

ABSTRACT

Title of Dissertation: Bayesian Hierarchical Meta-Analysis Using Individual Participant Data for Modeling Heterogeneous Dropout Patterns Across Multiple Clinical Trials (Part I)

AND

Efficacy of Randomized Clinical Trials in Adolescent Patients with Schizophrenia: A Network Meta Analysis (Part II)

Author: Eshetu Tefera

Advisor: Dr. Shuo Chen
School of Medicine ,University of Maryland

Co-Advisor: Dr. Chixiang Chen
School of Medicine ,University of Maryland

Missing data is a pervasive issue, notably in clinical trials and prospective studies employing a longitudinal design. This problem becomes particularly pronounced when dealing with data from multi-study clinical trials. Many established research teams undertake multiple clinical trials in closely related domains, each of which may exhibit distinct patterns of patient attrition/dropout. However, due to the shared group panel and study administration, these dropout patterns are often believed to exhibit similarities across these trials. While models addressing single dropout mechanism have been extensively investigated, the analysis of heterogeneous dropout patterns remains understudied. To leverage heterogeneous data and integrate information from multiple missing mechanisms, the first part of this dissertation discusses

a new meta-analysis strategy based on individual participant data (IPD) to model observational-level dropout patterns over multiple trials and improve statistical inference via a Bayesian Hierarchical Model (BHM). Extensive simulation studies were conducted to demonstrate the superiority of the new method over existing methods in terms of reduced bias, smaller estimation variability, and higher statistical power. The proposed method was also applied to 13 clinical trials for schizophrenia research exploring demographic and clinical determinants of dropout.

Most clinical trials enroll adults and treatment efficacy data in children and adolescents have become scarce. The second part of this dissertation deals with network meta-analysis in examining efficacy of schizophrenia treatments in underage population. Traditional meta-analysis is restricted to direct comparison of treatments in parallel study designs. However, network meta-analysis is a special kind of meta-analysis for the comparison of multiple treatments simultaneously in a single analysis by combining direct and indirect evidence. The direct evidence is obtained from randomized control trials with direct comparison of treatments in parallel study design. The indirect evidence is obtained from comparisons of one or more common comparators by transitivity. The direct and indirect pooled evidence is the network estimate. Network meta-analysis is vital in examining the effectiveness of different treatments in randomized control trials. Accordingly, it is gaining popularity.

Bayesian Hierarchical Meta-Analysis Using Individual Participant
Data for Modeling Heterogeneous Dropout Patterns Across Multiple
Clinical Trials (Part I)

AND

Efficacy of Randomized Clinical Trials in Adolescent Patients with
Schizophrenia: A Network Meta Analysis (Part II)

by

Eshetu Tefera

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2025

Committee:

Dr. Shuo Chen, Chair

Dr. Chixiang Chen

Dr. Tianzhou Ma

Dr. Lizhen Lin

Dr. Xin He

© Copyright by
Eshetu A Tefera
2025

Acknowledgments

First and foremost, I would like to thank the Almighty God who has given me strength and encouragement throughout my study and helping me throughout my life.

I express my deepest gratitude to my dissertation advisors, Prof. Shuo Chen and Dr. Chixiang Chen, for their continual advice, supervision, encouragement and unwavering support. Their insightful comments and valuable feedback have been instrumental for this accomplishment. This dissertation work would not have been achievable without their extraordinary patience, guidance and continued support. Further, I would like to extend my appreciation to my dissertation committee members, Dr. Tianzhou Ma, Prof. Lizhen Lin and Dr. Xin He.

The author expresses sincere gratitude to the Maryland Psychiatric Research Center (MPRC) for providing the necessary resources and data that significantly contributed to conduct this research. It is my pleasure to thank Dr. Paul Smith and Prof. Benjamin Kedem for their encouragement and support. I also thank all the faculty and staff of the Department of Mathematics of the University of Maryland College Park for their support throughout my studies at the University of Maryland. My sincere appreciation goes to Dr. David Li for his great support. I am also grateful to my colleagues Dr. Nawar Shara and Dr. Mihriye Mete who have supported me

during my PhD studies. In addition, I would like to thank Medstar Health Research Institute (MHRI) for providing support during my study.

Most of all, this dissertation is dedicated to the memory of my beloved father, whose love, blessing, prayer, and support continue to guide me, even in his absence. His belief in my potential fueled my PhD journey and this achievement is a tribute to his legacy. I would like to express my sincere gratitude to my mother for her love, support, and constant prayer. I never forget the reaction you showed me when I passed the qualification exams saying 'Elelelele...?!'. I am heartily thankful to my wife Dr. Tigist Kahsay, for her love, encouragement, support, and prayer.

I would like to thank all my family members: brothers, sisters, in-laws, nephews, nieces, cousins, relatives, and friends here and back home who have supported me in many ways during my study. Your support, encouragement and prayer contributed significantly to the completion of my study. Before I end, I offer my gratitude to all those who supported me in any way during the completion of my dissertation.

Table of Contents

Table of Contents	iv
I Bayesian Hierarchical Meta-Analysis Using Individual Participant Data for Modeling Heterogeneous Dropout Patterns Across Multiple Clinical Trials	1
Chapter 1: Introduction	1
Chapter 2: Method	5
2.1 Motivating data	5
2.2 Dropout model in the longitudinal data	6
2.3 WGEE	11
Chapter 3: Simulations	13
3.1 Data generation	13
3.2 Benchmarks	14
3.3 Evaluations	16
Chapter 4: Data Application	17
Chapter 5: Discussions	19
II Efficacy of Randomized Clinical Trials in Adolescent Patients with Schizophrenia: A Network Meta Analysis	32
Chapter 6: Introduction	33
Chapter 7: Method	35
7.1 Selection Criteria	35
7.2 Data Extraction	36
7.3 Intervention Description	38
7.4 Study Characteristics	38
7.5 Statistical Analysis	40
Chapter 8: Result	45

8.1	Primary Analysis: Combined analysis	45
8.2	Sensitivity analysis: Separate Analysis	55
Chapter 9: Discussion		69
9.1	Strengths	69
9.2	Limitations	70
9.3	Conclusions	70

Part I

Bayesian Hierarchical Meta-Analysis Using Individual Participant Data for
Modeling Heterogeneous Dropout Patterns Across Multiple Clinical Trials

Chapter 1: Introduction

Clinical trials are essential research tools for advancing medical knowledge and improving patient care, especially in the face of the devastating pandemic. They are critical means to develop vaccines and treatments to fight disease and save millions of lives [Han, 2015]. Clinical trials are also crucial for discovering new treatments for illness, inventing new ways to detect, diagnose, and reduce the chance of developing disease [NIH, 2020].

However, the dropout of participants from clinical trials seriously affects the credibility of results and significantly decreases the trial's capacity to impact clinical practice [Gillies et al., 2018]. High dropout rates during the study period can introduce bias and create problems in data analysis and result interpretation. Despite significant efforts to promote participant retention in clinical trials, the patient dropout continues to pose a significant challenge. Various factors contribute to participant dropout, including protocol deviations, inconvenience, lack of adherence, insufficient monitoring, withdrawal of consent, personal or familial obligations, adverse effects, non-compliance, and other reasons [Rohden et al., 2017]. Some participants may also withdraw from the trial for reasons that are not immediately identifiable. Lowering the dropout rate is crucial for the success of clinical trials. Therefore, identifying the

risk factors that contribute to participant dropout is immensely valuable for informing future trial designs and enhancing the overall quality of clinical research.

Furthermore, beyond the issue of dropout in a single trial, some study groups or centers conduct multiple clinical trials in related fields, each with its own dropout patterns. Although there is a common belief that dropout patterns are similar across these trials because of their shared group panel and study administration, aggregating all the missing data without considering study-specific dropout profiles can obscure important distinctions and lead to inaccurate inferences and decisions. However, the integration of information from heterogeneous dropout patterns over multiple trials has not been well explored. Meta-analysis is one of the most successful and widely applied integration tools in the literature [[Haidich, 2010](#), [Lin and Zeng, 2010](#)]. However, the traditional meta-analysis approach, which uses aggregate data (AD) obtained from published articles, conference abstracts, or trial registries like “clinicaltrials.gov” has several disadvantages and limitations [[Lee, 2019](#)]. These include heterogeneity, publication bias, inability to overcome subjectivity, and focus on only main effects [[Lee, 2019](#)]. These limitations may lead to inaccurate statistical inference [[Esterhuizen and Thabane, 2016](#)]. In the presence of individual-level data, we propose a new individual participant data (IPD) meta-analysis strategy, which involves analyzing the original clinical trial data obtained directly from individual studies [[Tierney et al., 2015](#)]. IPD meta-analysis is crucial for data transparency and reducing potential bias [[Tierney et al., 2015](#)]. In this study, we are able to obtain original data from multiple clinical trials conducted at the Maryland Psychiatric Research Center (MPRC). We included 13 phase II trials completed between 2009 to 2016, and our study aims to

explore demographic and clinical determinants of dropout in Schizophrenia trials.

We propose a new strategy for analyzing individual participant data (IPD) meta-analysis to model multiple dropouts and improve statistical inference. To address the challenges of leveraging heterogeneous data and incorporating information from multiple trials, we utilize a Bayesian hierarchical model (BHM) to estimate covariate effects based upon sampling posterior distributions. The Bayesian approach allows for inference on each trial and for the combined effect, making it more flexible under small sample sizes and better at accounting for estimation and prediction uncertainty. The key advantage of Bayesian methods over traditional frequentist methods is that Bayesian method provide flexible framework to aggregate multiple clinical trials, effect sizes vary across trials however the population level can better be captured by aggregating across different trials of effect sizes by the hierarchical structure which is challenging for traditional frequentist method. In addition, Bayesian methods naturally account for uncertainty by resampling strategy and could better capture potential heterogeneity. Traditional frequentist methods often only provide point estimates without capturing the full extent of uncertainty involved[Fornacon-Wood et al., 2022]. Our proposed IPD meta-analysis strategy models multiple dropout patterns in longitudinal clinical trials based on BHM, which has not been explored in the existing literature. Our proposed method is not limited to specific case studies like the current studies we used, it can be applied in many clinical trials including psychiatric research where the analysis of individual patient data (IPD) is conducted using an "intention-to-treat" (ITT) approach, which means all patients who are randomized are included in the analysis irrespective of whether they completely adhered to the

treatment assigned or dropped out of the study. We note that, while generalized linear mixed-effect models (GLMM) can be an alternative approach for modeling dropouts accounting for across-study variability [Li et al., 2022], their primary goal is to infer overall effects from covariates and account for multi-level variability among trials. In contrast, BHM directly estimates the heterogeneous covariate effects by integrating information from multiple heterogeneous trials. Furthermore, GLMM can be numerically challenging to implement when there are many random effects. In addition, our proposed method enables more accurate estimates of dropout probabilities at each time point, which is an essential intermediate step to make valid statistical inferences for main outcomes of interest in clinical trials. These outcomes could be modeled by applying weighted generalized estimating equations (WGEE) [Robins et al., 1995b, Chen et al., 2019]. Extensive simulation studies are implemented to demonstrate the superiority of our method over existing methods in terms of reduced bias, smaller estimation variability, and higher statistical power.

The remaining article is structured as follows: In the Methods section, we detail our motivating data and the proposed method. Moving on to the Simulations section, we delve into extensive numerical evaluations of our method in comparison to existing approaches. Subsequently, the Data Application section is dedicated to a comprehensive analysis and interpretation of our results. Finally, in the Discussion

section, we provide a summary and conclusion for the article.

Chapter 2: Method

2.1 Motivating data

We focus on 13 longitudinal clinical trials with individual level data for schizophrenia research. These trials were conducted within MPRC and share certain similarity. On the other hand, they were conducted by separate studies and there should be some heterogeneity among the trials. These are motivations to develop our BHM. In each trial, participants visit clinic repeatedly and may dropout the study due to various reasons. In order to unbiasedly analyze this data, it is imperative to address the issue of patients loss to follow-up (dropout). Dropout can significantly influence the outcomes of clinical trials, potentially leading to biased causal inferences. This study is dedicated to exploring potential factors that influence dropout rates in Schizophrenia trials conducted at MPRC. Identifying these factors can contribute to a deeper understanding of patient dropout mechanisms and provide valuable insights for the design of future trials. To achieve this, we analyzed data from 13 randomized clinical trials conducted between 2009 to 2016. Electronic trial data were retrieved from the MPRC database by a database administrator with access to all electronic medical records. Demographic and clinical characteristics of participants were extracted from the available data. Psychiatric symptoms were measured using the Brief Psychiatric

Rating Scale (BPRS), which assesses symptoms such as somatic concern, depression, anxiety, conceptual disorganization, hallucinations, and unusual behavior (Reference for BPRS may be added as needed). A BPRS score of one indicates the absence of symptoms, while a score of seven represents severe symptoms. We calculated BPRS total score and change in BPRS total score from baseline. Our descriptive data analysis revealed that dropout rates differed between trials (Table 1). The objective of this paper is to identify common and shared factors that are independently or in combination associated with dropout across trials, as well as factors specific to individual trials. By examining these factors, we aim to provide insights that could help reduce dropout rates and improve the design of future clinical trials.

2.2 Dropout model in the longitudinal data

In this section, we present a novel IPD meta-analysis strategy designed to address and leverage the diverse dropout patterns observed across multiple trials. Our approach involves the utilization of a Bayesian Hierarchical Model (BHM) for dropout modeling to enhance statistical inference. The BHM has been purposefully crafted to estimate posterior distribution parameters through Bayesian methods within hierarchical models, enabling trial-specific inferences while also deriving a overall effect estimate. With its capacity to accommodate various structures, the BHM offers flexibility in quantifying data variability across multiple levels, aiding in the comprehension of the heterogeneous dropout trajectories observed in various trials. Despite the variations in these trajectories, we posit that commonalities exist among them. These

commonalities can be effectively managed by incorporating a well-founded prior in the BHM. This prior serves to integrate information from across trials, resulting in more efficient and accurate parameter estimates. In the remainder of this section, we provide a detailed exposition of our utilization of the BHM in the context of longitudinal data to address the diverse dropout patterns observed across multiple trials within the realm of IPD meta-analysis.

Notations: Suppose there are J trials. For each trial $j \in 1, 2, \dots, J$, let i be the index of the subject, ranging from 1 to n_j . The patient sizes can vary across trials. Within each subject i of the j -th trial, let t denote the index of visit ranging from 2 to T_{ji} . We denote M_{jit} as a binary indicator for observing the subject i in the j -th trial at visit t . The covariates included in the vector x_{jit} could be in general all observed variables in studies. For illustration and based on our specific application, we consider the intercept, gender, age, visit time, and BPRS. Furthermore, we denote the vector of covariate effect sizes on patient dropout as β_j for the j -th trial, which is allowed to be slightly different among J trials, and assume that M_{jit} follows conditional Bernoulli distribution, i.e, $M_{jit} = 1 | M_{ji(t-1)} = 1, x_{jit} \sim \text{Bernoulli}(\pi_{jit})$. With all foregoing notations, the model of observing the subject i in the j -th trial at time t can be expressed as

$$\text{logit}(\pi_{jit}) = \beta_{j,1} + \beta_{j,2} \times \text{Gender}_{ji} + \beta_{j,3} \times \text{Age}_{ji} + \beta_{j,4} \times \text{Time}_{ji} + \beta_{j,5} \times \Delta \text{BPRS}_{ji(t-1)}, \quad (2.1)$$

where $\Delta \text{BPRS}_{i(t-1)}$ is the change in BPRS from $t - 1$ to the baseline.

We remark here that the conditional probability π_{jit} will be one if $t = 1$ holds

(always observed at the baseline), and zero if $M_{ji(t-1)} = 0$, which models the feature of dropout.

Likelihood. To build BHM, we need to first specify the likelihood of dropout data. In the context of multiple trial data, the joint likelihood can be written as

$$P(M_{jit} = m_{jit} | x_{jit}) = \prod_{j=1}^J \prod_{i=1}^{n_j} \prod_{t=2}^{t_i} \left\{ P(M_{jit} = m_{jit} | M_{ji(t-1)} = m_{ji(t-1)}, x_{jit})^{m_{ji(t-1)}} \times P(M_{ji1} = m_{ji1}) \right\}, \quad (2.2)$$

where $P(M_{jit} = m_{jit} | M_{ji(t-1)} = m_{ji(t-1)}, x_{jit})$ is equivalent to π_{jit} . The above likelihood contains the probability of dropout by using chain rule of conditional probabilities and acknowledging the feature of dropout, which is widely considered in existing literature [Robins et al., 1995b, Chen et al., 2019, Shen et al., 2023]. We hypothesize that the previous observed outcome can have an impact on dropout at the next time point, as patients with more severe symptoms may be more likely to drop out of the study [Najavits, 2015]. By incorporating the last observation of the outcome into the model, we assume that the missing mechanism is missing at random [Bhaskaran and Smeeth, 2014, Chen et al., 2021]. This assumption implies that there are no unmeasured confounders [Resseguier et al., 2011, Choi et al., 2019], which is reasonable in this paper since these clinical trials are well-designed and the research teams at MPRC will make every effort to record all possible factors relevant to patient outcomes

Priors. To calculate the posterior distribution, the next step is to specify multiple levels of priors. Specifically, the first one is $\beta_j \sim \mathcal{MVN}(\mu_\beta, \Sigma_\beta)$ (\mathcal{MVN} stands for multivariate normal distribution), which quantifies the variability of effect sizes across different trials. We assume that the vector of covariate effect sizes on

patient dropout, β_j , follows a multivariate normal distribution with mean vector μ_β and variance matrix Σ_β . The mean vector quantifies the overall effect across multiple trials, while the magnitude of variance reflects our belief about the similarity of dropout patterns across different trials. To make an objective and robust inference, we assign another level of prior to these variables, i.e., $\mu_\beta \sim \mathcal{MVN}(\mu, \Omega_1)$ and $\Sigma_\beta \sim W_5(\Omega_2, 5)$. There are 5 covariates for the overall effect (intercept, gender, age, time and BPRS) so that the covariance matrix Σ_β follows a Wishart distribution with 5 degrees of freedom, The covariate matrix Σ_β follows a Wishart distribution with 5 degrees of freedom, and the scale matrix which determines the variance structure is set as an identity matrix, meaning each diagonal element is 1 and all off-diagonal elements are 0; essentially representing a situation where all variables are independent and have the same variance. Note that we need to specify the hyperparameters in the priors, i.e., $\mu = [0, 0.88, 0, 0, 0]^T$, $\Omega_1 = \Omega_2 = (1/1000)I_5$, I_5 is identity matrix of size 5. For the hyperparameters, we set the value of zero for all variables of interest except for gender, for which we used the value obtained from our data since it remains unchanged.

Full conditional. Together with the given likelihood and priors, we can derive the full conditional distribution for the parameters of interest. Specifically, we first derive the full conditional of $\beta_j | M_{jit}, \mu_\beta, \Sigma_\beta, x_{jit}$, denoted by $P(\beta_j | M_{jit}, \mu_\beta, \Sigma_\beta, x_{jit})$ by using the likelihood function, i.e., $L(\beta_j | M_{jit}, x_{jit})$, and prior density function, i.e., $P(\beta_j | \mu_\beta, \Sigma_\beta)$:

$$\begin{aligned}
& P(\beta_j | M_{jit}, \mu_\beta, \Sigma_\beta, x_{jit}) \propto L(\beta_j | M_{jit}, x_{jit}) \times P(\beta_j | \mu_\beta, \Sigma_\beta) \\
& = \left(\prod_{j=1}^J \prod_{i=1}^{n_j} \prod_{t=2}^{t_i} P(M_{jit} = m_{jit} \mid M_{ji(t-1)} = m_{ji(t-1)}, x_{jit})^{m_{ji(t-1)}} \times P(M_{ji1} = m_{ji1}) \right) \times \\
& \quad (2\pi^{-5/2} \det(\Sigma_\beta)^{-1/2})^J \times \exp \left(\sum_{j=1}^J -1/2(\beta_j - \mu_\beta)^T \Sigma_\beta^{-1} (\beta_j - \mu_\beta) \right).
\end{aligned} \tag{2.3}$$

Second, we derive the full conditional of $\mu_\beta | \beta_j, \mu, \Omega_1$ i.e, $P(\mu_\beta | \beta_j, \mu, \Omega_1)$ by using two prior density functions, i.e., $P(\beta_j | \mu_\beta, \Sigma_\beta)$ and $P(\mu_\beta | \mu, \Omega_1)$, based on Bayesian theorem:

$$\begin{aligned}
& P(\mu_\beta | \beta_j, \mu, \Omega_1) \propto P(\beta_j | \mu_\beta, \Sigma_\beta) \cdot P(\mu_\beta | \mu, \Omega_1) \\
& \propto \exp \left(\sum_{j=1}^J -1/2(\beta_j - \mu_\beta)^T \Sigma_\beta^{-1} (\beta_j - \mu_\beta) \right) \times \\
& \exp \left(\sum_{\beta=1}^5 -1/2(\mu_\beta - \mu)^T \Omega_1^{-1} (\mu_\beta - \mu) \right).
\end{aligned} \tag{2.4}$$

Similarly, we derive the full conditional of $\Sigma_\beta | \beta_j, \Omega_2$ i.e, $P(\Sigma_\beta | \beta_j, \Omega_2)$ by using two prior density functions $P(\beta_j | \mu_\beta, \Sigma_\beta)$ and $P(\Sigma_\beta | \Omega_2)$ based on Bayesian theorem:

$$\begin{aligned}
P(\Sigma_\beta|\beta_j, \Omega_2) &\propto P(\beta_j|\mu_\beta, \Sigma_\beta).P(\Sigma_\beta|\Omega_2) \\
&\propto (2\pi^{-5/2} \det(\Sigma_\beta)^{-1/2})^J \times \exp\left(\sum_{j=1}^J -1/2(\beta_j - \mu_\beta)^T \Sigma_\beta^{-1}(\beta_j - \mu_\beta)\right) \times \\
&(|\Sigma_\beta|^{(5-5-1)/2})^5 \times \exp\left(\sum_{\beta=1}^5 \text{tr}(\Omega_2^{-1}, \Sigma_\beta)/2\right).
\end{aligned} \tag{2.5}$$

The parameter estimation and inference for the hierarchical model was carried out within a Bayesian framework using Markov chain Monte Carlo (MCMC) methods. Posterior mean, standard deviation, odds ratio, and 95 percent odds ratio credible intervals were provided. We considered 20,000 iterations in each MCMC with 5000 burn-ins and 1000 adaptations. Statistical analyses were conducted using rjags package of R statistical software.

