*, pages 885–903.

Burr, D. and Doss, H. (2005). A bayesian semiparametric model for random-effects meta-analysis. *Journal of the American Statistical Association*, 100(469):242–251.

Buuren, S. and Groothuis-Oudshoorn, K. (2011). MICE: Multivariate imputation by chained equations in r. *Journal of Statistical Software*, 45(3).

Carroll, R. J., Ruppert, D., Stefanski, L. A., and Crainiceanu, C. M. (2006). *Measurement error in nonlinear models: a modern perspective*. CRC press.

Carroll, R. J. and Stefanski, L. A. (1990). Approximate quasi-likelihood estimation in models with surrogate predictors. *Journal of the American Statistical Association*, 85(411):652–663.

Ceesay, S. M., Prentice, A. M., Cole, T. J., Foord, F., Poskitt, E. M., Weaver, L. T., and Whitehead, R. G. (1997). Effects on birth weight and perinatal mortality of maternal dietary supplements in rural gambia: 5 year randomised controlled trial. *BMJ*, 315(7111):786–790.

Chen, J., Zhang, D., and Davidian, M. (2002). A monte carlo em algorithm for generalized linear mixed models with flexible random effects distribution. *Biostatistics*, 3(3):347–360.

Cole, D. A. and Maxwell, S. E. (2003). Testing mediational models with longitudinal data: questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology*, 112(4):558.

Cole, S. R. and Frangakis, C. E. (2009). The consistency statement in causal inference: a definition or an assumption? *Epidemiology*, 20(1):3–5.

Conroy, A. L., Silver, K. L., Zhong, K., Rennie, M., Ward, P., Sarma, J. V., Molyneux, M. E., Sled, J., Fletcher, J. F., Rogerson, S., et al. (2013). Complement activation and the resulting placental vascular insufficiency drives fetal growth restriction associated with placental malaria. *Cell Host & Microbe*, 13(2):215–226.

Cusick, S. E., Opoka, R. O., Ssemata, A. S., Georgieff, M. K., and John, C. C. (2016). Comparison of iron status 28 d after provision of antimalarial

treatment with iron therapy compared with antimalarial treatment alone in ugandan children with severe malaria. *The American Journal of Clinical Nutrition*, 103(3):919–925.

Daniel, R., De Stavola, B., Cousens, S., and Vansteelandt, S. (2015). Causal mediation analysis with multiple mediators. *Biometrics*, 71(1):1–14.

Desai, M., Gutman, J., L'lanziva, A., Otieno, K., Juma, E., Kariuki, S., Ouma, P., Were, V., Laserson, K., Katana, A., et al. (2016). Intermittent screening and treatment or intermittent preventive treatment with dihydroartemisinin–piperaquine versus intermittent preventive treatment with sulfadoxine–pyrimethamine for the control of malaria during pregnancy in western kenya: an open-label, three-group, randomised controlled superiority trial. *The Lancet*, 386(10012):2507–2519.

Desai, M., Ter Kuile, F. O., Nosten, F., McGready, R., Asamoa, K., Brabin, B., and Newman, R. D. (2007). Epidemiology and burden of malaria in pregnancy. *The Lancet Infectious Diseases*, 7(2):93–104.

Dong, Y. and Peng, C.-Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus*, 2(1):222.

Fernando, D., De Silva, D., and Wickremasinghe, R. (2003). Short-term impact of an acute attack of malaria on the cognitive performance ofschoolchildren living in a malaria-endemic area of sri lanka. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 97(6):633–639.

Gallant, A. R. and Nychka, D. W. (1987). Semi-nonparametric maximum likelihood estimation. *Econometrica: Journal of the Econometric Society*, pages 363–390.

Ghidey, W., Lesaffre, E., and Eilers, P. (2004). Smooth random effects distribution in a linear mixed model. *Biometrics*, 60(4):945–953.

Golan, H., Lev, V., Hallak, M., Sorokin, Y., and Huleihel, M. (2005). Specific neurodevelopmental damage in mice offspring following maternal inflammation during pregnancy. *Neuropharmacology*, 48(6):903–917.

Gosling, R. D., Gesase, S., Mosha, J. F., Carneiro, I., Hashim, R., Lemnge, M., Mosha, F. W., Greenwood, B., and Chandramohan, D. (2009). Protective efficacy and safety of three antimalarial regimens for intermittent preventive treatment for malaria in infants: a randomised, double-blind, placebo-controlled trial. *The Lancet*, 374(9700):1521–1532.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60:549–576.

Graham, J. W., Olchowski, A. E., and Gilreath, T. D. (2007). How many imputations are really needed? some practical clarifications of multiple imputation theory. *Prevention Science*, 8(3):206–213.

Guyatt, H. L. and Snow, R. W. (2004). Impact of malaria during pregnancy on low birth weight in sub-saharan africa. *Clinical Microbiology Reviews*, 17(4):760–769.

Harrington, W. E., Mutabingwa, T. K., Kabyemela, E., Fried, M., and Duffy, P. E. (2011). Intermittent treatment to prevent pregnancy malaria does not confer benefit in an area of widespread drug resistance. *Clinical Infectious Diseases*, 53(3):224–230.

Heagerty, P. J. and Kurland, B. F. (2001). Misspecified maximum likelihood estimates and generalised linear mixed models. *Biometrika*, pages 973–985.

Heckman, J. and Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica: Journal of the Econometric Society*, pages 271–320.

Hendrix, N. and Berghella, V. (2008). Non-placental causes of intrauterine

growth restriction. In *Seminars in perinatology*, volume 32, pages 161–165. Elsevier.

Hoglund, R. M., Adam, I., Hanpithakpong, W., Ashton, M., Lindegardh, N., Day, N. P., White, N. J., Nosten, F., and Tarning, J. (2012). A population pharmacokinetic model of piperaquine in pregnant and non-pregnant women with uncomplicated plasmodium falciparum malaria in sudan. *Malaria Journal*, 11(1):398.

Horton, N. J., Lipsitz, S. R., and Parzen, M. (2003). A potential for bias when rounding in multiple imputation. *The American Statistician*, 57(4):229–232.

Iriemenam, N. C., Shah, M., Gatei, W., van Eijk, A. M., Ayisi, J., Kariuki, S., Eng, J. V., Owino, S. O., Lal, A. A., Omosun, Y. O., et al. (2012). Temporal trends of sulphadoxine-pyrimethamine (SP) drug-resistance molecular markers in plasmodium falciparum parasites from pregnant women in western kenya. *Malaria Journal*, 11(1):134.

John, C. C., Bangirana, P., Byarugaba, J., Opoka, R. O., Idro, R., Jurek, A. M., Wu, B., and Boivin, M. J. (2008). Cerebral malaria in children is associated with long-term cognitive impairment. *Pediatrics*, 122(1):e92–e99.

Jukes, M. C., Pinder, M., Grigorenko, E. L., Smith, H. B., Walraven, G., Bariau, E. M., Sternberg, R. J., Drake, L. J., Milligan, P., Cheung, Y. B., et al. (2006). Long-term impact of malaria chemoprophylaxis on cognitive abilities and educational attainment: follow-up of a controlled trial. *PLoS Clinical Trials*, 1(4):e19.

Kakuru, A., Jagannathan, P., Muhindo, M. K., Natureeba, P., Awori, P., Nakalembe, M., Opira, B., Olwoch, P., Ategeka, J., Nayebare, P., et al. (2016). Dihydroartemisinin–piperaquine for the prevention of malaria in pregnancy. *New England Journal of Medicine*, 374(10):928–939.

Kamya, M. R., Arinaitwe, E., Wanzira, H., Katureebe, A., Barusya, C., Kigozi, S. P., Kilama, M., Tatem, A. J., Rosenthal, P. J., Drakeley, C., et al. (2015). Malaria transmission, infection, and disease at three sites with varied transmission intensity in uganda: implications for malaria control. *The American Journal of Tropical Medicine and Hygiene*, 92(5):903–912.

