Assessing the Robustness of Network Meta-Analysis of Atypical Antipsychotics in the Presence of Heterogeneity
Claire M Ainsworth, Adrian D Vickers, Jean-Gabriel Le Moine
RTI Health Solutions, Manchester, United Kingdom

BACKGROUND
• Network meta-analysis (NMA) is a valuable tool for synthesis of results from different studies for multiple interventions within a given research area.1,2
  – It combines direct and indirect evidence on intervention effects, allowing an understanding of the relative efficacy of multiple interventions and their ranking.1,2
• However, criticisms of pairwise meta-analysis (MA) remain present, and the extension to NMA can present further challenges.1,4
  – Between-study heterogeneity and reporting bias can limit the validity of results obtained.1,4
  – Small-study effects (SSEs), defined as the tendency for small studies to show larger treatment effects than those in larger studies, can also influence results.8
• Numerous methods have been applied as sensitivity analyses for conventional MAs as potential solutions to such issues.9
  – Relative efficacy has been measured using mean difference (MD) in PANSS total score.10
  – Relative efficacy was measured using mean difference (MD) in PANSS total score.10
  – Meta-regression has been used to adjust NMA results for small-study effects.11
  – Differences in treatment outcomes between conventional NMA and each alternative analysis approach were assessed using the difference in MD (ΔMD), and by comparing probability of best treatment estimates, surface under the cumulative ranking curve (SUCRA), and rankings of each atypical antipsychotic.
  – A ΔMD greater than 0 indicates larger treatment outcomes with the MA of all trials than with the alternative strategy.

OBJECTIVE
• This research extends the methods proposed by Dechartres et al.10 to explore the difficulties presented by heterogeneity and reporting bias, in order to assess robustness of findings in NMA.

METHODS
• Data came from recently published Cochrane systematic reviews investigating the effect of atypical antipsychotics in patients with schizophrenia.10,11
  – Positive and Negative Syndrome Scale (PANSS) total score endpoint data for trials included in seven direct treatment comparisons (aripiprazole vs. clozapine, quetiapine, risperidone, olanzapine, and ziprasidone; risperidone vs. quetiapine, and olanzapine) were extracted and synthesized using Bayesian random-effects NMA.
  – The PANSS has 30 items, each scored on a 7-point scale (1 = absent; 7 = extreme),12 where a lower total score indicates lesser severity of schizophrenia symptoms.
  – Relative efficacy is measured using mean difference (MD) in average PANSS total score.
• Results from a conventional NMA of all trials were compared with results when
  1. NMA was conducted using the single most precise trial; that is, the trial with the narrowest confidence interval for each of the included direct treatment comparisons
  2. NMA was restricted to the largest trials, defined as those having the largest 25% of sample size within the NMA
  3. A meta-regression model was used to allow effect size to depend on its standard error (or precision), thereby adjusting NMA results for SSE.11
  4. NMA was restricted to trials at a low risk of bias (|∆MD| < 1; see handout).

RESULTS
• Figure 1 presents the networks of evidence under each analysis strategy.
• Pair-wise effect sizes among antipsychotics, probabilities of being the best antipsychotic, and ranking of antipsychotics differed among analysis strategies (Figure 2 and Figure 3).
  – Differences in results for comparisons with NMA restricted to the most precise trials and NMA restricted to trials at a low risk of bias (ΔMD < 1, see handout).
• No meaningful differences in effect sizes were observed in the comparison with a meta-regression model for SSE and NMA restricted to trials at a low risk of bias (ΔMD < 1, see handout).
  – Furthermore, differences were small, with all differences in MD (ΔMD) less than 4, which is unlikely to correspond to a clinically significant difference (see handout).
• Significant between-study heterogeneity (I2 = 45%-66%) was observed.
• Relative efficacy estimates differ depending on the strategy used; however, MDs were not statistically significant.

CONCLUSIONS
• Relative efficacy estimates differ depending on the strategy used; however, in this example such differences were small.
  – Nonetheless, this study highlights how heterogeneity can present a challenge in decision making with regard to variation in ranking and the identification of the best treatment, and demonstrates how confirmatory sensitivity techniques can be used to help validate NMA conclusions.
  – Sensitivity analyses based on statistical methods are widely used; however, this study illustrates the value for authors to also consider, conduct, and report sensitivity analyses directly motivated by heterogeneity and study selection bias.
  – Authors and readers should consider the results of sensitivity analyses and the robustness of conclusions in their interpretation of NMA.

REFERENCES
See handout for references.

CONTACT INFORMATION
Claire Ainsworth, MSc
Senior Statistician, Data Analytics and Design Strategy
RTI Health Solutions
2nd Floor, The Pavilion
Towers Business Park
Wilmot Road, Didsbury
Manchester, M20 2LS, UK
Phone: +44(0)1614750337
Fax: +44(0)1614348232
E-mail: cainsworth@rti.org