2.3 WGEE

One utility of our proposed method is to improve the performance of WGEE analyzing multi-trial data. WGEE is an estimation technique that is able to handle dropout missingness in longitudinal studies [Robins et al., 1995a]. The estimated inverse probabilities of observing outcomes (i.e., the subject has not dropped out) are used as weights to detect treatment effects. A certain weight is given to each study subject at each time point, and such an approach is called the “observation-specific weighted method.” There is also another approach in which a single weight is given to each study subject, which is called the “subject-specific weighted method” [Lin and Rodriguez, 2015a]. In this method, all observations of a study subject get

the same weight. Observation-specific weighted method is found to be more effective [Preisser et al., 2002]. Conventional strategy to construct weights is illustrated below: the conditional probability of being observed, i.e., $\hat{\pi}_{jis}$, will be firstly estimated using partial likelihood regression [Chen et al., 2019]. Similarly, we have $\hat{\pi}_{ji1} = 1$ and $\hat{\pi}_{jis} = 0$ if $M_{ji(t-1)} = 0$ for all $t = 3, \dots, T$. Then, we can calculate $\hat{w}_{jit} = M_{jit} / \{\prod_{s=1}^t \hat{\pi}_{jis}\}^T$, which is the inverse of the cumulative product of probabilities of the first time point to the time point t [Preisser et al., 2002]. We stack the weights and form the vector denoted by $\hat{w}_{ji} = (\hat{w}_{ji1}, \dots, \hat{w}_{jit})$. After obtaining \hat{w}_{jit} , one can use it to adjust for missing data issue when fitting GEE [LIANG and ZEGER, 1986]. Let $Y_{ji} = (Y_{ji1}, \dots, Y_{jiT})^T$ and $X_{ji} = (X_{ji1}^T, \dots, X_{jiT}^T)^T$ indicate the outcomes and covariates (treatment and other explanatory variables) obtained from trial j , $j = 1, \dots, J$ and individual i , $i = 1, \dots, n$, respectively, where Y_{jit} is the t th outcome and X_{jit} is a $p \times 1$ vector of covariates including the intercept. Suppose we are interested in estimating the parameter vector β involved in the conditional mean, i.e., $\mu_{ji} = E(Y_{ji}|X_{ji}) = g(X_{ji}^T\beta)$, where g is the inverse of certain link function. Then, WGEE obtains the estimate of β by solving

$$g(\beta) = \sum_{j=1}^J \sum_{i=1}^n g(X_{ji}, Y_{ji}, \beta; \hat{w}_{ji}) = \sum_{j=1}^J \sum_{i=1}^n D_{ji}^T V_{ji}^{-1} W_{ji} (Y_{ji} - \mu_{ji}) = 0, \quad (2.6)$$

where $V_{ji} = \text{Var}(Y_{ji}|X_{ji})$, $D_{ji} = \partial\mu_{ji}/\partial\beta^T$, and W_{ji} is the weighted matrix with diagonal elements equal to \hat{w}_{jit} , $t = 1, \dots, T$ and off-diagonal elements equal to zero.

Distinct from conventional strategies [Robins et al., 1995b, Chen et al., 2019, Shen

et al., 2023], the estimated probabilities π_{it} can be extracted from our developed BHM, which could be numerically more stable and closer to the underlying truth. Part of the reason explaining this is that traditional WGEE only uses the probability fitted by a single study (clinical trial), which may not be efficient and stable given a small sample size. However, our method uses the probability leveraging across studies (clinical trials), thus integrating information to achieve better efficiency. We assessed the performance of WGEE based on the weights calculated by our BHM method, which is illustrated in the Simulations section.

Chapter 3: Simulations

3.1 Data generation

To evaluate the proposed method and mimic real-world trial data, we considered the following data generation mechanism: first, we considered ten sets of data corresponding to ten trials (i.e., $j = 1, \dots, 10$), each of which contains data for 50 patients (i.e., $i = 1, \dots, 50$) and 6 time visits in total for each patient (i.e., $t = 1, \dots, 6$). Then, we generated patient-level factors in a vector for the dropout model $x_{jit} = (1, x_{jit1}, x_{ji(t-1)2})^T$, which consists of an intercept, gender denoted by x_{jit1} , and BPRS denoted by $x_{ji(t-1)2}$ from the time $t-1$, for $t = 2, \dots, T$. For notation convenience, we let x_{jit2} be zero when $t = 1$. Specifically, x_{jit1} followed a Bernoulli distribution with success probability 0.5, and x_{jit2} followed a normal distribution with

variance 0.4. We note here that BPRS is regarded as the primary outcome in each trial and might be affected by treatment assignment. Thus, to account for potential differences due to treatment effect, the means of the primary outcome were generated from a normal distribution with mean 25 and standard deviation 0.1 for the treatment group and with mean 30 and standard deviation 0.1 for the control group. Here, we allowed different treatment effect profiles across ten trials since different trials adopted different treatment strategies.

Afterward, we generated coefficients in the dropout model. Similar to the heterogeneous treatment effects, we allow different profiles of coefficients in different trials. To be specific, We sampled $\beta_j = (\beta_{j1}, \beta_{j2}, \beta_{j3})$ for intercept, gender and BPRS, where $\beta_{j1} \sim N(-0.19, 0.14)$, $\beta_{j2} \sim N(0.78, 0.13)$, $\beta_{j3} \sim N(0.50, 0.10)$. The values of hyperparameters were calculated from from the multi-trial data at MPRC, mimicking the real data application.

With generated covarites and outcomes, we dynamically generated the observation indicator in the dropout model, which follow a Bernoulli distribution given previous dropout status. The conditional probability of observing subject i at time t in trial j is $P(M_{jit} = 1 | M_{ji(t-1)} = 1, x_{jit}) = \{1 + \exp(-x_{jit}^T \beta_j)\}^{-1}$, where x_{jit} represents patient-level vector defined before. To generate the mechanism of dropout, we required that $M_{jit} = 0$ if $M_{ji(t-1)} = 0$ and $M_{ji1} = 1$ for all i and j .

3.2 Benchmarks

We compared our proposed method with existing methods from two perspec-

tives. Firstly, we evaluated the estimation and inference in the dropout model. Secondly, we assessed the performance of the main outcome model by applying estimated observing probabilities based on different methods. For the dropout model, we calculated bias, Monte Carlo standard deviation, and power for estimating and testing the coefficients β based on 100 Monte Carlo runs. We compared our proposed Bayesian hierarchical model (BHM) method with the competing method, which is the frequentist way by separately applying partial likelihood regression to each individual trial data. We do not consider GLMM as an alternative since it will only provide overall effect estimation instead of the trial specific estimation, which is of our primary interest. For the main outcome model, we considered the weighted generalized estimating equation (WGEE) method, owing to its ability to handle dropout missingness [Lin and Rodriguez, 2015b, Salazar et al., 2016, Chen et al., 2019]. To implement WGEE, we used the estimated probabilities of observing outcomes as weights to detect treatment effect over ten trials. We assessed the performance of WGEE based on the weights calculated by the proposed BHM method and the weights calculated by the frequentist regression method, respectively. Additionally, we also compared the performance between WGEE and the generalized estimating equation (GEE) without any weights. For each of the ten trials, we considered bias, Monte Carlo standard deviation, and power for estimating and testing treatment effects over 100 Monte Carlo runs.

3.3 Evaluations

We evaluated the robustness and performance of our method compared to the competing method (frequentist way by separately applying regression for each individual trial data) in terms of bias, Monte Carlo standard deviation, and power of inference in the dropout model. In Table 3, we observe that the existing methods led to higher bias for estimating β_j for all j s, whereas our proposed method led to little bias. This is because our method can leverage information across different trials, which improves statistical inference accuracy. However, for individual level analysis, it only uses one trial data each time and has little power to accurately estimate those parameters when the sample size in each trial is small. Our method in the dropout data model also showed much smaller standard deviation (estimation variability) compared to the existing one, which is attributed to the fact that our method has the ability to integrate information and synthesize similarity across different trials, while individual analysis fails to do this. In addition, we also found that the proposed method had a higher power of detecting a significant effect compared to the competing frequentist approach individual trial analysis (FITA) (Table 2), which also demonstrates the superiority of the BHM method.

For the main outcome model, we evaluated the performance of WGEE using weights from our BHM method and weights from frequentist regression method. We also compared the performance of WGEE using weights from our BHM method with that of GEE without any weights. Our results (Table 3) showed that the WGEE model using weights from our BHM method had smaller bias compared to both

competing methods. No significant difference was found between the three methods in terms of power. We also observed that the standard deviation for WGEE using weights from our BHM method was slightly smaller than that for WGEE using weights from frequentist regression. However, both methods had larger standard deviation compared to the GEE approach. These findings demonstrate the importance and necessity of accurately modeling dropout in practice.

Chapter 4: Data Application

We utilized records of 463 Schizophrenia patients who participated in 13 clinical trials conducted at the Maryland Psychiatric Research Center (MPRC) between 2009 and 2016. The average sample size was 36 with range 6 to 61. In this dataset, on average 8.0% of patients dropped out of the trials with range 0% to 22.7% over 13 trials. We used available demographic and clinical characteristics of the participants, including the Brief Psychiatric Rating Scale (BPRS) to measure psychiatric symptoms.

The aim of this data analysis was to identify common factors, either independently or in combination, that are associated with participant dropout in clinical trials, alongside trial-specific factors. Detecting these factors is essential for implementing remedial measures and designing more tailored early interventions specific to individual trials, thereby enhancing participant retention in future studies. We

first summarized patient characteristics in terms dropout status in Table 4. A higher proportion of patients who were dropout were female (57.0%, $p=0.003$). Patients who were dropout had a lower age compared to patients completed the trial (36.6 ± 10.9 versus 42.7 ± 12.8 ; $p=0.005$). No statistically significant difference in the proportion of dropout by race was found ($p=0.808$) based on Chi-squared test. We then applied our BHM to the data obtained from MPRC to explore demographic and clinical determinants of dropout in Schizophrenia trials. In order to estimate posterior distribution, we set out multiple levels of priors. The first one is $\beta_j \sim \mathcal{MVN}(\mu_\beta, \Sigma_\beta)$, that estimates the variability of effect sizes across different trials. We assume that the vector of covariate effect sizes on patient dropout β_j follows a multivariate normal distribution with mean vector μ_β and variance matrix Σ_β . In order to make a strong inference, we assign another level of prior to these variables, i.e., $\mu_\beta \sim \mathcal{MVN}(\mu, \Omega_1)$ and $\Sigma_\beta \sim W_5(\Omega_2, 5)$. There are 5 covariates for the overall effect (intercept, gender, age, time and BPRS) and Σ_β follows a Wishart distribution with 5 degrees of freedom. The hyperparameters $\mu = [0, 0.88, 0, 0, 0]^T$, and $\Omega_1 = \Omega_2 = (1/1000)I_5$, I_5 is identity matrix of size 5. For the hyperparameters, we set the value of zero for all variables of interest except for gender, for which we used the value obtained from our data since it remains unchanged.

In Table 5, we observe that younger age and female gender are common predictors of dropout for all trials, while a higher Brief Psychiatric Rating Scale (BPRS) score is a significant factor only in three trials. We also compared our method with frequentist-based partial likelihood method by separately analyzing each individual trial data. The frequentist-based method failed to detect all significant factors (Table

6). This might be attributed to the issue of small sample size in each trial, thus resulting in low statistical power to make inference. In addition, we do observe not stable outputs from the frequentist-based method. This comparison also signifies and validates the motivation of proposing our BHM approach.

Chapter 5: Discussions

Detecting a missing pattern and identifying factors associated with it are important because we can better develop a better retention plan at the stage of study design. In addition, it enables us to implement adaptive design of clinical trials based on missing rate [Dziura et al., 2013, O’Kelly and Ratitch, 2015]. In this paper, we examined the strength and performance of our method (BHM) compared to the competing method (individual trial data analysis) in terms of bias, standard deviation and power of estimation and inference in the missing data model. Our proposed method in the context of missing data leads to much smaller bias and standard deviation (estimation variability) compared to the competing methods. These demonstrate that our proposed method is more precise and robust compared to existing methods. This is in part due to ability of our method to leverage and integrate information across different trials that the individual trial analysis unable to do. It was also found that the proposed method has higher power compared to the competing method. Another advantage of using our proposed method is to predict and make inference for future’s

missingness of specific trials, an attractive feature of the Bayesian approach. The disadvantage of the competing individual trial analysis, on the other hand, is that it will result in underestimation or overestimation of parameters of our interest and fail to detect covariate effects due to small sample size issue in each individual trial.

Moreover, for the main outcome model, we have evaluated the performance of WGEE based on the weights calculated from our BHM method and weights calculated by frequentist regression method. We also have assessed the performance of WGEE based on the weights obtained from our BHM method and GEE without any weights. WGEE based on the weights calculated from our BHM method performs well in terms of bias compared to both GEE without weights and WGEE based on the weights calculated by frequentist regression method. No statistically significant difference was found between the three methods in terms of power.

In our data application, we observed that both younger age and gender are consistently influential predictors of dropout across all trials, signifying their widespread and significant impact. On the other hand, a higher BPRS score emerged as a significant predictor in only three individual trials. These findings emphasize the importance of giving due consideration to the specific subgroup of patients. For these particular three trials where a high BPRS score plays a significant role, it is essential for researchers to invest additional effort in the follow-up process, particularly for patients with elevated BPRS scores, to ensure the desired level of patient retention.

To further boost the performance of the proposed BHM, more robust prior could be considered [Jiang et al., 2021]. Not limited to the variables considered in this paper, more patient-level characteristics and study-level factors should be

evaluated in the dropout model. Since the current trials do not record extra variables, we regraded it as limitation in our application. Non-compliance in clinical trials makes the causal inference difficult to draw. [Tilbrook et al., 2014] found a significant beneficial treatment effect in the study subjects who were fully compliant. However, when all participants data were analyzed, a smaller treatment effect was estimated. The differing inferences obtained from the two analyses caused understanding of the study results about treatment effect hard. Thus, the analysis by considering both non-compliant and patient dropout deserve substantial effort and merit future work.

Trials	Sample size	Dropout rate(%)
Aripiprazole	43	14.0
Acamprostate	23	0.0
Olanzapine-clozapine	22	22.7
Fluvastatin	34	14.7
Galantamine-oxytocin	67	6.0
Gluten	19	15.8
L-THP	61	4.9
Minocycline	50	4.0
Nicotine	24	0.0
Dipyridamol	6	16.7
Oxytocin	41	7.3
Oxytocin-unblinded	20	0.0
Rasagiline	53	9.4

Table 1 Distribution of dropout rates across the trials

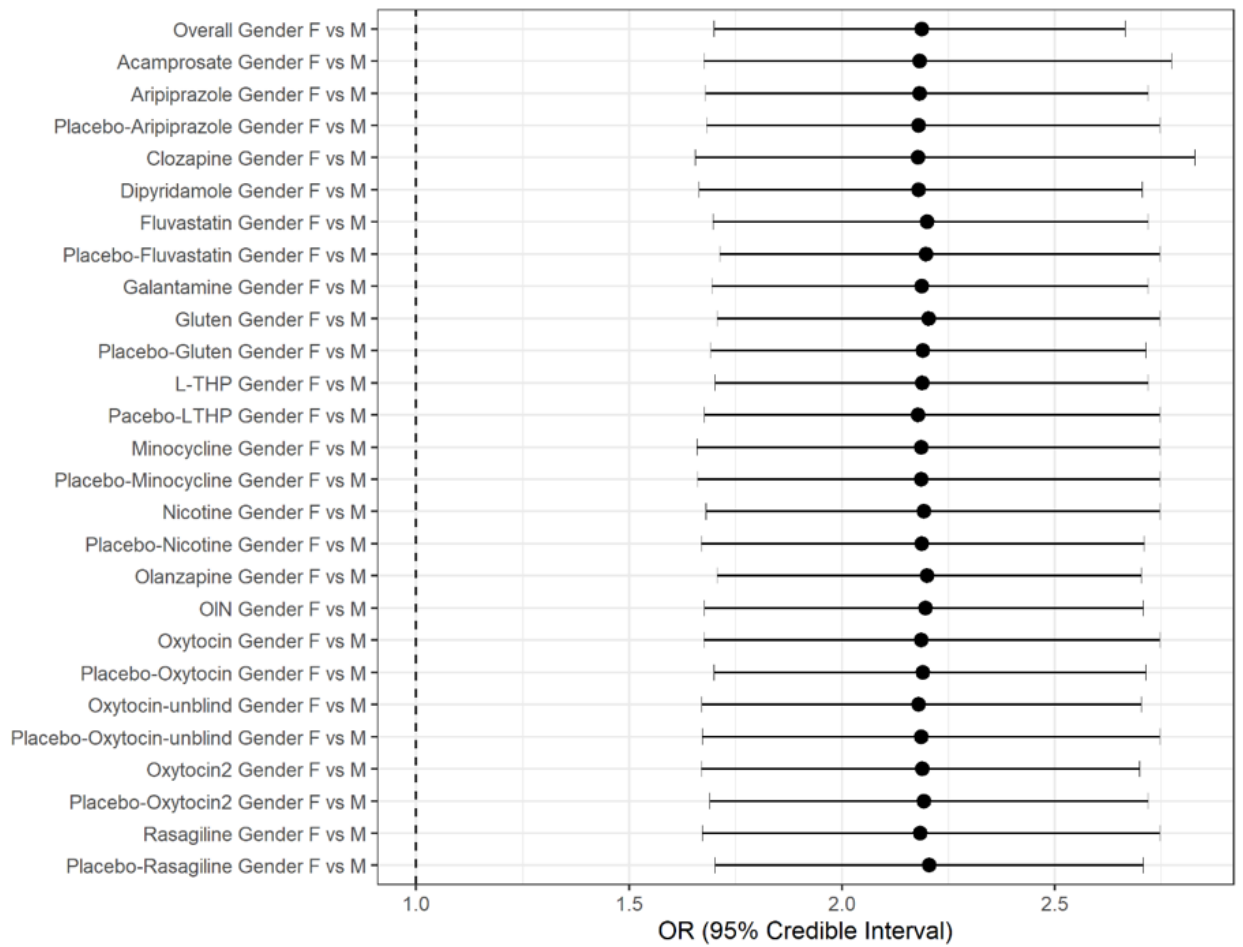


Figure 1 Posterior odds ratio with 95% credible intervals by gender. The outcome was patient dropout status and for the analysis, 20,000 MCMC draws from the posterior distribution were used (20,000 iterations, burn-in=5,000, and adapt=1000).