Kapisi, J., Kakuru, A., Jagannathan, P., Muhindo, M. K., Natureeba, P., Awori, P., Nakalembe, M., Ssekitoleko, R., Olwoch, P., Ategeka, J., et al. (2017). Relationships between infection with plasmodium falciparum during pregnancy, measures of placental malaria, and adverse birth outcomes. *Malaria Journal*, 16(1):400.

Kasumba, I., Nalunkuma, A., Mujuzi, G., Kitaka, F., Byaruhanga, R., Okong, P., and Egwang, T. (2000). Low birthweight associated with maternal anaemia and plasmodium falciparum infection during pregnancy, in a peri-urban/urban area of low endemicity in uganda. *Annals of Tropical Medicine & Parasitology*, 94(1):7–13.

Katz, J., Lee, A. C., Kozuki, N., Lawn, J. E., Cousens, S., Blencowe, H., Ezzati, M., Bhutta, Z. A., Marchant, T., Willey, B. A., et al. (2013). Mortality risk in preterm and small-for-gestational-age infants in low-income and middle-income countries: a pooled country analysis. *The Lancet*, 382(9890):417–425.

Kronfol, Z. and Remick, D. G. (2000). Cytokines and the brain: implications for clinical psychiatry. *American Journal of Psychiatry*, 157(5):683–694.

Lagerberg, R. E. (2008). Malaria in pregnancy: a literature review. *Journal of Midwifery & Womens Health*, 53(3):209–215.

Lee, K. J. and Thompson, S. G. (2008). Flexible parametric models for random-effects distributions. *Statistics in Medicine*, 27(3):418–434.

Little, R. J. and Rubin, D. (1987). Statistical analysis with missing data. series in probability and mathematical statistics.

Liu, M., Taylor, J. M., and Belin, T. R. (2000). Multiple imputation and posterior simulation for multivariate missing data in longitudinal studies. *Biometrics*, 56(4):1157–1163.

Lozoff, B. et al. (1998). Explanatory mechanisms for poorer development in iron-deficiency anemic infants. In *Nutrition, health, and child development. Research advances and policy recommendations*, pages 162–78. Pan American Health Organization.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Routledge.

MacKinnon, D. P., Fairchild, A. J., and Fritz, M. S. (2007). Mediation analysis. *Annu. Rev. Psychol.*, 58:593–614.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., and Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, 7(1):83.

MacKinnon, D. P., Lockwood, C. M., and Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, 39(1):99–128.

MacKinnon, D. P., Warsi, G., and Dwyer, J. H. (1995). A simulation study of mediated effect measures. *Multivariate Behavioral Research*, 30(1):41–62.

Magder, L. S. and Zeger, S. L. (1996). A smooth nonparametric estimate of a mixing distribution using mixtures of gaussians. *Journal of the American Statistical Association*, 91(435):1141–1151.

McCulloch, C. E. and Neuhaus, J. M. (2011). Misspecifying the shape of a random effects distribution: why getting it wrong may not matter. *Statistical Science*, pages 388–402.

Menendez, C., Ordi, J., Ismail, M., Ventura, P., Aponte, J., Kahigwa, E., Font, F., and Alonso, P. (2000). The impact of placental malaria on gestational age and birth weight. *The Journal of Infectious Diseases*, 181(5):1740–1745.

Meng, X.-L. (1994). Multiple-imputation inferences with uncongenial sources of input. *Statistical Science*, pages 538–558.

Meyer, U., Nyffeler, M., Engler, A., Urwyler, A., Schedlowski, M., Knuesel, I., Yee, B. K., and Feldon, J. (2006). The time of prenatal immune challenge determines the specificity of inflammation-mediated brain and behavioral pathology. *Journal of Neuroscience*, 26(18):4752–4762.

Mistler, S. A. (2013). A SAS ® macro for applying multiple imputation to multilevel data. In *Proceedings of the SAS Global Forum*.

Muehlenbachs, A., Mutabingwa, T. K., Edmonds, S., Fried, M., and Duffy, P. E. (2006). Hypertension and maternal–fetal conflict during placental malaria. *PLoS Med*, 3(11):e446.

Muhindo, M. K., Kakuru, A., Natureeba, P., Awori, P., Olwoch, P., Ategeka, J., Nayebare, P., Clark, T. D., Muehlenbachs, A., Roh, M., et al. (2016). Reductions in malaria in pregnancy and adverse birth outcomes following indoor residual spraying of insecticide in uganda. *Malaria Journal*, 15(1):437.

Mullen, E. M. et al. (1995). *Mullen scales of early learning*. AGS Circle Pines, MN.

Nankabirwa, J., Wandera, B., Kiwanuka, N., Staedke, S. G., Kamya, M. R., and Brooker, S. J. (2013). Asymptomatic plasmodium infection and cognition among primary schoolchildren in a high malaria transmission setting in uganda. *The American Journal of Tropical Medicine and Hygiene*, 88(6):1102–1108.

Nankabirwa, J. I., Conrad, M. D., Legac, J., Tukwasibwe, S., Tumwebaze, P., Wandera, B., Brooker, S. J., Staedke, S. G., Kamya, M. R., Nsobya, S. L.,

et al. (2016). Intermittent preventive treatment with dihydroartemisinin-piperaquine in ugandan schoolchildren selects for plasmodium falciparum transporter polymorphisms that modify drug sensitivity. *Antimicrobial Agents and Chemotherapy*, 60(10):5649–5654.

Nankabirwa, J. I., Wandera, B., Amuge, P., Kiwanuka, N., Dorsey, G., Rosenthal, P. J., Brooker, S. J., Staedke, S. G., and Kamya, M. R. (2014). Impact of intermittent preventive treatment with dihydroartemisinin-piperaquine on malaria in ugandan schoolchildren: a randomized, placebo-controlled trial. *Clinical Infectious Diseases*, 58(10):1404–1412.

Natureeba, P., Ades, V., Luwedde, F., Mwesigwa, J., Plenty, A., Okong, P., Charlebois, E. D., Clark, T. D., Nzarubara, B., Havlir, D. V., et al. (2014). Lopinavir/ritonavir-based antiretroviral treatment (art) versus efavirenz-based art for the prevention of malaria among hiv-infected pregnant women. *Journal of Infectious Diseases*, 210(12):1938–1945.

Nawa, H., Takahashi, M., and Patterson, P. (2000). Cytokine and growth factor involvement in schizophreniasupport for the developmental model. *Molecular Psychiatry*, 5(6):594.

Ndyomugyenyi, R., Clarke, S. E., Hutchison, C. L., Hansen, K. S., and Magnussen, P. (2011). Efficacy of malaria prevention during pregnancy in an area of low and unstable transmission: an individually-randomised placebo-controlled trial using intermittent preventive treatment and insecticide-treated nets in the kabale highlands, southwestern uganda. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 105(11):607–616.

Nelson, C. A., Collins, M. L., and Luciana, M. (2001). *Handbook of Developmental Cognitive Neuroscience*.

Neuhaus, J. M., Hauck, W. W., and Kalbfleisch, J. D. (1992). The effects of mixture distribution misspecification when fitting mixed-effects logistic models. *Biometrika*, 79(4):755–762.

Nevin, R. L. and Croft, A. M. (2016). Psychiatric effects of malaria and anti-malarial drugs: historical and modern perspectives. *Malaria Journal*, 15(1):332.

Ohlssen, D. I., Sharples, L. D., and Spiegelhalter, D. J. (2007). Flexible random-effects models using bayesian semi-parametric models: applications to institutional comparisons. *Statistics in Medicine*, 26(9):2088–2112.

Okoko, B. J., Enwere, G., and Ota, M. (2003). The epidemiology and consequences of maternal malaria: a review of immunological basis. *Acta Tropica*, 87(2):193–205.

Ozawa, K., Hashimoto, K., Kishimoto, T., Shimizu, E., Ishikura, H., and Iyo, M. (2006). Immune activation during pregnancy in mice leads to dopaminergic hyperfunction and cognitive impairment in the offspring: a neurodevelopmental animal model of schizophrenia. *Biological Psychiatry*, 59(6):546–554.

