BMJ Open

Comparative neurological outcomes and safety of anti-epileptic drugs during pregnancy and breastfeeding: a systematic review and network meta-analysis

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-017248
Article Type:	Research
Date Submitted by the Author:	10-Apr-2017
Complete List of Authors:	Veroniki, Areti Angeliki; Li Ka Shing Knowledge Institute, St. Michael's Hospital Rios, Patricia; Li Ka Shing Knowledge Institute, St. Michael's Hospital Cogo, Elise; Li Ka Shing Knowledge Institute, St. Michael's Hospital Straus, Sharon; Li Ka Shing Knowledge Institute, St. Michael's Hospital; University of Toronto, Department of Medicine Finkelstein, Yaron; The Hospital for Sick Children; University of Toronto, Department of Paediatrics Kealey, M.; Li Ka Shing Knowledge Institute, St. Michael's Hospital Reynen, Emily; Li Ka Shing Knowledge Institute, St. Michael's Hospital Soobiah, Charlene; Li Ka Shing Knowledge Institute, St. Michael's Hospital; University of Toronto, Institute for Health Policy Management & Evaluation Thavorn, Kednapa; University of Ottawa, School of Epidemiology, Public Health and Preventive Medicine, Faculty of Medicine; The Ottawa Hospital Research Institute, Clinical Epidemiology Program Hutton, Brian; University of Ottawa, School of Epidemiology, Public Health and Preventive Medicine, Faculty of Medicine; Ottawa Hospital Research Institute, Center for Practice Changing Research Hemmelgarn, BR; University of Calgary, Departments of Medicine and Community Health Sciences Yazdi, Fatemeh; Li Ka Shing Knowledge Institute, St. Michael's Hospital D'Souza, Jennifer; Li Ka Shing Knowledge Institute, St. Michael's Hospital MacDonald, Heather; Li Ka Shing Knowledge Institute, St. Michael's Hospital Tricco, Andrea; Li Ka Shing Knowledge Institute, St. Michael's Hospital; University of Toronto, Epidemiology Division, Dalla Lana School of Public Health
Primary Subject Heading:	Neurology
Secondary Subject Heading:	Obstetrics and gynaecology
Keywords:	multiple treatment meta-analysis, knowledge synthesis, Epilepsy < NEUROLOGY, pregnancy, infants, developmental delay

Comparative neurological outcomes and safety of anti-epileptic drugs during pregnancy and breastfeeding: a systematic review and network meta-analysis

Areti Angeliki Veroniki, PhD, MSc[1] Email: VeronikiA@smh.ca

Patricia Rios, MSc[1] Email: RiosP@smh.ca

Elise Cogo, ND, MLIS[1] Email: CogoE@smh.ca

Sharon E. Straus, MD, MSc[1,2] Email: Sharon.straus@utoronto.ca

Yaron Finkelstein, MD[3,4,5] Email: Yaron.Finkelstein@sickkids.ca

Ryan Kealey, PhD[1] Email: ryan.kealey@utoronto.ca

Emily Reynen, MD, CM, PharmD[1] Email: ereynen@gmail.com

Charlene Soobiah, PhD (Cand.)[1,6] Email: SoobiahC@smh.ca

Kednapa Thavorn, PhD[7,8,9] Email: kthavorn@ohri.ca

Brian Hutton, PhD, MSc[7,10] Email: bhutton@ohri.ca

Brenda R. Hemmelgarn, MD, PhD[11] Email: Bhemmelg@ucalgary.ca

Fatemeh Yazdi, MSc[1] Email: SabaghYazdiF@smh.ca

Jennifer D'Souza, HBSc[1] Email: jennifer.dsouza@mail.utoronto.ca

Heather MacDonald, MSc[1] Email: hrmacdonald@gmail.com

Andrea C. Tricco, PhD, MSc[1,12,*] Email: TriccoA@smh.ca

AUTHOR DETAILS

[1] Li Ka Shing Knowledge Institute, St. Michael's Hospital, 209 Victoria Street, East Building, Toronto, Ontario, M5B 1W8, Canada

[2] Department of Medicine, University of Toronto, 27 King's College Circle, Toronto, Ontario, M5S 1A1, Canada

[3] The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario, M5G 1X8, Canada

[4] Department of Paediatrics, University of Toronto, 172 St. George Street, Toronto, Ontario, M5R 0A3, Canada

[5] Department of Pharmacology and Toxicology, University of Toronto, Medical Sciences Building, Room 4207, 1 King's College Circle, Toronto, Ontario, M5S 1A8, Canada

[6] Institute for Health Policy Management & Evaluation, University of Toronto, 4th Floor, 155 College Street, Toronto, Ontario, M5T 3M6, Canada

[7] School of Epidemiology, Public Health and Preventive Medicine, Faculty of Medicine, University of Ottawa, Roger-Guindon Building, 451 Smyth Road, Ottawa, Ontario, K1H 8M5, Canada

[8] Clinical Epidemiology Program, Ottawa Hospital Research Institute, The Ottawa Hospital, 501 Smyth Road, Ottawa, Ontario, K1H 8L6, Canada

[9] Institute of Clinical and Evaluative Sciences (ICES uOttawa), 1053 Carling Avenue, Ottawa, Ontario, K1Y 4E9, Canada

[10] Ottawa Hospital Research Institute, Center for Practice Changing Research, The Ottawa Hospital–General Campus, 501 Smyth Road, PO Box 201B, Ottawa, Ontario, K1H 8L6, Canada.

[11] Departments of Medicine and Community Health Sciences, University of Calgary, TRW Building, 3rd Floor, 3280 Hospital Drive NW, Calgary, Alberta, T2N 4Z6, Canada

[12] Epidemiology Division, Dalla Lana School of Public Health, University of Toronto, 6th Floor, 155 College Street, Toronto, Ontario, M5T 3M7, Canada

***Corresponding author**

Prof. Andrea C. Tricco, PhD

Scientist, Knowledge Translation Program,

Li Ka Shing Knowledge Institute, St. Michael's Hospital,

209 Victoria Street, East Building, Toronto, Ontario, M5B 1W8, Canada

Phone: 416-864-6060, Fax: 416-864-5805, Email: TriccoA@smh.ca

Keywords: multiple treatment meta-analysis, knowledge synthesis, epilepsy, pregnancy, infants, developmental delay.

Word count: abstract (282 words); main text (3848 words); 1 table; 3 figures; 3 additional files; 41 references

ABSTRACT

Objectives: To compare the safety of Anti-epileptic drugs (AEDs) on neurodevelopment of infants/children exposed *in-utero* or during breastfeeding.

Design and Setting: Systematic review and Bayesian random-effects network meta-analysis (NMA).

Participants: 27 cohort studies including 4,841 infants/children.

Interventions: Mono- and poly-therapy AEDs were included, including first-generation (i.e., carbamazepine, clobazam, clonazepam, ethosuximide, phenobarbital, phenytoin, primidone, valproate) and newer-generation (i.e., gabapentin, lamotrigine, levetiracetam, oxcarbazepine, topiramate, vigabatrin) AEDs.

Primary and secondary Outcome measures: Cognitive developmental delay and autism/dyspraxia were primary outcomes. Attention deficit hyperactivity disorder, language delay, neonatal seizures, psychomotor developmental delay, and social impairment were secondary outcomes.

Results: The NMA on cognitive developmental delay 10 cohort studies, 748 children, 14 AEDs and control (no AED) suggested valproate (arm sample size (N)=160, odds ratio (OR)=8.63, 95% credible interval (CrI): 3.01-25.74) and the combination carbamazepine, phenobarbital, and valproate (N=3, OR=17.31, CrI: 1.02-434.50) were statistically significantly associated with more children experiencing cognitive developmental delay. A NMA was conducted on autism including 5 cohort studies, 2,551 children, 11 AEDs and control; only oxcarbazepine (N=321, OR=13.51, CrI: 1.28-221.40), valproate (N=485, OR=17.29, 95% CrI: 2.40-217.60), lamotrigine (N=745, OR=8.88, CrI: 1.28-112.00), and lamotrigine+valproate (N=6, OR=132.70, CrI: 7.41-3851.00) were associated with a

significantly greater risk of autism compared with control. Psychomotor developmental delay was the largest NMA of secondary outcomes (11 cohort studies, 1,145 children, 17 AEDs and control): valproate (N=137, OR=4.16, CrI: 2.04-8.75) and carbamazepine+phenobarbital+valproate (N=3, OR=19.12, CrI: 1.49-337.50) were associated with a significantly greater risk of psychomotor delay compared with control. **Conclusions**: Across all outcomes, valproate alone or combined with another AED is associated with the greatest risk, whereas oxcarbazepine and lamotrigine were associated with increased occurrence of autism. Counselling is advised for women considering pregnancy to tailor the safest regimen.

Registration: PROSPERO database (CRD42014008925).

Keywords: multiple treatment meta-analysis, knowledge synthesis, epilepsy, pregnancy, infants, developmental delay.

ARTICLE SUMMARY

Strengths and limitations of this study

- 27 cohort studies involving 4,841 children of women who took AEDs were included in this systematic review. More evidence from long-term follow-up studies is required.
- This study was the first that compared and ranked the safety of AEDs, including comparative safety of treatments that have not been directly compared.

- Across all neurological outcomes, valproate alone or combined with another AED is associated with the greatest risk.
- Oxcarbazepine and lamotrigine were associated with increased occurrence of autism.

INTRODUCTION

Anti-epileptic drugs (AEDs) are used by pregnant women for various conditions, such as epilepsy, pain syndromes, psychiatric disorders, and chronic migraine.[1] AED use during pregnancy is associated with risks to the fetus, as these drugs can cross the placenta or may be transferred to the infant through breastfeeding and may be associated with adverse neurodevelopment outcomes.[2-4] Two systematic reviews examined the association between AED exposure and neurodevelopment *in utero*, and reported that exposure to valproate was linked to significantly lower IQ scores and poorer overall neurodevelopmental outcomes in the children of women who used these medications.[5 6] No statistically significant associations were found between neurodevelopment and exposure to other AEDs such as carbamazepine, lamotrigine, or phenytoin.[5-8] However, there is a lack of sufficiently powered studies to assess the impact of AEDs on neurodevelopment in children of women exposed to these agents, especially for newer generation drugs, thus highlighting the need for a systematic review.[9 10]

The aim of this study was to compare the safety of AEDs and assess their impact on neurodevelopment in infants and children exposed *in-utero* or during breastfeeding, employing a systematic review and network meta-analysis (NMA).

METHODS

The methods are briefly described here; details can be found in the published protocol (Additional File 1).[11] This study was registered with PROSPERO (CRD42014008925). We followed the ISPOR[12] guidelines for our NMA, and reported our findings using the PRISMA extension for NMA (Additional File 2).[13]

Eligibility criteria

All randomized clinical trials (RCTs), quasi-RCTs, and observational studies were eligible. Included studies assessed infants or children ≤12 years of age whose mothers consumed AEDs during pregnancy and/or while breastfeeding. Both mono- and poly-therapy AEDs were eligible, including first-generation (i.e., carbamazepine, clobazam, clonazepam, ethosuximide, phenobarbital, phenytoin, primidone, valproate) and newer-generation (i.e., marketed after 1990 including: gabapentin, lamotrigine, levetiracetam, oxcarbazepine, topiramate, vigabatrin), with no restrictions on AED dosage. Placebo, no AED, other AEDs alone or in combination, were considered as comparators. Duplicate studies that used the same registry or population sample (i.e., companion studies) were used for supplementary information only. No language or other restrictions were imposed.

The primary neurological outcomes were cognitive developmental delay and autism/dyspraxia, and the secondary outcomes included attention deficit hyperactivity disorder (ADHD), language delay, neonatal seizures, psychomotor developmental delay, and social impairment (outcome measures and diagnostic scales used are provided in Additional File 3: Appendix A). Our initial intention was to evaluate all safety outcomes in infants and children who were exposed to AEDs *in-utero* or during breastfeeding in one publication. However, given the breadth of evidence we identified, we report results

related to risk of major congenital malformations, birth, and prenatal outcomes in a second paper (paper in preparation).

Searching, screening, abstraction, appraisal of methodological quality

We searched MEDLINE, EMBASE, and the Cochrane Central Register of Controlled Trials up to December 15, 2015, and identified additional studies from scanning references and contacting authors. Unpublished studies were sought by searching clinical trial registries and conference abstracts. After a calibration exercise, titles/abstracts and full-text papers were screened by two reviewers independently (further details reported in Additional File 3: Appendix B). Conflicts were resolved through discussion. The same approach was used for data abstraction and appraisal of methodological quality.

Observational studies were only identified, and their methodological quality was appraised with the Newcastle-Ottawa Scale (Additional File 3: Appendix C).[14] For each outcome with ≥10 studies and treatment comparisons with different total numbers of patients, the comparison-adjusted funnel plot was used to assess reporting bias,[15] where the overall treatment effect for each comparison was estimated under the fixed-effect meta-analysis model.

Synthesis of included studies

We used the odds ratio (OR) for each dichotomous outcome, and outcome data were pooled using hierarchical models and the Markov Chain Monte Carlo sampling method in a Bayesian framework. To account for anticipated methodological and clinical heterogeneity across studies, and to achieve the highest generalizability in the meta-analytical treatment effects, we applied a random-effects model.[16]

For connected evidence networks, we applied a random-effects NMA hierarchical model.[17] The review team pre-specified the network nodes. Due to the complexity of the data and the studies' underreporting, differences in drug dosages could not be accounted for, and it was assumed that different dosages of the same AED were equally effective. When a study reported multiple dosages for the same treatment, we combined the data for this treatment. We assessed the transitivity assumption for each outcome a priori using the treatment effect modifiers: age, baseline risk, treatment indication, timing, and methodological quality. The mean of each continuous effect modifier and the mode of each categorical effect modifier for each pairwise comparison were presented in tables for each outcome.[18] The consistency assumption was evaluated for the entire network of each outcome using the design-by-treatment interaction model.[19] If inconsistency was identified, further examination for local inconsistency in parts of the network was completed using the loop-specific method.[20 21] Common within-network between-study variance (τ^2) across treatment comparisons was assumed in the conventional meta-analysis, NMA, and design-by-treatment interaction model, so that treatment comparisons including a single study can borrow strength from the remaining network. This assumption was clinically reasonable, as the treatments included were of the same nature. In the loop-specific approach, common within-loop τ^2 was assumed.

For cognitive developmental delay and autism/dyspraxia outcomes, network meta-regression analyses for maternal age and baseline risk (i.e., using the control group) were conducted, when at least 10 studies provided relevant information, assuming a common fixed coefficient across treatment comparisons. Sensitivity analyses for cognitive developmental delay and autism/dyspraxia outcomes were performed for studies with the

treatment indication of epilepsy, large study size (i.e., >300), maternal alcohol intake, maternal tobacco use, only first-generation AEDs, and higher methodological quality for the two items of the Newcastle-Ottawa Scale that had the highest percentage of low methodological quality (adequacy of follow-up of cohorts and comparability of cohorts items for cohort studies). Severity of epilepsy, which may be a risk factor variable as the more severe the epilepsy the more necessary AED medications for the mother,[22] was not evaluated in our analyses since this was not commonly reported. For autism/dyspraxia, a sensitivity analysis on maternal IQ/psychiatric history was additionally conducted. We measured the goodness of fit using the posterior mean of the residual deviance, the degree of between-study heterogeneity, and the deviance information criterion. In a well-fitting model the posterior mean residual deviance should be close to the number of data points.[23 24] A difference of 3 units in the deviance information criterion was considered important and the lowest value of the deviance information criterion corresponded to the model with the best fit.[23 24]

All analyses were conducted in OpenBUGS[25] assuming non-informative priors for all model parameters and a half normal prior distribution for the between-study standard deviation ($\tau \sim N(0,1), \tau > 0$). The first 10,000 iterations were discarded and then 100,000 simulations were run with thinning of 10 values. Convergence was checked by visual inspection of the evaluation of the mixing of two chains. The median and 95% CrI were calculated for each parameter value, since medians are not overly influenced by outliers. The network command[26] was used to apply the design-by-treatment interaction model.

For NMA estimates, a 95% predictive interval (PrI) is also reported to capture the magnitude of the between-study variance (τ^2) and present the interval within which the

treatment effect of a future study is expected to lie.[27 28] The estimated safety of the included AED medications was ranked using the surface under the cumulative ranking (SUCRA) curve.[29] The larger the SUCRA for a treatment, the higher its safety rank among all the available treatment options. A steep gradient in the cumulative ranking curve suggests that the corresponding treatment is most likely the safest. SUCRA curve values are presented along with 95% CrIs to capture the uncertainty in the parameter values.

RESULTS

Literature search and included studies

Our literature search identified 5,305 titles and abstracts, which after the screening process yielded 642 articles potentially relevant for inclusion (Figure 1). After full-text review, 93 studies fulfilled eligibility criteria along with 17 studies identified through supplemental methods. Of the 110 total eligible studies in the complete review, 27 articles with nine companion reports or potentially overlapping studies included one or more relevant neurological outcomes (Additional File 3: Appendices D, E). Two of the included studies were conference abstracts with usable data, nine were non-English publications, and four studies, not captured in the original literature search, were identified through reference scanning. A table with the key excluded studies and a rationale for their exclusion is presented in Additional File 3: Appendix F.

Study and patient characteristics

We included 27 cohort studies (4,841 total patients) published between 1989 and 2015 (Table 1; Additional File 3: Appendices G, H). The number of patients included in each study ranged from 23 to 2,011 (median 69), and the number of arms compared in each study ranged from two to 12. Most studies (78%) were published after 2000, more than half of the studies (67%) included fewer than 100 patients, and 13 studies (48%) included a control group of pregnant/breastfeeding women with epilepsy who did not receive AEDs. The mean age of women ranged from 24 to 32 years. About half of the studies were funded through government/public research funding (52%).

Methodological quality results

Twenty-seven observational studies were appraised using the Newcastle Ottawa Scale (Additional File 3: Appendix I). All studies selected the non-exposed cohort from the same community as the exposed cohort, 25 (93%) included a representative (or somewhat representative) sample, 25 (93%) assessed outcomes independently, blindly or via a record linkage (e.g., identified through database records), and 21 (78%) ascertained exposure via secured records (e.g., database records) or structured interviews. No evidence for reporting bias was identified by the visual inspection of the comparison-adjusted funnel plots (Additional File 3: Appendix J).