(a)

		Our method (BHM)			Competing method (FITA)		
		Bias	Std	Power	Bias	Std	Power
Overall	BPRS	0.05	0.13	0.86	0.07	0.52	0.44
	Gender	0.01	0.21	0.95	0.05	0.66	0.39
Trial1	BPRS	0.02	0.20	0.62	0.05	0.50	0.35
	Gender	0.01	0.23	0.75	0.03	0.65	0.69
Trial2	BPRS	0.05	0.14	0.63	0.06	0.42	0.39
	Gender	0.03	0.24	0.78	0.04	0.62	0.33
Trial3	BPRS	0.05	0.33	0.51	0.05	0.49	0.62
	Gender	0.04	0.27	0.70	0.05	0.54	0.65
Trial4	BPRS	0.06	0.18	0.60	0.07	0.62	0.47
	Gender	0.03	0.26	0.79	0.05	0.75	0.25
Trial5	BPRS	0.04	0.19	0.53	0.05	0.58	0.46
	Gender	0.01	0.24	0.71	0.03	0.69	0.25
Trial6	BPRS	0.04	0.23	0.58	0.05	0.56	0.55
	Gender	0.10	0.18	0.68	0.07	0.70	0.50
Trial7	BPRS	0.05	0.23	0.60	0.06	0.46	0.35
	Gender	0.04	0.26	0.75	0.05	0.59	0.36
Trial8	BPRS	0.03	0.21	0.62	0.04	0.45	0.37
	Gender	0.06	0.20	0.66	0.08	0.70	0.15
Trial9	BPRS	0.03	0.23	0.64	0.05	0.63	0.38
	Gender	0.02	0.21	0.75	0.04	0.66	0.33
Trial10	BPRS	0.05	0.22	0.62	0.07	0.50	0.42
	Gender	0.04	0.21	0.81	0.05	0.65	0.40

(b)

		Our method (BHM)			Competing method (FITA)		
		Bias	Std	Power	Bias	Std	Power
Overall	BPRS	0.06	0.13	0.88	0.07	0.64	0.43
	Gender	0.01	0.21	0.96	0.06	0.72	0.38
Trial1	BPRS	0.04	0.22	0.63	0.05	0.64	0.33
	Gender	0.03	0.24	0.75	0.04	0.66	0.67
Trial2	BPRS	0.09	0.19	0.59	0.06	0.61	0.38
	Gender	0.04	0.25	0.80	0.04	0.68	0.31
Trial3	BPRS	0.06	0.32	0.50	0.07	0.80	0.60
	Gender	0.05	0.22	0.70	0.05	0.77	0.65
Trial4	BPRS	0.07	0.19	0.62	0.09	0.58	0.47
	Gender	0.03	0.26	0.86	0.06	0.72	0.24
Trial5	BPRS	0.06	0.22	0.55	0.08	0.62	0.45
	Gender	0.02	0.25	0.69	0.04	0.81	0.24
Trial6	BPRS	0.05	0.23	0.54	0.05	0.67	0.56
	Gender	0.10	0.21	0.69	0.07	0.80	0.51
Trial7	BPRS	0.07	0.23	0.61	0.06	0.68	0.35
	Gender	0.04	0.25	0.74	0.08	0.61	0.35
Trial8	BPRS	0.04	0.21	0.60	0.05	0.46	0.36
	Gender	0.10	0.22	0.87	0.11	0.67	0.14
Trial9	BPRS	0.05	0.23	0.63	0.06	0.78	0.34
	Gender	0.04	0.22	0.77	0.05	0.73	0.32
Trial10	BPRS	0.12	0.37	0.61	0.10	0.52	0.41
	Gender	0.04	0.27	0.79	0.05	0.74	0.39

Table 2 Comparison of our method and the competing method for the missing data model. Bias, Monte Carlo standard deviation, and power were calculated for estimating and testing the coefficients β based on 100 Monte Carlo runs for both the proposed method and the competing method. BHM=Bayesian hierarchical model; FITA= frequentist individual trial analysis; (a) and (b) are results with sample size of 50 and 26 respectively.

(a)

	Our method (WGEE)			Competing method (GEE without weight)			WGEE (weight based on frequentist way)		
	Bias	Std	Power	Bias	Std	Power	Bias	Std	Power
Trial1	0.03	0.60	1.00	0.05	0.24	1.00	0.08	0.67	1.00
Trial2	0.02	0.70	1.00	0.03	0.28	1.00	0.07	0.82	1.00
Trial3	0.02	0.75	1.00	0.04	0.23	1.00	0.05	0.75	1.00
Trial4	0.01	0.78	1.00	0.06	0.24	1.00	0.06	0.69	1.00
Trial5	0.05	0.57	1.00	0.05	0.22	1.00	0.057	0.40	1.00
Trial6	0.05	0.54	1.00	0.06	0.21	1.00	0.07	0.36	1.00
Trial7	0.03	0.61	1.00	0.05	0.27	1.00	0.04	0.61	0.98
Trial8	0.03	0.62	1.00	0.046	0.25	1.00	0.06	1.01	1.00
Trial9	0.01	0.53	1.00	0.06	0.25	1.00	0.05	0.50	1.00
Trial10	0.03	0.64	1.00	0.05	0.22	1.00	0.12	0.71	1.00

(b)

	Our method (WGEE)			Competing method (GEE without weight)			WGEE (weight based on frequentist way)		
	Bias	Std	Power	Bias	Std	Power	Bias	Std	Power
Trial1	0.03	0.59	0.99	0.05	0.35	1.00	0.07	0.57	1.00
Trial2	0.04	0.72	0.99	0.04	0.30	1.00	0.06	0.61	0.99
Trial3	0.03	0.56	1.00	0.04	0.35	1.00	0.04	0.69	0.99
Trial4	0.02	0.57	1.00	0.03	0.38	1.00	0.05	0.76	1.00
Trial5	0.04	0.55	0.99	0.05	0.31	1.00	0.06	0.55	1.00
Trial6	0.04	0.56	1.00	0.05	0.35	1.00	0.07	0.46	1.00
Trial7	0.02	0.68	0.99	0.03	0.33	1.00	0.04	0.56	1.00
Trial8	0.05	0.47	1.00	0.05	0.34	1.00	0.06	0.74	1.00
Trial9	0.01	0.54	0.99	0.03	0.35	1.00	0.03	0.63	0.96
Trial10	0.03	0.39	0.99	0.04	0.39	1.00	0.05	0.58	0.98

Table 3 Comparison of our method (WGEE) and the competing methods for the main outcome model. The performance of WGEE based on the weights calculated by our BHM method, the weights calculated by frequentist regression method and without weight were compared. Bias, Monte Carlo standard deviation, and power were calculated for estimating and testing treatment effects for ten trials over 100 Monte Carlo runs. (a) and (b) are results with sample size of 50 and 26 respectively.

Characteristics	Dropout N=37	No dropout N=426	P value
Age	36.6±10.9	42.7±12.8	0.005
Gender			0.003
Female	21(57%)	139(33%)	
Male	16(43%)	287(67%)	
Race			0.808
African American/Black	18(49%)	186(44%)	
White	17(46%)	209(49%)	
Other	2(5%)	31(7%)	

Table 4 Demographic characteristics of patients

Effect	Estimate	Std. Error	OR	2.5 %	97.5 %
Intercept	-0.192	0.140	0.825	0.657	1.084
Gender					
Female	0.783	0.130	2.188	1.699	2.665
Age (years)	-0.095	0.012	0.909	0.887	0.931
Time	-0.054	0.061	0.947	0.835	1.049
BPRS total change	0.064	0.022	1.066	1.019	1.116

Table 5 Posterior mean, standard deviation, odds ratio, and 95% odds ratio credible set for the regression coefficients of the model. The binary response variable is patient dropout status, and the predictors were gender, age and BPRS. For the analysis, 20,000 MCMC draws from the posterior distribution were used (20,000 iterations, burn-in=5,000, and adapt=1000).

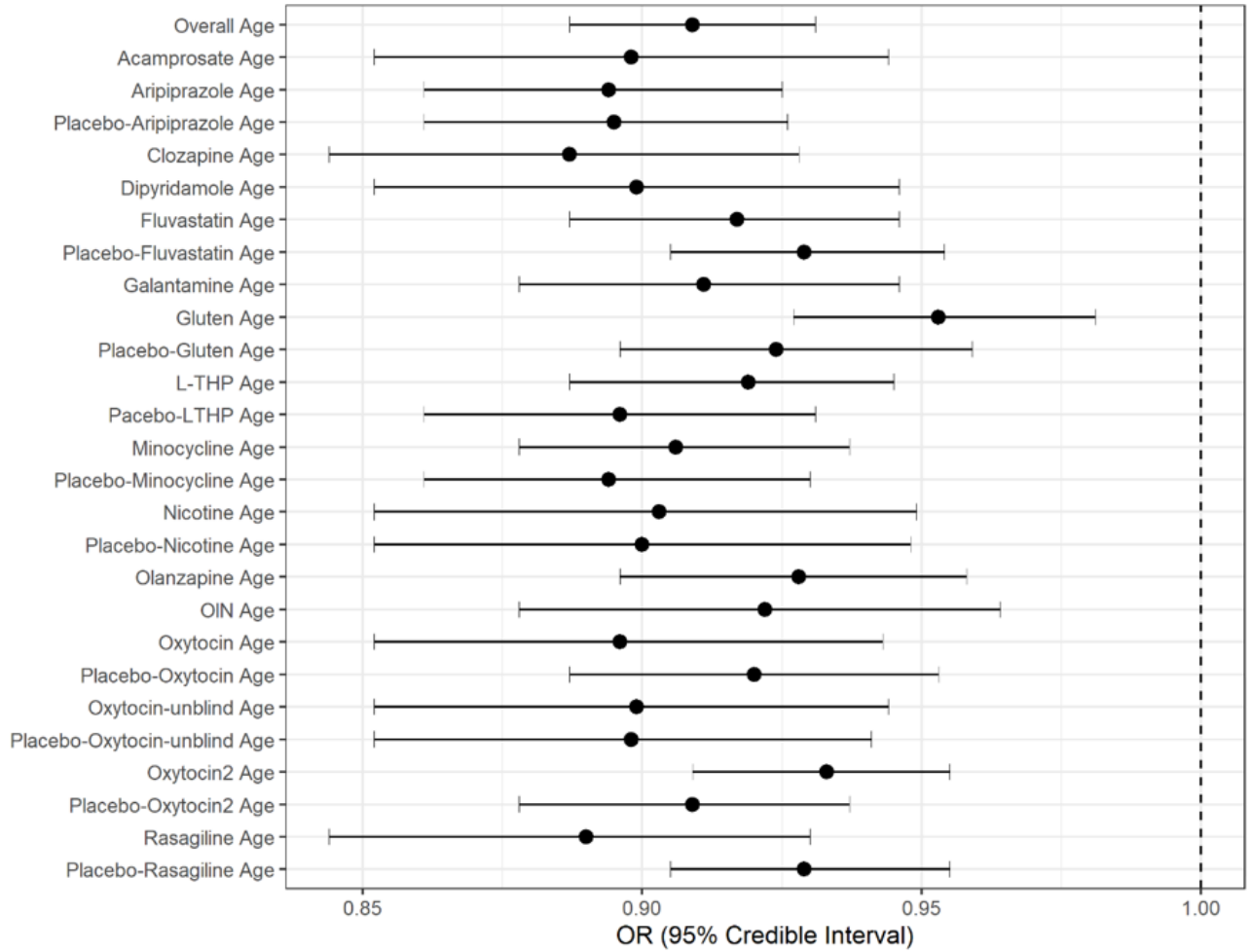


Figure 2 Posterior odds ratio with 95% credible intervals for a 1-year increase in age. The outcome was patient dropout status and for the analysis, 20,000 MCMC draws from the posterior distribution were used (20,000 iterations, burn-in=5,000, and adapt=1000).

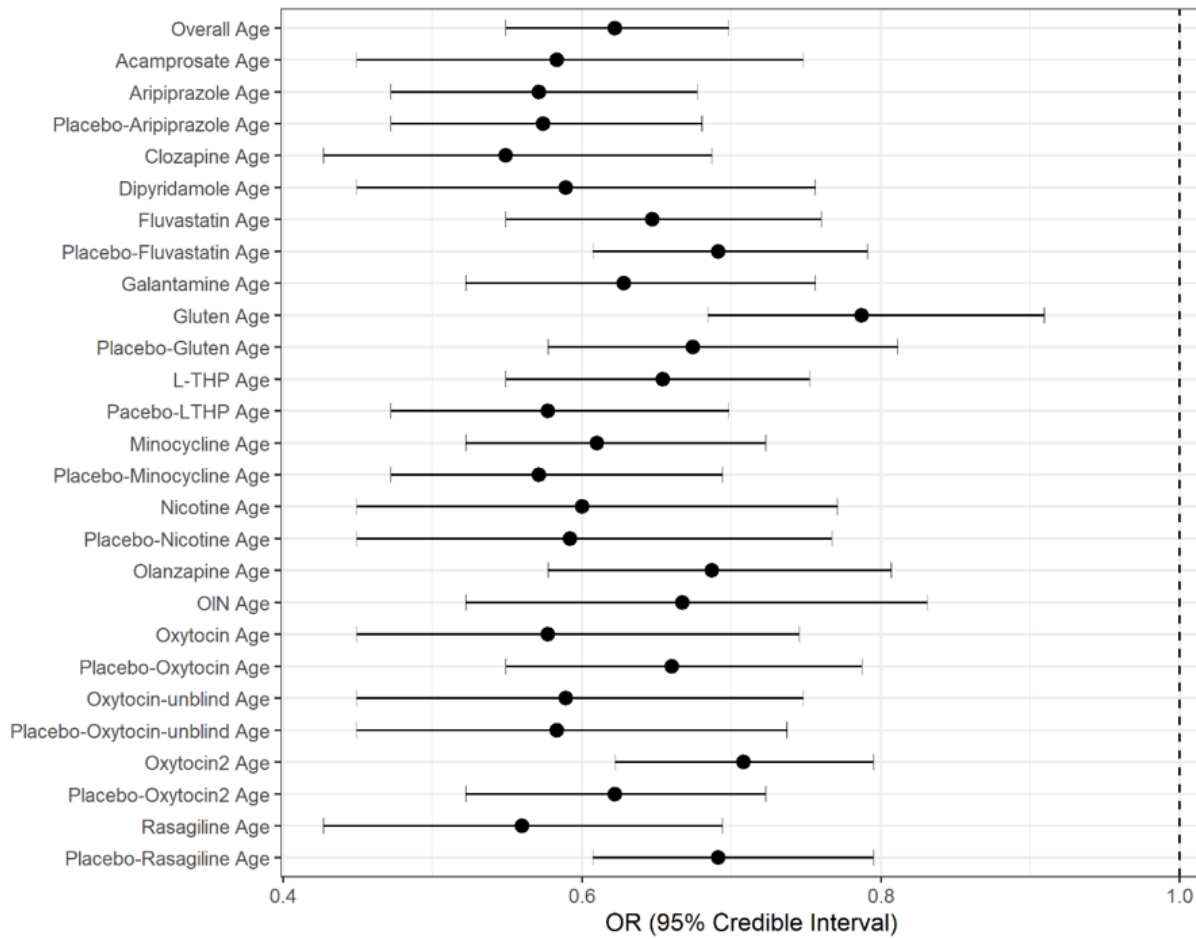


Figure 3 Posterior odds ratio with 95% credible intervals for a 5-years increase in age. The outcome was patient dropout status and for the analysis, 20,000 MCMC draws from the posterior distribution were used (20,000 iterations, burn-in=5,000, and adapt=1000).

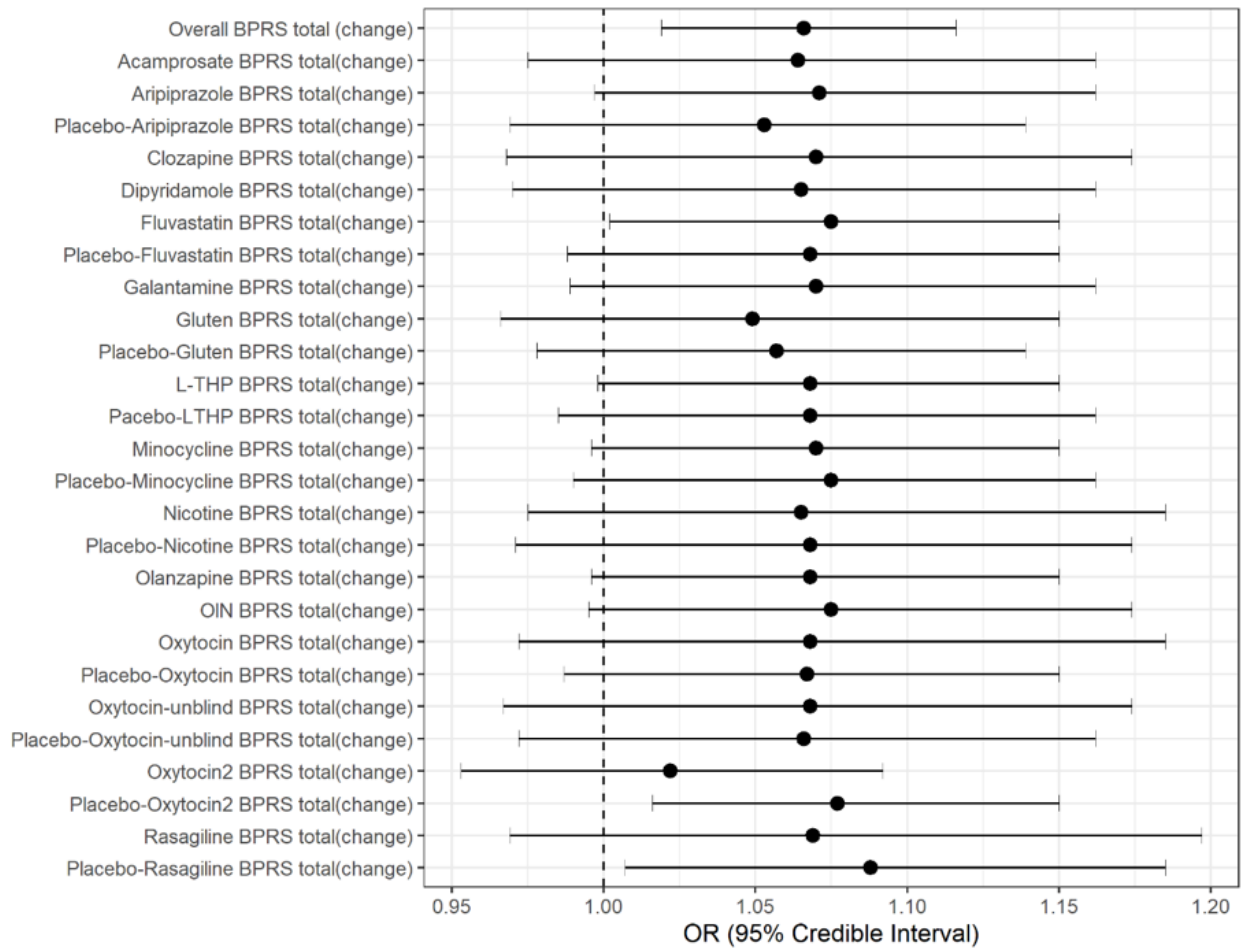


Figure 4 Posterior odds ratio with 95% credible intervals for a 1-unit increase in BPRS total(change). The outcome was patient dropout status and for the analysis, 20,000 MCMC draws from the posterior distribution were used (20,000 iterations, burn-in=5,000, and adapt=1000).