Piso, B., Zechmeister-Koss, I., and Winkler, R. (2014). Antenatal interventions to reduce preterm birth: an overview of cochrane systematic reviews. *BMC Research Notes*, 7(1):265.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raghunathan, T. E., Lepkowski, J. M., Van Hoewyk, J., and Solenberger, P. (2001). A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey Methodology*, 27(1):85–96.

Ramos-Martín, V., González-Martínez, C., Mackenzie, I., Schmutzhard, J., Pace, C., Lalloo, D. G., and Terlouw, D. J. (2014). Neuroauditory toxicity of artemisinin combination therapieshave safety concerns been addressed? *The American Journal of Tropical Medicine and Hygiene*, 91(1):62–73.

Rogerson, S. J., Hviid, L., Duffy, P. E., Leke, R. F., and Taylor, D. W. (2007).

Malaria in pregnancy: pathogenesis and immunity. *The Lancet Infectious Diseases*, 7(2):105–117.

Rogerson, S. J., Pollina, E., Getachew, A., Tadesse, E., Lema, V. M., and Molyneux, M. E. (2003). Placental monocyte infiltrates in response to plasmodium falciparum malaria infection and their association with adverse pregnancy outcomes. *The American Journal of Tropical Medicine and Hygiene*, 68(1):115–119.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3):581–592.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons.

Samet, J. H., Cheng, D. M., Libman, H., Nunes, D. P., Alperen, J. K., and Saitz, R. (2007). Alcohol consumption and hiv disease progression. *Journal of Acquired Immune Deficiency Syndromes (1999)*, 46(2):194.

Samet, J. H., Horton, N. J., Meli, S., Freedberg, K. A., and Palepu, A. (2004). Alcohol consumption and antiretroviral adherence among hiv-infected persons with alcohol problems. *Alcoholism: Clinical and Experimental Research*, 28(4):572–577.

Schafer, J. L. (1997). Imputation of missing covariates under a multivariate linear mixed model. Technical report, Dept. of Statistics, The Pennsylvania State University.

Schafer, J. L. (1999). Multiple imputation: a primer. *Statistical Methods in Medical Research*, 8(1):3–15.

Schafer, J. L. (2001). Multiple imputation with PAN.

Schafer, J. L. and Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods*, 7(2):147.

Schafer, J. L. and Olsen, M. K. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavioral Research*, 33(4):545–571.

Schafer, J. L. and Yucel, R. M. (2002). Computational strategies for multivariate linear mixed-effects models with missing values. *Journal of computational and Graphical Statistics*, 11(2):437–457.

Schmiegelow, C., Scheike, T., Oesterholt, M., Minja, D., Pehrson, C., Magistrado, P., Lemnge, M., Rasch, V., Lusingu, J., Theander, T. G., et al. (2012). Development of a fetal weight chart using serial trans-abdominal ultrasound in an east african population: a longitudinal observational study. *PLoS One*, 7(9):e44773.

Scifres, C. M. and Nelson, D. M. (2009). Intrauterine growth restriction, human placental development and trophoblast cell death. *The Journal of Physiology*, 587(14):3453–3458.

Shah, P. et al. (2011). Long-term neurocognitive outcomes of SGA/IUGR infants. *Obstetrics, Gynaecology and Reproductive Medicine*, 21(5):142–146.

Shi, L., Tu, N., and Patterson, P. H. (2005). Maternal influenza infection is likely to alter fetal brain development indirectly: the virus is not detected in the fetus. *International Journal of Developmental Neuroscience*, 23(2):299–305.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, 13(1982):290–312.

Steketee, R. W., Nahlen, B. L., Parise, M. E., and Menendez, C. (2001). The burden of malaria in pregnancy in malaria-endemic areas. *The American Journal of Tropical Medicine and Hygiene*, 64(1_suppl):28–35.

Taljaard, M., Donner, A., and Klar, N. (2008). Imputation strategies for missing continuous outcomes in cluster randomized trials. *Biometrical Journal*, 50(3):329–345.

Tarning, J., Ashley, E., Lindegardh, N., Stepniewska, K., Phaiphun, L., Day, N., McGready, R., Ashton, M., Nosten, F., and White, N. (2008). Population pharmacokinetics of piperaquine after two different treatment regimens with dihydroartemisinin-piperaquine in patients with plasmodium falciparum malaria in thailand. *Antimicrobial Agents and Chemotherapy*, 52(3):1052–1061.

Urakubo, A., Jarskog, L. F., Lieberman, J. A., and Gilmore, J. H. (2001). Prenatal exposure to maternal infection alters cytokine expression in the placenta, amniotic fluid, and fetal brain. *Schizophrenia Research*, 47(1):27–36.

Valeri, L., Lin, X., and VanderWeele, T. (2014). Mediation analysis when a continuous mediator is measured with error and the outcome follows a generalized linear model. *Statistics in Medicine*.

Van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research*, 16(3):219–242.

Van Buuren, S., Brand, J. P., Groothuis-Oudshoorn, C., and Rubin, D. B. (2006). Fully conditional specification in multivariate imputation. *Journal of Statistical Computation and Simulation*, 76(12):1049–1064.

van der Laan, M. J. and Petersen, M. L. (2004). Estimation of direct and indirect causal effects in longitudinal studies. Technical Report 155, UC Berkeley Division of Biostatistics Working paper.

VanderWeele, T. J. (2009). Concerning the consistency assumption in causal inference. *Epidemiology*, 20(6):880–883.

VanderWeele, T. J. (2010). Direct and indirect effects for neighborhood-based clustered and longitudinal data. *Sociological Methods & Research*, 38(4):515–544.

VanderWeele, T. J. (2016). Mediation analysis: A practitioner's guide. *Annual Review of Public health*, 37:17–32.

VanderWeele, T. J. and Tchetgen Tchetgen, E. J. (2017). Mediation analysis with time varying exposures and mediators. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(3):917–938.

Verbeke, G. and Lesaffre, E. (1996). A linear mixed-effects model with heterogeneity in the random-effects population. *Journal of the American Statistical Association*, 91(433):217–221.

Wang, N., Lin, X., and Guttierrez, R. G. (1999). A bias correction regression calibration approach in generalized linear mixed measurement error models. *Communications in Statistics-Theory and Methods*, 28(1):217–232.

White, I. R., Royston, P., and Wood, A. M. (2011). Multiple imputation using chained equations: issues and guidance for practice. *Statistics in Medicine*, 30(4):377–399.

WHO (2015). *World malaria report 2015*. World Health Organization.

WHO (2016). *WHO recommendations on antenatal care for a positive pregnancy experience*. World Health Organization.

WHO (2017). *World malaria report 2017*. World Health Organization, Geneva.

Wu, W. and Jia, F. (2013). A new procedure to test mediation with missing data through nonparametric bootstrapping and multiple imputation. *Multivariate Behavioral Research*, 48(5):663–691.

Yucel, R. M. (2015). R mlmmm package: Fitting multivariate linear mixed effects models with missing values. *Turkiye Klinikleri Journal of Biostatistics*, 7(1).

Yucel, R. M. and Demirtas, H. (2010). Impact of non-normal random effects on inference by multiple imputation: a simulation assessment. *Computational Statistics & Data Analysis*, 54(3):790–801.

Zhang, D. and Davidian, M. (2001). Linear mixed models with flexible distributions of random effects for longitudinal data. *Biometrics*, 57(3):795–802.

Zhang, Z., Wang, L., and Tong, X. (2015). Mediation analysis with missing data through multiple imputation and bootstrap. In *Quantitative Psychology Research*, pages 341–355. Springer.

Zhao, J. H. and Schafer, J. L. (2016). *PAN: Multiple imputation for multivariate panel or clustered data*. R package version 1.4.

Zhou, J., Connella, L. E., and Graham, J. W. (2014). Handling missing data in cluster randomized trials: A demonstration of multiple imputation with PAN through SAS. *The Quantitative Methods of Psychology*, 10(2):153–166.