Statistical analysis results

No important concerns were raised regarding the violation of the transitivity assumption when mean maternal age, mean baseline risk, treatment indication, and timing were assessed (Additional File 3: Appendix K). However, the average methodological quality appraisal across treatment comparisons varied across treatment comparisons. The evaluation of the consistency assumption using the design-by-treatment interaction model suggested that there was no evidence of significant inconsistency across all outcomes (Additional File 3: Appendix K).

In the following sections, we present the overall results of the NMA analyses for each outcome, while the SUCRA curve values from all outcomes are presented in Additional File 3: Appendix L and depicted in a rank-heat plot (http://rh.ktss.ca/)[30] in Additional File 3: Appendix M.

Cognitive developmental delay

The NMA for cognitive developmental delay (definitions in Additional File 3: Appendix A) included ten cohort studies, 748 children, and examined 13 AEDs plus control (i.e., no

exposure to AEDs). One study included children exposed to AEDs both *in-utero* and through breastfeeding, and nine included children exposed to AEDs *in-utero*. Overall, 6% of the treatment comparisons in the network reached statistical significance (Figure 2a; Additional File 3: Appendices K, N). Valproate (OR=8.63, 95% CrI: 3.01-25.74) and carbamazepine+phenobarbital+valproate (OR=17.31, 95% CrI: 1.02-434.50) were statistically significantly associated with greater risk in children experiencing developmental delay compared with control (Figure 3).

Restricting the NMA to 9 cohort studies including 725 offspring of women only with epilepsy as their treatment indication, comparing 13 AEDs plus control produced results that were generally in agreement with the overall results. The same was observed in a network meta-regression model of baseline risk for offspring of women with epilepsy who were not exposed to AEDs (estimated regression coefficient on OR scale: 0.94, 95% CrI: 0.67-2.19; τ^2=0.17, 95% CrI: 0.00-1.38; residual deviance= 40.26, data points= 42, deviance information criterion= 71). Restricting the analysis to 6 cohort studies including 480 children comparing 11 first-generation AEDs, we found that valproate was statistically significantly more harmful than control, phenytoin, and carbamazepine, yet carbamazepine+phenobarbital+valproate was no longer statistically significant versus control. The results were no longer statistically significant when restricted to two studies of 319 offspring of women with a history of alcohol and tobacco use comparing 3 AEDs and control. This result was consistent in sensitivity analyses including only higher methodological quality studies in the 'comparability of cohorts' item on the Newcastle-Ottawa Scale (2 studies, 181 children, 3 AEDs plus control) and the 'adequacy of follow-up of cohorts' (4 studies, 283 children, 11 AEDs plus control).

Autism/dyspraxia

The NMA on autism/dyspraxia (definitions in Additional File 3: Appendix A) included five cohort studies, 2,551 children exposed *in utero*, and examined 11 AEDs plus control (i.e., no AED exposure). Overall, 9% of the treatment comparisons in the network reached statistical significance (Figure 2b; Additional File 3: Appendices K, N). Compared with control, only valproate (OR=17.29, 95% CrI: 2.40-217.60), oxcarbazepine (OR= 13.51, 95% CrI: 1.28-221.40), lamotrigine (OR= 8.88, 95% CrI: 1.28-112.00), and lamotrigine+valproate (OR=132.70, 95% CrI: 7.41-3851.00) were significantly associated with increased occurrence of autism/dyspraxia (Figure 3).

Restricting the NMA to studies including only women with epilepsy as their treatment indication produced results that were generally in agreement with the overall results, except that oxcarbazepine was no longer in the network (4 cohort studies, 540 children, 9 AEDs plus control). Two cohort studies of 404 offspring of women with a history of tobacco use compared 3 AEDs and control and found similar results except that oxcarbazepine and lamotrigine+valproate were no longer in the network. The results were in agreement in sensitivity analyses including only higher methodological quality studies in the 'comparability of cohorts' item on the Newcastle-Ottawa Scale (4 studies, 2395 children, 11 AEDs plus control) and the 'adequacy of follow-up of cohorts' (3 studies, 2244 children, 9 AEDs plus control), except that lamotrigine was no longer statistically significant than control for the latter.

Neonatal Seizure

One cohort study included 72 children who were exposed to AEDs *in-utero* as well as through breastfeeding reported on the incidence of neonatal seizures. The study compared

valproate against lamotrigine and found no statistically significant difference in neonatal seizures between the two drugs (OR=0.18, 95% CrI: 0.01-3.70).

Psychomotor developmental delay

The NMA on psychomotor developmental delay (definitions in Additional File 3: Appendix A) included 11 cohort studies, 1,145 children exposed *in utero*, and examined 17 AEDs plus control (i.e., no AED exposure). Overall, 4% of treatment comparisons in the network reached statistical significance (Figure 2c; Additional File 3: Appendices K, N). Valproate (OR=4.16, 95% CrI: 2.04-8.75) and carbamazepine+phenobarbital+valproate (OR=19.12, 95% CrI: 1.49-337.50) were statistically significantly more harmful than control (Figure 3).

Language delay

The NMA on language delay (definitions in Additional File 3: Appendix A) included five cohort studies, 509 children, and examined four AEDs plus control (i.e., no AED exposure; Figure 2d; Additional File 3: Appendices K, N). One study included children exposed to AEDs *in-utero* and through breastfeeding, and four included children exposed to AEDs *in-utero*. Compared with control, valproate was the only treatment significantly associated with increased risk of language delay (OR=7.95, 95% CrI: 1.50-49.13; Figure 3).

Attention deficit hyperactivity disorder

The NMA on ADHD (definitions in Additional File 3: Appendix A) included four cohort studies, 750 children, and examined five AEDs plus control (i.e., no AED exposure). One study included children exposed to AEDs *in-utero* and through breastfeeding, while three studies included children exposed to AEDs *in-utero*. None of the treatment comparisons reached statistical significance (Figure 2e; Additional File 3: Appendices K, N).

Social Impairment

One cohort study included 422 children exposed to AEDs *in-utero* as well as through breastfeeding. The children were exposed to carbamazepine (n=48), lamotrigine (n=71), valproate (n=27) and control (n=278). No significant differences in social impairment were identified.[31]

DISCUSSION

Our results suggest that AEDs generally pose a risk for infants and children exposed *in-utero* or during breastfeeding. Valproate was statistically significantly associated with more children experiencing autism/dyspraxia, language, cognitive and psychomotor developmental delays versus children who were not exposed to AEDs. Oxcarbazepine, lamotrigine and lamotrigine+valproate were associated with increased occurrence of autism/dyspraxia, whereas for the cognitive developmental delay and psychomotor developmental delay outcomes, children exposed to the combination of carbamazepine, phenobarbital, and valproate were at a greater risk of harm than those who were not exposed to AEDs. However, due to the lack of data identified in these studies, we were unable to consider a number of factors, such as anticonvulsant dosing, severity of epilepsy, duration of exposure, serum concentrations of exposure, mother's IQ/education, which may all influence outcomes, and hence these results should be interpreted with caution. In addition, our subsequent analyses may be underpowered due to missing data (e.g., maternal age is not reported in 17 of the 27 studies, alcohol use is not reported in 23 of 27 studies, tobacco use is not reported in 22 of 27 studies, and epileptic control group was not included in 14 of 27 studies).

NMA is a particularly useful tool for decision-makers because it allows the ranking of treatments for each outcome. However, the results of our SUCRA curves should be interpreted with caution, especially due to the small number of studies and children included in each NMA, which is also reflected in the high uncertainty around the SUCRA values (Additional File 3: Appendix L).[32] The probability that a top AED is actually among

the worst one is likely high,[32] as the SUCRA findings were unstable with overlapping uncertainty intervals overlap.

Our results are consistent with a longitudinal study of 311 children that found exposure to lamotrigine was associated with significantly higher IQ scores and verbal function at six years of age compared to children exposed to valproate (Additional File 3: Appendix F).[8] As indicated in Additional File 3: Appendix F, we were unable to include this study because the outcome was reported as a continuous measure, where we focused on dichotomous outcomes to ease interpretation. Our results are supported by findings from a cohort study in the UK, which found that children exposed to levetiracetam were not at increased risk for delayed development compared to unexposed children (Additional File 3: Appendix F).[33] As indicated in Additional File 3: Appendix F, we were unable to include this study due to the same reason as above. A NMA of 195 RCTs and 28,013 patients (including both males and females) showed that gabapentin and levetiracetam showed the best tolerability profile compared with other AEDs, whereas oxcarbazepine and topiramate had a higher withdrawal rate, and lamotrigine an intermediate withdrawal rate.[34]

Across all outcomes, valproate alone or combined with another AED (even with a newer-generation agent, e.g., lamotrigine) was associated with the greatest risk. Similarly, two previous systematic reviews that did not conduct a NMA found valproate was associated with significantly lower IQ scores and poorer overall neurodevelopmental outcomes when compared to an unexposed control group.[5 6] Also consistent with our results, a 2014 Cochrane review (with a meta-analysis of 10 studies) concluded that AED polytherapy led to poorer developmental outcomes and IQ compared to healthy controls, epileptic controls, and unspecified monotherapy.[5] This Cochrane review also concluded that insufficient data

exist for newer anti-epileptic drugs. These risks must be balanced with the need to control seizure activity in pregnancy and thus informed decision making by patients and clinicians is critical.

Strengths of our study include a comprehensive systematic review methodology that followed the Cochrane Handbook[35] and ISPOR[12] guidelines, and reported using the PRISMA extension for NMA.[13] To the best of our knowledge, our study was the first that compared and ranked the safety of AEDs. We evaluated the comparative safety of treatments that have not been directly compared head-to-head before. In addition, we calculated predictive intervals, which account for between-study variation and provide a predicted range for the treatment effect estimate, should a future study be conducted. On average, the predictive intervals suggested that our results are robust.

Our systematic review has a few limitations worth noting. First, when multiple doses were reported for the same treatment, we lumped the dosages because this information was not consistently reported across the included studies. However, this is common for cohort studies, which report on a number of different types of exposures amongst patients. Second, several polytherapies had high SUCRA estimates but with very wide CrIs, which is due to the small number of studies included for each drug combination with underpowered sample sizes. Evidence suggests that ranking probabilities for a treatment of being the best may be biased toward the treatments with the smallest number of studies, which may have influenced our SUCRA results.[32 36] As such, the effect sizes need to be taken into account when considering the SUCRA values. Third, due to the absence of evidence from RCTs, our conclusions were based on evidence from observational studies only, and inherent biases because of confounding and shortcomings of these studies may have impacted our findings.

For example, the included studies often failed to report important confounding variables, such as family history of autism, ADHD, and maternal IQ, making it impossible for us to control these variables through subgroup analysis and meta-regression. Recent research papers have explored methods to incorporate non-randomized with randomized evidence in a NMA and have highlighted the need to carefully explore the level of confidence in the non-randomized evidence.[37 38] However, the use of observational studies allows the assessment of the safety profile of AED treatments and offers the opportunity to evaluate effects in pregnancy.[39] Future large-scale observational studies are needed to allow the evaluation of rare adverse events that otherwise cannot be adequately evaluated in RCTs, especially during pregnancy. Fourth, although no intransitivity for most treatment effect modifiers assessed was evident, there was an imbalance in the methodological study quality appraisal across treatment comparisons and most outcomes, which may impact our results. However, the assessment of consistency suggested no disagreement between the different sources of evidence in the network. Fifth, although the tendency towards publication bias is greater with observational studies than with randomized trials,[40] the assessment of publication bias and small-study effects using adjusted funnel plots suggested no evidence for their prevalence. Also, the majority of the included studies in this review compared multiple treatments inducing correlations in each funnel plot, which may mask asymmetry. Although we plotted data points corresponding to the study-specific basic parameters to reduce correlations, this issue may still exist. Sixth, we were unable to conduct sub-group analysis by type of exposure (breastfeeding versus *in utero*) due to the small number of studies included in the NMA and due to the poor reporting; 22 studies did

not report whether exposure was also in breastfeeding (additional to *in utero*). Hence, we included all studies in the analysis irrespective of the type of exposure.

More evidence from long-term follow-up studies is required to further delineate neurodevelopmental risks in children. Registries should aim to include a suitable control group and collect information on potential confounders, such as alcohol and tobacco use, allowing researchers to identify the safest agents for different patient-level covariates, and enhance decision-making for healthcare providers and patients. An individual patient data NMA would likely provide further clarity to the field, which allows the tailoring of management to specific patient characteristics.[41]

CONCLUSION

Across all outcomes, valproate alone or combined with another AED was associated with the greatest risk, whereas oxcarbazepine and lamotrigine were associated with increased occurrence of autism. Counselling is advised for women considering pregnancy to tailor the safest regimen.

LIST OF ABBREVIATIONS

AEDs: Anti-epileptic drugs; CrI: Credible interval; NMA: Network Meta-analysis; OR: Odds ratio; PrI: Predictive interval; SUCRA curve: Surface under the cumulative ranking curve

ADDITIONAL FILES

Additional File 1: Protocol

Additional File 2: PRISMA NMA Checklist

Additional File 3: Supplementary Online Content (Appendices A-O)

Appendix A. Outcome measures and diagnostic scales used in analysis

Appendix B. Additional Information on Methods

Appendix C. Newcastle-Ottawa Scale scoring guide

Appendix D. List of included studies

Appendix E. Additional information on search results

Appendix F. Key Excluded Studies

Appendix G. Table of Individual Study Characteristics

Appendix H. Table of Patient Characteristics

Appendix I. Methodological quality of observational studies – Newcastle Ottawa Scale

Appendix J. Comparison-adjusted funnel plots

Appendix K. Statistically significant network meta-analysis results along with meta-analysis results, transitivity, and inconsistency assessments

Appendix L. Frequencies, events and samples sizes, SUCRA values, and total group risks per treatment and outcome

Appendix M. Rank-heat plot of cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcomes

Appendix N. Number of studies and treatments per outcome

FIGURE LEGENDS

Figure 1. Study flow diagram

Figure 2. Network diagrams for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcomes

Each treatment node is weighted according to the number of patients that have received the particular treatment, and each edge is weighted according to the number of studies comparing the treatments it connects.

Abbreviations: carbam - carbamazepine, clobaz - clobazam, clonaz - clonazepam, ethos - ethosuximide, gabap - gabapentin, lamot - lamotrigine, levet - levetiracetam, oxcar - oxcarbazepine, pheno - phenobarbital, pheny - phenytoin, primid - primidone, topir - topiramate, valpro - valproate, vigab – vigabatrin

Figure 3. Forest plots for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcomes

DECLARATIONS

CONTRIBUTORS

AAV analysed the data, interpreted the results, and drafted the manuscript. ACT and SES conceived and designed the study, helped obtain funding, interpreted the results, and helped write sections of the manuscript. PR and EC coordinated the review, screened citations and full-text articles, abstracted data, appraised quality, resolved discrepancies, contacted authors, and edited the manuscript. CS provided methodological support and screened citations and full-text articles and edited the manuscript. RK, ER, FY, JDS, KT, and HM screened citations and full-text articles, abstracted data, and/or appraised quality. BH, BRH and YF helped conceive the study and edited the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

We thank Dr. David Moher for providing his feedback on our protocol. We thank Dr. Laure Perrier for conducting the literature searches, Becky Skidmore for peer-reviewing the MEDLINE search, and Alissa Epworth for obtaining the full-text articles. We thank Alistair Scott, Wing Hui, and Geetha Sanmugalingham for screening some of the citations and/or abstracting some of the data for a few of the included studies, Misty Pratt and Mona Ghannad for helping scan reference lists, and Ana Guzman, Susan Le, and Inthuja Selvaratnam for contacting authors and formatting the manuscript.

FUNDING

This systematic review was funded by the Canadian Institutes for Health Research/Drug Safety and Effectiveness Network (CIHR/DSEN). AAV is funded by the Banting Postdoctoral Fellowship Program from the CIHR. SES is funded by a Tier 1 Canada Research Chair in

Knowledge Translation. BH is funded by a CIHR/DSEN New Investigator Award in Knowledge Synthesis. BRH receives funding from the Alberta Heritage Foundation for Medical Research. ACT is funded by a Tier 2 Canada Research Chair in Knowledge Synthesis. The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

COMPETING INTERESTS

None declared.

ETHICS APPROVAL

Not applicable.

PROVENANCE AND PEER REVIEW

Not commissioned; externally peer reviewed.

DATA SHARING STATEMENT

All datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

OPEN ACCESS

REFERENCES

1. Spina E, Perugi G. Antiepileptic drugs: indications other than epilepsy. *Epileptic Disord*. 2004;6(2):57-75.
2. Harden CL, Pennell PB, Koppel BS, et al. Management issues for women with epilepsy--focus on pregnancy (an evidence-based review): III. vitamin K, folic acid, blood levels, and breast-feeding: report of the quality standards subcommittee and therapeutics and technology assessment subcommittee of the American Academy of Neurology and the American Epilepsy Society. *Epilepsia*. 2009;50(5):1247-55.
3. Samren EB, van Duijn CM, Koch S, et al. Maternal use of antiepileptic drugs and the risk of major congenital malformations: a joint European prospective study of human teratogenesis associated with maternal epilepsy. *Epilepsia*. 1997;38(9):981-90.
4. Meador K, Reynolds MW, Crean S, Fahrbach K, Probst C. Pregnancy outcomes in women with epilepsy: a systematic review and meta-analysis of published pregnancy registries and cohorts. *Epilepsy Res*. 2008;81(1):1-13.
5. Bromley R, Weston J, Adab N, et al. Treatment for epilepsy in pregnancy: neurodevelopmental outcomes in the child. *The Cochrane database of systematic reviews*. 2014(10):Cd010236.
6. Banach R, Boskovic R, Einarson T, Koren G. Long-term developmental outcome of children of women with epilepsy, unexposed or exposed prenatally to antiepileptic drugs: a meta-analysis of cohort studies. *Drug Saf*. 2010;33(1):73-9.
7. Christensen J, Gronborg TK, Sorensen MJ, et al. Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. *JAMA*. 2013;309(16):1696-703.
8. Meador KJ, Baker GA, Browning N, et al. Fetal antiepileptic drug exposure and cognitive outcomes at age 6 years (NEAD study): a prospective observational study. *Lancet Neurol*. 2013;12(3):244-52.
9. Wlodarczyk BJ, Palacios AM, George TM, Finnell RH. Antiepileptic drugs and pregnancy outcomes. *Am J Med Genet A*. 2012;158a(8):2071-90.
10. Velez-Ruiz NJ, Meador KJ. Neurodevelopmental effects of fetal antiepileptic drug exposure. *Drug Saf*. 2015;38(3):271-8.
11. Tricco AC, Cogo E, Veroniki AA, et al. Comparative safety of anti-epileptic drugs among infants and children exposed in utero or during breastfeeding: protocol for a systematic review and network meta-analysis. *Syst Rev*. 2014;3:68.
12. Jansen JP, Trikalinos T, Cappelleri JC, et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. *Value Health*. 2014;17(2):157-73.
13. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med*. 2015;162(11):777-84.
14. Wells G, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses Ottawa, Canada: Ottawa Hospital Research Institute; 2000. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
15. Chaimani A, Higgins JP, Mavridis D, Spyridonos P, Salanti G. Graphical tools for network meta-analysis in STATA. *PLoS One*. 2013;8(10):e76654.
16. Furukawa TA, Guyatt GH, Griffith LE. Can we individualize the 'number needed to treat'? An empirical study of summary effect measures in meta-analyses. *Int J Epidemiol*. 2002;31(1):72-6.

17. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med*. 2004;23(20):3105-24.
18. Jansen JP, Naci H. Is network meta-analysis as valid as standard pairwise meta-analysis? It all depends on the distribution of effect modifiers. *BMC Med*. 2013;11:159.
19. White IR, Barrett JK, Jackson D, Higgins JP. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Res Synth Methods*. 2012;3(2):111-25.
20. Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ*. 2003;326(7387):472.
21. Veroniki AA, Vasiliadis HS, Higgins JP, Salanti G. Evaluation of inconsistency in networks of interventions. *Int J Epidemiol*. 2013;42(1):332-45.
22. Dalessio DJ. Seizure Disorders and Pregnancy. *N Engl J Med*. 1985;312(9):559-63.
23. Welton NJ, Sutton AJ, Cooper N, Abrams KR, Ades A. Evidence synthesis for decision making in healthcare. New York: Wiley; 2012.
24. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc Ser B Stat Methodol*. 2002;64(4):583-639.
25. Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: Evolution, critique and future directions. *Stat Med*. 2009;28(25):3049-67.
26. Palmer T, Sterne J. Meta-Analysis in Stata: An Updated Collection from the Stata Journal. White I, editor. Texas: Stata Press; 2016.
27. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ*. 2011;342:d549.
28. Higgins JP, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *J R Stat Soc Ser A Stat Soc*. 2009;172(1):137-59.
29. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol*. 2011;64(2):163-71.
30. Veroniki AA, Straus SE, Fyraridis A, Tricco AC. The rank-heat plot is a novel way to present the results from a network meta-analysis including multiple outcomes. *J Clin Epidemiol*. 2016.
31. Veiby G, Engelsen BA, Gilhus NE. Early child development and exposure to antiepileptic drugs prenatally and through breastfeeding: a prospective cohort study on children of women with epilepsy. *JAMA neurology*. 2013;70(11):1367-74.
32. Trinquart L, Attiche N, Bafeta A, Porcher R, Ravaud P. Uncertainty in Treatment Rankings: Reanalysis of Network Meta-analyses of Randomized Trials. *Ann Intern Med*. 2016;164(10):666-73.
33. Shallcross R, Bromley RL, Irwin B, Bonnett LJ, Morrow J, Baker GA. Child development following in utero exposure: levetiracetam vs sodium valproate. *Neurology*. 2011;76(4):383-9.
34. Zaccara G, Giovannelli F, Giorgi FS, Franco V, Gasparini S, Benedetto U. Tolerability of new antiepileptic drugs: a network meta-analysis. *Eur J Clin Pharmacol*. 2017.
35. Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. 5.1.0 ed: The Cochrane Collaboration; 2009.
36. Kibret T, Richer D, Beyene J. Bias in identification of the best treatment in a Bayesian network meta-analysis for binary outcome: a simulation study. *Clin Epidemiol*. 2014;6:451-60.
37. Efthimiou O, Mavridis D, Debray TP, et al. Combining randomized and non-randomized evidence in network meta-analysis. *Stat Med*. 2017.
38. Schmitz S, Adams R, Walsh C. Incorporating data from various trial designs into a mixed treatment comparison model. *Stat Med*. 2013;32(17):2935-49.
39. Cameron C, Fireman B, Hutton B, et al. Network meta-analysis incorporating randomized controlled trials and non-randomized comparative cohort studies for assessing the safety and effectiveness of medical treatments: challenges and opportunities. *Syst Rev*. 2015;4:147.

40. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet*. 1991;337(8746):867-72.
41. Veroniki AA, Straus SE, Soobiah C, Elliott MJ, Tricco AC. A scoping review of indirect comparison methods and applications using individual patient data. *BMC Med Res Methodol*. 2016;16(1):47.

Table 1. Summary characteristics of included studies

Table 1. Summary Characteristics of included studies		
Study/Patient Characteristic	**# of Studies (n=27)**	**% of Studies**
Year of publication		
1980-1989	1	3.70
1990-1999	6	22.22
2000-2009	5	18.52
2010-2015	14	51.85
NR	1	3.70
Continent (of country of study conduct)		
Europe	18	66.67
North America	5	18.52
Asia	1	3.70
Australia	2	7.41
Trans-Continental	1	3.70
Study design		
Observational cohort	27	100.00
Case-control	0	0.00
Randomized clinical trial	0	0.00
Registry study		
Yes	10	37.04
No	17	62.96
Sample size		
0-99	17	62.96
100-299	8	29.63
300-499	1	3.70
500-699	0	0.00
700-999	0	0.00
1000+	1	3.70
Number of interventions		
2	4	14.81
3	5	18.52
4	8	29.63
5-7	7	25.93
8-10	2	7.41
11+	1	3.70
Outcomes[*, †]		
Cognitive Developmental Delay	11	40.74
Autism/Dyspraxia	5	18.52
Language Delay	5	18.52

Table 1. Summary Characteristics of included studies

Study/Patient Characteristic	# of Studies (n=27)	% of Studies
ADHD	4	14.81
Psychomotor Developmental Delay	11	40.74
Neonatal Seizures	2	7.41
Social Impairment	1	3.70
Funding		
Public	14	51.85
Private	0	0.00
Mixed public and private	4	14.81
NR/Unclear	9	33.33
Treatment indication		
Epilepsy	21	77.78
Mixed indications‡	0	0.00
Not reported	6	22.22
Epileptic control group§		
Yes	13	48.15
No/NR/NA	14	51.85
Mean maternal age		
24-26 y	2	7.41
27-29 y	5	18.52
30-32 y	3	11.11
Not reported	17	62.96
AED exposure during pregnancy		
Reported as during 1st trimester	6	22.22
Reported as any time during pregnancy	6	22.22
Not reported	15	55.56
Alcohol use during pregnancy		
Yes	4	14.81
NR	23	85.19
Tobacco use during pregnancy		
Yes	5	18.52
NR	22	81.48

Abbreviations: ADHD - Attention Deficit Hyperactivity Disorder; AED - anti-epileptic drug(s); NA - Not applicable; NR - Not reported

* Values in this category do not match totals as some studies report more than one outcome

† Percentage of total number of included studies (n=27)

‡ Includes individuals taking AEDs for psychiatric disorders, migraine, and neuropathic/neurological pain

§ Consisted of women with Epilepsy who did not take AEDs during pregnancy

626

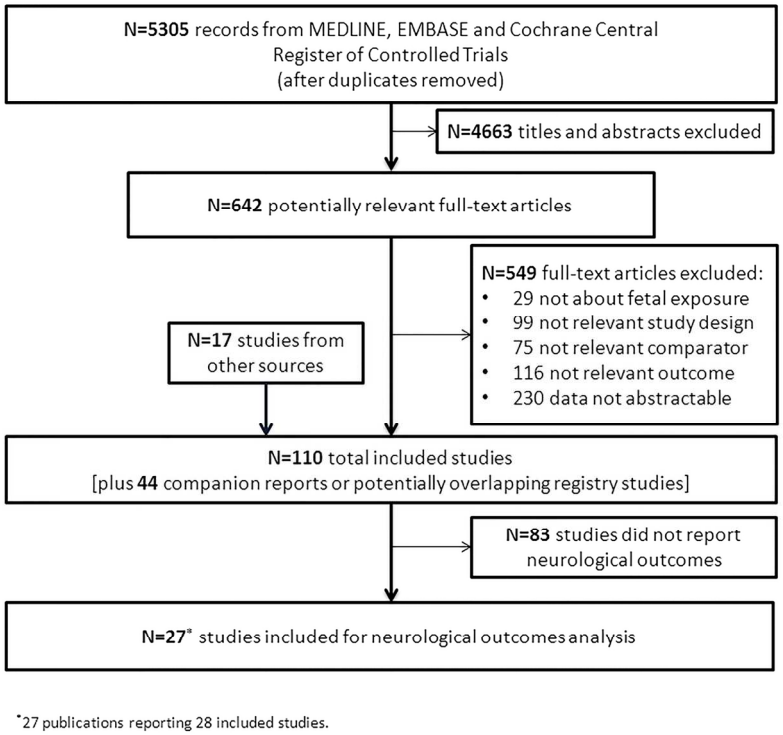


Figure 1. Study flow diagram

171x128mm (300 x 300 DPI)

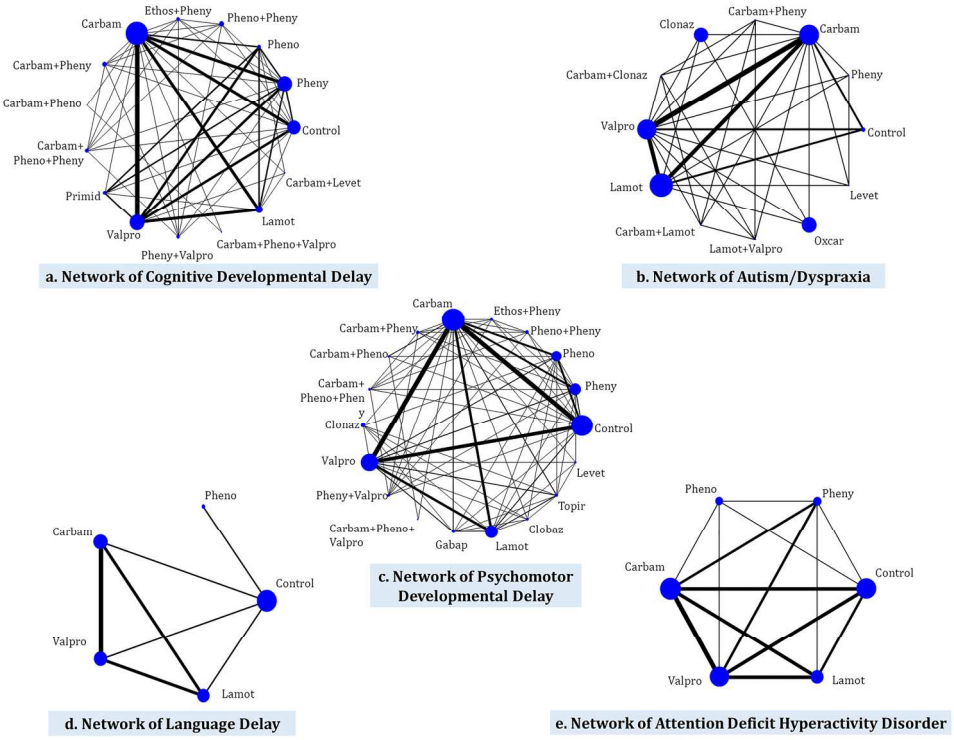


Figure 2. Network diagrams for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcomes.

‼ ┼ ‼ ┼ Each treatment node is weighted according to the number of patients that have received the particular treatment, and each edge is weighted according to the number of studies comparing the treatments it connects.‼ ┼ ‼ ┼

Abbreviations: carbam - carbamazepine, clobaz - clobazam, clonaz - clonazepam, ethos - ethosuximide, gabap - gabapentin, lamot - lamotrigine, levet - levetiracetam, oxcar - oxcarbazepine, pheno - phenobarbital, pheny - phenytoin, primid - primidone, topir - topiramate, valpro - valproate, vigab – vigabatrin‼ ┼

171x128mm (300 x 300 DPI)

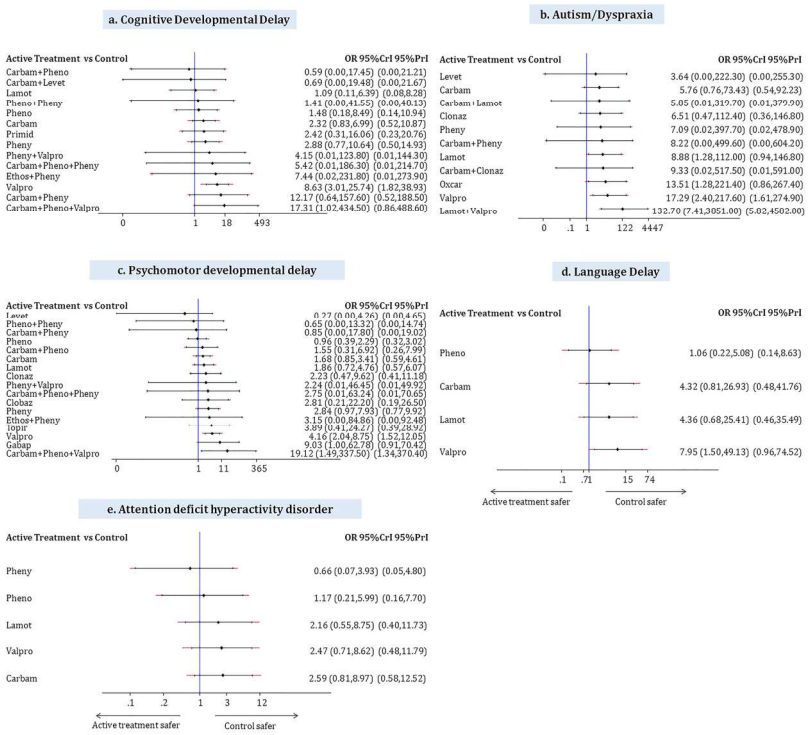


Figure 3. Forest plots for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcomes

171x128mm (300 x 300 DPI)

Tricco et al. Systematic Reviews 2014, 3:68
http://www.systematicreviewsjournal.com/content/3/1/68

SYSTEMATIC REVIEWS

PROTOCOL Open Access

Comparative safety of anti-epileptic drugs among infants and children exposed *in utero* or during breastfeeding: protocol for a systematic review and network meta-analysis

Andrea C Tricco[1], Elise Cogo[1], Veroniki A Angeliki[1], Charlene Soobiah[1,2], Brian Hutton[3], Brenda R Hemmelgarn[4], David Moher[3], Yaron Finkelstein[5,6,7] and Sharon E Straus[1,8*]

Abstract

Background: Epilepsy affects about 1% of the general population. Anti-epileptic drugs (AEDs) prevent or terminate seizures in individuals with epilepsy. Pregnant women with epilepsy may continue taking AEDs. Many of these agents cross the placenta and increase the risk of major congenital malformations, early cognitive and developmental delays, and infant mortality. We aim to evaluate the comparative safety of AEDs approved for chronic use in Canada when administered to pregnant and breastfeeding women and the effects on their infants and children through a systematic review and network meta-analysis.

Methods: Studies examining the effects of AEDs administered to pregnant and breastfeeding women regardless of indication (e.g., epilepsy, migraine, pain, psychiatric disorders) on their infants and children will be included. We will include randomized clinical trials (RCTs), quasi-RCTs, non-RCTs, controlled before-after, interrupted time series, cohort, registry, and case-control studies. The main literature search will be executed in MEDLINE, EMBASE, and the Cochrane Central Register of Controlled Trials. We will seek unpublished literature through searches of trial protocol registries and conference abstracts. The literature search results screening, data abstraction, and risk of bias appraisal will be performed by two individuals, independently. Conflicts will be resolved through discussion. The risk of bias of experimental and quasi-experimental studies will be appraised using the Cochrane Effective Practice and Organization of Care Risk-of-Bias tool, methodological quality of observational studies will be appraised using the Newcastle-Ottawa Scale, and quality of reporting of safety outcomes will be conducted using the McMaster Quality Assessment Scale of Harms (McHarm) tool. If feasible and appropriate, we will conduct random effects meta-analysis. Network meta-analysis will be considered for outcomes that fulfill network meta-analysis assumptions.

The primary outcome is major congenital malformations (overall and by specific types), while secondary outcomes include fetal loss/miscarriage, minor congenital malformations (overall and by specific types), cognitive development, psychomotor development, small for gestational age, preterm delivery, and neonatal seizures.

(Continued on next page)

* Correspondence: sharon.straus@utoronto.ca
[1]Li Ka Shing Knowledge Institute, St. Michael's Hospital, 209 Victoria Street, East Building, Toronto, Ontario M5B 1 T8, Canada
[8]Department of Geriatric Medicine, University of Toronto, 172 St. George Street, Toronto, Ontario M5R 0A3, Canada
Full list of author information is available at the end of the article

(Continued from previous page)

Discussion: Our systematic review will address safety concerns regarding the use of AEDs during pregnancy and breastfeeding. Our results will be useful to healthcare providers, policy-makers, and women of childbearing age who are taking anti-epileptic medications.

Systematic review registration: PROSPERO CRD42014008925.

Keywords: Anti-epileptic drug, Breastfeeding, Comparative safety, Congenital malformation, Epilepsy, Fetus, Infant, Network meta-analysis, Pregnancy, Systematic review

Background

Epilepsy is the most common chronic neurological condition, affecting 0.6 to 1% of the population [1,2]. Individuals with uncontrolled epilepsy experience recurrent seizures, which can have psychosocial and physical consequences, including a compromised life expectancy [3,4]. The goal of anti-epileptic treatment is to improve quality of life and health outcomes by reducing the frequency of seizures [4].

Anti-epileptic medications decrease seizures by reducing excitation and enhancing inhibition of neurons [5-7]. Many of these medications target different channels, including calcium, sodium, and glutamate, and are broadly classified as first generation agents (e.g., phenobarbitone, phenytoin, carbamazepine, sodium valproate, ethosuximide) and second generation agents (e.g., lamotrigine, levetiracetam, topiramate, gabapentin, vigabatrin, oxcarbazepine, clobazam, clonazepam, zonisamide, lacosamide, rufinamide, primidone) [8]. Due to the broad and varied mechanisms of action, the indications for some of these medications also include pain syndromes, psychiatric disorders, and migraine headaches [8].