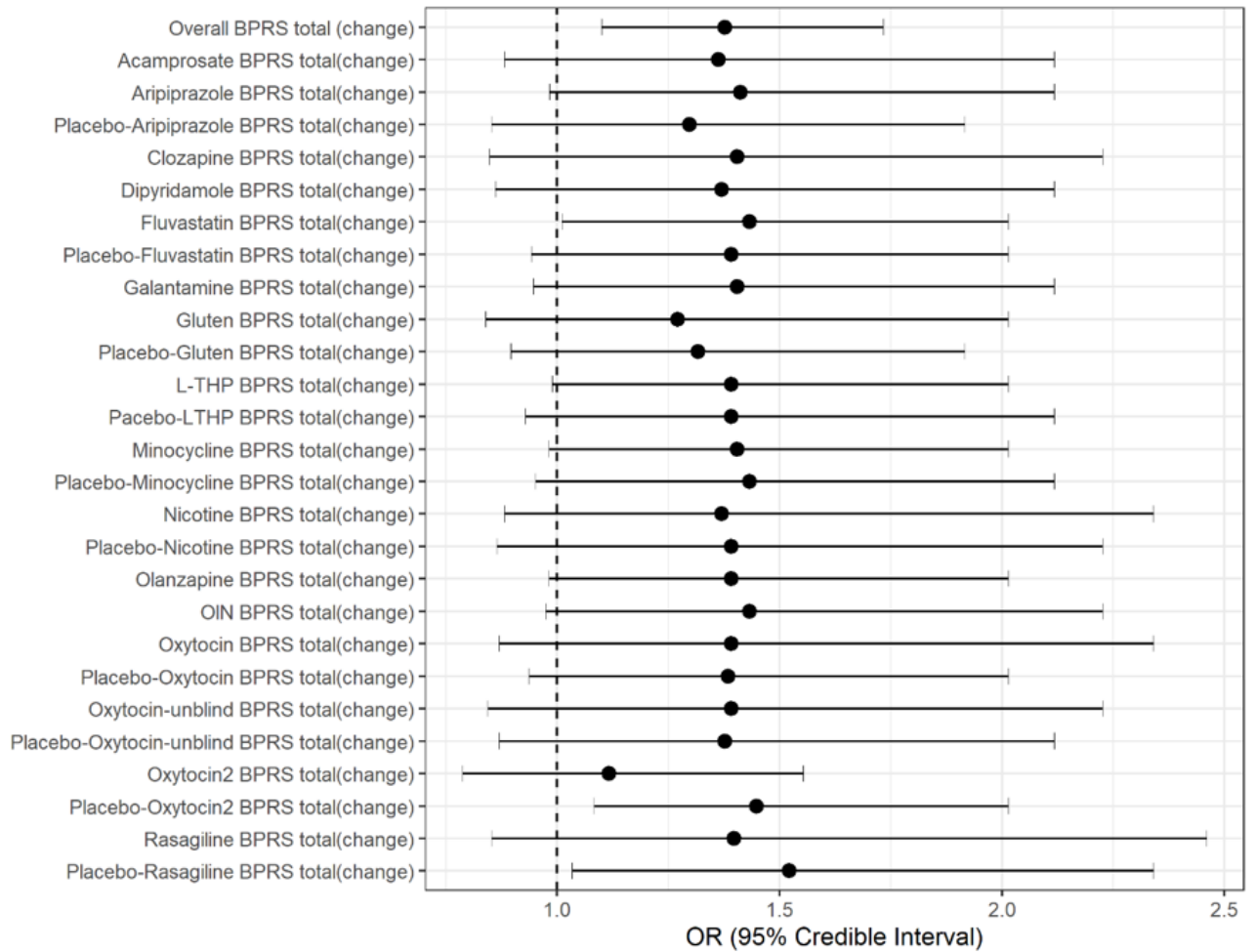


Figure 5 Posterior odds ratio with 95% credible intervals for a 5-unit increase in BPRS total(change). The outcome was patient dropout status and for the analysis, 20,000 MCMC draws from the posterior distribution were used (20,000 iterations, burn-in=5,000, and adapt=1000).

Effect	OR	2.5 %	97.5 %
Overall Gender F vs M	2.174	1.286	3.677
Aripiprazole Gender F vs M	<.001	<.001	>100
Placebo-Aripiprazole Gender F vs M	<.001	<.001	>100
Fluvastatin Gender F vs M	1.241	0.078	19.791
Placebo-Fluvastatin Gender F vs M	1.791	0.189	16.997
Galantamine Gender F vs M	<.001	<.001	>100
Gluten Gender F vs M	>.001	<.001	>100
Placebo-Gluten Gender F vs M	289.7	0.161	>100
L-THP Gender F vs M	8.859	0.448	175.3
Placebo-LTHP Gender F vs M	<.001	<.001	>100
Minocycline Gender F vs M	3.347	0.146	76.9
Placebo-Minocycline Gender F vs M	<.001	<.001	>100
Olanzapine Gender F vs M	10.21	0.701	148.6
Placebo-Oxytocin Gender F vs M	<.001	<.001	>.100
Oxytocin2 Gender F vs M	2.962	0.678	12.94
Placebo-Oxytocin2 Gender F vs M	4.709	0.505	43.92
Placebo-Rasagiline Gender F vs M	0.066	<.001	7.675
Overall Age	0.974	0.952	0.997
Aripiprazole Age	0.960	0.857	1.076
Placebo-Aripiprazole Age	0.981	0.873	1.103
Fluvastatin Age	0.965	0.835	1.114
Placebo-Fluvastatin Age	0.989	0.881	1.110
Galantamine Age	0.958	0.755	1.215
Gluten Age	1.061	0.917	1.227
Placebo-Gluten Age	1.094	0.869	1.378
L-THP Age	1.028	0.918	1.151
Placebo-LTHP Age	0.839	0.559	1.257
Minocycline Age	1.056	0.934	1.194
Placebo-Minocycline Age	0.153	0.002	14.152
Olanzapine Age	1.053	0.908	1.221
Placebo-Oxytocin Age	0.974	0.890	1.065
Oxytocin2 Age	0.949	0.871	1.034
Placebo-Oxytocin2 Age	0.873	0.733	1.039
Placebo-Rasagiline Age	0.704	0.497	0.997
Overall BPRS total(change)	1.048	1.013	1.085
Aripiprazole BPRS total(change)	1.067	0.896	1.270
Placebo-Aripiprazole BPRS total(change)	1.038	0.822	1.310
Fluvastatin BPRS total(change)	1.151	0.923	1.436
Placebo-Fluvastatin BPRS total(change)	1.152	0.908	1.461
Galantamine BPRS total(change)	1.800	0.480	6.754
Gluten BPRS total(change)	1.171	0.791	1.734
Placebo-Gluten BPRS total(change)	1.588	0.751	3.358
L-THP BPRS total(change)	1.040	0.910	1.188
Placebo-LTHP BPRS total(change)	1.424	0.628	3.230
Minocycline BPRS total(change)	1.000	0.845	1.185
Placebo-Minocycline BPRS total(change)	14.80	0.021	>100
Olanzapine BPRS total(change)	1.082	0.911	1.285
Placebo-Oxytocin BPRS total(change)	1.027	0.753	1.401
Oxytocin2 BPRS total(change)	0.955	0.878	1.038
Placebo-Oxytocin2 BPRS total(change)	1.080	0.989	1.180
Placebo-Rasagiline BPRS total(change)	1.896	1.101	3.265

Table 6 Odds ratios and 95% Confidence intervals of odds ratios obtained from logistic regression model of the Frequentist-based individual trial analyses. The binary response variable is patient dropout status, and the predictors were gender, age and BPRS, inactive drugs were used for placebo.

Part II

Efficacy of Randomized Clinical Trials in Adolescent Patients with Schizophrenia: A Network Meta Analysis

Chapter 6: Introduction

Schizophrenia is a mental health issue which affects overall life situations including family, educational, personal, social, and working functioning [WHO, 2022]. In the general population, the lifetime prevalence of schizophrenia is 1% [Kendhari et al., 2016]. Most clinical trials enroll adults and there is scarcity of treatment efficacy data in children and adolescents [Joseph et al., 2015]. Though there is no agreement about the prevalence of schizophrenia among underage population, a meta-analysis study showed that the estimated prevalence of psychotic symptoms among children and adolescents reach 5% [Kelleher et al., 2012]. Accordingly, much more attention should be given to younger population and more studies need to be performed to examine efficacy of schizophrenia or psychotic symptoms treatments in underage population. Meta-analysis uses aggregate data (AD) obtained from published articles, conference abstracts, or trial registries like “clinicaltrials.gov” to determine overall effect of treatments by integrating results from individual studies. However, the traditional meta-analysis is restricted to direct comparison of treatments in parallel study designs [Rouse et al., 2017]. Network meta-analysis (NMA) is a special kind of meta-analysis for the comparison of multiple treatments simultaneously in a single analysis by combining direct and indirect evidence. The direct evidence is obtained from randomized

control trials (RCTs) with direct comparison of treatments in parallel study design. The indirect evidence is obtained from comparisons of one or more common comparators by transitivity[Rouse et al., 2017]. The direct and indirect pooled evidence is called network estimate. Network meta-analysis is vital in examining the effectiveness of different treatments or interventions in randomized control trials (RCTs). Accordingly, NMA is gaining popularity. Meanwhile, inferences might be biased unless the NMA is conducted properly and interpreted accurately. Therefore, Network meta-analysis need to be designed thoroughly and conducted appropriately to generate valid results. We conducted a network meta-analysis on randomized control clinical trials for schizophrenia treatments using PANSS (positive and negative syndrome score) as an outcome to examine efficacy of schizophrenia treatments in adolescents (underage patients). The PANSS is a 30-items rating scale to measure the severity of symptoms of schizophrenia patients. The PANSS is composed of the following three sub scales. Positive Scale which used to measure positive symptoms (excess of normal functions), Negative Scale which used to measure negative symptoms (loss of normal functions) and General Psychopathology Scale. Each item is rated with scores of 1 to 7 points, 1 for absence of symptom and 7 for extreme symptom. The total PANSS score is the sum of the scores of all the three sub scales with range 30 to 210. We also examined treatment effects using another outcome clinical global impression (CGI-S). The CGI-S is a 7- point rating scale to measure the severity of symptoms of schizophrenia patients. It also used to measure treatment response and efficacy of treatments in patients with schizophrenia. The CGI-S scores range from 1 to 7 points, 1 for normal and 7 for extreme symptom. Literature searches were conducted through electronic

databases MEDLINE, PsycINFO, PsycARTICLES and PubMed up to September 2023. The following terms: schizophrenia treatment or schizophrenia therapy or Psychotherapy or psychological intervention, PANSS or positive and negative syndrome score, Clinical trials or randomized controlled trials or controlled clinical trials, and Efficacy or effectiveness/ efficacy or effectiveness or impact or benefits or outcomes or success were used for the search. The inclusion criteria were: Randomized Controlled Trials (RCT) to examine efficacy of schizophrenia treatments, with PANSS and CGI outcomes, included patients with schizophrenia, age 0-18 years, male or female, any race/ethnicity, and published in English.

Chapter 7: Method

7.1 Selection Criteria

Three electronic databases (Medline, PubMed, and Psycinfo) were searched. The search strategy eligibility criteria were RCTs for Schizophrenia treatment or schizophrenia therapy, published in scientific peer reviewed papers, written in English, have outcomes PANSS and/or CGI-S, published up to September 2023, included cohort of patients with age 0-18 years, male or female, and any race/ethnicity. Conference abstracts and reports were excluded.

7.2 Data Extraction

Symptom data for outcomes Positive and Negative Syndrome Scale (PANSS) and CGI-S were extracted. The data extracted were treatments effects, within-group mean change score with standard deviation. Treatments effects and their corresponding standard deviations of each study were obtained. 22 studies have PANSS outcome, and 15 studies have CGI-S outcome, and 10 studies have both PANSS and CGI-S outcomes and these 10 studies were used for the primary combined analysis. For the final analyses of our NMA, we identified 27 relevant RCTs with available symptom outcomes data (Figure 7.2.1). The 27 RCTs included 4195 randomized participants with age 18 years or less.

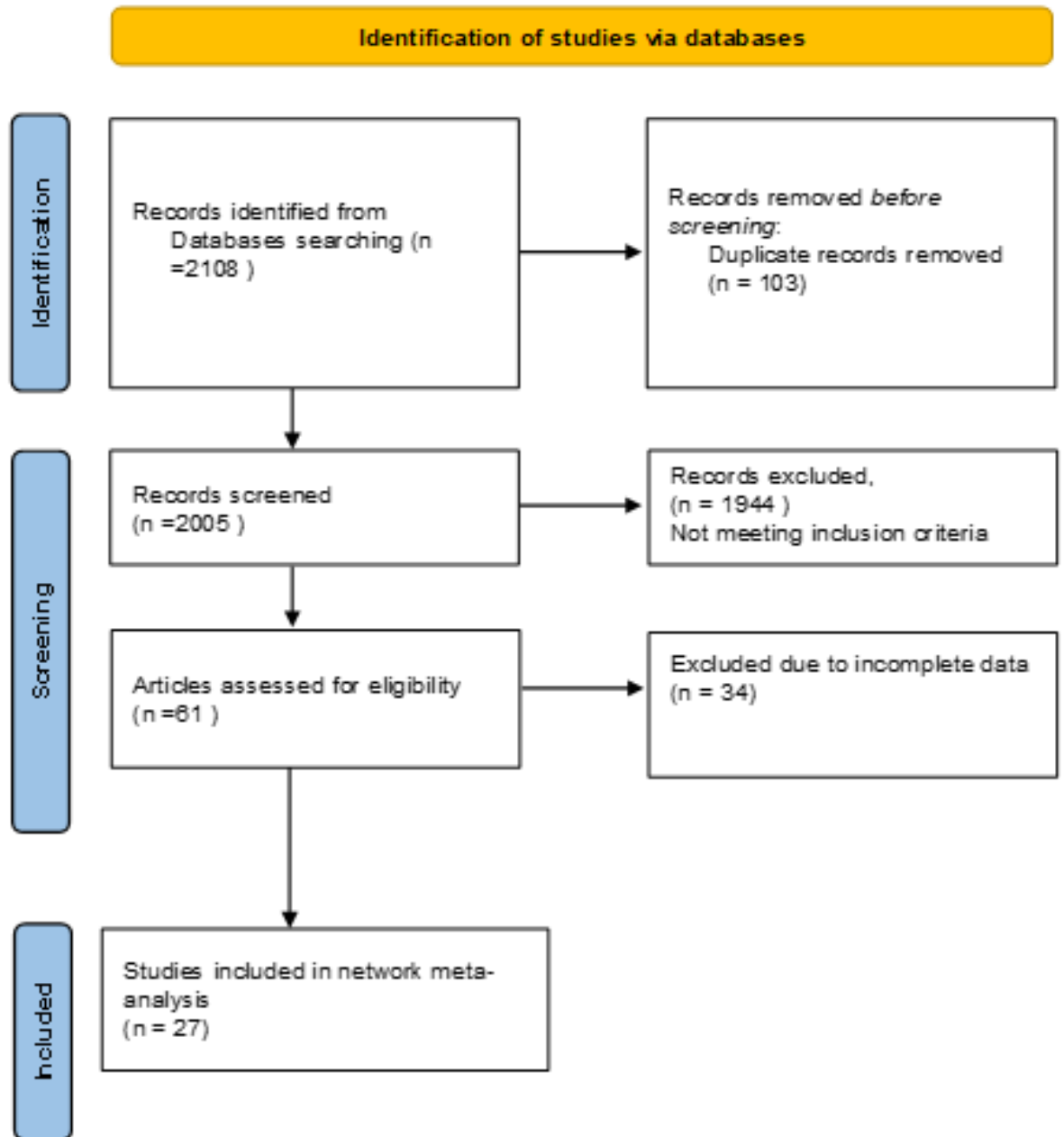


Figure 7.2.1 PRISMA diagram of network meta-analysis

7.3 Intervention Description

The interventions comprised both medical and psychological treatments for Schizophrenia, including placebo and treatment as usual (TAU). The Pharmacological interventions were Aripiprazole, Asenapine, Clozapine, Haloperidol, Iloperidone, Lurasidone, Methylphenidate, Molindone, Olanzapine, Paliperidone, Quetiapine, Risperidone, Ziprasidone and Venlafaxine. The Psychological interventions were Cognitive remediation and Psychotherapy.

7.4 Study Characteristics

In this paper, we used 27 Randomized Controlled Trials (RCT) conducted to examine efficacy of schizophrenia treatments in underage patients. There were 25 two-arm trials, 2 three-arm trials and 4195 randomized participants with age 18 years or less. The onset of schizophrenia frequently occurs in adolescence age and adolescent patients can start treatments after diagnosis [Hollis and Rapoport, 2011]. However, most randomized controlled clinical trials for schizophrenia treatments include only adult patients in their study. In this study we are interested in focusing on Schizophrenia clinical trials for children and the adolescent population. Outcomes of interest included PANSS (positive and negative syndrome score) and CGI-S (clinical global impression for severity). The PANSS is a 30-items rating scale to measure the severity of schizophrenia patients' symptoms. The PANSS is comprised of positive scale which measures positive symptoms (excess of normal functions), negative scale which

measures negative symptoms (loss of normal functions) and general psychopathology scale. Each item has score of 1 to 7 points, 1 for absence of symptom and 7 for extreme symptom. The total PANSS score, which is the sum of the scores of all the three sub scales ranges 30 to 210. The CGI-S is a 7- point rating scale to measure the severity of symptoms of schizophrenia patients. It is also used to measure treatment response and efficacy of treatments in patients with schizophrenia. The CGI-S score ranges from 1 to 7 points, 1 for normal and 7 for extreme symptoms

7.5 Statistical Analysis

NMA using frequentist framework was performed to examine the efficacy of Randomized Clinical Trials in Adolescent Patients with Schizophrenia. The parameters that represent the characteristics of the study population were inferred using the likelihood of the observed data. The method calculates the probability under the assumption that the observed data repeats infinitely. Results are provided as a point estimate (effect measures, mean difference) with a 95% CI. Our NMA utilized random-effects approach. Random-effects methods assumes that the observed difference in the effect size take into account sampling errors, true effect size variations across studies and between direct and indirect comparisons.

Assumptions validation: To lower risks of bias, the assumptions of homogeneity, transitivity and consistency were assessed. Homogeneity (similarity), transitivity and consistency assumptions apply to direct comparisons, indirect comparisons, and mixed comparisons (direct and indirect comparisons) respectively. Cochran's Q statistic was used for assessing heterogeneity within and between designs. Consistency achieved when the relative effect based on direct comparison does not differ from indirect comparison [[Schwarzer et al., 2015](#)]. Global consistency under the assumption of a full design-by-treatment interaction random effects model were performed to assess consistency between direct and indirect comparison overall. Q statistics was used to test global consistency. Local consistency tests were also conducted for studies with both direct and indirect comparisons. Z-test was used to test local consistency.

Local consistency model: Suppose we have treatments A, B and C

$$\mu_{AC} = \mu_{AB} - \mu_{BC} \quad (7.1)$$

Inconsistency model:

$$\mu_{AC} = \mu_{AB} - \mu_{BC} + \omega_{ABC} \quad (7.2)$$

Where, μ_{AB} is the treatment effect of AB, μ_{BC} is the treatment effect of BC, μ_{AC} is the treatment effect of AC, and ω_{ABC} is the inconsistency variable. The distribution of the inconsistent parameter is $\omega \sim \mathcal{N}(0, \sigma^2)$

Egger's test was used to assess potential publication bias via funnel plot asymmetry.

NMA model::

Let K, M and n denote number of studies, number of pairwise comparisons and number of treatments respectively. We calculated the effect size $\hat{\beta}_m$ for each pairwise comparison m and all effect sizes are collected in a vector $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_M)$ The following model was used to generate the vector of observed effect sizes $\hat{\beta}$.

$$\hat{\beta} = \mathbf{X}\beta_{treat} + \epsilon \quad (7.3)$$

Where \mathbf{X} is a $m \times n$ design matrix, in which the rows represent m pairwise comparisons and the columns represent the n different treatments. In our network, the true effects of the n unique treatments (the mean value of outcome given different treatments)

was represented by the vector β_{treat} . This vector is the vector which was estimated by our NMA model. β_{treat} enables us to determine the best performing or most effective treatments in our network. The sampling errors of each pairwise comparison is represented by ϵ_m and they are included in the vector ϵ .

The distribution of the sampling errors of each pairwise comparison is assumed to be a Gaussian normal with a mean of zero and variance σ_m^2 .

$$\epsilon_m \sim \mathcal{N}(0, \sigma_m^2) \tag{7.4}$$

Primary Analysis: Combined analysis

The performance of treatments within each outcome was compared using studies with both outcomes. In addition, treatments between PANSS and CGI-S outcomes were compared.

For the combined analysis, there were seven treatments such as Placebo, Aripiprazole, Asenapine, Olanzapine, Paliperidone, Risperidone and Ziprasidone. Let these treatments be denoted by A, B, C, D, E, F, and G respectively for model illustration.

The objective is to estimate the true effect sizes of the treatments in our network,

$$\beta_{treat} = (\beta_A, \beta_B, \beta_C, \beta_D, \beta_E, \beta_F, \beta_G)^T$$

The effect sizes of the pairwise comparisons are a vector:

$$\hat{\beta} = (\hat{\beta}_{AB}, \hat{\beta}_{AC}, \hat{\beta}_{AD}, \hat{\beta}_{AE}, \hat{\beta}_{AF}, \hat{\beta}_{AG}, \hat{\beta}_{BE})^T$$

We get the following equation when we put the above parameters in our network

$$\text{model: } \hat{\beta} = \mathbf{X}\beta_{treat} + \epsilon$$

$$\begin{bmatrix} \hat{\beta}_{A,B} \\ \hat{\beta}_{A,C} \\ \hat{\beta}_{A,D} \\ \hat{\beta}_{A,E} \\ \hat{\beta}_{A,F} \\ \hat{\beta}_{A,G} \\ \hat{\beta}_{B,E} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \beta_A \\ \beta_B \\ \beta_C \\ \beta_D \\ \beta_E \\ \beta_F \\ \beta_G \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \\ \epsilon_7 \end{bmatrix} \quad (7.5)$$

Secondary analysis: Separate analysis

To test the robustness of our findings and to validate our conclusion, we performed two sensitivity analyses (separate analysis for each outcome) based on full sets of literature (each outcome is separately analyzed). By doing so, we finally compared whether the sensitivity analyses support or violate our findings in the primary analysis.