Appendix A

Chapter 2 Appendix

A.1 Time-varying ACME and ADE derivation

Assume that Y_{ij} is the outcome for subject i at time $j = 1, 2, \ldots, s$, M_{ij} is a continuous mediator repeatedly measured with error, X_i is the binary exposure/treatment variable (1=Treatment, 0=Control) but can be continuous, G_i is a baseline covariate, C_{ij} is a mediator-outcome confounder, and t_j is the time indicator i.e. $t_j = I(t = j)$. The derivation assumes the data generation mechanism for the binary exposure coefficient in the mediator calibration model is fixed.

$$
\begin{aligned}
Y_{ij} &= u_{0i}^{Y^*} + \beta_0 + \beta_x X_i + \beta_{t_2} t_2 + \cdots + \beta_{t_s} t_s + \beta_g G_i + \beta_c C_{ij} + e_{ij} && \text{(A.1)} \\
Y_{ij} &= u_{0i}^{Y} + \gamma_0 + \gamma_x X_i + \gamma_m M_{ij}^* + \gamma_{xm} X_i M_{ij}^* + \gamma_{t_2} t_2 + \cdots + \gamma_{t_s} t_s \\
&+ \gamma_g G_i + \gamma_c C_{ij} + r_{ij} && \text{(A.2)} \\
M_{ij} &= u_{0i}^{M} + \zeta_0 + \zeta_x X_i + \zeta_{t_2} X_i t_2 + \cdots + \zeta_{t_s} X_i t_s + \zeta_g G_i + \epsilon_{ij} && \text{(A.3)}
\end{aligned}
$$

where $e_{ij} \sim N(0, \sigma_e^2)$, $\epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$, $r_{ij} \sim N(0, \sigma_r^2)$, $u_{0i}^{Y^*} \sim N(0, \sigma^2_{u_i^{Y^*}})$, $u_{0i}^{Y} \sim N(0, \sigma^2_{u_i^{Y}})$ and $u_{0i}^{M} \sim N(0, \sigma^2_{u_i^{M}})$; e_{ij} is independent of $u_{0i}^{Y^*}$, ϵ_{ij} is independent of u_{0i}^{M}, r_{ij} is independent of u_{0i}^{Y}, u_{0i}^{M} is independent of u_{0i}^{Y}, ϵ_{ij} and r_{ij} are independent, and $M_{ij}^* = u_{0i}^{M} + \zeta_0 + \zeta_x X_i + \zeta_{t_2} X_i t_2 + \cdots + \zeta_{t_s} X_i t_s + \zeta_g G_i$ is the mediator measured without error. The total effect is given by β_x in model

(A.1). Using model (A.2), for values $x, \tilde{x}, x'$ of X_i and the calibrated mediator $\widehat{M}^*_{ij}$, the direct effect (ADE) is given by

$$\begin{aligned} ADE_j &= \mathbb{E}\big\{Y_j(x, M_j^*(x')) - Y_j(\tilde{x}, M_j^*(x')) \mid G_i = g_i, C_{ij} = c_{ij}, t_j\big\} \\ &= (\gamma_x + \gamma_{xm} m_{ij}^*)(x - \tilde{x}) \end{aligned}$$

Using equations (A.2) and (A.3) and the product method (Baron and Kenny, 1986), the time-varying average causal mediation effect (indirect effect) is given by

$$\begin{aligned} ACME_j &= \mathbb{E}\big\{Y_j(x', M_j^*(x)) - Y_j(x', M_j^*(\tilde{x})) \mid G_i = g_i, C_{ij} = c_{ij}, t_j\big\} \\ &= \gamma_0 + \gamma_x x' + \gamma_m\big\{\zeta_0 + \zeta_x x + \zeta_{t_j} x\big\} + \gamma_{xm} x'\big\{\zeta_0 + \zeta_x x + \zeta_{t_j} x\big\} \\ &- \gamma_0 - \gamma_x x' - \gamma_m\big\{\zeta_0 + \zeta_x \tilde{x} + \zeta_{t_j} \tilde{x}\big\} - \gamma_{xm} x'\big\{\zeta_0 + \zeta_x \tilde{x} + \zeta_{t_j} \tilde{x}\big\} \\ &= \gamma_m\big\{\zeta_x + \zeta_{t_j}\big\}(x - \tilde{x}) + \gamma_{xm} x'\big\{\zeta_x + \zeta_{t_j}\big\}(x - \tilde{x}) \end{aligned}$$

Assuming $x = 1$ and $\tilde{x} = 0$, then

$$ACME_j = \gamma_m \zeta_x + \gamma_m \zeta_{t_j} + \gamma_{xm} \zeta_x x' + \gamma_{xm} x' \zeta_{t_j}$$

where x' can take on values 0 or 1. When there is no exposure-mediator interaction in the outcome model (A.2) but there is exposure-time interaction in the mediator model (A.3), then ACME at time j is given by

$$ACME_j = \gamma_m \zeta_x + \gamma_m \zeta_{t_j}$$

When there is exposure-mediator interaction in the outcome model (A.2) but no exposure-time interaction in the mediator model (A.3), then ACME at time j is given by

$$ACME_j = \gamma_m \zeta_x + \gamma_{xm} \zeta_{t_j}$$

When there is no interaction in the outcome model (A.2) and mediator model (A.3),

$$ACME_j = \gamma_m \zeta_x$$

A.2 Checking identification assumptions for data application examples.

Identification assumptions were checked using subject matter knowledge. Using directed acyclic graphs (DAG), we assessed whether a covariate was a confounder or not. The assumption of no mediator-outcome confounder associated with the exposure was assessed using both subject knowledge and empirically.

A.2.1 Severe malaria and cognition data

(i) Exposure-mediator confounder was: Any child education.

(ii) Exposure-outcome confounders were: Age and any child education.

(iii) Mediator-outcome confounders were: Age and any child education.

(iv) No mediator-outcome confounders associated with the exposure: The exposure (malaria group) was regressed on baseline age and then on any child education. The p-values were 0.410 and 0.477 for baseline age and any child education respectively.

(v) No time-varying confounding with respect to the mediator: Since the exposure group (malaria group) for each subject is fixed throughout the study, we do not expect the mediator (hemoglobin level) at time j to affect the exposure at time $j + 1$. We also judge that the cognitive outcome at time j does not affect the mediator (hemoglobin level) at time $j + 1$.

A.2.2 HIV-LIVE (Longitudinal Interrelationships of Viruses and Ethanol) data

(i) Exposure-mediator confounders were: Literacy score, baseline age, and HIV quality of life scale.

(ii) Exposure-outcome confounders were: Baseline age, homeless in the last 6 months, and HIV quality of life scale.

(iii) Mediator-outcome confounders were: Homeless in the last 6 months and HIV quality of life scale.

(iv) No mediator-outcome confounders associated with the exposure: First, the exposure (at risk of drinking) was regressed on HIV quality of life at the 6 time points (1, 3, 5, 6, 7, 8) and p-values were 0.996, 0.453, 0.614, 0.201, 0.092, 0.049 respectively. Second, the exposure was regressed on homeless status in the last 6 months at the 6 time points (1, 3, 5, 6, 7, 8) and p-values were 0.214, 0.545, 0.918, 0.390, 0.564, 0.161 respectively.

(v) Time-varying confounding with respect to the mediator: In this analysis, the exposure status for a subject throughout followup is fixed at baseline. We do not expect the mediator at time j to affect the exposure of any subject at subsequent time points. We acknowledge that if the exposure is time-varying, which is a more likely and practical behavior by study subjects during followup, this assumption can be violated. It is also possible that the outcome (CD4) at time j affects the mediator (adherence) at subsequent time points.

A.3 Simulation results with only three (3) time points.

Results from this simulation with three time points are very similar to the simulation results with six time points (Table A.1). Though the mean square errors under the different mediator random effect distribution assumptions are slightly higher than when there were six time points, they are similar within distribution assumption for normal calibration and SNP calibration (Table A.2).

Table A.1 Simulation results: average causal mediation effect (ACME) and average direct effect (ADE) estimates either assuming normality or SNP approach for the random effects in the mediator calibration model for three time points.