Many clinical practice guidelines recommend that women of childbearing age continue to take their anti-epileptic medications; however, medications with lower risk of teratogenic events are advised [9,10] since anti-epileptic drugs (AEDs) cross the placenta or transfer through breast milk, posing risks to the fetus and infant [9,11,12].

Some AEDs have been associated with increased risk of harm to the fetus and infants. For example, exposure to valproate has led to increased risk of major congenital malformations [10], cognitive delay, and minor congenital abnormalities [13-16]. Phenobarbital has been associated with minor congenital abnormalities and developmental delay [17,18]. Carbamazepine and lamotrigine have been associated with minor congenital abnormalities [19-22]. However, other than studies of the use of valproate, many studies have produced inconsistent findings regarding harm to the fetus and infant with use of other agents [23]. As such, our objective is to evaluate the comparative safety of AEDs for infants and children who were exposed *in utero* or during breastfeeding through a systematic review and network meta-analysis.

Methods/Design

Protocol

A systematic review protocol was developed and registered with the PROSPERO database (CRD42014008925, available at: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014008925). It was revised with feedback from the decision-makers who posed the query within Health Canada, healthcare practitioners, content experts, and research methodologists. The reporting of our systematic review protocol was guided by the Preferred Reporting Items for Systematic Reviews and Meta-analyses Protocols [24].

Eligibility criteria

We will include studies examining the effects of AEDs on infants and children who were exposed *in utero* or during breastfeeding. We will include experimental studies (randomized clinical trials [RCTs], quasi-RCTs, non-RCTs), quasi-experimental studies (controlled before and after studies, interrupted time series), and observational studies (cohort, case-control, registry studies) of pregnant women at any stage of pregnancy and breastfeeding women and their infants/children. The rationale for including other study designs in addition to RCTs is that there are ethical issues in conducting RCTs of AEDs in pregnancy, so RCT evidence might not exist for some or all of these drugs. Given that our review includes rare outcomes, including observational evidence is crucial. In contrast to efficacy evaluation, safety assessment usually requires very large sample sizes to be able to detect adverse events. Therefore, while RCTs have lower risk of bias, they usually do not have the statistical power needed to adequately evaluate uncommon/rare safety outcomes due to Type II (i.e., false negative) error [25]. Given that our review includes rare outcomes, including observational evidence is crucial [26]. Additionally, observational studies can often provide more generalizable evidence due to the strict participant inclusion criteria in most RCTs [27]. Real-world safety evidence that has external validity is important for the assessment of the possible risks of AEDs in pregnant and breastfeeding women.

The diagnosis of neurodevelopmental delay related to *in utero* exposure is made before adolescence, and

hence, we will limit inclusion to children up to 12 years of age. AEDs that are approved for chronic use in Canada will be included. Drugs that are only used acutely or those that are not currently approved for use in Canada will be excluded, as the focus of this review is on the Canadian setting [28-32]. However, most of the medications we will examine are available in other countries as well. The relevant 16 medications and their synonyms are listed in Additional file 1, and the excluded drugs are listed in Additional file 2. Studies of all combinations and doses of these medications are eligible for inclusion. Since we are only interested in exposures that occur *in utero* or during breastfeeding, studies examining AEDs administered directly to infants or children will be excluded. All indications for AEDs will be included such as epilepsy, migraine, pain, and psychiatric disorders.

In order to be included, studies must compare an anti-epileptic medication against another included anti-epileptic medication, placebo, a 'no intervention' control group, or combinations of two or more anti-epileptic medications. Only studies providing results for our outcomes of interest will be included. Our primary outcome is major congenital malformations (overall and by specific type, such as craniofacial defects and neural tube defects). Secondary outcomes include minor congenital malformations (overall and by specific type, such as epicanthal folds and microstomia), cognition (e.g., global cognitive functioning and specific cognitive domains such as attention), psychomotor development (e.g., autism, dyspraxia), small for gestational age, preterm delivery, neonatal seizures, and fetal loss/miscarriage. No other limitations will be imposed on the eligibility criteria, including published/unpublished material, language of dissemination, duration of follow-up, or year of publication. The draft eligibility criteria can be found in Additional file 3.

Information sources and literature search

Our main literature search will be executed in the MEDLINE database. The search terms were drafted by an experienced librarian and can be found in Additional file 4. The search was peer reviewed by another librarian using the Peer Review of Electronic Search Strategies checklist [33].

In addition to MEDLINE, we will also search the EMBASE and the Cochrane Central Register of Controlled Trials databases. We will follow the MEDLINE search strategy for these databases, and the search terms will be adjusted accordingly. The electronic database search will be supplemented by searching for unpublished literature [34]. This will be accomplished through exploring conference abstracts, clinical trial registries, and contacting manufacturers of AEDs. We will also scan the reference lists of included studies and previous reviews in the area [23,35,36].

Study selection process

The eligibility criteria screening form will be pilot-tested by the team and is presented in Additional file 3. We will calculate inter-rater reliability from the pilot-test and screening will only commence after high agreement (e.g., kappa statistic ≥60%) is observed [37]. Subsequently, two reviewers will screen each title/abstract and potentially relevant full-text articles from the literature search results, independently. Conflicts will be resolved through discussion. All screening will occur using our online screening software (synthesi.SR) [38].

Data items and data collection process

We will abstract data on the PICOS elements [39], including patient characteristics (e.g., age of the mother and infant/child, indication for anti-epileptic treatment, co-morbidities, concomitant medications), intervention details (e.g., type of anti-epileptic treatment, dose, route of administration, duration of treatment, timing [trimester] of treatment during pregnancy), comparator details (e.g., comparator agent, dose, route of administration), outcome results (e.g., major congenital abnormality, minor congenital abnormality, cognitive function, psychomotor development) at the longest duration of follow-up, and study characteristics (e.g., study design, country of conduct, year of conduct, sample size, setting). These characteristics will be abstracted using a data abstraction form created in Excel with an accompanying "cheat sheet" that will guide the reviewers with this process. The data abstraction form and cheat sheet will be pilot-tested and data abstraction will only commence when high agreement (e.g., kappa statistic ≥60%) [37] is observed. Each included study will be abstracted by two team members, independently, who will resolve disagreements through discussion.

Methodological quality/risk of bias appraisal

We will use various tools to assess the methodological quality/risk of bias of each of the studies that fulfill our eligibility criteria. This will be conducted by two reviewers, independently, and conflicts will be resolved through discussion. First, we will appraise the risk of bias of experimental and quasi-experimental studies using the Cochrane Effective Practice and Organization of Care Risk-of-Bias tool [40]. Second, we will assess the methodological quality of observational studies using the Newcastle-Ottawa Scale [41]. Third, the quality of reporting of harms will be appraised using the McMaster Quality Assessment Scale of Harms (McHarm) tool [42].

Synthesis of included studies

A narrative summary of study results will be presented along with evidence summary tables. When sufficient data are available, we will conduct random effects meta-

analysis to calculate pooled odds ratios for dichotomous data and pooled mean differences for continuous data [43,44]. Direct (pairwise) meta-analysis will be performed with RCTs alone in order to examine whether the data are consistent between direct and indirect evidence. If the large majority of included studies are observational, we will also conduct additional meta-analyses including observational studies alone. Analyses will be stratified by treatment indication (e.g., epilepsy, pain, etc.) to reduce clinical heterogeneity between different study populations whenever possible; for example, epilepsy itself in pregnant women is related to an increased baseline risk of certain neonatal adverse outcomes. Statistical, clinical, and methodological heterogeneity will be examined prior to conducting the meta-analysis. Funnel plots will be drawn for outcomes including at least 10 studies to explore asymmetry that might be explained by clinical, statistical, and methodological heterogeneity. The proportion of statistical heterogeneity will be examined using the I^2 measure [45] and the magnitude of statistical heterogeneity will be calculated using the restricted maximum likelihood [46]. Meta-regression will be conducted for clinically relevant subgroups or when extensive statistical heterogeneity is observed (e.g., $I^2 \geq 75\%$) [47]. This will allow the examination of the impact of important factors on our results, such as maternal age, dose, duration and timing (e.g., trimester) of antiepileptic treatment, co-morbidities, concomitant medications, risk of bias results, and sample size (due to Type II statistical power errors with rare adverse events). To ensure the meta-regression analysis is intuitive, the number of covariates examined will be less than 10% of the number of studies included in the meta-analysis for the particular outcome.

We anticipate that many of these outcomes will be rare. To deal with studies reporting zero events in one treatment arm, 0.5 will be added to the numerator and 1 will be added to the denominator. We will exclude studies reporting zero events in all treatment arms for a particular outcome [48,49]. We also anticipate that we will encounter missing data in the included studies. We will contact the study authors for this data and if we are unable to receive the data, we will impute missing data (e.g., measures of variance) using established methods [50]. To ensure that our imputations do not bias our results, we will conduct a sensitivity analysis [51]. The meta-analysis and meta-regression will be analyzed in R using the *metafor* command [52].

A random-effects network meta-analysis will be conducted to make inferences regarding the comparative safety of the various AEDs [15], as well as rank their safety using rankograms and the surface under the cumulative ranking curve [53]. We will ensure the following factors are present prior to conducting network meta-analysis: i) transitivity (i.e., comparable distribution of effect modifiers across comparisons), which will be examined using boxplots or percentages to visually inspect potential effect modifiers of treatment effect [54]; ii) consistency between direct and indirect data, which will be examined locally (i.e., in certain paths of the network) using the loop-specific method [55,56] and the node-splitting method [57], and globally (i.e., evaluating the network as a whole), using the design-by-treatment interaction model [58]; and iii) we will quantify the amount of variability attributed to heterogeneity and inconsistency rather than sampling error, by calculating the I^2 [59]. We will estimate the amount of heterogeneity using the restricted maximum likelihood method and assuming common within-network heterogeneity. We will compare the magnitude of heterogeneity between consistency and inconsistency models, as well as between meta-regression and network meta-analysis models to determine how much heterogeneity will be explained by inconsistency or the explanatory variable, respectively. We will first use the design-by-treatment model for the evaluation of inconsistency in a network as a whole and then, if inconsistency is detected, we will employ the loop-specific and node-splitting methods to identify which piece of evidence is responsible for inconsistency. As mentioned above, analyses will be stratified by treatment indication when clinically appropriate. Important heterogeneity and inconsistency will be explored using network meta-regression using the same methods as described above, as necessary.

Prior to conducting the network meta-analysis, we will hold a team meeting to finalize which treatment nodes will be included in the analysis since we are unclear about the indications, dosages, patient populations, and outcomes reported in all of the studies. We will discuss issues, including conducting a class versus independent drug analysis, inclusion of drug routes of administration and dosages, as well as timing of drug administration. These decisions will be examined through a sensitivity analysis in which we will classify treatment nodes using a different classification to see how stable our results are. The network meta-analysis results will be presented as summary treatment effects for each pair of treatments. Network meta-analysis will be conducted in Stata with the *mvmeta* routine [60].

A sequential approach will be used for the network meta-analysis. We will first restrict our analysis to RCTs, which will be the primary analysis of interest. We will then include data from quasi-experimental studies, and finally, data from observational studies. This will provide an understanding of the contribution of each type of study design to our summary estimates, providing us with information on how these agents work above and beyond clinical trials.

Discussion

Epilepsy is the most common chronic neurological condition, affecting 0.6 to 1% of the population [1,2]. Given that approximately a third of patients receiving AEDs are of reproductive age and almost half of pregnancies are unplanned [61], the fetus may be exposed to these in the first trimester of pregnancy, including during the critical stage of embryogenesis [62].

The comparative safety of these agents is currently unknown and our results will be important for policy-makers, healthcare providers, and women of childbearing age. To ensure our results have wide dissemination and uptake, we will publish our results in open access journals, present our findings at scientific conferences, conduct dissemination meetings with key stakeholders (including policy-makers and healthcare providers), and produce policy briefs for Health Canada, the organization that posed this query.

Additional files

Additional file 1: List of relevant medications.
Additional file 2: Excluded drugs.
Additional file 3: Draft eligibility criteria.
Additional file 4: MEDLINE literature search.

Abbreviations

AEDs: Anti-epileptic drugs; RCTs: Randomized clinical trials.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

ACT conceived and designed the study, helped obtain funding for the study, and helped write the draft protocol. EC registered the protocol with the PROSPERO database and edited the draft protocol. AV helped write the draft protocol. CS edited the draft protocol. BH, BRH, DM, and YF provided input into the design, helped obtain funding for the study, and edited the draft protocol. SES conceived the study, designed the study, obtained the funding, and helped write the draft protocol. All authors read and approved the final protocol.

Acknowledgements

This systematic review was funded by the Canadian Institutes of Health Research/Drug Safety and Effectiveness Network (CIHR/DSEN). ACT and BH are funded by a CIHR/DSEN New Investigator Award in Knowledge Synthesis. BRH receives funding from the Alberta Heritage Foundation for Medical Research. DM is funded by a University of Ottawa Research Chair. SES is funded by a Tier 1 Canada Research Chair in Knowledge Translation.
We thank Laure Perrier for conducting the literature searches and Becky Skidmore for peer reviewing the MEDLINE search strategy. We also thank Dr. Joseph Beyene for providing feedback on our original proposal and Wing Hui and Judy Tran for formatting the paper.

Author details

[1]Li Ka Shing Knowledge Institute, St. Michael's Hospital, 209 Victoria Street, East Building, Toronto, Ontario M5B 1 T8, Canada. [2]Institute of Health Policy Management and Evaluation, University of Toronto, Health Sciences Building, 155 College Street, Suite 425, Toronto, Ontario M5T 3 M6, Canada. [3]Clinical Epidemiology Program, Centre for Practice-Changing Research, Ottawa Hospital Research Institute, The Ottawa Hospital – General Campus and University of Ottawa, 501 Smyth Road, Box 711, Ottawa, Ontario K1H 8 L6, Canada. [4]Departments of Medicine and Community Health Sciences, University of Calgary, TRW Building, 3rd Floor, 3280 Hospital Drive NW, Calgary, Alberta T2N 4Z6, Canada. [5]The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada. [6]Department of Pediatrics, University of Toronto, 172 St. George Street, Toronto, Ontario M5R 0A3, Canada. [7]Department of Pharmacology and Toxicology, University of Toronto, Medical Sciences Building, 1 King's College Circle, Room 4207, Toronto, Ontario M5S 1A8, Canada. [8]Department of Geriatric Medicine, University of Toronto, 172 St. George Street, Toronto, Ontario M5R 0A3, Canada.

Received: 9 April 2014 Accepted: 17 June 2014
Published: 25 June 2014

References

1. Hauser WA, Hesdorffer D: *Epilepsy, Frequency, Causes and Consequences.* New York: Demos Publications; 1990.
2. Wiebe S, Bellhouse DR, Fallahay C, Eliasziw M: **Burden of epilepsy: the Ontario Health Survey.** *Can J Neurol Sci* 1999, **26**(4):263–270.
3. Sperling MR: **The consequences of uncontrolled epilepsy.** *CNS Spectr* 2004, **9**(2):98–101. 106–109.
4. Jones MW: **Consequences of epilepsy: why do we treat seizures?** *Can J Neurol Sci* 1998, **25**(4):S24–S26.
5. Dickenson AH, Ghandehari J: **Anti-convulsants and anti-depressants.** *Handb Exp Pharmacol* 2007, **177:**145–177.
6. Stefani A, Spadoni F, Bernardi G: **Voltage-activated calcium channels: targets of antiepileptic drug therapy?** *Epilepsia* 1997, **38**(9):959–965.
7. Snutch TP, Reiner PB: **Ca^{2+} channels: diversity of form and function.** *Curr Opin Neurobiol* 1992, **2**(3):247–253.
8. Spina E, Perugi G: **Antiepileptic drugs: indications other than epilepsy.** *Epileptic Disord* 2004, **6**(2):57–75.
9. Harden CL, Pennell PB, Koppel BS, Hovinga CA, Gidal B, Meador KJ, Hopp J, Ting TY, Hauser WA, Thurman D, Kaplan PW, Robinson JN, French JA, Wiebe S, Wilner AN, Vazquez B, Holmes L, Krumholz A, Finnell R, Shafer PO, Le Guen CL, American Academy of Neurology; American Epilepsy Society: **Management issues for women with epilepsy–focus on pregnancy (an evidence-based review): III. Vitamin K, folic acid, blood levels, and breast-feeding: report of the quality standards subcommittee and therapeutics and technology assessment subcommittee of the American Academy of Neurology and the American Epilepsy Society.** *Epilepsia* 2009, **50**(5):1247–1255.
10. Harden CL, Meador KJ, Pennell PB, Hauser WA, Gronseth GS, French JA, Wiebe S, Thurman D, Koppel BS, Kaplan PW, Robinson JN, Hopp J, Ting TY, Gidal B, Hovinga CA, Wilner AN, Vazquez B, Holmes L, Krumholz A, Finnell R, Hirtz D, Le Guen C, American Academy of Neurology; American Epilepsy Society: **Management issues for women with epilepsy-Focus on pregnancy (an evidence-based review): II. Teratogenesis and perinatal outcomes: Report of the Quality Standards Subcommittee and Therapeutics and Technology Subcommittee of the American Academy of Neurology and the American Epilepsy Society.** *Epilepsia* 2009, **50**(5):1237–1246.
11. Samren EB, van Duijn CM, Koch S, Hiilesmaa VK, Klepel H, Bardy AH, Mannagetta GB, Deichl AW, Gaily E, Granstrom ML, Meinardi H, Grobbee DE, Hofman A, Janz D, Lindhout D: **Maternal use of antiepileptic drugs and the risk of major congenital malformations: a joint European prospective study of human teratogenesis associated with maternal epilepsy.** *Epilepsia* 1997, **38**(9):981–990.
12. Meador K, Reynolds MW, Crean S, Fahrbach K, Probst C: **Pregnancy outcomes in women with epilepsy: a systematic review and meta-analysis of published pregnancy registries and cohorts.** *Epilepsy Res* 2008, **81**(1):1–13.
13. Adab N, Jacoby A, Smith D, Chadwick D: **Additional educational needs in children born to mothers with epilepsy.** *J Neurol Neurosurg Psychiatry* 2001, **70**(1):15–21.
14. Adab N, Kini U, Vinten J, Ayres J, Baker G, Clayton-Smith J, Coyle H, Fryer A, Gorry J, Gregg J, Mawer G, Nicolaides P, Pickering L, Tunnicliffe L, Chadwick DW: **The longer term outcome of children born to mothers with epilepsy.** *J Neurol Neurosurg Psychiatry* 2004, **75**(11):1575–1583.
15. Gaily E, Kantola-Sorsa E, Hiilesmaa V, Isoaho M, Matila R, Kotila M, Nylund T, Bardy A, Kaaja E, Granstrom ML: **Normal intelligence in children with prenatal exposure to carbamazepine.** *Neurology* 2004, **62**(1):28–32.
16. Meador KJ, Baker GA, Browning N, Clayton-Smith J, Combs-Cantrell DT, Cohen M, Kalayjian LA, Kanner A, Liporace JD, Pennell PB, Privitera M, Loring DW, for the NEAD Study Group: **Cognitive function at 3 years of age after fetal exposure to antiepileptic drugs.** *N Engl J Med* 2009, **360**(16):1597–1605.