Rank probabilities: Based on the results of NMA, rank probabilities were calculated for each treatment to be the first rank (best treatment), second, third, fourth, fifth... and last rank (worst). The cumulative ranking probabilities were also calculated and demonstrated by cumulative ranking curves. P scores of the treatments are estimated and presented in tables. P score is the average proportion of treatments worse than it [Rücker and Schwarzer, 2015]. In other words, it is the

average proportion of certainty that the treatment is better than all the competing treatments. The probability that a treatment is among the best options is represented by P-score. A treatment with a larger P-score have a higher rank in the network. Forest plots were used to provide a visual summary of the analysis results. The plots graphically represent the effect size estimates and the corresponding 95% confidence intervals for each study. Forest plots were used for visual evaluation of the individual studies included in the analyses. Network Estimate Forest plot for treatments which had both direct and indirect evidence are provided. Network graphs which are visual presentation of evidence are also provided. In the graphs, each treatment is represented by one node, size of the node indicates the sample size of subjects randomized to that treatment, and thickness of width of the edge shows the number of clinical trials comparing two treatments[Rouse et al., 2017].Network graph provides illustrations of the overall structure of comparisons in our network that enables us to understand the treatment comparisons conducted in the main data. Statistical analyses were conducted using R Statistical Software Package version 4.3.2 (R Core Team, 2023) for Microsoft Windows. Statistical significance is defined as p-value ≤ 0.05 .

Chapter 8: Result

8.1 Primary Analysis: Combined analysis

The systematic review identified 10 relevant RCTs with available PANSS and CGI-S outcome data that examined the efficacy of schizophrenia treatments in underage patients. There were 10 two-arm trials, 7 treatments and 1868 randomized participants with age 18 years or less. The performance of treatments within each outcome, and between outcomes PANSS and CGI-S were compared. Based on effect sizes, Olanzapine and Paliperidone were found to be the best treatment for PANSS and CGI respectively. Although the effect size shows a decreasing trend, if we are taking 95% confidence intervals into account Olanzapine, Risperidone and Paliperidone are nested within each other. Since 95% confidence intervals indicate the extent of evidence, Olanzapine, Risperidone and Paliperidone are good to use, and they are found to be very good across studies. Aripiprazole works good for PANSS but not that good for CGI as the effect size is close to zero. Overall Olanzapine, Risperidone and Paliperidone perform very well in both outcome analysis and they are robust.

Network Diagram:

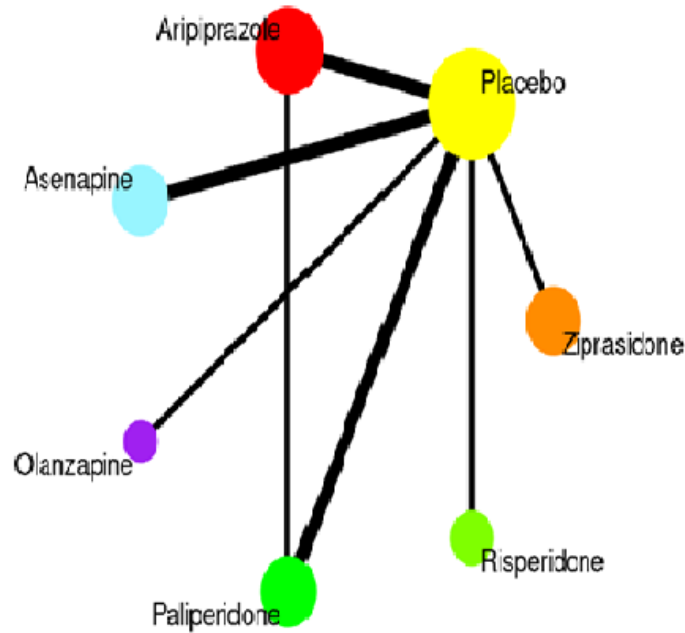


Figure 8.1.1 Network graph, each treatment is represented by one node, the size of the node indicates the sample size of subjects randomized to that treatment, and the thickness of width of the edge shows the number of clinical trials comparing two treatments. Rouse et al, 2017. This network graph provides illustrations of the overall structure of comparisons in our network that enables us to understand the treatment comparisons conducted in the main analysis

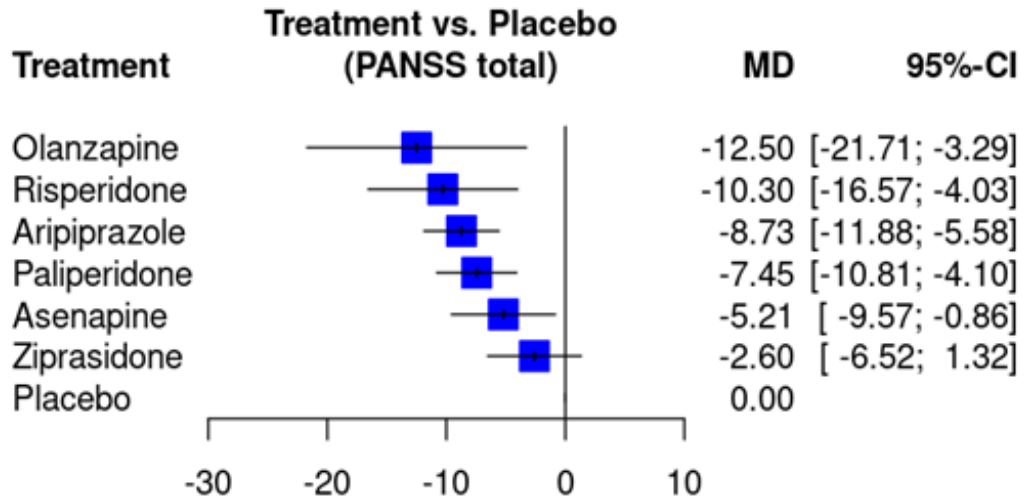


Figure 8.1.2 Treatment effect estimates with 95% confidence interval for the comparison of treatments using PANSS total score.

Rank	Effects of treatments						
	Aripiprazole	Asenapine	Olanzapine	Paliperidone	Placebo	Risperidone	Ziprasidone
1	0.055	0.007	0.605	0.023	0.00	0.310	0.000
2	0.293	0.031	0.194	0.107	0.00	0.374	0.001
3	0.432	0.085	0.071	0.273	0.00	0.130	0.009
4	0.189	0.165	0.068	0.442	0.00	0.106	0.030
5	0.031	0.484	0.043	0.142	0.00	0.062	0.238
6	0.0	0.215	0.017	0.013	0.11	0.016	0.629
7	0.0	0.013	0.002	0.000	0.89	0.002	0.093

Table 8.1.1 Rank probabilities for the comparison of treatments using PANSS total score

Each row represent the probability of each treatment to be the 1st rank (the best), second, third, fourth, fifth, . . . and last rank(worst). The probability of Olanzapine being the best treatment (1st rank) is 60.5%. The probability of Risperidone being the best treatment (1st rank) is 31.1%. The probability of Aripiprazole being the best treatment (1st rank) is 5.5%.

Composite line chart of rank probability on PANSS.random

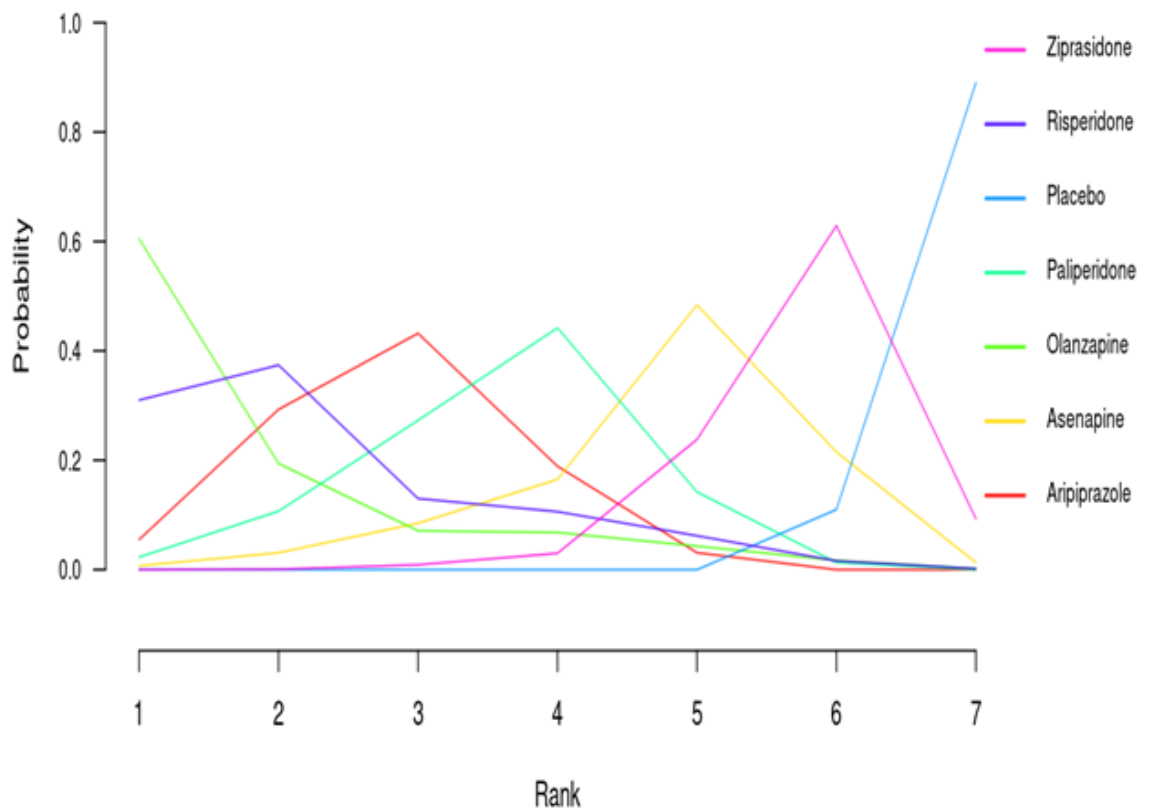


Figure 8.1.3 Composite line chart for probabilities of treatments on each rank using PANSS

Accumulative bar chart of rank probability on PANSS.random

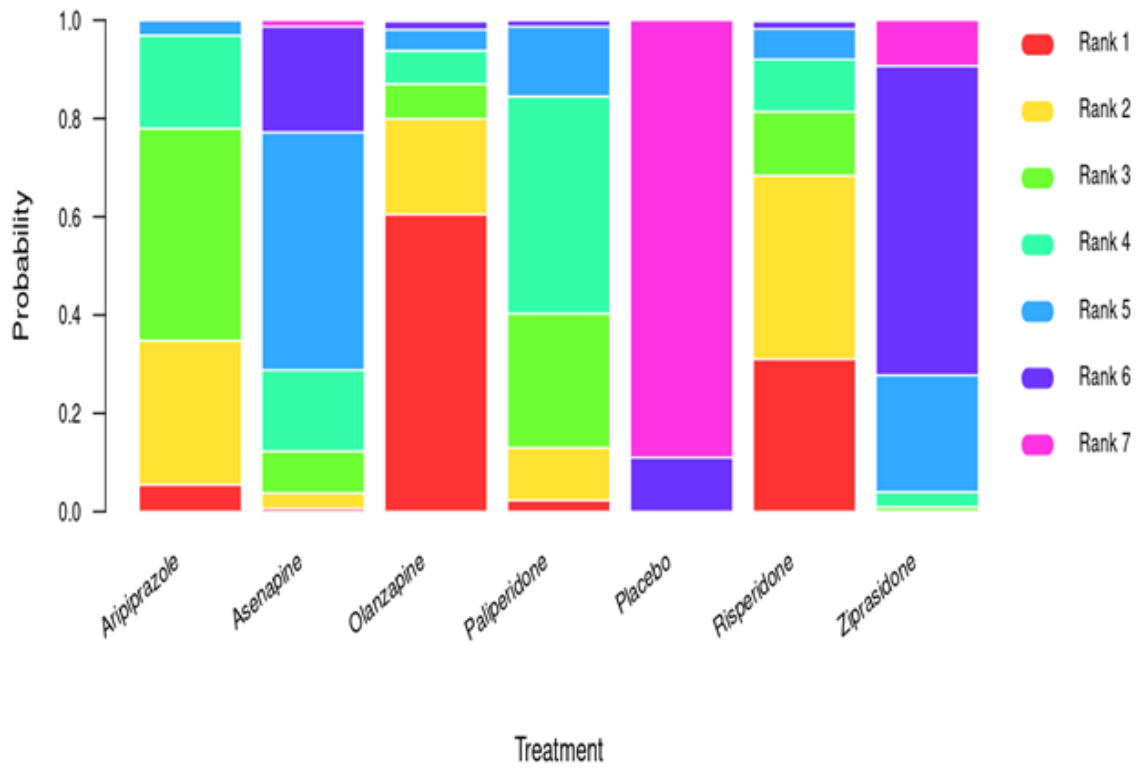


Figure 8.1.4 Accumulative bar chart for probabilities of treatments on each rank using PANSS

Treatments	P score
Olanzapine	0.8599
Risperidone	0.7809
Aripiprazole	0.7251
Paliperidone	0.5373
Asenapine	0.3813
Ziprasidone	0.1970
Placebo	0.0185

Table 8.1.2 P score of treatments for the comparison of treatments using PANSS total score

As can be seen from the above table, we are 85.99% certain that the treatment Olanzapine is the best treatment compared to all other treatments. In other words, 85.99% of treatments are worse than Olanzapine.

	Aripiprazole	Asenapine	Olanzapine	Paliperidone	Placebo	Risperidone	Ziprasidone
Aripiprazole	0.00	-3.52	3.77	-1.28	-8.73	1.57	-6.13
Asenapine	3.52	0.00	7.29	2.24	-5.21	5.09	-2.61
Olanzapine	-3.77	-7.29	0.00	-5.05	-12.5	-2.2	-9.9
Paliperidone	1.28	-2.24	5.05	0.00	-7.45	2.85	-4.85
Placebo	8.73	5.21	12.5	7.45	0.00	10.3	2.6
Risperidone	-1.57	-5.09	2.2	-2.85	-10.3	0.00	-7.7
Ziprasidone	6.13	2.61	9.9	4.85	-2.6	7.7	0.00

Table 8.1.3 Network estimate result matrix for the comparison of treatments using PANSS

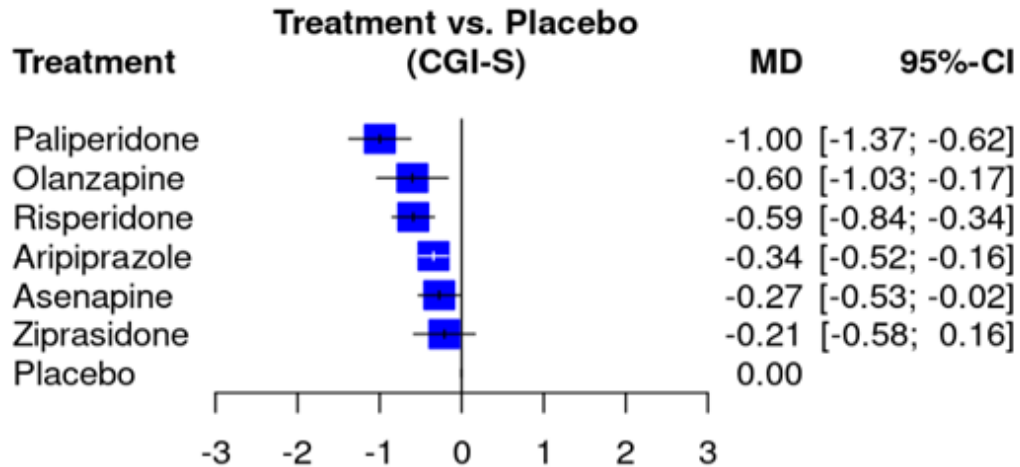


Figure 8.1.5 Treatment effect estimates with 95% confidence interval for the comparison of treatments using CGI-S score

Ranking	Effects of treatments						
	Aripiprazole	Asenapine	Olanzapine	Paliperidone	Placebo	Risperidone	Ziprasidone
1	0.000	0.000	0.064	0.908	0.000	0.027	0.001
2	0.012	0.008	0.442	0.076	0.000	0.436	0.026
3	0.105	0.076	0.300	0.016	0.000	0.427	0.076
4	0.439	0.201	0.116	0.000	0.000	0.074	0.170
5	0.343	0.381	0.043	0.000	0.001	0.030	0.202
6	0.101	0.318	0.033	0.000	0.133	0.006	0.409
7	0.000	0.016	0.002	0.000	0.866	0.000	0.116

Table 8.1.4 Rank probabilities of for the comparison of treatments using CGI-S score

The probability of Paliperidone being the best treatment (1st rank) is 90.8%.

The probability of Olanzapine being the best treatment (1st rank) is 6.4%. The

probability of Risperidone being the best treatment (1st rank) is 2.7%.

Composite line chart of rank probability on CGI.random

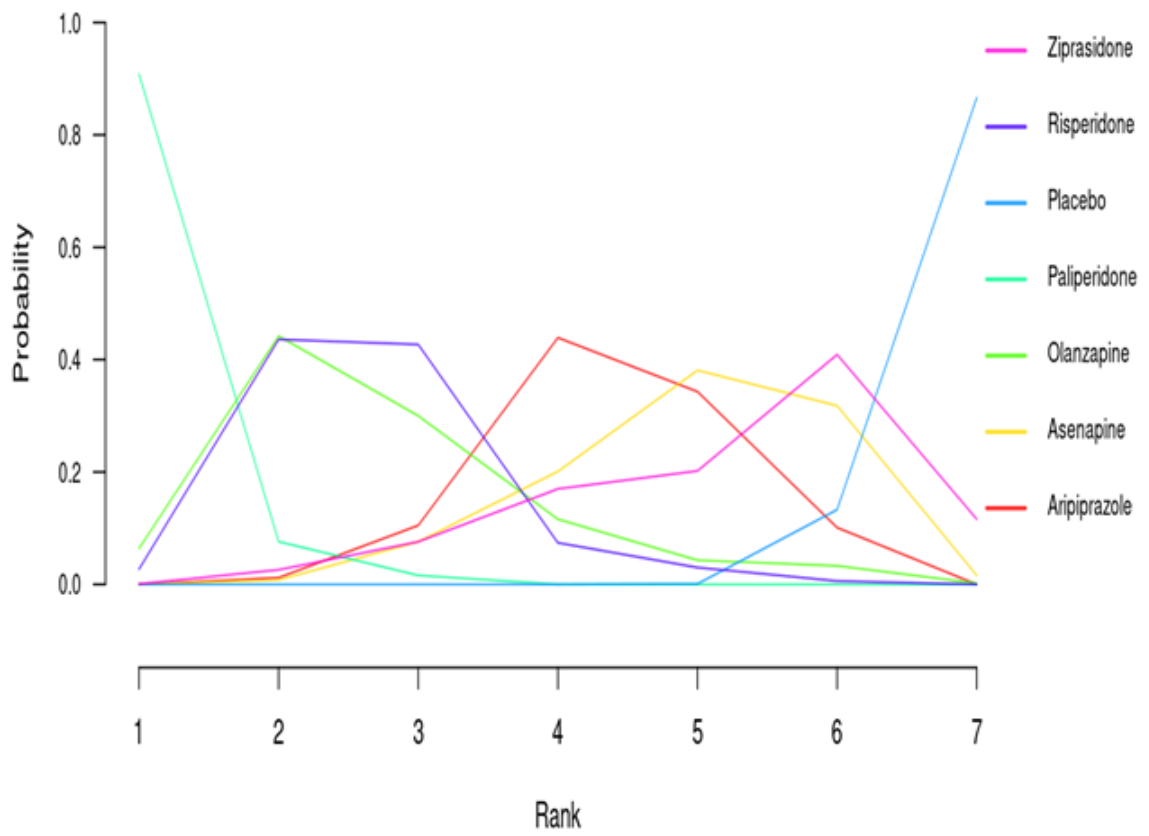


Figure 8.1.6 Composite line chart for probabilities of treatments on each rank using CGI-S