Distribution of RE	Average causal mediation effect (ACME)								
	Normal calibration model			SNP calibration model			Using mediator with error		
	Estimate (-1.360)	Bias	MSE	Estimate (-1.360)	Bias	MSE	Estimate (-1.360)	Bias	MSE
Normal	-1.361	-0.001	0.021	-1.361	-0.001	0.021	-0.501	0.859	0.742
$\chi^2(4)$	-1.358	0.002	0.035	-1.358	0.002	0.035	-0.778	0.582	0.354
Mixture of Normals	-1.363	-0.003	0.025	-1.365	-0.005	0.025	-0.551	0.809	0.661
	Average direct effect (ADE)								
	Estimate (-0.600)	Bias	MSE	Estimate (-0.600)	Bias	MSE	Estimate (-0.600)	Bias	MSE
Normal	-0.604	-0.004	0.014	-0.604	-0.004	0.014	-1.463	-0.863	0.763
$\chi^2(4)$	-0.603	-0.003	0.012	-0.603	-0.003	0.012	-1.183	-0.583	0.360
Mixture of Normals	-0.602	-0.002	0.014	-0.601	-0.001	0.014	-1.414	-0.814	0.683

RE, Random effects; MSE, Mean square error; SNP, Seminonparametric approach. Number Monte-Carlo data sets = 2000.

Table A.2 Simulation results: mean square errors and standard deviation comparing true random effects distribution and estimated random effects distribution assuming normality or using the SNP approach for three time points.

Distribution of Random effects	Normal		χ^2_4		Mixture of Normals	
	MSE	SD	MSE	SD	MSE	SD
Normal calibration	0.456	0.030	0.495	0.037	0.466	0.031
SNP calibration	0.459	0.034	0.490	0.054	0.456	0.052

RE, Random effects; MSE, Mean square error; SD, Standard deviation. The mediator and outcome are independent. Number of Monte-Carlo data sets = 2000.

A.4 Simulation results when the outcome and mediator are correlated

This simulation focused on when the mediator model random effects are normally distributed. Its objective was to investigate changes in the ACME and ADE at different levels of intra-class correlation (ICC) and at different levels of mediator-outcome correlations. The model assumptions were the same as the simulation in the paper except that $Cov(\epsilon_{Y_{ij}}, \epsilon_{M_{ij}})$ was varied to (0, 0.3062, 0.6124, 0.9798), inducing mediator-outcome correlations (0, 0.25, 0.5, 0.8) respectively. Also, the mediator random effect variance was varied such that $b_{0i}^{M} \sim N(0, \sigma^2_{b^M})$, where $\sigma^2_{b^M} = (0.25, 1.00, 2.25, 4.00, 6.25, 9.00, 16.00, 25.00)$, so that the ICC changes for each $\sigma^2_{b^M}$ to (0.14, 0.40, 0.60, 0.73, 0.81, 0.86, 0.91, 0.94) respectively. This yielded 32 different scenarios and 300 Monte-Carlo data sets were generated for each scenario. Results are displayed in Figure A.1.

Results from this smaller simulation show that as within subject variability becomes small relative to between subject variability, the estimates (ACME and ADE) tend towards their true values irrespective of the mediator-outcome correlation. When within subject variability is large relative to between subject variability, higher mediator outcome correlations indicated higher deviation of estimates from their true values. More research towards bias correction in estimates is needed, especially when mediator and outcome are correlated.

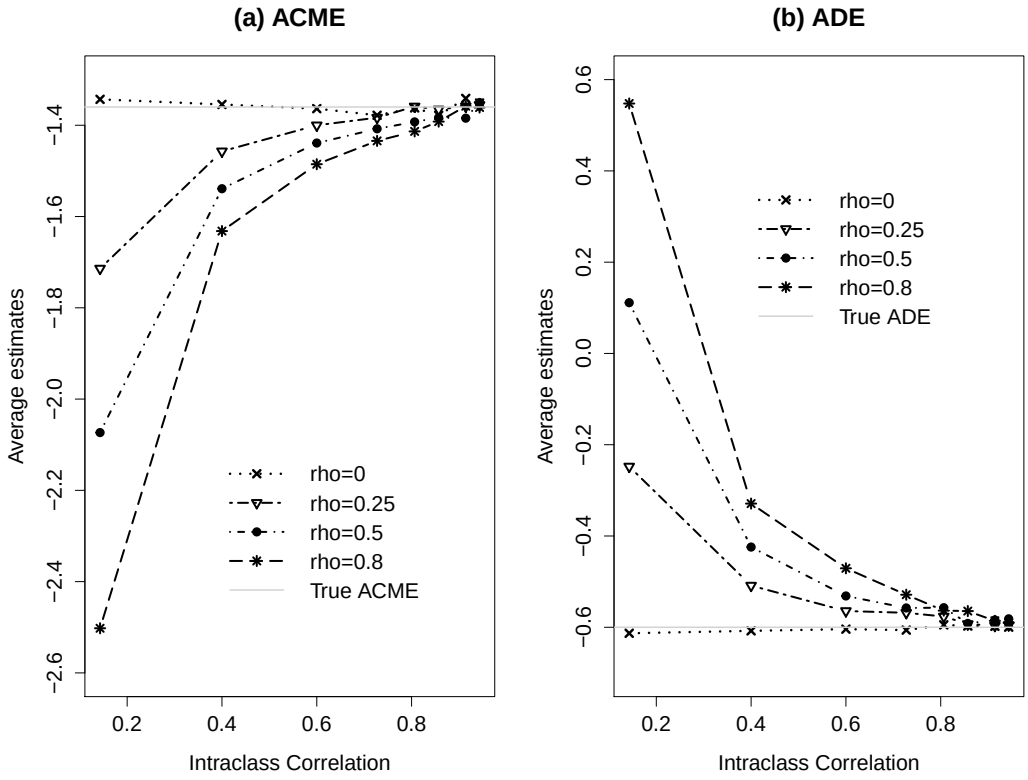

Figure A.1 ACME and ADE estimate averages by intraclass correlation with differing mediator-outcome correlation (rho). Number of Monte-Carlo data sets per average = 300.

Appendix B

Chapter 3 Appendix

B.1 Mediator difference and then impute approach

In this simulation, instead of imputing missing data for the outcome Y_{ij} and mediator M_{ij}, the deviation from the baseline mediator value ($M_{ij} - M_{i1}$) at each time point was first calculated before imputation. If the mediator was missing at a time point, $M_{ij} - M_{i1}$ was missing too at that time point. The outcome and $M_{ij} - M_{i1}$ were then imputed ignoring clustering, considering clustering, and using the linear model (LM). 2000 complete datasets were generated and data was set to missing according to missing completely at random (MCAR) and missing at random (MAR) for each data set, at rates of 10%, 20%, 30% and 40%. 10 data sets were imputed for missing values in each data set using the three methods, ACME and ADE estimated and pooled. Comparisons of the methods on bias, mean square error, mean Sobel standard error (MS-SE), and the coverage probability for the true values were similar with results when the mediator was imputed first and the deviation from baseline computed afterwards.

Table B.1 Simulation results: average causal mediation effect (ACME) estimates when no imputation is conducted and after multiple imputation using different approaches in longitudinal data for differing proportions of missing data. Computing the mediator difference first, followed by multiple imputation.