17. Holmes LB, Wyszynski DF, Lieberman E: **The AED (antiepileptic drug) pregnancy registry: a 6-year experience.** *Arch Neurol* 2004, **61**(5):673–678.
18. Reinisch JM, Sanders SA, Mortensen EL, Rubin DB: **In utero exposure to phenobarbital and intelligence deficits in adult men.** *JAMA* 1995, **274**(19):1518–1525.
19. Morrow J, Russell A, Guthrie E, Parsons L, Robertson I, Waddell R, Irwin B, McGivern RC, Morrison PJ, Craig J: **Malformation risks of antiepileptic drugs in pregnancy: a prospective study from the UK Epilepsy and Pregnancy Register.** *J Neurol Neurosurg Psychiatry* 2006, **77**(2):193–198.
20. Meador KJ, Baker GA, Finnell RH, Kalayjian LA, Liporace JD, Loring DW, Mawer G, Pennell PB, Smith JC, Wolff MC, NEAD Study Group: **In utero antiepileptic drug exposure: fetal death and malformations.** *Neurology* 2006, **67**(3):407–412.
21. Vajda FJ, Hitchcock A, Graham J, Solinas C, O'Brien TJ, Lander CM, Eadie MJ: **Foetal malformations and seizure control: 52 months data of the Australian Pregnancy Registry.** *Eur J Neurol* 2006, **13**(6):645–654.
22. Holmes LB, Baldwin EJ, Smith CR, Habecker E, Glassman L, Wong SL, Wyszynski DF: **Increased frequency of isolated cleft palate in infants exposed to lamotrigine during pregnancy.** *Neurology* 2008, **70**(22 Pt 2):2152–2158.
23. Meador KJ, Penovich P, Baker GA, Pennell PB, Bromfield E, Pack A, Liporace JD, Sam M, Kalayjian LA, Thurman DJ, Moore E, Loring DW, NEAD Study Group: **Antiepileptic drug use in women of childbearing age.** *Epilepsy Behav* 2009, **15**(3):339–343.
24. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L: **Reporting Guidelines for Systematic Review Protocols.** In *19th Cochrane Colloquium: 19–22 October 2011; Madrid, Spain.*
25. Ross CJ, Visscher H, Sistonen J, Brunham LR, Pussegoda K, Loo TT, Rieder MJ, Koren G, Carleton BC, Hayden MR, CPNDS Consortium: **The Canadian Pharmacogenomics Network for Drug Safety: a model for safety pharmacology.** *Thyroid* 2010, **20**(7):681–687.
26. Eypasch E, Lefering R, Kum CK, Troidl H: **Probability of adverse events that have not yet occurred: a statistical reminder.** *BMJ* 1995, **311**(7005):619–620.
27. Atkins D: **Creating and synthesizing evidence with decision makers in mind: integrating evidence from clinical trials and other study designs.** *Med Care* 2007, **45**(10 Supl 2):S16–S22.
28. *Health Canada: Drug Product Database.* http://www.hc-sc.gc.ca/dhp-mps/prodpharma/databasdon/index-eng.php.
29. *United States National Library of Medicine's ChemIDPlus Lite Database.* http://chem.sis.nlm.nih.gov/chemidplus/chemidlite.jsp.
30. *Canadian Pharmacists Association: E-CPS (Compendium of Pharmaceuticals and Specialties).* http://www.e-therapeutics.ca/home.whatsnew.action.
31. *Epilepsy Canada: Anticonvulsant Medications.* http://www.epilepsy.ca/en-CA/Diagnosis-and-Treatment/Anticonvulsant-Medications.html.
32. *Epilepsy Ontario: Anticonvulsant/Anti-Seizure Medication from A to Z.* http://epilepsyontario.org/anticonvulsantanti-seizure-medication-from-a-to-z/.
33. Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, Lefebvre C: **An evidence-based practice guideline for the peer review of electronic search strategies.** *J Clin Epidemiol* 2009, **62**(9):944–952.
34. *Canadian Agency for Drugs and Technologies in Health: Grey Matters: A Practical Search Tool for Evidence-Based Medicine.* http://www.cadth.ca/resources/grey-matters.
35. Adab N, Tudur SC, Vinten J, Williamson P, Winterbottom J: **Common antiepileptic drugs in pregnancy in women with epilepsy.** *Cochrane Database Syst Rev* 2004, **3**:CD004848.
36. Banach R, Boskovic R, Einarson T, Koren G: **Long-term developmental outcome of children of women with epilepsy, unexposed or exposed prenatally to antiepileptic drugs: a meta-analysis of cohort studies.** *Drug Saf* 2010, **33**(1):73–79.
37. Landis JR, Koch GG: **The measurement of observer agreement for categorical data.** *Biometrics* 1977, **33**(1):159–174.
38. *Synthesi.SR.* http://knowledgetranslation.ca/sysrev/login.php.
39. Stone PW: **Popping the (PICO) question in research and evidence-based practice.** *Appl Nurs Res* 2002, **15**(3):197–198.
40. *Cochrane Effective Practice and Organization of Care Group Draft Risk of Bias Tool.* http://epoc.cochrane.org/epoc-author-resources.
41. *The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-Analyses.* http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
42. Santaguida PL, Raina P, Ismaila A: *The Development of the McHarm Quality Assessment Scale for Adverse Events.* Hamilton, Ontario: McMaster University; 2008.
43. Raudenbush SW: **Analyzing Effect Sizes: Random Effects Models.** In *The Handbook of Research Synthesis and Meta-analysis.* 2nd edition. Edited by Cooper H, Hedges LV, Valentine JC. New York: Russell Sage Foundation; 2009:295–315.
44. Viechtbauer W: **Bias and efficiency of meta-analytic variance estimators in the random-effects model.** *J Educ Behav Stat* 2005, **30**(3):261–293.
45. Higgins JP, Thompson SG: **Quantifying heterogeneity in a meta-analysis.** *Stat Med* 2002, **21**(11):1539–1558.
46. Viechtbauer W: **Confidence intervals for the amount of heterogeneity in meta-analysis.** *Stat Med* 2007, **26**(1):37–52.
47. Higgins JPT, Green S: *Cochrane Handbook for Systematic Reviews of Interventions.* The Cochrane Collaboration; 2009. http://www.cochrane.org/handbook.
48. Sweeting MJ, Sutton AJ, Lambert PC: **What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data.** *Stat Med* 2004, **23**(9):1351–1375.
49. Bradburn MJ, Deeks JJ, Berlin JA, Russell Localio A: **Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events.** *Stat Med* 2007, **26**(1):53–77.
50. Littell JH, Corcoran J, Pillai V: *Systematic Reviews and Meta-Analysis.* New York: Oxford University Press; 2008.
51. Carpenter J, Rucker G, Schwarzer G: **Assessing the sensitivity of meta-analysis to selection bias: a multiple imputation approach.** *Biometrics* 2011, **67**(3):1066–1072.
52. *Conducting Meta-Analyses in R with the metafor Package.* http://www.jstatsoft.org/v36/i03/.
53. Salanti G, Ades AE, Ioannidis JP: **Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial.** *J Clin Epidemiol* 2011, **64**(2):163–171.
54. Salanti G: **Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool.** *Res Synth Methods* 2012, **3**(2):80–97.
55. Song F, Altman DG, Glenny AM, Deeks JJ: **Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses.** *BMJ* 2003, **326**(7387):472.
56. Veroniki AA, Vasiliadis HS, Higgins JP, Salanti G: **Evaluation of inconsistency in networks of interventions.** *Int J Epidemiol* 2013, **42**(1):332–345.
57. Dias S, Welton NJ, Caldwell DM, Ades AE: **Checking consistency in mixed treatment comparison meta-analysis.** *Stat Med* 2010, **29**(7–8):932–944.
58. White IR, Barrett JK, Jackson D, Higgins JPT: **Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression.** *Res Synth Methods* 2012, **3**(2):111–125.
59. Jackson D, Barrett JK, Stephen R, White IR, Higgins JPT: **A design-by-treatment interaction model for network meta-analysis with random inconsistency effects.** *Stat Med* 2014, In press.
60. White IR: **Multivariate random-effects meta-regression: updates to mvmeta.** *Stata J* 2011, **11**(2):255–270.
61. Centers for Disease Control and Prevention: *Unintended Pregnancy Prevention.* http://www.cdc.gov/reproductivehealth/unintendedpregnancy/.
62. Yerby MS: **Pregnancy, teratogenesis, and epilepsy.** *Neurol Clin* 1994, **12**(4):749–771.

doi:10.1186/2046-4053-3-68
Cite this article as: Tricco *et al.*: Comparative safety of anti-epileptic drugs among infants and children exposed *in utero* or during breastfeeding: protocol for a systematic review and network meta-analysis. *Systematic Reviews* 2014 **3**:68.

PRISMA NMA Checklist

Section/Topic	Item #	Checklist Item*	Reported on Page #
TITLE			
Title	1	Identify the report as a systematic review *incorporating a network meta-analysis (or related form of meta-analysis).*	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: **Background:** main objectives **Methods:** data sources; study eligibility criteria, participants, and interventions; study appraisal; and *synthesis methods, such as network meta-analysis.* **Results:** number of studies and participants identified; summary estimates with corresponding confidence/credible intervals; *treatment rankings may also be discussed. Authors may choose to summarize pairwise comparisons against a chosen treatment included in their analyses for brevity.* **Discussion/Conclusions:** limitations; conclusions and implications of findings. **Other:** primary source of funding; systematic review registration number with registry name.	4-5
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known*, including mention of why a network meta-analysis has been conducted.*	7
Objectives	4	Provide an explicit statement of questions being addressed, with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	7
METHODS			
Protocol and registration	5	Indicate whether a review protocol exists and if and where it can be accessed (e.g., Web address); and, if available, provide registration information,	8

		including registration number.	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. *Clearly describe eligible treatments included in the treatment network, and note whether any have been clustered or merged into the same node (with justification).*	8-9
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	9 (see also Appendix C)
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Additional File 1
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	9 (see also Appendix C)
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	9 (see also Appendix C)
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	Additional File 1
Geometry of the network	**S1**	Describe methods used to explore the geometry of the treatment network under study and potential biases related to it. This should include how the evidence base has been graphically summarized for presentation, and what characteristics were compiled and used to describe the evidence base to readers.	9-12 (see also Appendix C)
Risk of bias within individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	9
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means). *Also describe the use of additional summary measures assessed, such as treatment rankings and surface under the cumulative ranking curve (SUCRA) values, as well as modified approaches used to present*	9-12

		summary findings from meta-analyses.	
Planned methods of analysis	14	Describe the methods of handling data and combining results of studies for each network meta-analysis. This should include, but not be limited to: • *Handling of multi-arm trials;* • *Selection of variance structure;* • *Selection of prior distributions in Bayesian analyses; and* • *Assessment of model fit.*	9-12 (see also Appendix C)
Assessment of Inconsistency	**S2**	Describe the statistical methods used to evaluate the agreement of direct and indirect evidence in the treatment network(s) studied. Describe efforts taken to address its presence when found.	10 (see also Appendix C)
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	9 (see also Appendix C)
Additional analyses	16	Describe methods of additional analyses if done, indicating which were pre-specified. This may include, but not be limited to, the following: • Sensitivity or subgroup analyses; • Meta-regression analyses; • *Alternative formulations of the treatment network; and* • *Use of alternative prior distributions for Bayesian analyses (if applicable).*	10-12 (see also Appendix C)
RESULTS[†]			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	13 and Figure 1
Presentation of network structure	**S3**	Provide a network graph of the included studies to enable visualization of the geometry of the treatment network.	Figure 2
Summary of network geometry	**S4**	Provide a brief overview of characteristics of the treatment network. This may include commentary on the abundance of trials and randomized patients for the different interventions and pairwise comparisons in the network, gaps of evidence in the treatment network, and potential biases reflected by the network structure.	14-18 (see also Appendix L)
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS,	Table 1, Appendices H

		follow-up period) and provide the citations.	and I
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment.	Appendix J
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: 1) simple summary data for each intervention group, and 2) effect estimates and confidence intervals. *Modified approaches may be needed to deal with information from larger networks.*	N/A (data can be provided by the corresponding author)
Synthesis of results	21	Present results of each meta-analysis done, including confidence/credible intervals. *In larger networks, authors may focus on comparisons versus a particular comparator (e.g. placebo or standard care), with full findings presented in an appendix. League tables and forest plots may be considered to summarize pairwise comparisons.* If additional summary measures were explored (such as treatment rankings), these should also be presented.	Figure 3, Appendices L, M, O
Exploration for inconsistency	**S5**	Describe results from investigations of inconsistency. This may include such information as measures of model fit to compare consistency and inconsistency models, *P* values from statistical tests, or summary of inconsistency estimates from different parts of the treatment network.	14 (see also Appendix L)
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies for the evidence base being studied.	14 (see also Appendix K)
Results of additional analyses	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression analyses*, alternative network geometries studied, alternative choice of prior distributions for Bayesian analyses,* and so forth).	Appendix M
DISCUSSION			
Summary of evidence	24	Summarize the main findings, including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy-makers).	19-21
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review level (e.g., incomplete retrieval of identified research, reporting bias). *Comment on the validity of the assumptions, such as transitivity and consistency. Comment on any concerns regarding network*	21-23

		geometry (e.g., avoidance of certain comparisons).	
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	23
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. This should also include information regarding whether funding has been received from manufacturers of treatments in the network and/or whether some of the authors are content experts with professional conflicts of interest that could affect use of treatments in the network.	27-28

Abbreviations: PICOS - population, intervention, comparators, outcomes, study design
* Text in italics indicates wording specific to reporting of network meta-analyses that has been added to guidance from the PRISMA statement.
† Authors may wish to plan for use of appendices to present all relevant information in full detail for items in this section.

Supplementary Online Content

Appendix A. Outcome measures and diagnostic scales used in analysis

Cognitive developmental delay	
Bayley Scales of Infant Development (children ≤42 mo.)	Score ≥2 standard deviations below the mean
Griffiths Scale of Infant Development (children >42 mo.)	Score ≥2 standard deviations below the mean
McCarthy Scales of Children's Abilities (children >30 mo.)	Score ≥1 standard deviations below the mean
Stanford-Binet IV Intelligence scale for children	Intelligence quotient ≤80
Touwen's Test	Above average number of items rated abnormal in one or more domains
Wechsler Scale of Preschool and Primary Intelligence	Intelligence quotient <90
Wechsler Intelligence Scale for Children - III	Intelligence quotient <80; verbal intelligence quotient <69
Developmental Assessment	Confirmed diagnosis by developmental pediatrician or pediatric neurologist
Autism/dyspraxia	
Developmental Assessment	Diagnosis confirmed by developmental specialists at 2 years of age
Medical Records	Confirmed diagnosis recorded in medical history; registry records (ICD-10 codes F84.0, F84.1, F84.5, F84.8, and F84.9)
Modified checklist for autism in toddlers	Scored positive for ≥2 out of 6 critical items OR ≥3 any items of the total scale
Psychomotor developmental delay	
Ages and Stages Questionnaire	>3 standard deviations from the test mean
Bayley Scales of Infant Development – Psychomotor Index	>2 standard deviations below the standardized mean for the test
Touwen's Test	Demonstrated dysfunctions in fine motor balance, fine motor functions, and coordination of extremities
Schedule of Growing Skills II	Scored as 'delayed' in ≥1 domain of the test
Developmental Assessment	Infant scored >2 negative items (administered by general practitioner or pediatrician); diagnosis of neuromotor deficit confirmed by a trained nurse practitioner; infant failing to sit by 10 months or walk by 18 months
Health/Medical Records	Diagnosis of psychomotor delay recorded in medical records

Language Delay	
Ages and Stages Questionnaire	>3 standard deviations from the test mean
Clinical Evaluation of Language Fundamentals – 4th Edition	Score <70 in core language domain; score <84 overall
Learning Accomplishment Profile	Below average performance in expressive speech (adjusted for age)
Comprehensive Language Assessment (Peabody Picture Vocabulary Test; Receptive Expressive Emergent Language Scale; Expressive One Word Picture Vocabulary Test, or Sequenced Inventory of Communication Development)	Scores/assessment indicate a >6 moth delay in age appropriate language development
ADHD	
Attention Problems and Hyperactivity Scales	Score >1 standard deviations from the test mean
Child Behaviour Checklist	≥6 positive items on checklist
Diagnostic and Statistical Manual – IV	≥5 positive items on checklist
Medical Records	Confirmed diagnosis in hospital/medical records made by a pediatrician or child psychiatrist
Neonatal Seizure	
Medical records	Record of seizures during 1st year; confirmation of neonatal seizure by electroencephalography or diagnosis
Social Impairment	
Developmental Assessment (Ages and Stages Questionnaire [6 and 18 months]; Child Behaviour Checklist [36 months])	Scores dichotomized into 'normal' or 'adverse' range based on pre-defined values used by scale, for scales without pre-defined values cut-off was set at a score >2 standard deviations outside the test mean

Appendix B. Additional Information on the Methods

Information sources

An experienced librarian executed search strategies for MEDLINE, EMBASE, and the Cochrane Central Register of Controlled Trials. The search strategy for MEDLINE was peer reviewed by another librarian using the PRESS checklist,[1] and is available in the protocol.[2] The literature search was initially conducted from inception to March 18, 2014, and a rapid search update of these databases was subsequently conducted on December 15, 2015. Authors of AED studies were contacted for unpublished study data, and the reference lists of all included studies were scanned to identify additional studies.