Accumulative bar chart of rank probability on CGI.random

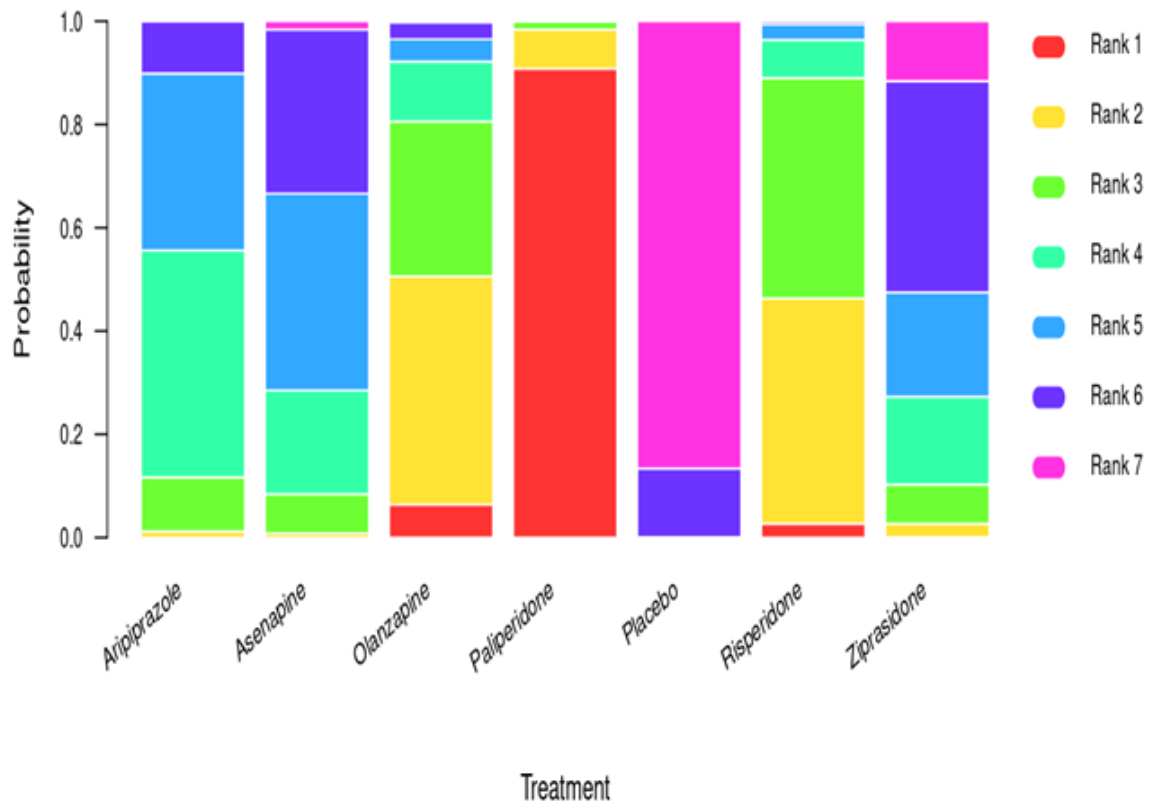


Figure 8.1.7 Accumulative bar chart for probabilities of treatments on each rank using CGI-S

Treatments	P-score
Paliperidone	0.9785
Risperidone	0.7288
Olanzapine	0.7115
Aripiprazole	0.4345
Asenapine	0.3441
Ziprasidone	0.2769
Placebo	0.0258

Table 8.1.5 P score of treatments for the comparison of treatments using CGI-S score

We are 97.85% certain that the treatment Paliperidone is the best treatment compared to all other treatments. In other words, 97.85% of treatments are worse than Paliperidone.

	Aripiprazole	Asenapine	Olanzapine	Paliperidone	Placebo	Risperidone	Ziprasidone
Aripiprazole	0.00	-0.07	0.26	0.65	-0.34	0.25	-0.13
Asenapine	0.07	0.00	0.33	0.72	-0.27	0.32	-0.06
Olanzapine	-0.26	-0.33	0.00	0.40	-0.60	-0.01	-0.39
Paliperidone	-0.65	-0.72	-0.40	0.00	-1.00	-0.41	-0.79
Placebo	0.34	0.27	0.60	1.00	0.00	0.59	0.21
Risperidone	-0.25	-0.32	0.01	0.41	-0.59	0.00	-0.38
Ziprasidone	0.13	0.06	0.39	0.79	-0.21	0.38	0.00

Table 8.1.6 Network estimate result matrix for the comparison of treatments using CGI-S score

Based on effect sizes, Olanzapine and Paliperidone were found to be the best treatment for PANSS and CGI respectively. Although the effect size shows a decreasing trend, if we are taking 95% confidence intervals into account Olanzapine, Risperidone and Paliperidone are nested within each other. Since 95% confidence intervals indicates the extent of evidence, Olanzapine, Risperidone and Paliperidone are good to use, and they are found being very good across the studies. Aripiprazole works good for PANSS but not that good for CGI as the effect size is close to zero. Overall Olanzapine, Risperidone and Paliperidone perform very well in both outcome

analysis and they are robust.

8.2 Sensitivity analysis: Separate Analysis

To test the robustness of our findings and to validate our conclusion, we conducted two sensitivity analyses (separate analysis for each outcome) based on full sets of literature. By doing so we finally compared whether those sensitivity analyses support or violate our findings in the primary analysis.

Outcome: PANNS total score

The systematic review identified 22 relevant RCTs with available PANSS outcome data (Figure2) that examined the efficacy of schizophrenia treatments in underage patients. There were 21 two-arm trials, 1 three-arm trial, 14 treatments and 3703 randomized participants with age 18 years or less. The performance of treatments was compared. Based on effect sizes, Haloperidol was found to be the best treatment for PANSS. Although the effect size shows a decreasing trend, if we are taking 95% confidence intervals into account Haloperidol, Iloperidone, Molindone, Risperidone, Olanzapine and Aripiprazole are nested within each other. Since 95% confidence intervals indicate the extent of evidence, Haloperidol, Iloperidone, Molindone, Risperidone, Olanzapine and Aripiprazole are good to use, and they are found to be very good across studies. Based on the rank probabilities, the probability of Haloperidol, Iloperidone and Molindone being the best treatment (1st rank) for PANSS was found to be 45.7%, 36.1% and 7.9% respectively. The P score also assures Haloperidol was the best treatment for PANSS, P-score=92.28%. We are 92.28% cer-

tain that the treatment Haloperidol is the best treatment for PANSS compared to all other treatments.

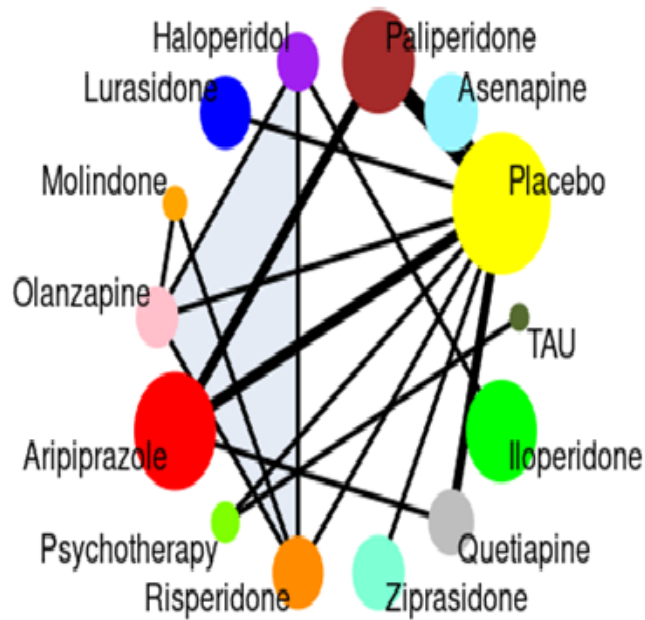


Figure 8.2.1 Network graph, each treatment is represented by one node, the size of the node indicates the sample size of subjects randomized to that treatment, and the thickness of width of the edge shows the number of clinical trials comparing two treatments. Rouse et al, 2017. This network graph provides illustrations of the overall structure of comparisons in our network that enables us to understand the treatment comparisons conducted in the main data

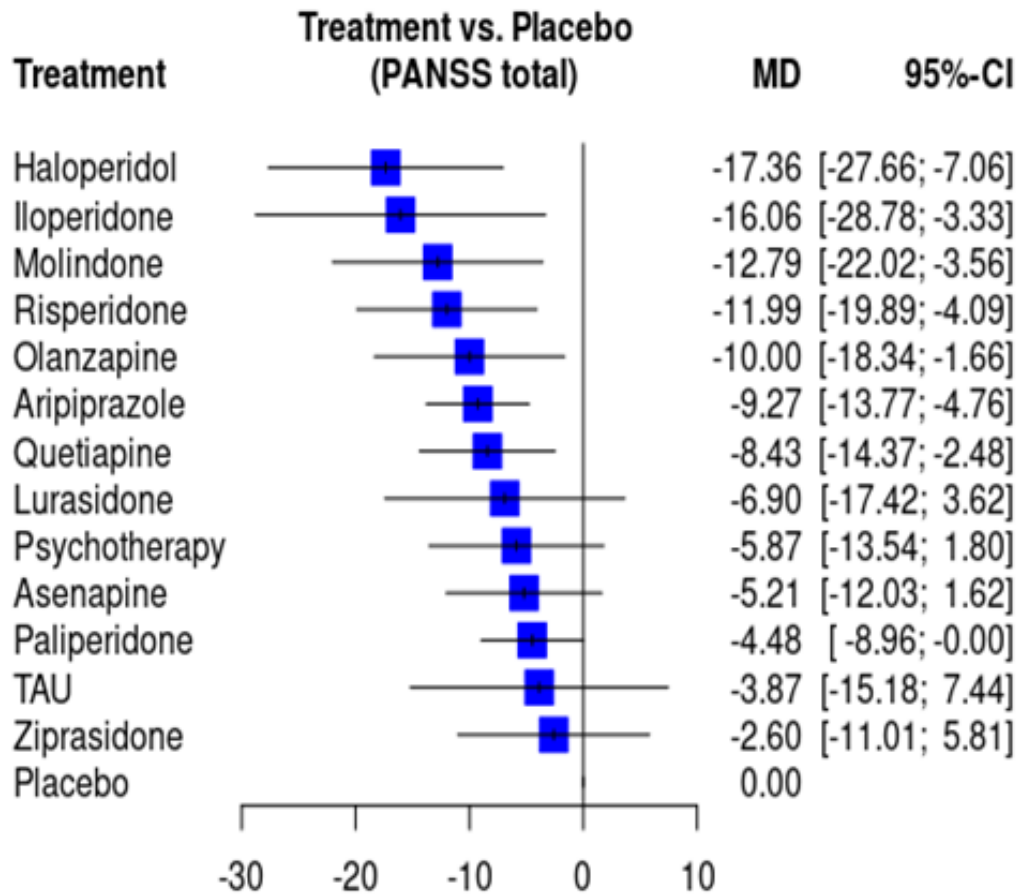


Figure 8.2.2 Treatment effect estimates with 95% confidence interval for the comparison of treatments using PANSS total score

R	Efficacy of treatments													
	Aripip razole	Asen apine	Halop ridol	loperi done	Luras idone	Molin done	Olanza pine	Palipe ridone	Plac ebo	Psych therapy	Queti apine	Risper done	TAU	Zipras idone
1	0.004	0.001	0.457	0.361	0.026	0.079	0.023	0.000	0.000	0.000	0.006	0.033	0.008	0.002
2	0.016	0.004	0.281	0.258	0.039	0.170	0.047	0.001	0.000	0.010	0.030	0.123	0.017	0.004
3	0.048	0.020	0.116	0.100	0.045	0.210	0.118	0.002	0.000	0.017	0.060	0.227	0.031	0.006
4	0.096	0.032	0.063	0.074	0.075	0.149	0.162	0.002	0.000	0.038	0.092	0.168	0.032	0.017
5	0.151	0.039	0.035	0.054	0.082	0.118	0.132	0.012	0.000	0.061	0.111	0.147	0.037	0.021
6	0.190	0.050	0.014	0.038	0.067	0.086	0.143	0.022	0.000	0.075	0.143	0.110	0.041	0.021
7	0.162	0.085	0.021	0.031	0.066	0.064	0.120	0.043	0.000	0.087	0.133	0.094	0.058	0.036
8	0.140	0.082	0.005	0.032	0.091	0.050	0.102	0.087	0.000	0.111	0.115	0.049	0.074	0.062
9	0.099	0.116	0.003	0.015	0.091	0.031	0.079	0.143	0.000	0.127	0.103	0.025	0.081	0.085
10	0.053	0.150	0.002	0.012	0.088	0.024	0.035	0.194	0.011	0.140	0.095	0.013	0.101	0.082
11	0.030	0.141	0.002	0.015	0.105	0.010	0.024	0.194	0.064	0.121	0.072	0.006	0.110	0.106
12	0.011	0.119	0.001	0.004	0.080	0.006	0.011	0.188	0.162	0.111	0.030	0.004	0.120	0.154
13	0.000	0.098	0.000	0.006	0.073	0.002	0.003	0.093	0.358	0.066	0.010	0.001	0.118	0.172
14	0.000	0.063	0.000	0.000	0.072	0.001	0.001	0.019	0.403	0.037	0.000	0.000	0.172	0.232

Table 8.2.1 Rank probabilities of for the comparison of treatments using PANSS total score

The probability of Haloperidol being the best treatment (1st rank) is 45.7%. The probability of Iloperidone being the best treatment (1st rank) is 36.1%. The probability of Molindone being the best treatment (1st rank) is 7.9%.

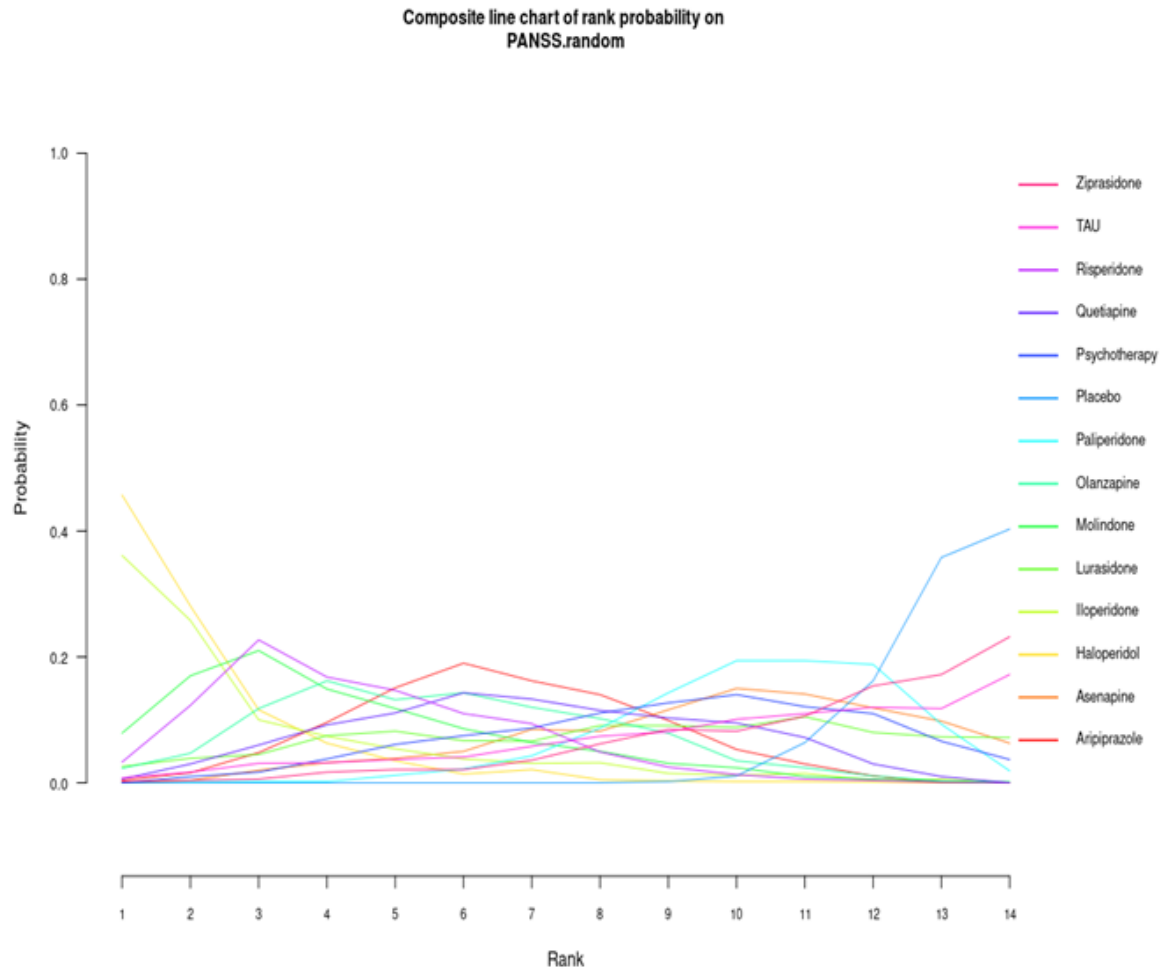


Figure 8.2.3 Composite line chart for probabilities of treatments on each rank using PANSS

Accumulative bar chart of rank probability on PANSS.random

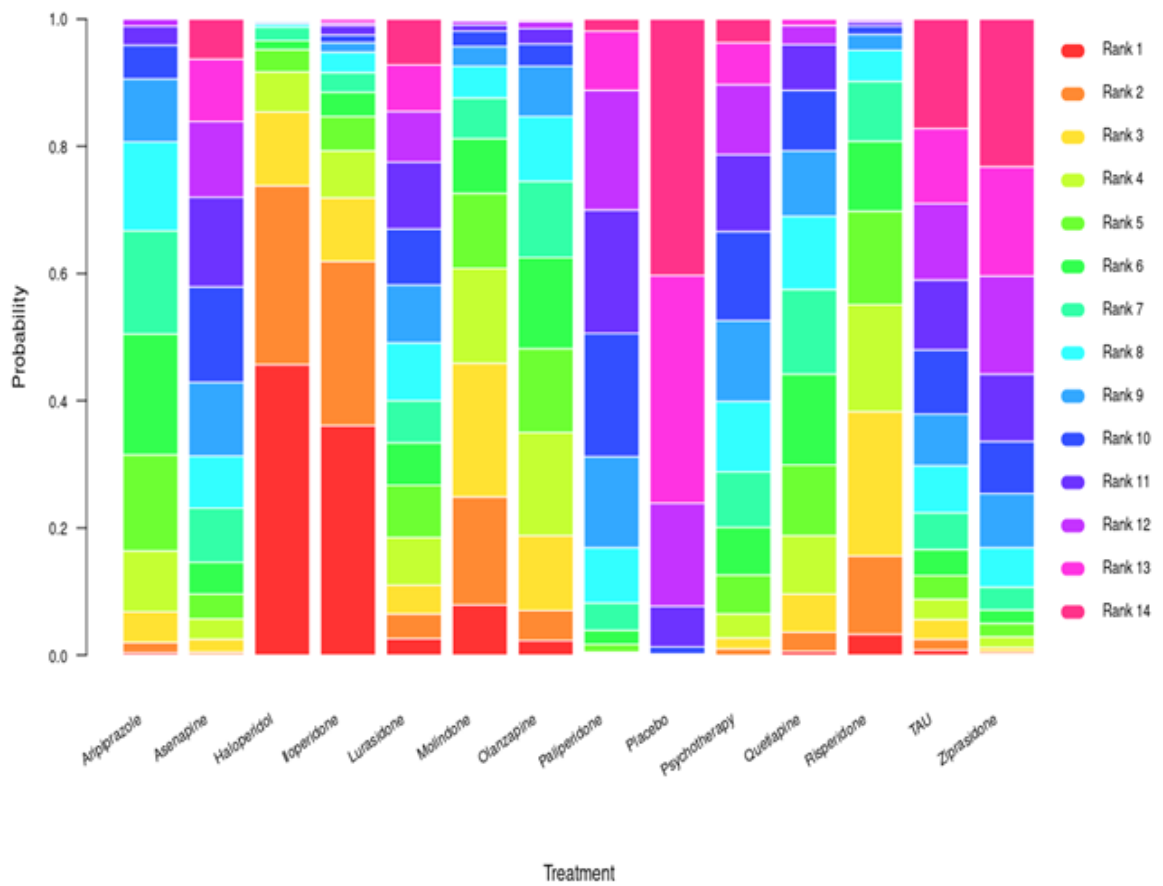


Figure 8.2.4 Accumulative bar chart for probabilities of treatments on each rank using PANSS

Treatments	P-score
Haloperidol	0.9228
Iloperidone	0.8411
Molindone	0.7487
Risperidone	0.7098
Aripiprazole	0.5994
Olanzapine	0.5842
Quetiapine	0.5380
Lurasidone	0.4485
Psychotherapy	0.3923
Asenapine	0.3463
TAU	0.2972
Paliperidone	0.2885
Ziprasidone	0.2211
Placebo	0.0620

Table 8.2.2 P score of treatments for the comparison of treatments using PANSS total score

We are 92.28% certain that the treatment Haloperidol is the best treatment compared to all other treatments. We can also say that 92.28% of treatments are worse than Haloperidol.