Prop.	Method	Missing completely at random (MCAR)						Missing at random (MAR)					
		Estimate	Bias	MSE	MC-SD	MS-SE	CP	Estimate	Bias	MSE	MC-SD	MS-SE	CP
Complete		-2.17	-0.01	0.17	0.41	0.40	0.95	-2.17	-0.01	0.17	0.41	0.40	0.95
10%	No imputation	-2.17	-0.01	0.19	0.43	0.42	0.94	-2.25	-0.09	0.19	0.43	0.42	0.94
	IC	-2.19	-0.03	0.19	0.44	0.42	0.94	-2.26	-0.10	0.20	0.44	0.42	0.94
	LM	-2.15	0.01	0.19	0.43	0.42	0.94	-2.24	-0.08	0.19	0.43	0.42	0.95
	CC	-2.17	-0.01	0.18	0.43	0.41	0.94	-2.21	-0.05	0.18	0.42	0.41	0.95
20%	No imputation	-2.17	-0.01	0.20	0.44	0.45	0.95	-2.30	-0.14	0.22	0.45	0.45	0.95
	IC	-2.23	-0.07	0.22	0.46	0.45	0.94	-2.32	-0.16	0.24	0.46	0.45	0.94
	LM	-2.14	0.02	0.20	0.45	0.45	0.95	-2.27	-0.11	0.21	0.45	0.45	0.95
	CC	-2.17	-0.01	0.19	0.43	0.43	0.94	-2.23	-0.07	0.19	0.43	0.43	0.95
30%	No imputation	-2.16	-0.00	0.24	0.49	0.48	0.95	-2.34	-0.18	0.28	0.49	0.49	0.93
	IC	-2.28	-0.12	0.28	0.52	0.48	0.92	-2.39	-0.23	0.33	0.52	0.49	0.90
	LM	-2.11	0.05	0.24	0.49	0.48	0.94	-2.29	-0.13	0.26	0.49	0.48	0.94
	CC	-2.17	-0.01	0.22	0.47	0.44	0.94	-2.25	-0.09	0.23	0.47	0.45	0.93
40%	No imputation	-2.17	-0.01	0.28	0.53	0.52	0.95	-2.37	-0.21	0.33	0.53	0.53	0.94
	IC	-2.34	-0.18	0.36	0.58	0.51	0.90	-2.47	-0.31	0.43	0.58	0.53	0.89
	LM	-2.09	0.07	0.29	0.53	0.51	0.93	-2.27	-0.11	0.30	0.53	0.52	0.93
	CC	-2.17	-0.01	0.25	0.50	0.46	0.93	-2.27	-0.11	0.26	0.50	0.48	0.94

CC, Considering clustering; CP, Coverage probability; IC, Ignoring clustering; LM, Linear model; MC-SD, Monte-Carlo standard deviation; MSE, Mean square error; MS-SE, Mean Sobel standard error; Prop, Proportion of missing data. True ACME = -2.16, sample size = 300, and number of Monte-Carlo data sets = 2000.

Table B.2 Simulation results: average direct effect (ADE) estimates when no imputation is conducted and after multiple imputation using different approaches in longitudinal data for differing proportions of missing data. Computing the mediator difference first, followed by multiple imputation.

Prop.	Method	Missing completely at random (MCAR)						Missing at random (MAR)					
		Estimate	Bias	MSE	MC-SD	Avg-SE	CP	Estimate	Bias	MSE	MC-SD	Avg-SE	CP
Complete		-1.50	0.00	0.07	0.27	0.27	0.96	-1.50	0.00	0.07	0.27	0.27	0.96
10%	No imputation	-1.51	-0.01	0.08	0.29	0.29	0.95	-1.37	0.13	0.10	0.28	0.28	0.93
	IC	-1.51	-0.01	0.08	0.29	0.29	0.95	-1.37	0.13	0.10	0.28	0.28	0.93
	LM	-1.51	-0.01	0.08	0.29	0.29	0.95	-1.37	0.13	0.10	0.28	0.28	0.93
	CC	-1.51	-0.01	0.08	0.27	0.28	0.95	-1.51	-0.01	0.08	0.28	0.28	0.95
20%	No imputation	-1.51	-0.01	0.09	0.30	0.31	0.96	-1.29	0.21	0.13	0.29	0.29	0.88
	IC	-1.51	-0.01	0.09	0.30	0.31	0.96	-1.29	0.21	0.13	0.29	0.30	0.89
	LM	-1.51	-0.01	0.09	0.30	0.31	0.95	-1.29	0.21	0.13	0.30	0.30	0.89
	CC	-1.51	-0.01	0.08	0.28	0.28	0.95	-1.50	0.00	0.08	0.28	0.29	0.95
30%	No imputation	-1.50	0.00	0.11	0.33	0.33	0.94	-1.23	0.27	0.17	0.31	0.32	0.86
	IC	-1.50	0.00	0.11	0.33	0.33	0.94	-1.24	0.26	0.17	0.31	0.32	0.87
	LM	-1.50	0.00	0.11	0.34	0.33	0.94	-1.23	0.27	0.17	0.32	0.32	0.86
	CC	-1.50	0.00	0.08	0.29	0.29	0.95	-1.50	0.00	0.08	0.29	0.29	0.96
40%	No imputation	-1.51	-0.01	0.12	0.35	0.35	0.95	-1.16	0.34	0.23	0.35	0.34	0.82
	IC	-1.51	-0.01	0.12	0.35	0.36	0.95	-1.17	0.33	0.22	0.34	0.35	0.83
	LM	-1.51	-0.01	0.13	0.36	0.36	0.94	-1.17	0.33	0.24	0.36	0.35	0.82
	CC	-1.51	-0.01	0.09	0.29	0.30	0.95	-1.49	0.01	0.09	0.29	0.31	0.96

CC, Considering clustering; CP, Coverage probability; IC, Ignoring clustering; LM, Linear model; MC-SD, Monte-Carlo standard deviation; MSE, Mean square error; Avg-SE, Average standard error; Prop, Proportion of missing data. True ADE = -1.50, sample size = 300, and number of Monte-Carlo data sets = 2000.

Appendix C

Chapter 4 Appendix

C.1 Effect of intermittent preventive treatment in pregnancy (IPTp) on potential mediators

Table C.1 Effect of intermittent preventive treatment in pregnancy (IPTp) on different mediators (maternal malaria, placental malaria, adverse birth outcomes, cord blood hemoglobin, and child malaria at 12 and 24 months).

Exposure	Maternal malaria		Placental malaria		Adverse birth outcome		Cord blood hemoglobin		Child malaria at 12mo		Child malaria at 24mo	
	Effect (SE)	P	Effect (SE)	P	Effect (SE)	P	Effect (SE)	P	Effect (SE)	P	Effect (SE)	P
Three-dose DP	-0.94 (0.37)	0.01	-0.64 (0.31)	0.04	-0.25 (0.32)	0.44	-0.63 (0.35)	0.07	-0.35 (0.66)	0.60	-0.78 (0.38)	0.04
Monthly DP	-1.38 (0.40)	<0.01	-1.02 (0.32)	<0.01	-0.10 (0.31)	0.75	0.22 (0.34)	0.51	0.20 (0.58)	0.73	-0.37 (0.35)	0.28

Reference exposure group: Three-dose sulfadoxine-pyremethamine (SP); DP, Dihydroartenisinin-piperaquine; mo, Months.
To estimate the IPTp effects on mediators, we fitted a linear regression model for cord blood hemoglobin and logistic regression models for the rest of the mediators with only IPTp as the predictor.
Effects are log odds ratios i.e. model coefficients.

C.2 Effect of potential mediators on child neurodevelopment (ND) outcomes at 12 and 24 months

Table C.2 Effect of mediators (maternal malaria, placental malaria, adverse birth outcome, and cord blood hemoglobin) on neurodevelopment outcomes at 12 and 24 months.