Study selection and data collection

Prior to screening, two pilot-tests were conducted to assess the eligibility criteria and screening form. Once reviewers reached ≥80% agreement, pairs of reviewers independently screened titles and abstracts (level 1) and later screened potentially relevant full-text articles (level 2). Upon completion of title and abstract screening, 6% of citations were discrepant between reviewer pairs and had to be resolved by discussion or a third reviewer. At the conclusion of full-text screening, 16% of articles were discrepant and had to be resolved by discussion. The same process was used for data abstraction and quality appraisal. Three rounds of pilot testing were conducted prior to data abstraction to train reviewers and refine the data abstraction form. For studies published in the last 10 years, authors were contacted to request clarification or additional data.

Appraisal of selection bias using the comparison-adjusted funnel plot

All eligible medications were ordered from oldest to newest using their international market approval dates. To overcome some of the correlations induced by multi-arm studies, which may possibly cause overestimation and mask funnel plot asymmetry, we plotted data points corresponding to the study-specific basic parameters (treatment comparisons with common comparator). In each study, we used the control group as the common comparator or if this was missing, we used the oldest treatment comparator against the remaining AEDs. We used the fixed-effect model, as the random-effects model gives more weight to smaller studies and this may impact the assessment of small-study effects. We planned to explore observed asymmetry through subgroup analysis or meta-regression.

Additional details on the synthesis of included studies

Due to the complexity of the data and the studies' underreporting, differences in drug dosages could not be accounted for, and it was assumed that different dosages of the same AED were equally effective. When a study reported multiple dosages for the same treatment, we combined the data for this treatment. We assessed the transitivity assumption for each outcome *a priori* using the treatment effect modifiers: age, baseline risk, treatment indication, timing, and methodological quality. The mean of each continuous effect modifier and the mode of each categorical effect modifier for each pairwise comparison were presented in tables for each outcome.[3] The consistency assumption was evaluated for the entire network of each outcome

using the design-by-treatment interaction model.[4] If inconsistency was identified, further examination for local inconsistency in parts of the network was completed using the loop-specific method.[5 6] Common within-network between-study variance (τ^2) across treatment comparisons was assumed in the conventional meta-analysis, NMA, and design-by-treatment interaction model, so that treatment comparisons including a single study can borrow strength from the remaining network. This assumption was clinically reasonable, as the treatments included were of the same nature. In the loop-specific approach, common within-loop τ^2 was assumed.

For cognitive developmental delay and autism/dyspraxia outcomes, network meta-regression analyses for maternal age and baseline risk (i.e., using the control group) were conducted, when at least 10 studies provided relevant information, assuming a common fixed coefficient across treatment comparisons. Sensitivity analyses for cognitive developmental delay and autism/dyspraxia outcomes were performed for studies with the treatment indication of epilepsy, large study size (i.e., >300), maternal alcohol intake, maternal tobacco use, only first-generation AEDs, and higher methodological quality for the two items of the Newcastle-Ottawa Scale that had the highest percentage of low methodological quality (adequacy of follow-up of cohorts and comparability of cohorts items for cohort studies). Severity of epilepsy, which may be a risk factor variable, was not evaluated in our analyses since this was not commonly reported. For autism/dyspraxia, a sensitivity analysis on maternal IQ/psychiatric history was additionally conducted. We measured the goodness of fit using the posterior mean of the residual deviance, the degree of between-study heterogeneity, and the deviance information criterion. In a well-fitting model the posterior mean residual deviance should be close to the number of data points.[7 8] A difference of 3 units in the deviance information criterion was considered important and the lowest value of the deviance information criterion corresponded to the model with the best fit.[7 8]

All analyses were conducted in OpenBUGS,[9] assuming non-informative priors for all model parameters and a half normal prior distribution for the between-study standard deviation ($\tau \sim N(0,1), \tau > 0$). The first 10,000 iterations were discarded and then 100,000 simulations were run with thinning of 10 values. Convergence was checked by visual inspection of the evaluation of the mixing of two chains. The median and 95% CrI were calculated for each parameter value, since medians are not overly influenced by outliers. The *network* command[10] was used to apply the design-by-treatment interaction model.

References

1. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *J Clin Epidemiol*. 2016;75:40-6.
2. Tricco AC, Cogo E, Angeliki VA, et al. Comparative safety of anti-epileptic drugs among infants and children exposed in utero or during breastfeeding: protocol for a systematic review and network meta-analysis. *Syst Rev*. 2014;3:68.
3. Jansen JP, Naci H. Is network meta-analysis as valid as standard pairwise meta-analysis? It all depends on the distribution of effect modifiers. *BMC Med*. 2013;11:159.
4. White IR, Barrett JK, Jackson D, Higgins JP. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Res Synth Methods*. 2012;3(2):111-25.
5. Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ*. 2003;326(7387):472.

6. Veroniki AA, Vasiliadis HS, Higgins JP, Salanti G. Evaluation of inconsistency in networks of interventions. *Int J Epidemiol*. 2013;42(1):332-45.
7. Welton NJ, Sutton AJ, Cooper N, Abrams KR, Ades A. Evidence synthesis for decision making in healthcare. New York: Wiley; 2012.
8. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc Ser B Stat Methodol*. 2002;64(4):583-639.
9. Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: Evolution, critique and future directions. *Stat Med*. 2009;28(25):3049-67.
10. Palmer T, Sterne J. Meta-Analysis in Stata: An Updated Collection from the Stata Journal. White I, editor. Texas: Stata Press; 2016.

Appendix C. Newcastle-Ottawa Scale scoring guide

COHORT Studies

Excel Column	NOS* Answer Options**	NOS Coding Manual*
RefID	Enter the report's RefID.	
DA	Enter your initials.	
First author	Enter the first author's last name.	
Year of publication	Enter the year of the publication.	
SELECTION:		
1) Representative-ness of the exposed cohort	a) truly representative of the average pregnant woman taking AEDs in the community b) somewhat representative of the average pregnant woman taking AEDs in the community c) selected group of users e.g., nurses, volunteers d) no description of the derivation of the cohort	Item is assessing the representativeness of exposed individuals in the community, not the representativeness of the sample of women from some general population. For example, subjects derived from groups likely to contain middle class, better educated, health oriented women are likely to be representative of postmenopausal estrogen users while they are not representative of all women (e.g. members of a health maintenance organisation (HMO) will be a representative sample of estrogen users. While the HMO may have an under-representation of ethnic groups, the poor, and poorly educated, these excluded groups are not the predominant users of estrogen). Note: Truly representative (A) is a population-based cohort at the provincial or national levels (e.g., a sample from 2 cities is not enough). We need very 'broad' sample of the population. Somewhat representative (B) includes private clinics, hospital-based, or

		community-based.
2) Selection of the non-exposed cohort	a) drawn from the same community as the exposed cohort b) drawn from a different source c) no description of the derivation of the non-exposed cohort	Note: In our review of mostly multi-arm studies, this question pertains to the study's comparator group(s) – including "active" controls (for example, a less teratogenic AED). Therefore, this will often be 'A' for our studies.
3) Ascertainment of exposure	a) secure record (e.g., surgical records) b) structured interview c) written self-report d) no description	Note: Option 'A' includes patient hospital records, prescription drug database, or hospital/clinic visits (e.g., patient is asked about "current" AED use during a visit with their doctor). Option 'B' includes a hospital/clinic visit, but the patients are asked to remember their AED use during pregnancy (e.g., retrospectively ascertained exposure). If a study used both medical records and interviews for everyone, select 'A'.
4) Demonstration that outcome of interest was not present at start of study	a) yes b) no	In the case of mortality studies, outcome of interest is still the presence of a disease/incident, rather than death. That is to say that a statement of 'no history of disease or incident' earns a star (i.e. option 'A'). Note: Since our review is on pregnant women, this question is 'A' for all. **Please email us if a study involves breastfeeding women.**
COMPARABILITY:		
1) Comparability of cohorts on the basis of the design or analysis	a) answer is BOTH B & C (i.e. study controls for age and one other important factor) b) study controls for age of the women	Either exposed and non-exposed individuals must be matched in the design and/or confounders must be adjusted for in the analysis. Statements of no differences between groups or that differences were not statistically significant are not sufficient for establishing comparability.

	c) study controls for any other important factor d) study does not control for any important factor or it is not described	Note: If the relative risk for the exposure of interest is adjusted for the confounders listed, then the groups will be considered to be comparable on each variable used in the adjustment. There may be multiple ratings for this item for different categories of exposure (e.g., ever vs. never, current vs. previous or never). [A maximum of 2 stars can be allotted in this category]. Note: The study should have initially matched the groups or presented adjusted odds ratios, AND in addition, since in our review we are analyzing each AED arm separately (instead of the whole exposed cohort), the study must also report the factor of interest **for 'each AED arm'** (or state that **'each AED arm'** is matched). **Thus, there are 2 parts to this question:** 1) The study should have matched/adjusted for age at whatever level of groups they were focused on (even if they aren't our abstracted AED arms); AND 2) Then the study should also have reported the age for each AED arm. If they haven't done both of these 2 things, it's a 'D' here (unless they happen to combine these by reporting adjusted ORs for each of our AED arms). For our review, this generally pertains to **the comparability of the MOTHERS.** The exception here is in studies of cognitive/psychomotor development disorders in children - when age of the children should be comparable. The "other important factors" here are any one of these:

		• history of congenital malformations (CMs), fetal losses, preterm deliveries or small babies. • family history of genetic problems or CMs. • alcohol use. • nutritional deficiencies (e.g., lack of folic acid). Example: - Option 'B' indicates that the study initially matched groups based on the women's age (or reported adjusted ORs) AND they report the mean women's age for EACH of our arms (e.g., for Tx1, Tx2, etc.).
OUTCOME:		
1) Assessment of outcome	a) independent OR blind assessment b) record linkage c) self-report d) no description	For some outcomes (e.g. fractured hip), reference to the medical record is sufficient to satisfy the requirement for confirmation of the fracture. This would not be adequate for vertebral fracture outcomes where reference to x-rays would be required. a) Independent or blind assessment stated in the paper, or confirmation of the outcome by reference to secure records (x-rays, medical records, etc.) b) Record linkage (e.g. identified through ICD codes on database records) c) Self-report (i.e. no reference to original medical records or x-rays to confirm the outcome) d) No description. Note: Blind (A) is if they tell us that the outcome assessors were blinded to exposures; or if the outcome is objective. For our purposes, we will focus on the primary outcome of interest of our systematic review, which is major malformations (an objective outcome).

		So most of ours will be A, unless the study is only on a secondary outcome (e.g., cognitive development) and is based on the mother's self-report of their child (e.g., not a clinical examination).
2) Was follow-up long enough for outcomes to occur	a) yes b) no	An acceptable length of time should be decided before quality assessment begins (e.g. 5 yrs. for exposure to breast implants) Note: For this component, focus only on the outcomes that are reported in the results. For our purposes, we will focus on the primary outcome of interest of our systematic review, which is major malformations. • For studies focusing on 'birth' outcomes (i.e. malformations, preterm, fetal losses, born small), the answer is 'A' if they follow the groups until birth. • For studies focusing on cognitive developmental disorders, an adequate follow-up period (i.e. child's age) is 4 years. • For studies focusing on psychomotor delays, an adequate follow-up period is the earliest point of detection of the disorder. • For studies focusing on neonatal seizures, an adequate follow-up period (i.e. infant's age) is 6 months.
3) Adequacy of follow up of cohorts	a) complete follow up - all subjects accounted for b) subjects lost to follow up unlikely to introduce bias - small number lost (see 'Note'), or description provided of those lost c) follow up rate is inadequate (see 'Note') and no description of those lost d) no statement	This item assesses the follow-up of the exposed and non-exposed cohorts to ensure that losses are not related to either the exposure or the outcome. Note: **Especially check ones that start their total sample size (or figure diagram) with only the ones who had "complete" data (or only those who they had "successfully" recruited), as these are often a 'D' (since they don't report on the ones NOT followed up).** • For a prospective study, ≥90% follow-up rate per year is adequate (e.g., 10% dropout or less for 1 year, 20% for 2 years of follow-up, etc.). This includes missing or incomplete data, etc.

		• For a retrospective cohort study, ≥80% follow-up rate is adequate; including the ones that they could NOT recruit or who would NOT participate. • For a survey/mail questionnaire, ≥75% response rate is adequate. (For a survey, a dropout rate is congruent to a survey response rate).

CASE-CONTROL Studies

Excel Column	NOS* Answer Options**	NOS Coding Manual*
RefID	Enter the report's RefID.	
DA	Enter your initials.	
First author	Enter the first author's last name.	
Year of publication	Enter the year of the publication.	
SELECTION:		
1) Is the case definition adequate?	a) yes, with independent validation b) yes, e.g., record linkage or based on self-reports c) no description	a) Requires some independent validation (e.g. >1 person/record/time/ process to extract information, or reference to primary record source such as x-rays or medical/hospital records) b) Record linkage (e.g. ICD codes in database) or self-report with no reference to primary record c) No description Note: This question is assessing the group of infants that have the outcome of interest (e.g., CMs) – i.e. the "cases" in a case-control study design.
2) Representative-ness of the cases	a) consecutive or obviously representative series of cases b) potential for selection biases, or not stated	a) All eligible cases with outcome of interest over a defined period of time, all cases in a defined catchment area, all cases in a defined hospital or clinic, group of hospitals, health maintenance organisation, or an appropriate sample of those cases (e.g. random sample) b) Not satisfying requirements in part (a), or not stated. Note: Option 'A' is a population-based sample.
3) Selection of controls	a) community controls b) hospital controls c) no description	This item assesses whether the control series used in the study is derived from the same population as the cases and essentially would have been cases had the outcome been present.

		a) Community controls (i.e. same community as cases and would be cases if had outcome) b) Hospital controls, within same community as cases (i.e. not another city) but derived from a hospitalised population c) No description Note: This question is assessing the group of infants that don't have the outcome (e.g., CMs) – i.e. the "controls" in a case-control study design. Community controls (A) includes a population-based sample.
4) Definition of controls	a) no history of disease (endpoint) b) no description of source	a) If cases are first occurrence of outcome, then it must explicitly state that controls have no history of this outcome. If cases have new (not necessarily first) occurrence of outcome, then controls with previous occurrences of outcome of interest should not be excluded. b) No mention of history of outcome Note: Since our review is on fetal effects, this question is 'A' for all studies. **Please email us if a study involves exposure during breastfeeding.**
COMPARABILITY:		
1) Comparability of cases and controls on the basis of the design or analysis	a) answer is BOTH B & C (i.e. study controls for age and one other important factor) b) study controls for age of the women c) study controls for any other important factor d) study does not control for any important factor or it is not described	Either cases and controls must be matched in the design and/or confounders must be adjusted for in the analysis. Statements of no differences between groups or that differences were not statistically significant are not sufficient for establishing comparability. Note: If the odds ratio for the exposure of interest is adjusted for the confounders listed, then the groups will be considered to be comparable on each variable used in the adjustment. There may be multiple ratings for this item for different categories of exposure (e.g. ever vs. never, current vs. previous or never). [A maximum

		of 2 stars can be allotted in this category]. Note: The study should have initially matched the groups, AND in addition, since in our review we are analyzing each AED arm separately (instead of the whole cases group), the study must also report the factor of interest **for 'each AED arm'** (or state that **'each AED arm'** is matched). For our review, this generally pertains to **the comparability of the MOTHERS of the cases and controls.** The exception here is in studies of cognitive/psychomotor development disorders in children - when age of the children should be comparable. The "other important factors" here are any one of these: • history of congenital malformations (CMs), fetal losses, preterm deliveries or small babies. • family history of genetic problems or CMs. • alcohol use. • nutritional deficiencies (e.g., lack of folic acid). For example, Option 'B' indicates that the study initially matched groups based on the women's age AND they report the mean women's age for EACH arm (e.g., for Tx1, Tx2, etc.).
EXPOSURE:		
1) Assessment of exposure	a) secure record (e.g., surgical records) b) structured interview where blind to case/control status c) interview not blinded to case/control status d) written self-report or medical	Note: Option 'A' includes patient hospital records, prescription drug database, or hospital/clinic visits (e.g., patient is asked about "current" AED use during a visit with their doctor). "Interview" here includes a hospital/clinic visit, but the patients are asked to remember their AED use during pregnancy (e.g., retrospectively

	record only e) no description	ascertained exposure).
2) Same method of ascertainment for cases and controls	a) yes b) no	Note: This question is asking whether the method of ascertainment of exposure was the same for 'cases' (with the outcome) and 'controls' (without the outcome; in this case-control study design).
3) Non-response rate	a) same rate for both groups b) non-respondents described c) rate different and no designation	Note: For our review, this pertains to either the infants or the mothers of the case and control groups. We're allowing 10% dropout per year for a prospective study – e.g., 10% for 1 year, 20% for 2 years of follow-up, etc. For a survey, we allow for a 75% response rate in order for it be adequate. For a survey, a dropout rate is congruent to a survey response rate.

*Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available at: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

In the **"NOS Coding Manual" column, the first section for each item is copied straight from the NOS documentation while the lower portions in each item are our "Notes" tailored for the AED review.