	Arip	Asen	Halo	Ilop	Lura	Moli	Olan	Palip	Plac	Quet	Risp	TAU	Ther	Zipra
Arip	0.00	- 4.06	8.09	6.79	- 2.37	3.52	0.73	- 4.78	- 9.27	- 0.84	2.72	-5.4	-3.4	- 6.67
Asen	4.06	0.00	12.15	10.85	51.69	7.59	4.8	- 0.72	- 5.21	3.22	6.78	- 1.34	0.66	- 2.61
Halo	- 8.09	- 12.15	0.00	- 1.30	- 10.5	- 4.56	- 7.36	- 12.87	- 17.4	- 8.93	- 5.37	- 13.5	- 11.5	- 14.8
Ilop	- 6.79	- 10.85	1.30	0.00	- 9.16	- 3.26	- 6.06	- 11.57	- 16.1	- 7.63	- 4.07	- 12.2	- 10.2	- 13.5
Lura	2.37	- 1.69	10.5	9.16	0.00	5.89	3.1	- 2.42	-6.9	1.53	5.09	- 3.03	- 1.03	- 4.30
Moli	- 3.52	- 7.59	4.56	3.26	- 5.89	0.00	- 2.79	- 8.31	- 12.8	- 4.36	-0.8	- 8.92	- 6.92	- 10.2
Olan	- 0.73	- 4.8	7.36	6.06	-3.1	2.79	0.00	- 5.52	- 10.0	- 1.57	1.99	- 6.13	- 4.13	-7.4
Palip	4.78	0.72	12.87	11.57	72.42	8.31	5.52	0.00	- 4.48	3.94	7.5	- 0.61	1.39	- 1.88
Plac	9.27	5.21	17.4	16.1	6.9	12.8	10.00	4.48	0.00	8.43	11.9	3.87	5.87	2.60
Quet	0.84	- 3.22	8.93	7.63	- 1.53	4.36	1.57	- 3.94	- 8.43	0.00	3.56	- 4.56	- 2.56	- 5.83
Risp	- 2.72	- 6.78	5.37	4.07	- 5.09	0.8	- 1.99	-7.5	- 11.9	- 3.56	0.00	- 8.12	- 6.12	- 9.39
TAU	5.4	1.34	13.5	12.2	3.03	8.92	6.13	0.61	- 3.87	4.56	8.12	0.00	2.00	- 1.27
Ther	3.40	- 0.66	11.5	10.2	1.03	6.92	4.13	- 1.39	- 5.87	2.56	6.12	- 2.00	0.00	- 3.27
Zipr	6.67	2.61	14.8	13.5	4.30	10.2	7.4	1.88	- 2.60	5.83	9.39	1.27	3.27	0.00

Table 8.2.3 Network estimate result matrix for the comparison of treatments using PANSS

Outcome: CGI-S

The systematic review identified 15 relevant RCTs with available CGI-S outcome data that examined the efficacy of schizophrenia treatments in underage patients. There were 14 two-arm trials, 1 three-arm trial, 11 treatments, and 2360 randomized participants with age 18 years or less. The performance of treatments was compared. Based on effect sizes, Methylphenidate was found to be the best treatment for CGI-S (Figure 10). Although the effect size shows a decreasing trend, if we are taking 95% confidence intervals into account Methylphenidate, Paliperidone, Clozapine, Risperidone, Venlafaxine and Olanzapine are nested within each other. Since 95% confidence intervals indicate the extent of evidence, Methylphenidate, Paliperidone, Clozapine, Risperidone, Venlafaxine and Olanzapine are good to use, and they are found to be very good across studies for CGI-S. Based on the rank probabilities, the probability of Methylphenidate, Paliperidone and Clozapine being the best treatment (1st rank) for CGI-S was found to be 52.6%, 31.7% and 15.4% respectively. Based on P-score, Paliperidone was the best treatment for CGI-S, P-score=89.93%. We are 89.93% certain that Paliperidone is the best treatment compared to all other treatments.

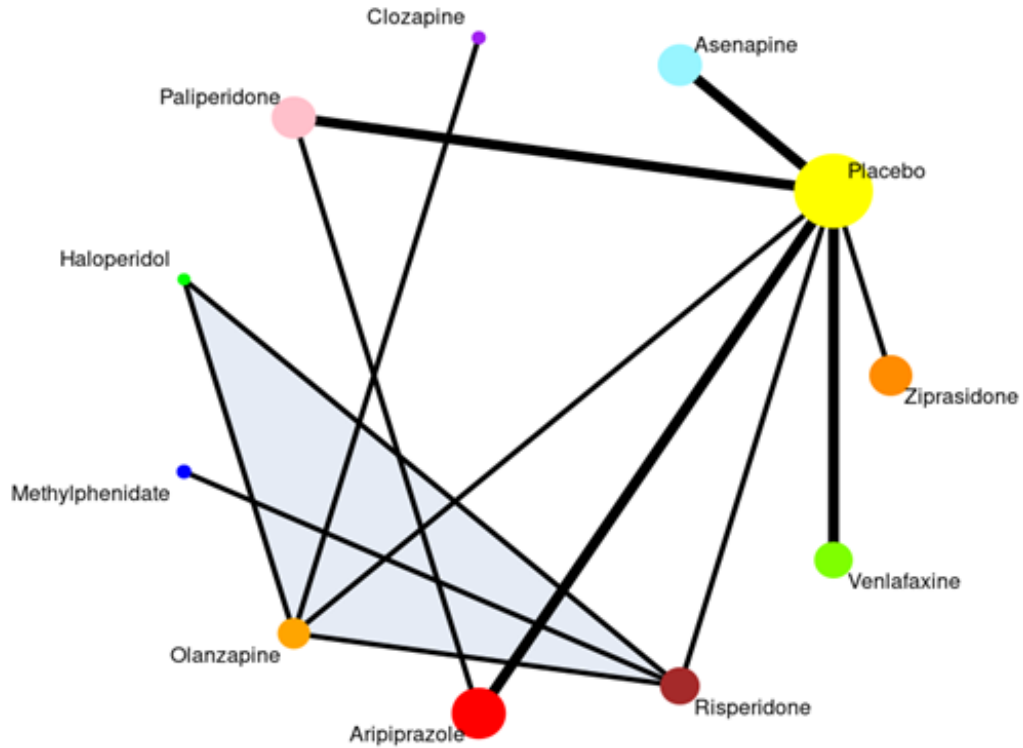


Figure 8.2.5 Network graph, each treatment is represented by one node, the size of the node indicates the sample size of subjects randomized to that treatment, and the thickness of width of the edge shows the number of clinical trials comparing two treatments. Rouse et al, 2017. This network graph provides illustrations of the overall structure of comparisons in our network that enables us to understand the treatment comparisons conducted in the main data

Rank	Efficacy of treatments										
	Aripiprazole	Asenapine	Clozapine	Haloperidol	Methyphenidate	Olanzapine	Paliperidone	Placebo	Risperidone	Venlafaxine	Ziprasidone
1	0.000	0.000	0.154	0	0.526	0.000	0.317	0.000	0.003	0.000	0.000
2	0.000	0.001	0.282	0.001	0.218	0.005	0.437	0.000	0.052	0.003	0.001
3	0.002	0.001	0.315	0.011	0.135	0.035	0.208	0.000	0.25	0.036	0.007
4	0.015	0.021	0.113	0.009	0.044	0.085	0.025	0.000	0.51	0.159	0.019
5	0.079	0.057	0.052	0.014	0.025	0.281	0.008	0.000	0.126	0.322	0.036
6	0.177	0.089	0.041	0.020	0.011	0.273	0.002	0	0.046	0.273	0.068
7	0.319	0.163	0.022	0.025	0.016	0.197	0.002	0.000	0.010	0.145	0.101
8	0.276	0.307	0.017	0.044	0.012	0.092	0.001	0.007	0.003	0.048	0.193
9	0.113	0.281	0.004	0.1	0.008	0.031	0.000	0.138	0.000	0.013	0.312
10	0.019	0.065	0.000	0.153	0.003	0.001	0.000	0.576	0.000	0.001	0.182
11	0.000	0.015	0.000	0.623	0.002	0.000	0.000	0.279	0.000	0.000	0.081

Table 8.2.4 Rank probabilities of for the comparison of treatments using CGI-S score

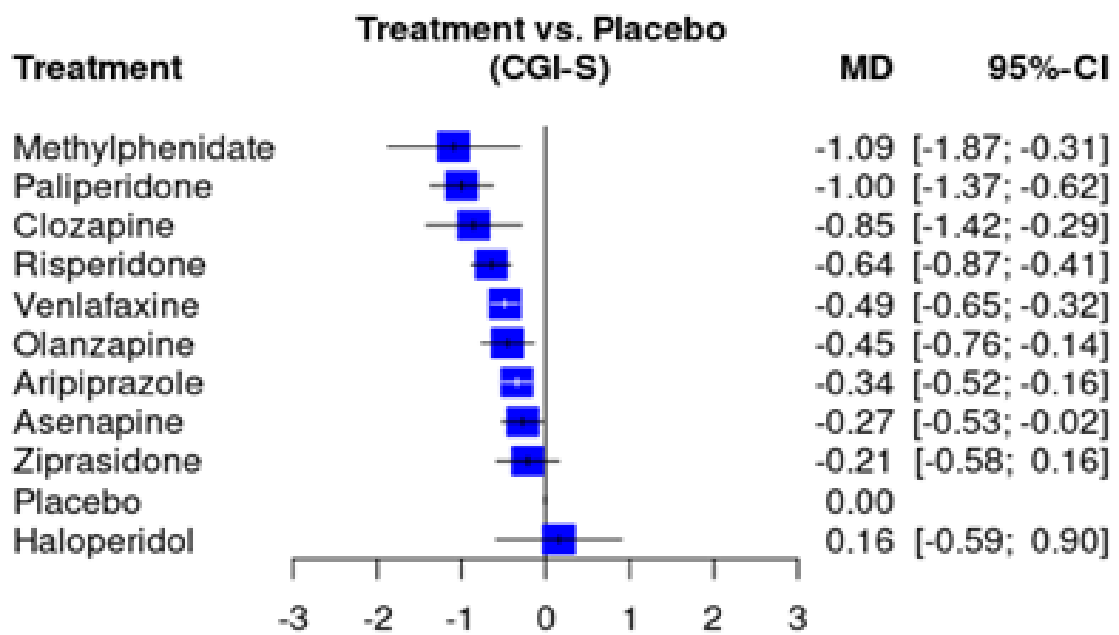


Figure 8.2.6 Treatment effect estimates with 95% confidence interval for the comparison of treatments using CGI-S score

The probability of Methylphenidate being the best treatment (1st rank) is 52.6%. The probability of Paliperidone being the best treatment (1st rank) is 31.7%. The probability of Clozapine being the best treatment (1st rank) is 15.4%.

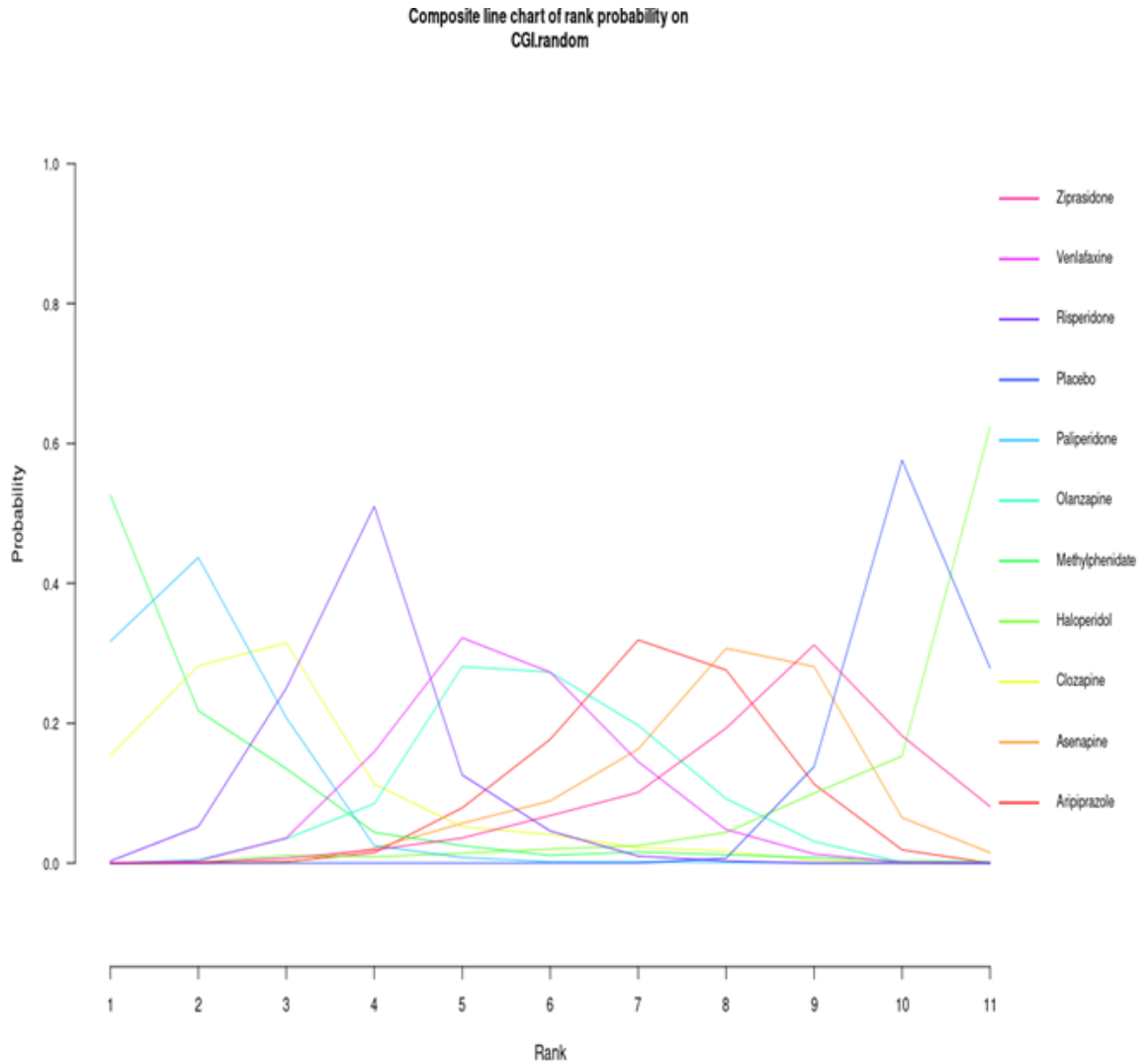


Figure 8.2.7 Composite line chart for probabilities of treatments on each rank using CGI-S

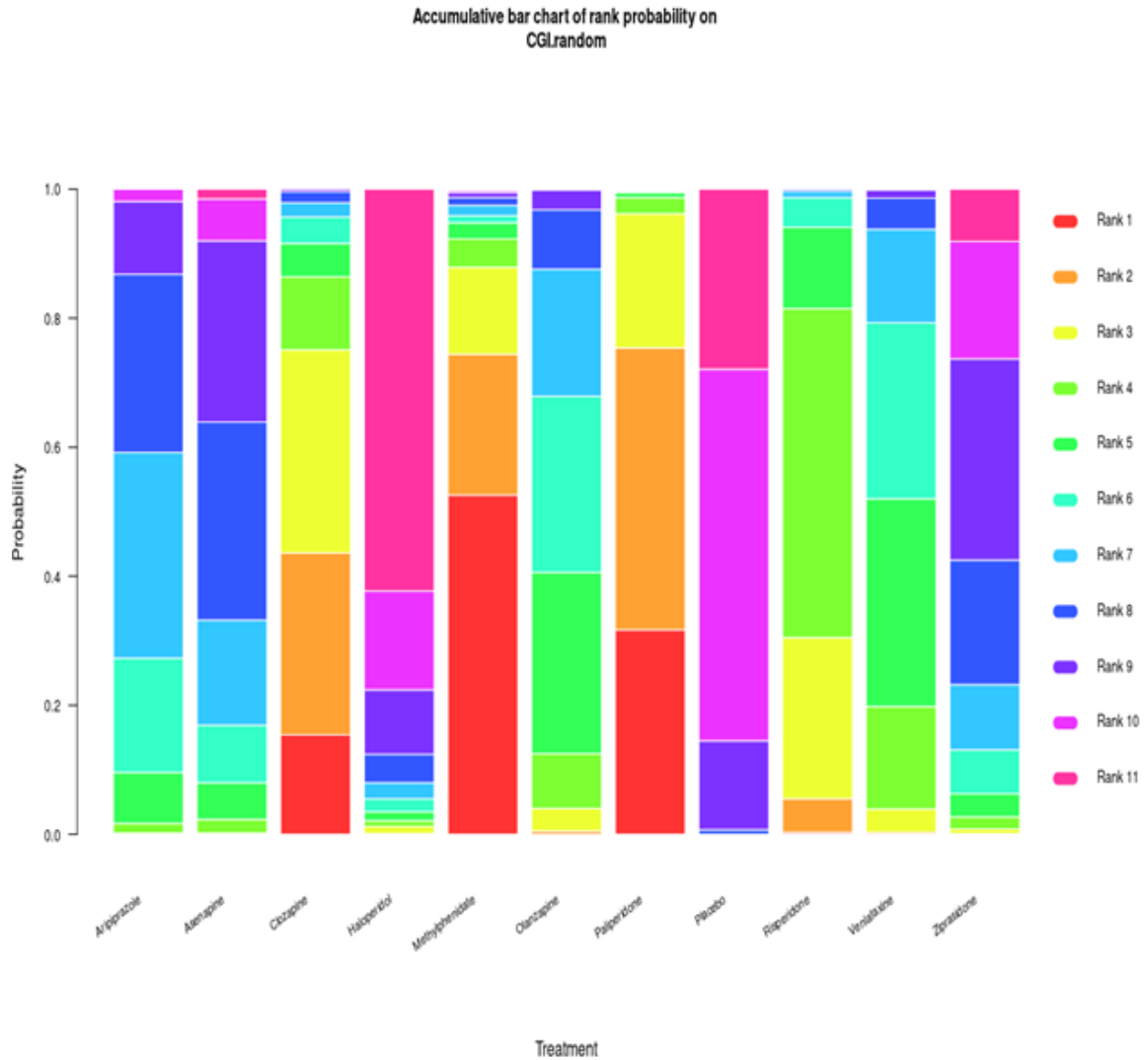


Figure 8.2.8 Accumulative bar chart for probabilities of treatments on each rank using CGI-S

Treatments	P score
Paliperidone	0.8993
Methylphenidate	0.8938
Clozapine	0.8135
Risperidone	0.7073
Venlafaxine	0.5553
Olanzapine	0.4979
Aripiprazole	0.3809
Asenapine	0.3133
Ziprasidone	0.2666
Haloperidol	0.0902
Placebo	0.0819

Table 8.2.5 P score of treatments for the comparison of treatments using CGI-S score

We are 89.93% certain that the treatment Paliperidone is the best treatment compared to all other treatments. We can also say that 89.93% of treatments are worse than Paliperidone.

	Aripip	Asena	Cloz	Halo	Methyl	Olan	Palip	Plac	Risp	Venla	Zipra
Aripip	0.00	-0.07	0.51	-0.50	0.75	0.11	0.65	-0.34	0.30	0.14	-0.13
Asena	0.07	0.00	0.58	-0.43	0.82	0.18	0.72	-0.27	0.37	0.21	-0.06
Cloz	-0.51	-0.58	0.00	-1.01	0.24	-0.40	0.15	-0.85	-0.21	-0.37	-0.64
Halo	0.50	0.43	1.01	0.00	1.25	0.61	1.16	0.16	0.80	0.64	0.37
Methyl	-0.75	-0.82	-0.24	-1.25	0.00	-0.64	-0.09	-1.09	-0.45	-0.61	-0.88
Olan	-0.11	-0.18	0.40	-0.61	0.64	0.00	0.55	-0.45	0.19	0.03	-0.24
Palip	-0.65	-0.72	-0.15	-1.16	0.09	-0.55	0.00	-1.00	-0.36	-0.51	-0.79
Plac	0.34	0.27	0.85	-0.16	1.09	0.45	1.00	0.00	0.64	0.49	0.21
Risp	-0.30	-0.37	0.21	-0.80	0.45	-0.19	0.36	-0.64	0.00	-0.16	-0.43
Venla	-0.14	-0.21	0.37	0.64	0.61	-0.03	0.51	-0.49	0.16	0.00	-0.28
Zipra	0.13	0.06	-0.64	-0.37	0.88	0.24	0.79	-0.21	0.43	0.28	0.00

Table 8.2.6 Network estimate result matrix for the comparison of treatments using CGI-S score

Chapter 9: Discussion

In the combined analysis, the performance of treatments within each outcome, and between outcomes PANSS and CGI-S were compared. Olanzapine was found to be the best treatment for lowering PANSS total score. Paliperidone was found being the best treatment in lowering CGI-S score. However, considering the 95% confidence intervals, the effect sizes of Olanzapine, Risperidone and Paliperidone are nested within each other and performed great for both outcomes. Accordingly, these three treatments were found being very good across the studies. Aripiprazole was not found being good for lowering CGI-S score as its effect size is close to zero, but it was good for lowering PANSS total score. To examine the robustness of our findings and to validate our conclusion, two sensitivity analyses were conducted for each outcome using full sets of literature. Result of the sensitivity analyses support our findings in the primary analysis.