Mediators	Cognitive ability		Expressive language		Receptive language		Fine motor		Gross motor	
	Effect (SE)	P	Effect (SE)	P	Effect (SE)	P	Effect (SE)	P	Effect (SE)	P
At 12 months										
Maternal malaria	0.31 (0.38)	0.41	0.41 (0.25)	0.11	0.11 (0.34)	0.74	0.27 (0.41)	0.51	0.38 (0.33)	0.25
Placental malaria	-0.17 (0.31)	0.58	0.03 (0.21)	0.88	-0.17 (0.28)	0.54	0.17 (0.34)	0.62	-0.01 (0.28)	0.97
Maternal or Placental malaria	-0.16 (0.30)	0.59	0.03 (0.20)	0.88	-0.09 (0.27)	0.75	0.32 (0.33)	0.33	0.2 (0.27)	0.47
Adverse birth outcome	-0.32 (0.32)	0.32	0.01 (0.22)	0.97	-0.15 (0.29)	0.61	-0.44 (0.35)	0.21	-0.51 (0.29)	0.08
Cord blood hemoglobin	0.04 (0.07)	0.56	-0.08 (0.05)	0.08	-0.05 (0.06)	0.39	0.07 (0.08)	0.33	-0.08 (0.06)	0.19
Child malaria	-0.15 (0.62)	0.81	-0.31 (0.42)	0.46	-0.10 (0.56)	0.86	-0.56 (0.68)	0.41	-0.95 (0.55)	0.09
At 24 months										
Maternal malaria	-0.01 (0.26)	0.98	0.33 (0.25)	0.18	0.35 (0.33)	0.29	-0.26 (0.26)	0.32	-0.42 (0.43)	0.33
Placental malaria	-0.29 (0.22)	0.18	-0.19 (0.20)	0.35	-0.45 (0.27)	0.10	-0.04 (0.22)	0.84	0.24 (0.36)	0.50
Maternal or Placental malaria	-0.15 (0.21)	0.48	-0.05 (0.20)	0.82	-0.23 (0.27)	0.39	-0.13 (0.21)	0.54	-0.04 (0.35)	0.90
Adverse birth outcome	-0.35 (0.23)	0.12	-0.55 (0.21)	0.01*	-0.52 (0.29)	0.07	-0.55 (0.23)	0.02*	-0.56 (0.37)	0.14
Cord blood hemoglobin	0.00 (0.05)	0.94	0.02 (0.04)	0.72	-0.06 (0.06)	0.35	0.02 (0.05)	0.73	-0.02 (0.08)	0.80
Child malaria	0.33 (0.37)	0.37	-0.32 (0.25)	0.19	0.14 (0.33)	0.68	-0.21 (0.4)	0.61	-0.39 (0.33)	0.23

SE, Standard error; *P < 0.05

To estimate the mediator effects on neurodevelopment outcomes, we fitted linear regression models with each mediator as the only predictor.

C.3 Mediation effects of intermittent preventive treatment in pregnancy on neurodevelopment outcomes at 12 and 24 months (tabular form)

Table C.3 displays results for Figures 4.2 and 4.3 in tabular format.

Table C.3 Mediation effects of maternal malaria prophylaxis on child neurodevelopment outcomes through different mediators among children whose mothers received either three-dose DP or monthly DP.

Mediators	Time point	Prophylaxis	Neurodevelopment outcome measures				
			Cognitive ability Effect (95% CI)	Expressive language Effect (95% CI)	Receptive language Effect (95% CI)	Fine motor Effect (95% CI)	Gross motor Effect (95% CI)
Maternal malaria	12 Months	3 dose DP	0.10 (-1.12, 1.32)	-1.28 (-1.95, -0.67)*	0.03 (-1.05, 1.28)	-0.89 (-2.09, 0.27)	-0.85 (-1.88, 0.29)
		Monthly DP	-0.52 (-2.58, 1.74)	0.52 (-0.22, 1.33)	-0.43 (-1.55, 0.88)	0.88 (-0.90, 2.78)	-0.73 (-1.83, 0.30)
	24 Months	3 dose DP	0.52 (-0.58, 1.70)	-0.38 (-1.30, 0.61)	0.11 (-0.88, 1.19)	0.79 (0.15, 1.56)*	1.00 (-0.06, 2.22)
		Monthly DP	-0.16 (-0.68, 0.46)	0.24 (-0.67, 1.23)	0.13 (-0.93, 1.28)	0.06 (-1.24, 1.56)	0.83 (-0.55, 2.04)
Placental malaria	12 Months	3 dose DP	0.87 (0.20, 1.54)*	0.42 (-0.17, 1.09)	0.44 (-0.14, 1.02)	-0.19 (-0.99, 0.66)	0.51 (-0.29, 1.36)
		Monthly DP	-0.28 (-1.06, 0.53)	-0.50 (-1.06, -0.05)*	-0.42 (-1.19, 0.27)	-0.31 (-1.26, 0.63)	-0.62 (-1.32, 0.07)
	24 Months	3 dose DP	0.45 (-0.11, 1.09)	0.68 (0.19, 1.20)*	0.76 (0.19, 1.48)*	0.40 (-0.05, 0.88)	0.05 (-0.88, 0.99)
		Monthly DP	0.01 (-0.55, 0.56)	-0.17 (-0.65, 0.29)	0.16 (-0.57, 0.86)	-0.08 (-0.63, 0.48)	-0.7 (-1.94, 0.33)
Adverse birth outcome	12 Months	3 dose DP	0.50 (-0.43, 1.46)	0.17 (-0.47, 0.93)	0.13 (-0.5, 0.75)	0.31 (-0.63, 1.26)	0.55 (-0.46, 1.69)
		Monthly DP	0.31 (-0.46, 1.19)	0.06 (-0.41, 0.54)	-0.04 (-0.69, 0.60)	0.28 (-0.65, 1.15)	0.25 (-0.43, 0.86)
	24 Months	3 dose DP	0.50 (-0.07, 1.07)	0.77 (0.24, 1.33)*	0.89 (0.18, 1.64)*	0.60 (0.10, 1.11)*	0.18 (-0.92, 1.21)
		Monthly DP	0.00 (-0.39, 0.41)	-0.14 (-0.67, 0.35)	-0.23 (-0.96, 0.47)	0.59 (-0.09, 1.27)	0.78 (-0.06, 1.63)
Cord blood hemoglobin	12 Months	3 dose DP	-0.03 (-0.21, 0.16)	0.05 (-0.05, 0.19)	0.01 (-0.09, 0.09)	0.02 (-0.14, 0.16)	0.09 (-0.03, 0.28)
		Monthly DP	-0.02 (-0.16, 0.08)	-0.02 (-0.15, 0.04)	-0.02 (-0.17, 0.06)	0.03 (-0.07, 0.21)	-0.01 (-0.14, 0.07)
	24 Months	3 dose DP	0.00 (-0.09, 0.11)	-0.04 (-0.19, 0.05)	0.06 (-0.02, 0.16)	-0.04 (-0.13, 0.02)	0.01 (-0.10, 0.11)
		Monthly DP	-0.03 (-0.14, 0.06)	-0.04 (-0.20, 0.05)	-0.07 (-0.33, 0.09)	-0.04 (-0.21, 0.07)	-0.04 (-0.25, 0.10)
Child malaria	24 Months	3 dose DP	-0.03 (-0.21, 0.16)	0.05 (-0.05, 0.19)	0.01 (-0.09, 0.09)	0.02 (-0.14, 0.16)	0.09 (-0.03, 0.28)
		Monthly DP	-0.02 (-0.16, 0.08)	-0.02 (-0.15, 0.04)	-0.02 (-0.17, 0.06)	0.03 (-0.07, 0.21)	-0.01 (-0.14, 0.07)

CI, 95% percentile confidence interval from 1000 bootstrap samples; DP, Dihydroartenisinin-piperaquine.

†Includes children who either had a preterm birth, were small for gestation age (<10th percentile), or had a low birth weight (< 2.5 kilograms).

*Significant i.e. 95% percentile confidence interval excludes the null (zero).

C.4 Effects of IPTp on child ND outcomes at 12 and 24 months mediated by any maternal malaria (maternal or placental malaria) (tabular form)

Table C.4 displays results for Figure 4.4 in tabular format.

Table C.4 Indirect effects of any maternal malaria (maternal malaria or placental malaria or both) on the effect of malaria prophylaxis on child neurodevelopment outcomes at 12 and 24 months among children whose mothers received either three-dose DP or monthly DP.