Appendix D. List of included studies

1. Adab N, Kini U, Vinten J, et al. The longer term outcome of children born to mothers with epilepsy. *J Neurol Neurosurg Psychiatry*. 2004;75(11):1575-83.
2. Vinten J, Bromley RL, Taylor J, Adab N, Kini U, Baker GA. The behavioral consequences of exposure to antiepileptic drugs in utero. *Epilepsy Behav*. 2009;14(1):197-201.
3. Mawer G, Clayton-Smith J, Coyle H, Kini U. Outcome of pregnancy in women attending an outpatient epilepsy clinic: adverse features associated with higher doses of sodium valproate. *Seizure*. 2002;11(8):512-8.
4. Arkilo D, Hanna J, Dickens D, et al. Pregnancy and neurodevelopmental outcomes with in-utero antiepileptic agent exposure. A pilot study. *Eur J Paediatr Neurol*. 2015;19(1):37-40.
5. Bromley R, Baxter N, Calderbank R, Mawer G, Clayton-Smith J, Baker G. A comprehensive review of the language abilities of children exposed to valproate or carbamazepine in utero. American Epilepsy Society; Texas2010.
6. Bromley RL, Mawer GE, Briggs M, et al. The prevalence of neurodevelopmental disorders in children prenatally exposed to antiepileptic drugs. *J Neurol Neurosurg Psychiatry*. 2013;84(6):637-43.
7. Bromley RL, Mawer G, Clayton-Smith J, Baker GA. Autism spectrum disorders following in utero exposure to antiepileptic drugs. *Neurology*. 2008;71(23):1923-4.
8. Christensen J, Gronborg TK, Sorensen MJ, et al. Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. *JAMA*. 2013;309(16):1696-703.
9. Cohen MJ, Meador KJ, Browning N, et al. Fetal antiepileptic drug exposure: Adaptive and emotional/behavioral functioning at age 6 years. *Epilepsy Behav*. 2013;29(2):308-15.
10. Cummings C, Stewart M, Stevenson M, Morrow J, Nelson J. Neurodevelopment of children exposed in utero to lamotrigine, sodium valproate and carbamazepine. *Arch Dis Child*. 2011;96(7):643-7.
11. Tomson T, Battino D, Bonizzoni E, et al. Antiepileptic drugs and intrauterine death: A prospective observational study from EURAP. *Neurology*. 2015;85(7):580-8.
12. Dean JCS, Hailey H, Moore SJ, Lloyd DJ, Turnpenny PD, Little J. Long term health and neurodevelopment in children exposed to antiepileptic drugs before birth. *J Med Genet*. 2002;39(4):251-9.
13. Rasalam AD, Hailey H, Williams JH, et al. Characteristics of fetal anticonvulsant syndrome associated autistic disorder. *Dev Med Child Neurol*. 2005;47(8):551-5.
14. D'Souza SW, Robertson IG, Donnai D, Mawer G. Fetal phenytoin exposure, hypoplastic nails, and jitteriness. *Arch Dis Child*. 1991;66(3):320-4.
15. Eriksson K, Viinikainen K, Mönkkönen A, et al. Children exposed to valproate in utero—Population based evaluation of risks and confounding factors for long-term neurocognitive development. *Epilepsy Res*. 2005;65(3):189-200.
16. Viinikainen K, Eriksson K, Monkkonen A, et al. The effects of valproate exposure in utero on behavior and the need for educational support in school-aged children. *Epilepsy Behav*. 2006;9(4):636-40.
17. Gaily E. Development and growth in children of epileptic mothers: a prospective controlled study. Helsinki, Finland: University of Helsinki; 1990.
18. Gaily EK, Granstrom ML, Hiilesmaa VK, Bardy AH. Head circumference in children of epileptic mothers: contributions of drug exposure and genetic background. *Epilepsy Res*. 1990;5(3):217-22.
19. Hiilesmaa V. A prospective study on maternal and fetal outcome in 139 women with epilepsy. Helsinki: University of Helsinki; 1982.
20. Hiilesmaa VK, Bardy A, Teramo K. Obstetric outcome in women with epilepsy. *Am J Obstet Gynecol*. 1985;152(5):499-504.
21. Gogatishvili N, Ediberidze T, Lomidze G, Tatishvili N, Kasradze S. PO-0834 Long-term Developmental Outcome Of Children Prenatally Exposed To Antiepileptic Drugs. *Arch Dis Child*. 2014;99(Suppl 2):A526.
22. Gogatishvili N, Ediberidze T, Lomidze G, Tatishvili N, Kasradze S. Cognitive outcomes of children with fetal antiepileptic drug exposure at the age of 3-6 years-preliminary data. 1st Congress of the European Academy of Neurology; Berlin: European Journal of Neurology; 2015. p. 329.
23. Jones KL, Lacro RV, Johnson KA, Adams J. Pattern of malformations in the children of women treated with carbamazepine during pregnancy. *N Engl J Med*. 1989;320(25):1661-6.

24. Katz JM, Pacia SV, Devinsky O. Current Management of Epilepsy and Pregnancy: Fetal Outcome, Congenital Malformations, and Developmental Delay. *Epilepsy Behav*. 2001;2(2):119-23.
25. Koch S, Jager-Roman E, Losche G, Nau H, Rating D, Helge H. Antiepileptic drug treatment in pregnancy: drug side effects in the neonate and neurological outcome. *Acta Paediatr*. 1996;85(6):739-46.
26. Lacroix I, Hurault-Delarue C, Guitard C, et al. Psychomotor effects of in utero exposure to psychotropic medications: A comparative study in EFEMERIS database. Congrès de Physiologie Pharmacologie et Thérapeutique; Dijon, France2012.
27. Miskov S, Juraski RG, Fucic A, et al. Croatian Pregnant Women with Epilepsy and Effects of Antiepileptic Drugs Exposure in their Offspring - seven years of prospective surveillance. American Epilepsy Society; Texas2010.
28. Nadebaum C, Anderson VA, Vajda F, Reutens DC, Barton S, Wood AG. Language skills of school-aged children prenatally exposed to antiepileptic drugs. *Neurology*. 2011;76(8):719-26.
29. Rihtman T, Parush S, Ornoy A. Developmental outcomes at preschool age after fetal exposure to valproic acid and lamotrigine: cognitive, motor, sensory and behavioral function. *Reprod Toxicol*. 2013;41:115-25.
30. Scolnik D, Nulman I, Rovet J, et al. Neurodevelopment of children exposed in utero to phenytoin and carbamazepine monotherapy. *JAMA*. 1994;271(10):767-70.
31. Shankaran S, Woldt E, Nelson J, Bedard M, Delaney-Black V. Antenatal phenobarbital therapy and neonatal outcome. II: Neurodevelopmental outcome at 36 months. *Pediatrics*. 1996;97(5):649-52.
32. van der Pol MC, Hadders-Algra M, Huisjes HJ, Touwen BC. Antiepileptic medication in pregnancy: late effects on the children's central nervous system development. *Am J Obstet Gynecol*. 1991;164(1 Pt 1):121-8.
33. Veiby G, Engelsen BA, Gilhus NE. Early child development and exposure to antiepileptic drugs prenatally and through breastfeeding: a prospective cohort study on children of women with epilepsy. *JAMA neurology*. 2013;70(11):1367-74.
34. Veiby G, Daltveit AK, Schjolberg S, et al. Exposure to antiepileptic drugs in utero and child development: a prospective population-based study. *Epilepsia*. 2013;54(8):1462-72.
35. Wood AG, Nadebaum C, Anderson V, et al. Prospective assessment of autism traits in children exposed to antiepileptic drugs during pregnancy. *Epilepsia*. 2015;56(7):1047-55.

Appendix E. Additional information on search results

Of the included 110 studies, nine were written in languages other than English and three were conference abstracts or letter to the editor with usable data. Scanning of reference lists of included articles and related reviews identified 13 additional studies. Forty-eight percent of contacted authors (22/46) provided clarification or additional data. Additionally, 29% (13/45) of authors of conference abstracts provided additional unpublished data. We were unable to contact 11 authors due to non-working email addresses. One author provided a manuscript and three authors provided unpublished data that were included in the analysis.

Appendix F. Key excluded studies

Author, Year	Research Group	Title	Reason for Exclusion
Meador, 2009[1]	Neurodevelopmental Effects of Antiepileptic Drug (NEAD) Study Group	Cognitive Function at 3 Years of Age after Fetal Exposure to Antiepileptic Drugs	Outcomes only reported as continuous variables
Meador, 2010[2]	Neurodevelopmental Effects of Antiepileptic Drug (NEAD) Study Group	Effects of breastfeeding in children of women taking antiepileptic drugs	Outcomes only reported as continuous variables
Meador, 2011[3]	Neurodevelopmental Effects of Antiepileptic Drug (NEAD) Study Group	Foetal antiepileptic drug exposure and verbal versus non-verbal abilities at three years of age	Outcomes only reported as continuous variables
Meador, 2012[4]	Neurodevelopmental Effects of Antiepileptic Drug (NEAD) Study Group	Effects of fetal antiepileptic drug exposure: Outcomes at age 4.5 years	Outcomes only reported as continuous variables
Meador, 2013[5]	Neurodevelopmental Effects of Antiepileptic Drug (NEAD) Study Group	Fetal antiepileptic drug exposure and cognitive outcomes at age 6 years (NEAD study): a prospective observational study	Outcomes only reported as continuous variables
Shallcross, 2011[6]	Liverpool and Manchester Neurodevelopment Group and The UK Epilepsy and Pregnancy Register	Child development following in utero exposure: Levetiracetam vs. sodium valproate	Outcomes only reported as continuous variables
Shallcross, 2014[7]	Liverpool and Manchester	In utero exposure to levetiracetam vs. valproate: Development and language at 3 years of age	Outcomes only reported as continuous variables

	Neurodevelopment Group and The UK Epilepsy and Pregnancy Register		

References

1. Meador KJ, Baker GA, Browning N, et al. Cognitive function at 3 years of age after fetal exposure to antiepileptic drugs. *N Engl J Med*. 2009;360(16):1597-605.
2. Meador KJ, Baker GA, Browning N, et al. Effects of breastfeeding in children of women taking antiepileptic drugs. *Neurology*. 2010;75(22):1954-60.
3. Meador KJ, Baker GA, Browning N, et al. Foetal antiepileptic drug exposure and verbal versus non-verbal abilities at three years of age. *Brain*. 2011;134(Pt 2):396-404.
4. Meador KJ, Baker GA, Browning N, et al. Effects of fetal antiepileptic drug exposure: outcomes at age 4.5 years. *Neurology*. 2012;78(16):1207-14.
5. Meador KJ, Baker GA, Browning N, et al. Fetal antiepileptic drug exposure and cognitive outcomes at age 6 years (NEAD study): a prospective observational study. *Lancet Neurol*. 2013;12(3):244-52.
6. Shallcross R, Bromley RL, Irwin B, Bonnett LJ, Morrow J, Baker GA. Child development following in utero exposure: levetiracetam vs sodium valproate. *Neurology*. 2011;76(4):383-9.
7. Shallcross R, Bromley RL, Cheyne CP, et al. In utero exposure to levetiracetam vs valproate: development and language at 3 years of age. *Neurology*. 2014;82(3):213-21.

Appendix G. Table of Individual Study characteristics

Author, Year	Country of conduct	Registry or Setting	Study period	Interventions	Outcomes	Funding
Adab, 2004[1] [CR: Vinten, 2009[2]; Mawer, 2002[3*]]	UK	Mersey Regional Epilepsy Clinic; Epilepsy Clinic at the Manchester Royal Infirmary; Antenatal clinic at St Mary's Hospital, Manchester	2000-2001	Carbam, Control, Valpro	Cognitive Developmental Delay, Psychomotor Developmental Delay	NR
Arkilo, 2015[4]	USA	Minnesota Epilepsy Group	2006-2011	Carbam, Lamot, Levet, Pheny, Valpro	Autism/Dyspraxia, Psychomotor Developmental Delay	NR
Bromley, 2010[5]	UK	Liverpool and Manchester Neurodevelopment Group	NR	Carbam, Valpro	Language Delay	NR
Bromley, 2013[6] [CR: Bromley, 2008[7]]	UK	Liverpool and Manchester Neurodevelopment group	2000-2004	Carbam, Control, Lamot, Valpro	Autism/Dyspraxia, ADHD	mixed public & private
Christensen, 2013[8]†	Denmark	Danish Civil Registration System; Danish Prescription Register; Danish Psychiatric Central Register; Danish Birth Register; Danish National Hospital Register	1996-2006	Carbam, Clonaz, Lamot, Oxcar, Valpro	Autism/Dyspraxia	public

Cohen, 2013[9]	USA;UK	Neurodevelopmental Effects of Antiepileptic Drugs Study Group	1999-2004	Carbam, Lamot, Pheny, Valpro,	ADHD	public
Cummings, 2011[10]† [CR: Tomson, 2015[11]]	Northern Ireland	UK Epilepsy and Pregnancy Register (Northern Ireland); Northern Ireland Child Health System	1996-2005	Carbam, Lamot, Valpro,	Cognitive Developmental Delay	public
Dean, 2002[12] [CR: Rasalam, 2005[13]]	Scotland	Aberdeen Maternity Hospital	1976-2000	Carbam, Carbam+Pheno, Carbam+Pheny, Carbam+Valpro, Control, Ethos, Pheno, Pheno+Pheny, Pheno+Valpro, Pheny, Primid, Valpro	Psychomotor Developmental Delay, ADHD	NR
D'Souza, 1991[14]	United Kingdom	St Mary's Hospital	1980-1982	Carbam, Control, Pheno, Pheny, Valpro	Cognitive Developmental Delay	public
Eriksson, 2005[15]† [CR: Viinikainen, 2006[16]]	Finland	Kuopio University Hospital	1989-2000	Carbam, Control, Valpro	Cognitive Developmental Delay, Psychomotor Developmental Delay	public
Gaily, 1990[17] [CR: Gaily, 1990[18]; Hiilesmaa, 1982[19]; Hiilesmaa, 1985[20]]	Finland	Helsinki University Central Hospital	1975-1979	Carbam, Carbam+Pheno+Pheny, Carbam+Pheny, Carbam+Valpro, Control, Ethos+Pheny, Pheno+Pheny, Pheny, Pheny+Primid, Pheny+Valpro	Cognitive Developmental Delay , Psychomotor Developmental Delay	mixed public & private
Gogatishvili, 2014[21]	Georgia	Georgian National AED-Pregnancy Registry	NR	Carbam, Lamot, Valpro	Cognitive Developmental	public

					Delay	
Gogatishvili, 2015[22]	Georgia	Georgian National AED-Pregnancy Registry	NR	Carbam, Carbam+Levet, Lamot, Pheno, Valpro	Language Delay	public
Jones, 1989[23]†	US	California Teratogen Registry	1979-1988	Carbam, Carbam+Pheno, Carbam+Pheno+Valpro, Carbam+Primid	Cognitive Developmental Delay , Psychomotor Developmental Delay	public
Katz, 2001[24]	USA	Mount Sinai Comprehensive Epilepsy Center	1990-2000	Carbam, Control, Lamot, Pheno, Pheny, Primid, Valpro	Cognitive Developmental Delay	NR
Koch, 1996[25]	Germany	NR	1976-1983	Pheno, Pheny, Primid, Valpro	Cognitive Developmental Delay	public
Lacroix, 2012[26]	France	EFEMERIS database - Caisse Primaire d'Assurance Maladie of Haute-Garonne and Maternal and Infant Protection Service; Antenatal Diagnostic Centre	2004-2008	Carbam, Clobaz, Clonaz, Gabap, Lamot, Pheno, Topir, Valpro	Psychomotor Developmental Delay	NR
Mawer, 2002[3]	England	Manchester Royal Infirmary	1990-1999	Carbam, Lamot, Pheny, Valpro	Cognitive Developmental Delay	NR

Miskov, 2010[27]	Croatia	NR	2003-2010	Carbam, Control, Gabap, Lamot, Valpro	Psychomotor Developmental Delay, Neonatal Seizures	NR
Nadebaum, 2011[28]†	Australia	Australian Registry of Antiepileptic Drug Use in Pregnancy	2007-2009	Carbam, Lamot, Valpro	Language Delay	mixed public & private
Rihtman, 2013[29]	Israel	Israeli Teratogen Information Service	NR	Lamot, Valpro	Neonatal Seizure	mixed public & private
Scolnik, 1994[30]	Canada	Hospital for Sick Children - Motherisk Program; North York General Hospital; Toronto Hospital; Oshawa General Hospital	1987-1992	Carbam, Pheny	Cognitive Developmental Delay	public
Shankaran, 1996[31]	USA	Children's Hospital of Michigan	NR	Control, PHENO,	Psychomotor Developmental Delay, Language Delay	public
Van der Pol, 1991[32]	Netherlands	Groningen University Hospital	1973-1981	Carbam, Carbam+Pheno, Control, Pheno	Psychomotor Developmental Delay	public
Veiby, 2013a[33]†	Norway	Norwegian Institute of Public Health- Mother and Child Cohort Study	1999-2009	Carbam, Control, Lamot, Valpro	Social Impairment	public
Veiby, 2013b[34]†	Norway	Medical Birth Registry of Norway	1999-2008	Carbam, Control, Lamot, Valpro	Psychomotor Developmental Delay, Autism/Dyspraxia,	public

					Language Delay, ADHD	
Wood, 2015[35]†	Australia	Australian Registry of Antiepileptic Drug Use in Pregnancy	2007-2010	Carbam, Carbam+Clonaz, Carbam+Lamot, Carbam+Pheny, Lamot+Valpro, Valpro	Autism/Dyspraxia	public

Abbreviations: ADHD – Attention Deficit Hyperactivity Disorder; NR – Not Reported

Carbam = Carbamazepine; Clobaz = Clobazam; Clonaz = Clonazepam; Ethos = Ethosuximide; Gabap = Gabapentin; Lamot = Lamotrigine; Levet = Levetiracetam; Oxcar = Oxcarbazepine; Pheno = Phenobarbital; Pheny = Phenytoin; Pridmid = Primidone; Topir = Topiramate; Valpro = Valproate; Vigab = Viagabatrin

*Single publication reporting on two separate cohorts
†Registry Studies

Appendix H. Table of Patient characteristics

Author, Year	Indication	Sample Size*	Mean Age (Women)	Mean Age (Children)/ Follow-up period†	AED Exposure Timing	Maternal Alcohol Use n/N‡	Maternal Tobacco Use n/N‡
Adab, 2004a[1]§ [CR: Vinten, 2009[2]; Mawer, 2002[3]]	Epilepsy	177	26.1	9-10.5	NR	24/279‡	68/249‡
Adab, 2004b[1]§ [CR: Vinten, 2009[2]; Mawer, 2002[3]]	Epilepsy	81	26.1	3-3.33	NR	24/279‡	68/249‡
Arkilo, 2015[4]	Epilepsy	59	NR	NA	First trimester	NR	NR
Bromley, 2010[5]	NR	60	NR	6-7	Whole pregnancy	NR	NR
Bromley, 2013[6] [CR: Bromley, 2008[7]]	Epilepsy	156	28	6	NR	28/156	42/156
Christensen, 2013[8]	NR	2011	NR	NR	Whole pregnancy	NR	NR
Cohen, 2013[9]	Epilepsy	108	30	6	NR	12/192‡	NR
Cummings, 2011[10] [CR: Tomson, 2015[11]]	Epilepsy	142	NR	2-3	Whole pregnancy	32/108‡	19/108‡
Dean, 2002[12] [CR: Rasalam, 2005[13]]	Epilepsy	287	27	3.75-15.5	First trimester	NR	NR
D'Souza, 1991[14]	Epilepsy	42	26.5	2.5-3.5	Whole pregnancy	NR	NR
Eriksson, 2005[15] [CR: Viinikainen, 2006[16]]	Epilepsy	39	28.2	NR	NR	NR	NR