9.1 Strengths

This study utilized different parameters such as effect sizes, rank probabilities and P-scored for examining the performance of the treatments. Graphical illustrations were provided for the network of treatments and comparisons. Consistency

between direct and indirect comparisons were assessed. Publication bias was examined. Proper statistical analysis methods were used. Measures of uncertainty for the results were reported. Two sensitivity analyses to test the robustness of our findings and to validate our conclusion. The formulated conclusions were fair and appropriate. No conflict of interest was declared by the authors.

9.2 Limitations

Even though there is no agreement on the number of RCTs required to perform network meta-analysis, it is important to include as many RCTs as possible to obtain robust evidence. Meanwhile, for our study there weren't many RCTs that included adolescents for the treatment of schizophrenia. Due to lack of data, the influence of patient characteristics on treatment effects were not examined. This study didn't include articles published in languages other than English.

9.3 Conclusions

This analysis indicates that Olanzapine, Risperidone and Paliperidone potentially being more effective for treatments for adolescent patients with Schizophrenia. Though this finding is based on studies included in this NMA and additional RCTs and network meta-analyses with more data are needed to produce more conclusive evidence. This analysis obtained evidence from combined analysis and sperate two sensitivity analyses. The results from the sensitivity analyses support our findings in the primary analysis that overall Olanzapine, Risperidone and Paliperidone per-

form very well for both outcomes, and they are robust. This study provides further evidence to support patient care and clinician decision-making about treatments for young and adolescent patients with schizophrenia.

Bibliography

- AE Ades, NJ Welton, S Dias, DM Phillippo, and DM Caldwell. Twenty years of network meta-analysis: Continuing controversies and recent developments. *Res Synth Methods*, 2024. doi: 10.1002/jrsm.1700.
- E. Ahn and H. Kang. Concepts and emerging issues of network meta-analysis. *Korean J Anesthesiol*, 2021. doi: 10.4097/kja.21358.
- Richard Barker, Matthew Schofield, Doug Armstrong, and R. Davidson. *Bayesian Hierarchical Models for Inference About Population Growth*, pages 3–17. 01 2009. ISBN 978-0-387-78150-1. doi: 10.1007/978-0-387-78151-8_1.
- Krishnan Bhaskaran and Liam Smeeth. What is the difference between missing completely at random and missing at random? *International journal of epidemiology*, 43(4):1336–1339, 2014.
- Chixiang Chen, Biyi Shen, Lijun Zhang, Yuan Xue, and Ming Wang. Empirical-likelihood-based criteria for model selection on marginal analysis of longitudinal data with dropout missingness. *Biometrics*, 75(3):950–965, 2019.
- Chixiang Chen, Biyi Shen, Aiyi Liu, Rongling Wu, and Ming Wang. A multiple robust propensity score method for longitudinal analysis with intermittent missing data. *Biometrics*, 77(2):519–532, 2021.
- Chen Chieh-feng, Chuang Yu-Chieh, Chan Edwin, Chen Jin-Hua, Hou Wen-Hsuan, and Kang Enoch. Beading plot: A novel graphics for ranking interventions. *NA*, 2023. doi: 10.21203/rs.3.rs-3370844.
- Jungyeon Choi, Olaf M. Dekkers, and Saskia Cessie. A comparison of different methods to handle missing data in the context of propensity score analysis. *European Journal of Epidemiology*, 34:23–36, 01 2019. doi: 10.1007/s10654-018-0447-z.
- Edward Cook, M.P. Tingley, E. Wahl, and E. Zorita. Bayesian hierarchical models for climate field reconstruction. *PAGES news*, 2011. doi: 10.22498/pages.19.2.78.
- CU Correll, RL Findling, M Tocco, A Pikalov, L Deng, and R Goldman. Safety and effectiveness of lurasidone in adolescents with schizophrenia: results of a 2-year, open-label extension study. *CNS Spectr*, 2022. doi: 10.1017/S1092852920001893.

- James D. Dziura, Lori A. Post, Qing Zhao, Zhixuan Fu, and Peter Peduzzi. Strategies for dealing with missing data in clinical trials: From design to analysis. *Yale Journal of Biology and Medicine*, 2013.
- Tonya M. Esterhuizen and Lehana Thabane. Con: Meta-analysis: some key limitations and potential solutions. *Nephrology Dialysis Transplantation*, 31(6):882–885, 05 2016. ISSN 0931-0509. doi: 10.1093/ndt/gfw092. URL <https://doi.org/10.1093/ndt/gfw092>.
- RL Findling, A Robb, M Nyilas, RA Forbes, N Jin, S Ivanova, R Marcus, RD McQuade, T Iwamoto, and WH Carson. A multiple-center, randomized, double-blind, placebo-controlled study of oral aripiprazole for treatment of adolescents with schizophrenia. *Am J Psychiatry*, 2008. doi: 10.1176/appi.ajp.2008.07061035.
- RL Findling, K McKenna, WR Earley, J Stankowski, and S Pathak. Efficacy and safety of quetiapine in adolescents with schizophrenia investigated in a 6-week, double-blind, placebo-controlled trial. *J Child Adolesc Psychopharmacol*, 2012. doi: 10.1089/cap.2011.0092.
- RL Findling, I Cavuş, E Pappadopulos, DG Vanderburg, JH Schwartz, BK Gundapaneni, and MP DelBello. Ziprasidone in adolescents with schizophrenia: results from a placebo-controlled efficacy and long-term open-extension study. *J Child Adolesc Psychopharmacol*, 2013. doi: 10.1089/cap.2012.0068.
- RL Findling, RP Landbloom, M Mackle, W Pallozzi, S Braat, C Hundt, MZ Wamboldt, and M Mathews. Safety and efficacy from an 8 week double-blind trial and a 26 week open-label extension of asenapine in adolescents with schizophrenia. *J Child Adolesc Psychopharmacol*, 2015. doi: 10.1089/cap.2015.0027.
- Isabella Fornacon-Wood, Hitesh Mistry, Corinne Johnson-Hart, Corinne Faivre-Finn, James P.B. O'Connor, and Gareth J and Price. Understanding the differences between bayesian and frequentist statistics. *International Journal of Radiation Oncology, Biology, Physics*, 2022.
- Katie Gillies, Peter Bower, Jim Elliott, Graeme MacLennan, Rumana SN Newlands, Margaret Ogden, Shaun P Treweek, Mary Wells, Miles D Witham, Bridget Young, et al. Systematic techniques to enhance retention in randomised controlled trials: the steer study protocol. *Trials*, 19(1):1–10, 2018.
- D Gothelf, A Apter, J Reidman, A Brand-Gothelf, Y Bloch, G Gal, L Kikinzon, S Tyano, R Weizman, and G Ratzoni. Olanzapine, risperidone and haloperidol in the treatment of adolescent patients with schizophrenia. *J Neural Transm (Vienna)*, 2003. doi: 10.1007/s00702-002-0803-7.
- Sandeep K Gupta. Use of bayesian statistics in drug development: Advantages and challenges. *International journal of applied and basic medical research*, pages 3–6, 01 2012. doi: 10.4103/2229-516X.96789.

- M Haas, M Eerdeken, S Kushner, J Singer, I Augustyns, J Quiroz, G Pandina, and V Kusumakar. Efficacy, safety and tolerability of two dosing regimens in adolescent schizophrenia: double-blind study. *Br J Psychiatry*, 2009. doi: 10.1192/bjp.bp.107.046177.
- Anna-Bettina Haidich. Meta-analysis in medical research. *Hippokratia*, 14(Suppl 1): 29, 2010.
- Seunghoon Han. Clinical vaccine development. *Clin Exp Vaccine Res*, 2015. doi: 10.7774/cevr.2015.4.1.46.
- C Hollis and J Rapoport. Child and adolescent schizophrenia. *Schizophrenia (3rd ed., pp. 24-46)*, 2011. doi: 10.1002/9781444327298.ch3.
- Sae-rom Jeon, Dongwoo Nam, and Tae-Hun Kim. Dropouts in randomized clinical trials of korean medicine interventions: a systematic review and meta-analysis. *Trials*, 22(1):1–12, 2021.
- Liyun Jiang, Lei Nie, and Ying Yuan. Elastic priors to dynamically borrow information from historical data in clinical trials. *Biometrics*, 2021.
- PD Joseph, JC Craig, and PH Caldwell. Clinical trials in children. *Br J Clin Pharmacol*, 2015. doi: 10.1111/bcp.12305.
- JM Kane, J Lauriello, E Laska, M Di Marino, and CD Wolfgang. Long-term efficacy and safety of iloperidone: results from 3 clinical trials for the treatment of schizophrenia. *J Clin Psychopharmacol*, 2008. doi: 10.1097/JCP.0b013e318169cca7.
- I Kelleher, D Connor, MC Clarke, N Devlin, M Harley, and M Cannon. Prevalence of psychotic symptoms in childhood and adolescence: a systematic review and meta-analysis of population-based studies. *Psychol Med*, 2012. doi: 10.1017/S0033291711002960.
- J. Kendhari, R. Shankar, and Young-Walker L. A review of childhood-onset schizophrenia. *FOCUS The Journal of Lifelong Learning in Psychiatry*, 2016. doi: 10.1176/appi.focus.20160007.
- L Kryzhanovskaya, SC Schulz, C McDougale, J Frazier, R Dittmann, C Robertson-Plouch, T Bauer, W Xu, W Wang, J Carlson, and M Tohen. Olanzapine versus placebo in adolescents with schizophrenia: a 6-week, randomized, double-blind, placebo-controlled trial. *J Am Acad Child Adolesc Psychiatry*, 2009. doi: 10.1097/CHI.0b013e3181900404.
- S Kumra, H Kranzler, G Gerbino-Rosen, HM Kester, C De Thomas, V Kafantaris, CU Correll, and JM Kane. Clozapine and "high-dose" olanzapine in refractory early-onset schizophrenia: a 12-week randomized and double-blind comparison. *Biol Psychiatry*, 2008. doi: 10.1016/j.biopsych.2007.04.043.

- Jack J. Lee and Caleb T. Chu. Bayesian clinical trials in action. *Statistics in medicine*, pages 2955–2972, 06 2012. doi: 10.1002/sim.5404.
- Young Lee. Strengths and limitations of meta-analysis. *The Korean Journal of Medicine*, 94:391–395, 10 2019. doi: 10.3904/kjm.2019.94.5.391.
- Yihan Li, Dai Feng, Yunxia Sui, Hong Li, Yanna Song, Tianyu Zhan, Greg Cicconetti, Man Jin, Hongwei Wang, Ivan Chan, and Xin Wang. Analyzing longitudinal binary data in clinical studies. *Contemporary Clinical Trials*, 115:106717, 2022. ISSN 1551-7144. doi: <https://doi.org/10.1016/j.cct.2022.106717>. URL <https://www.sciencedirect.com/science/article/pii/S155171442200043X>.
- KUNG-YEE LIANG and SCOTT L. ZEGER. Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13–22, 1986.
- Dan-Yu Lin and Daniel Zeng. On the relative efficiency of using summary statistics versus individual-level data in meta-analysis. *Biometrika*, 97(2):321–332, 2010.
- Guixian Lin and Robert N. Rodriguez. Weighted methods for analyzing missing data with the gee procedure. *Paper SAS166-2015*, pages 1–8, 2015a.
- Guixian Lin and Robert N. Rodriguez. Weighted methods for analyzing missing data with the gee procedure. SAS Institute, 2015b.
- G Masi, A Manfredi, G Nieri, P Muratori, C Pfanner, and A Milone. A naturalistic comparison of methylphenidate and risperidone monotherapy in drug-naive youth with attention-deficit/hyperactivity disorder comorbid with oppositional defiant disorder and aggression. *J Clin Psychopharmacol*, 2017. doi: 10.1097/JCP.0000000000000747.
- J McGrath, S Saha, D Chant, and J Welham. Schizophrenia: a concise overview of incidence, prevalence, and mortality. *Epidemiol Rev*, 2008. doi: NA.
- Lisa M Najavits. The problem of dropout from "gold standard" ptsd therapies. *F1000Prime Rep*, 2015. doi: 10.12703/P7-43.
- National Institute on Aging NIH. What are clinical trials and studies? online, 2020. URL <https://www.nia.nih.gov/health/what-are-clinical-trials-and-studies>.
- A Nikolakopoulou, D Mavridis, V Chiocchia, T Papakonstantinou, TA Furukawa, and G Salanti. Network meta-analysis results against a fictional treatment of average performance: Treatment effects and ranking metric. *Res Synth Method*, 2021. doi: 10.1002/jrsm.1463.
- M. O’Kelly and B. Ratitch. Clinical trials with missing data: A guide for practitioners. *Biometrical Journal*, 2015. doi: 10.1002/bimj.201500058.

- AK Pagsberg, P Jeppesen, DG Klauber, KG Jensen, D Rudå, M Stentebjerg-Olesen, P Jantzen, S Rasmussen, EA Saldeen, MG Lauritsen, N Bilenberg, AD Stenstrøm, L Nyvang, S Madsen, TM Werge, T Lange, C Gluud, M Skoog, P Winkel, JRM Jepsen, B Fagerlund, CU Correll, and A Fink-Jensen. Quetiapine extended release versus aripiprazole in children and adolescents with first-episode psychosis: the multicentre, double-blind, randomised tolerability and efficacy of antipsychotics (tea) trial. *Lancet Psychiatry*, 2017. doi: 10.1016/S2215-0366(17)30166-9.
- John Preisser, Kurt Lohman, and Paul Rathouz. Performance of weighted estimating equations for longitudinal binary data with drop-outs missing at random. *Statistics in medicine*, 21:3035–54, 10 2002. doi: 10.1002/sim.1241.
- O Puig, R Penadés, I Baeza, E De la Serna, V Sánchez-Gistau, M Bernardo, and J Castro-Fornieles. Cognitive remediation therapy in adolescents with early-onset schizophrenia: a randomized controlled trial. *J Am Acad Child Adolesc Psychiatry*, 2014. doi: 10.1016/j.jaac.2014.05.012.
- Noémie Resseguier, Roch Giorgi, and Xavier Paoletti. Sensitivity analysis when data are missing not-at-random. *Epidemiology*, 22(2):282, 2011.
- James M. Robins, Andrea Rotnitzky, and Lue Ping Zhao. Analysis of semi-parametric regression models for repeated outcomes in the presence of missing data. *JASA*, 90(429):106–121, 1995a.
- James M Robins, Andrea Rotnitzky, and Lue Ping Zhao. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the american statistical association*, 90(429):106–121, 1995b.
- A. I. Rohden, M. C. Benchaya, R. S. Camargo, T. C. Moreira, H. M. T. Barros, and M. Ferigolo. Dropout Prevalence and Associated Factors in Randomized Clinical Trials of Adolescents Treated for Depression: Systematic Review and Meta-analysis. *Clin Ther*, 39(5):971–992, May 2017.
- B Rouse, A Chaimani, and T Li. Network meta-analysis: an introduction for clinicians. *Intern Emerg Med*, 2017. doi: 10.1007/s11739-016-1583-7.
- MA Rynn, MA Riddle, PP Yeung, and NR Kunz. Efficacy and safety of extended-release venlafaxine in the treatment of generalized anxiety disorder in children and adolescents: two placebo-controlled trials. *Am J Psychiatry*, 2007. doi: 10.1176/ajp.2007.164.2.290.
- G Rücker and G Schwarzer. Ranking treatments in frequentist network meta-analysis works without resampling methods. *BMC Med Res Methodol*, 2015. doi: 10.1186/s12874-015-0060-8.
- Alejandro Salazar, Begoña Ojeda, María Dueñas, Fernando Palacín, and Inmaculada Failde. Simple generalized estimating equations (gees) and weighted generalized estimating equations (wgees) in longitudinal studies with dropouts: Guidelines and implementation in r. *Statistics in medicine*, 35, 04 2016. doi: 10.1002/sim.6947.

- AJ Savitz, R Lane, I Nuamah, S Gopal, and Hough D. Efficacy and safety of paliperidone extended release in adolescents with schizophrenia: a randomized, double-blind study. *J Am Acad Child Adolesc Psychiatry*, 2015. doi: 10.1016/j.jaac.2014.11.009.
- G Schwarzer, J Carpenter, and G Rücker. Meta-analysis with r. *BOOK*, 2015. doi: 10.1007/978-3-319-21416-0.
- M Seo, TA Furukawa, AA Veroniki, T Pillinger, A Tomlinson, G Salanti, A Cipriani, and O Efthimiou. The kilim plot: A tool for visualizing network meta-analysis results for multiple outcomes. *Res Synth Methods*, 2021. doi: 10.1002/jrsm.1428.
- P She, H Zeng, and B Yang. Effect of self-consistency group intervention for adolescents with schizophrenia: An inpatient randomized controlled trial. *J Psychiatr Res*, 2016. doi: 10.1016/j.jpsychires.2015.11.006.
- Biyi Shen, Haoyu Ren, Michelle Shardell, Jason Falvey, and Chixiang Chen. A semi-parametric model simultaneously handling unmeasured confounding, informative cluster size, and truncation by death with a data application in medicare claims. *arXiv preprint arXiv:2303.03502*, 2023.
- L Sikich, RM Hamer, RA Bashford, BB Sheitman, and JA Lieberman. A pilot study of risperidone, olanzapine, and haloperidol in psychotic youth: a double-blind, randomized, 8-week trial. *Neuropsychopharmacology*, 2004. doi: 10.1038/sj.npp.1300327.
- L Sikich, JA Frazier, J McClellan, RL Findling, B Vitiello, L Ritz, D Ambler, M Puglia, AE Maloney, E Michael, S De Jong, K Slifka, N Noyes, S Hlastala, L Pierson, NK McNamara, D Delpuerto-Bedoya, R Anderson, RM Hamer, and JA Lieberman. Double-blind comparison of first- and second-generation antipsychotics in early-onset schizophrenia and schizo-affective disorder: findings from the treatment of early-onset schizophrenia spectrum disorders (teoss) study. *Am J Psychiatry*, 2008. doi: 10.1176/appi.ajp.2008.08050756.
- J Singh, A Robb, U Vijapurkar, I Nuamah, and D Hough. A randomized, double-blind study of paliperidone extended-release in treatment of acute schizophrenia in adolescents. *Biol Psychiatry*, 2011. doi: 10.1016/j.biopsych.2011.06.021.
- Zoë C Skea, Rumana Newlands, and Katie Gillies. Exploring non-retention in clinical trials: a meta-ethnographic synthesis of studies reporting participant reasons for drop out. *BMJ Open*, 9(6), 2019. ISSN 2044-6055. doi: 10.1136/bmjopen-2018-021959. URL <https://bmjopen.bmj.com/content/9/6/e021959>.
- J. F. Tierney, C. Vale, R. Riley, C. T. Smith, L. Stewart, M. Clarke, and M. Rovers. Individual participant data (ipd) meta-analyses of randomised controlled trials: Guidance on their use. *PLoS medicine*, 2015. doi: 10.1371/journal.pmed.1001855.

H.E. Tilbrook, C.E. Hewitt, J.D. Aplin, A. Semlyen, A. Trehela, I. Watt, and D.J. Torgerson. Compliance effects in a randomised controlled trial of yoga for chronic low back pain: a methodological study. *Physiotherapy*, 100(3):256–262, 2014. ISSN 0031-9406. doi: <https://doi.org/10.1016/j.physio.2013.10.001>. URL <https://www.sciencedirect.com/science/article/pii/S0031940613001181>.

WHO. Schizophrenia. *WHO*, 2022. URL <https://www.who.int/news-room/fact-sheets/detail/schizophrenia#:~:text=Schizophrenia%20causes%20psychosis%20and%20is,as%20jumbled%20or%20irrelevant%20speech>.