ND outcomes	12 months		24 months	
	Three-dose DP Effect (95% CI)	Monthly DP Effect (95% CI)	Three-dose DP Effect (95% CI)	Monthly DP Effect (95% CI)
Cognitive ability	0.62 (0.04, 1.27)*	-0.32 (-1.06, 0.43)	0.28 (-0.16, 0.78)	-0.04 (-0.46, 0.40)
Expressive language	0.13 (-0.36, 0.64)	-0.23 (-0.64, 0.15)	0.39 (0.00, 0.83)*	-0.08 (-0.47, 0.31)
Receptive language	0.40 (-0.08, 0.95)	-0.51 (-1.13, 0.06)	0.55 (0.07, 1.12)*	0.18 (-0.38, 0.79)
Fine motor	-0.41 (-1.09, 0.27)	-0.02 (-0.80, 0.75)	0.35 (-0.01, 0.75)	-0.07 (-0.57, 0.51)
Gross motor	0.16 (-0.53, 0.82)	-0.53 (-1.13, 0.06)	0.21 (-0.50, 0.92)	-0.16 (-1.05, 0.67)

CI, 95% percentile confidence interval from 1000 bootstrap samples; DP, Dihydroartenisinin piperaquine; ND, Neurodevelopment.
*Significant i.e. 95% percentile confidence interval excludes the null (zero).

C.5 Derivation of the mediation effect when there is one mediator, a mediator-treatment interaction, and adjusting for child prophylaxis

Assuming Y, M, A, and X are a neurodevelopment outcome, a binary mediator, a binary exposure (IPTp), and child prophylaxis respectively.

$$\begin{aligned} E(M|A) &= \frac{exp(\theta_0 + \theta_a A)}{1 + exp(\theta_0 + \theta_a A)} \\ E(Y|A, M, X) &= \beta_0 + \beta_a A + \beta_m M + \beta_{am} AM + \beta_x X \end{aligned}$$

The average causal mediation effect (ACME) or indirect effect is given by,

$$\begin{aligned} ACME &= E\{Y(1, M(1), x) - Y(1, M(0), x)\} \\ &= \beta_m \left[\frac{exp(\theta_0 + \theta_a)}{1 + exp(\theta_0 + \theta_a)} - \frac{exp(\theta_0)}{1 + exp(\theta_0)} \right] \\ &+ \beta_{am} \left[\frac{exp(\theta_0 + \theta_a)}{1 + exp(\theta_0 + \theta_a)} - \frac{exp(\theta_0)}{1 + exp(\theta_0)} \right] \end{aligned}$$

C.6 Derivation of the mediation effect with two mediators

Let Y be a neurodevelopment outcome, M_1 be a binary any maternal malaria (placental or maternal malaria) mediator, M_2 be a binary child malaria mediator, A be a binary maternal malaria prophylaxis, and X be any covariate that affects M_2 and Y but not M_1. The derivation assumes that M_1 affects M_2 but not the other way around (Figure C.1). Our analysis involving two mediators in Chapter 4 does not include X. Then,

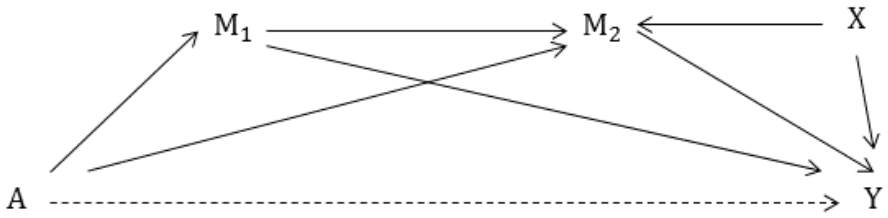

Figure C.1 Directed acyclic graph (DAG) for two mediators.

$$
\begin{aligned}
E(M_1|A) &= \frac{exp(\theta_0 + \theta_a A)}{1 + exp(\theta_0 + \theta_a A)} \\
E(M_2|A, M_1, X) &= \frac{exp(\alpha_0 + \alpha_a A + \alpha_3 M_1 + \alpha_4 X)}{1 + exp(\alpha_0 + \alpha_a A + \alpha_3 M_1 + \alpha_4 X)} \\
E(Y|A, M_1, M_2, X) &= \beta_0 + \beta_a A + \beta_3 M_1 + \beta_4 M_2 + \beta_5 X
\end{aligned}
$$

Indirect effects can then be partitioned into indirect effects through only M_1, through only M_2, and through both M_1 and M_2. These effects together with the direct effect of A to Y have been found to sum to the total effect (Daniel et al., 2015).

1. Indirect effect through only M_1

$$\begin{aligned} IE_{M_1} &= E\{Y(1, M_1(1), M_2(0, M_1(0), x), x) \\ &- Y(1, M_1(0), M_2(0, M_1(0), x), x)\} \\ &= \beta_3 \left[\frac{exp(\theta_0 + \theta_a)}{1 + exp(\theta_0 + \theta_a)} - \frac{exp(\theta_0)}{1 + exp(\theta_0)} \right] \end{aligned}$$

2. Indirect effect through only M_2

$$\begin{aligned} IE_{M_2} &= E\{Y(1, M_1(1), M_2(1, M_1(0), x), x) \\ &- Y(1, M_1(1), M_2(0, M_1(0), x), x)\} \\ &= \beta_4 \left[\frac{exp(\alpha_0 + \alpha_a + \alpha_3 \left(\frac{exp(\theta_0)}{1+exp(\theta_0)} \right) + \alpha_4 x)}{1 + exp(\alpha_0 + \alpha_a + \alpha_3 \left(\frac{exp(\theta_0)}{1+exp(\theta_0)} \right) + \alpha_4 x)} \right. \\ &- \left. \frac{exp(\alpha_0 + \alpha_3 \left(\frac{exp(\theta_0)}{1+exp(\theta_0)} \right) + \alpha_4 x)}{1 + exp(\alpha_0 + \alpha_3 \left(\frac{exp(\theta_0)}{1+exp(\theta_0)} \right) + \alpha_4 x)} \right] \end{aligned}$$

3. Indirect effect through both M_1 and M_2

$$\begin{aligned} IE_{M_1,M_2} &= E\{Y(1, M_1(1), M_2(1, M_1(1), x), x) \\ &- Y(1, M_1(1), M_2(1, M_1(0), x), x)\} \\ &= \beta_4 \left[\frac{exp(\alpha_0 + \alpha_a + \alpha_3 \left(\frac{exp(\theta_0+\theta_a)}{1+exp(\theta_0+\theta_a)} \right) + \alpha_4 x)}{1 + exp(\alpha_0 + \alpha_a + \alpha_3 \left(\frac{exp(\theta_0+\theta_a)}{1+exp(\theta_0+\theta_a)} \right) + \alpha_4 x)} \right. \\ &- \left. \frac{exp(\alpha_0 + \alpha_a + \alpha_3 \left(\frac{exp(\theta_0)}{1+exp(\theta_0)} \right) + \alpha_4 x)}{1 + exp(\alpha_0 + \alpha_a + \alpha_3 \left(\frac{exp(\theta_0)}{1+exp(\theta_0)} \right) + \alpha_4 x)} \right] \end{aligned}$$

4. Total mediation effect (through M_1, M_2, and through both M_1 and M_2).

$$
\begin{aligned}
TME &= E\{Y(1, M_1(1), M_2(1, M_1(1), x), x) \\
&- Y(1, M_1(0), M_2(0, M_1(0), x), x)\} \\
&= \beta_3 \left[\frac{exp(\theta_0 + \theta_a)}{1 + exp(\theta_0 + \theta_a)} - \frac{exp(\theta_0)}{1 + exp(\theta_0)} \right] \\
&+ \beta_4 \left[\frac{exp(\alpha_0 + \alpha_a + \alpha_3 \left(\frac{exp(\theta_0 + \theta_a)}{1 + exp(\theta_0 + \theta_a)} \right) + \alpha_4 x)}{1 + exp(\alpha_0 + \alpha_a + \alpha_3 \left(\frac{exp(\theta_0 + \theta_a)}{1 + exp(\theta_0 + \theta_a)} \right) + \alpha_4 x)} \right. \\
&- \left. \frac{exp(\alpha_0 + \alpha_3 \left(\frac{exp(\theta_0)}{1 + exp(\theta_0)} \right) + \alpha_4 x)}{1 + exp(\alpha_0 + \alpha_3 \left(\frac{exp(\theta_0)}{1 + exp(\theta_0)} \right) + \alpha_4 x)} \right]
\end{aligned}
$$