Gaily, 1990[17] [CR: Gaily, 1990[18]; Hiilesmaa, 1982[19]; Hiilesmaa, 1985[20]	Epilepsy	134	27.8	5.5	First trimester	NR	NR
Gogatishvili, 2014[21]	NR	39	NR	2 to 4	NR	NR	NR
Gogatishvili, 2015[22]	NR	23	NR	3 to 6	NR	NR	NR
Jones, 1989[23]	Epilepsy	63	NR	NR	Whole pregnancy	NR	NR
Katz, 2001[24]	Epilepsy	51	31	NR	NR	NR	NR
Koch, 1996[25]	Epilepsy	40	NR	6	First trimester	NR	NR
Lacroix, 2012[26]	NR	109	NR	0.75	NR	NR	NR
Mawer, 2002[3]	Epilepsy	52	NR	NR	NR	NR	NR
Miskov, 2010[27]	Epilepsy	55	NR	NR	NR	NR	NR
Nadebaum, 2011[28]	Epilepsy	66	31.6	7.4	First trimester	NR	5/66
Rihtman, 2013[29]	Epilepsy	72	NR	NR	Whole pregnancy	NR	NR
Scolnik, 1994[30]	Epilepsy	75	NR	1.5-3	1st trimester	NR	NR
Shankaran, 1996[31]	NR	96	NR	NR	NR	NR	NR
Van der Pol, 1991[32]	Epilepsy	57	NR	6-13	NR	NR	NR
Veiby, 2013a[33]	Epilepsy	422	NR	0.5	NR	NR	NR
Veiby, 2013b[34]	Epilepsy	248	28.9	3	NR	NR	68/726‡
Wood, 2015[35]	Epilepsy	77	NR	6-8	NR	NR	NR

Abbreviations: NA – Not applicable; NR – Not reported

* Sample size used for analysis; ineligible treatment arms (i.e. treatment arms with excluded drugs or unspecified polytherapy) are not included in the count

† The mean age for children/follow-up period data were only collected for outcomes related to cognitive and/or psychomotor development

‡ Total sample size is based on the number of women enrolled in the study; may differ from the sample size used for analysis

§ Single publication reporting on two separate cohorts

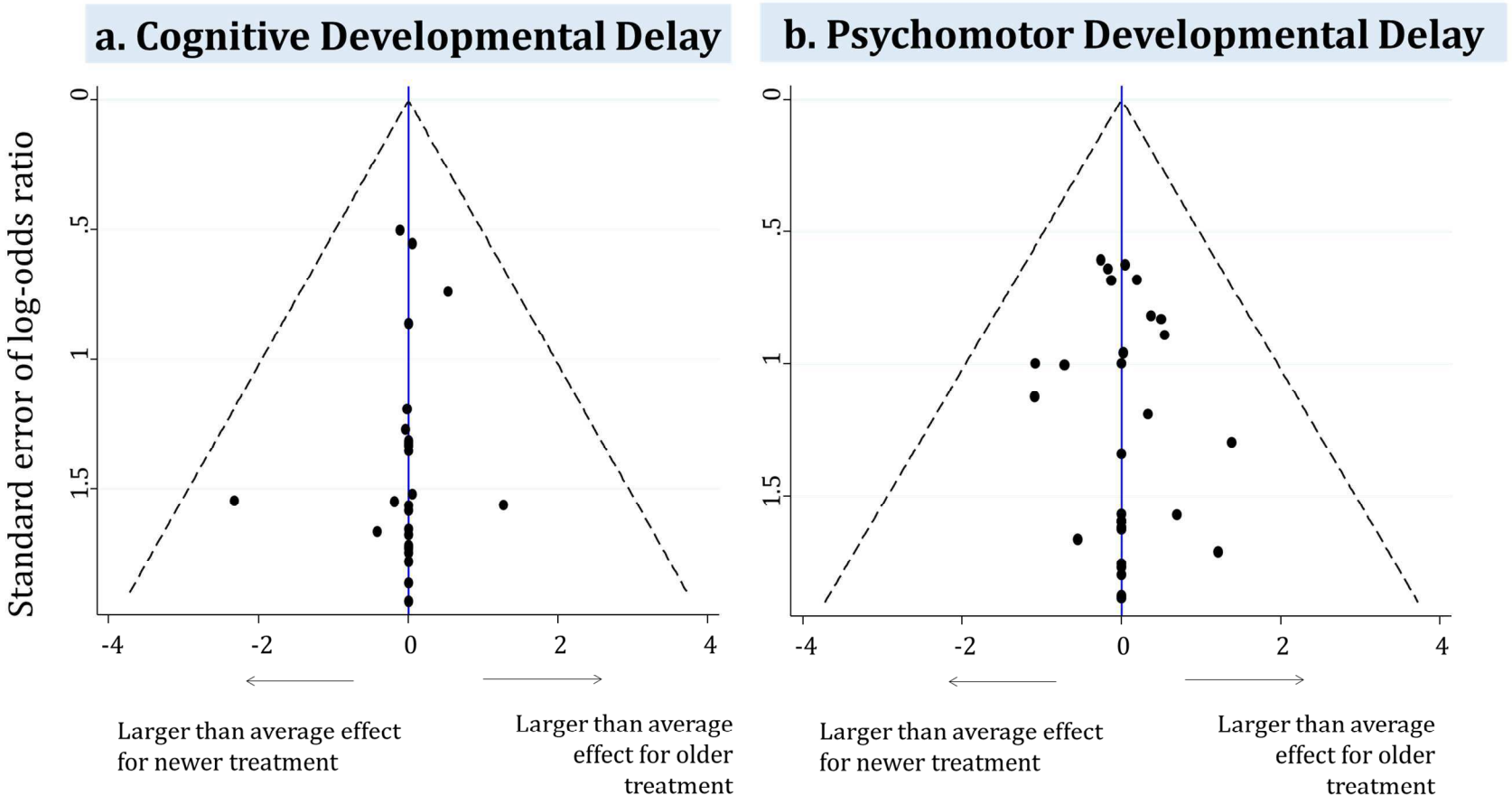
Appendix I. Methodological quality of observational studies – Newcastle Ottawa Scale results

First Author, Year	Representativeness of the exposed cohort	Selection of the non-exposed cohort	Ascertainment of exposure	Demonstration that outcome of interest was not present at start of study	Comparability of cohorts on the basis of the design or analysis	Assessment of outcome	Was follow-up long enough for outcomes to occur	Adequacy of follow up of cohorts
Adab, 2004[1]	B	A	A	A	C	A	A	C
Arkilo, 2015[4]	B	A	B	A	D	A	A	C
Bromley, 2010[5]	D	A	D	A	D	D	B	D
Bromley, 2013[6]	A	A	A	A	A	A	A	C
Christensen, 2013[8]	A	A	A	A	A	B	A	B
Cohen, 2013[9]	A	A	D	A	A	A	A	C
Cummings, 2011[10]	A	A	A	A	A	A	A	C
Dean, 2002[12]	B	A	A	A	D	A	A	C
D'Souza, 1991[14]	B	A	A	A	D	A	A	A
Eriksson, 2005[15]	B	A	A	A	B	A	A	D
Gaily, 1990[17]	B	A	A	A	D	A	A	A
Gogatishvili, 2014[21]	A	A	D	A	D	A	A	D

Gogatishvili, 2015[22]	A	A	D	A	D	A	A	D
Jones, 1989[23]	A	A	B	A	D	A	A	B
Katz, 2001[24]	B	A	A	A	D	A	A	D
Koch, 1996[25]	B	A	B	A	D	A	A	C
Lacroix, 2012[26]	A	A	A	A	A	A	A	A
Mawer, 2002[3]	B	A	A	A	D	A	A	B
Miskov, 2010[27]	D	A	D	A	D	D	A	D
Nadebaum, 2011[28]	A	A	A	A	A	A	A	B
Rihtman, 2013[29]	A	B	A	A	A	A	A	C
Scolnik, 1994[30]	B	A	A	A	D	A	A	A
Shankaran, 1996[31]	B	A	A	A	D	A	A	B
Van der Pol, 1991[32]	B	A	D	A	A	A	A	B
Veiby, 2013a[33]	A	A	A	A	A	A	A	D
Veiby, 2013b[34]	A	A	A	A	A	A	A	C
Wood, 2015[35]	A	A	A	A	D	A	A	C

Abbreviations: A – low risk; B – moderate risk; C – high risk; D – unclear risk

Appendix J. Comparison-adjusted funnel plots*



* Funnel plots have been produced only for outcomes with ≥10 studies. For multi-arm studies we plot data points from each study-specific basic parameter (treatment comparisons with a study-specific common comparator)

Appendix K. Statistically significant network meta-analysis results along with meta-analysis results, transitivity, and inconsistency assessments

Treatment Comparison	Number of Studies (Mean Baseline Risk)	Number of patients (Mean Age)	Treatment Indication	Timing	Comparability of cohorts	Adequacy of follow up of cohorts	MA Odds Ratio (95% CrI)	NMA Odds Ratio (95% CrI) (95% PrI)
Cognitive Developmental Delay (10 studies, 748 patients, 14 treatments)								
Carbam+Pheno+Valpro vs Control	NA	NR	NR	NR	NR	NR	NA	17.31 (1.03-434.50) (0.86-488.60)
Lamot vs Valpro	4 (NA)	140 (31.00)	Epilepsy	NR	H	H	0.17 (0.02-0.80)	0.13 (0.02-0.57) (0.01-0.79)
Valpro vs Carbam	1 (NA)	23 (31.00)	Epilepsy	NR	H	H	0.44 (0.00-11.07)	3.69 (1.72-7.63) (0.94-13.88)
Valpro vs Control	3 (0.06)	165 (28.80)	Epilepsy	NR	H	H	10.45 (3.42-33.73)	8.63 (3.01-25.74) (1.82-38.93)
Valpro vs Pheno	3 (NA)	36 (31.00)	Epilepsy	1st trimester	H	H	4.41 (0.79-38.91)	5.87 (1.26-42.27) (0.93-56.07)
Valpro vs Pheny	3 (NA)	58 (31.00)	Epilepsy	1st trimester	H	H	3.28 (0.81-14.38)	3.01 (1.06-9.18) (0.71-14.27)
Common between-study variance across treatment comparisons							0.13 (0.00-1.01)	0.15 (0.00-1.25) (NA)
Residual deviance: 40 Data points: 42								

Treatment Comparison	Number of Studies (Mean Baseline Risk)	Number of patients (Mean Age)	Treatment Indication	Timing	Comparability of cohorts	Adequacy of follow up of cohorts	MA Odds Ratio (95% CrI)	NMA Odds Ratio (95% CrI) (95% PrI)
DIC: 69								
Evaluation of consistency using the design-by-treatment interaction model				Chi-square test: 13.33 Degrees of Freedom: 16		P- value: 0.64 Heterogeneity: 0		
Autism Dyspraxia (5 studies, 2551 patients, 12 treatments)								
Lamot vs Control	2 (0.00)	254 (27.75)	Epilepsy	1st trimester	H	H	13.77 (2.06-188.00)	8.88 (1.29-112.00) (0.94-146.80)
Lamot+Valpro vs Carbam	1 (NA)	40 (NR)	Epilepsy	NR	L	L	15.02 (2.04-171.90)	22.89 (2.58-219.00) (1.90-282.20)
Lamot+Valpro vs Clonaz	NA	NR	NR	NR	NR	NR	NA	20.21 (1.48-351.30) (1.15-455.00)
Lamot+Valpro vs Control	NA	NR	NR	NR	NR	NR	NA	132.70 (7.41-3,9 x 10^3) (5.82-4.6 x 10^3)
Lamot+Valpro vs Lamot	NA	NR	NR	NR	NR	NR	NA	14.61 (1.51-149.10) (1.14-196.80)
Oxcar vs Control	NA	NR	NR	NR	NR	NR	NA	13.51 (1.28-221.40) (0.86-267.40)
Valpro vs Carbam	5 (NA)	1003 (27.83)	Epilepsy	1st trimest	L	L	3.20 (1.20-8.68)	3.02 (1.09-8.40) (0.57-

Treatment Comparison	Number of Studies (Mean Baseline Risk)	Number of patients (Mean Age)	Treatment Indication	Timing	Comparability of cohorts	Adequacy of follow up of cohorts	MA Odds Ratio (95% CrI)	NMA Odds Ratio (95% CrI) (95% PrI)
				er				14.31)
Valpro vs Control	2 (0.00)	249 (27.75)	Epilepsy	1st trimester	H	H	9.19 (1.14-132.10)	17.29 (2.40-217.60) (1.61-274.90)
Common between-study variance across treatment comparisons							0.12 (0.00-1.37)	0.16 (0.00-1.95) (NA)
Residual deviance: 24 Data points: 24 DIC: 44								
Evaluation of consistency using the design-by-treatment interaction model				Chi-square test: 3.79 Degrees of Freedom: 5		P- value: 0.57 Heterogeneity: 0		
Psychomotor Developmental Delay (11 studies, 1145 patients, 18 treatments)								
Carbam+Pheno+Valpro vs Control	NA	NR	NR	NR	NR	NR	NA	19.12 (1.49-337.50) (1.34-370.40)
Carbam+Pheno+Valpro vs Pheno	NA	NR	NR	NR	NR	NR	NA	19.86 (1.38-393.60) (1.26-423.30)
Levet vs Carbam+Pheno+Valpro	NA	NR	NR	NR	NR	NR	NA	0.01 (0.00-0.58) (0.00-0.62)
Valpro vs Carbam	7 (NA)	331 (27.80)	Epilepsy	1st trimester	H	H	2.72 (1.39-5.67)	2.45 (1.27-4.88) (0.95-6.77)
Valpro vs Control	5 (0.07)	331	Epilepsy	1st	H	H	3.53 (1.60-	4.16

Treatment Comparison	Number of Studies (Mean Baseline Risk)	Number of patients (Mean Age)	Treatment Indication	Timing	Comparability of cohorts	Adequacy of follow up of cohorts	MA Odds Ratio (95% CrI)	NMA Odds Ratio (95% CrI) (95% PrI)
		(28.38)		trimester			8.64)	(2.04-8.75) (1.52-12.05)
Valpro vs Pheno	2 (NA)	141 (NR)	Epilepsy	1st trimester	H	H	3.68 (1.17-12.30)	4.32 (1.72-11.20) (1.34-14.51)
Common between-study variance across treatment comparisons							0.05 (0.00-0.49)	0.06 (0.00-0.63) (NA)
Residual deviance: 45 Data points: 51 DIC: 78								
Evaluation of consistency using the design-by-treatment interaction model				Chi-square test: 13.46 Degrees of Freedom: 21		P- value: 0.89 Heterogeneity: 0		
Language Delay (5 studies, 509 patients, 5 treatments)								
Valpro vs Control	1 (0.03)	173 (28.90)	Epilepsy	NR	L	H	6.96 (1.14-37.03)	7.95 (1.50-49.13) (0.96-74.52)
Common between-study variance across treatment comparisons							0.15 (0.00-1.85)	0.16 (0.00-2.15) (NA)
Residual deviance: 12 Data points: 14 DIC: 23								
Evaluation of consistency using the design-by-treatment interaction model				Chi-square test: 2.33 Degrees of Freedom: 3		P- value: 0.50 Heterogeneity: 0		

Treatment Comparison	Number of Studies (Mean Baseline Risk)	Number of patients (Mean Age)	Treatment Indication	Timing	Comparability of cohorts	Adequacy of follow up of cohorts	MA Odds Ratio (95% CrI)	NMA Odds Ratio (95% CrI) (95% PrI)
ADHD (4 studies, 750 patients, 6 treatments)								
Residual deviance:12 Data points: 17 DIC: 22								

Abbreviations: ADHD - Attention Deficit Hyperactivity Disorder; CrI - Credible Interval; DIC - Deviance Information Criterion; H-high risk of bias; L - low risk of bias; MA - Meta-analysis; NA - Not applicable; NMA - Network Meta-analysis; NR- Not Reported; PrI - Predictive Interval; ROB - Risk of Bias

Carbam = Carbamazepine; Clobaz = Clobazam; Clonaz = Clonazepam; Ethos = Ethosuximide; Gabap = Gabapentin; Lamot = Lamotrigine; Levet = Levetiracetam; Oxcar = Oxcarbazepine; Pheno = Phenobarbital; Pheny = Phenytoin; Pridmid = Primidone; Topir = Topiramate; Valpro = Valproate; Vigab = Viagabatrin

Appendix L. Frequencies, events and samples sizes, SUCRA values, and total group risks per treatment and outcome

Treatment	Frequency of treatment in network	Total # events/Total Sample size	Median SUCRA (95% CrI)	Total Group Risk Median (IQR)
Cognitive Developmental Delay				
Carbamazepine+Levetiracetam	1	0/2	0.86 (0.07-1.00)	0.00 (0.00-0.00)
Carbamazepine+Phenobarbital	1	0/3	0.86 (0.14-1.00)	0.00 (0.00-0.00)
Control	4	8/125	0.79 (0.43-0.93)	0.04 (0.00-0.09)
Lamotrigine	4	0/43	0.71 (0.29-1.00)	0.00 (0.00-0.00)
Phenobarbital+Phenytoin	1	0/15	0.71 (0.07-1.00)	0.00 (0.00-0.00)
Phenobarbital	3	1/12	0.64 (0.21-1.00)	0.00 (0.00-0.33)
Carbamazepine	9	29/238	0.50 (0.29-0.79)	0.15 (0.05-0.20)
Primidone	2	2/13	0.50 (0.14-0.93)	0.09 (0.00-0.18)
Phenytoin	5	11/111	0.43 (0.21-0.79)	0.10 (0.00-0.21)
Phenytoin+Valproate	1	0/5	0.36 (0.00-1.00)	0.00 (0.00-0.00)
Carbamazepine+Phenobarbital+Phenytoin	1	0/4	0.29 (0.00-1.00)	0.00 (0.00-0.00)
Ethosuximide+Phenytoin	1	0/3	0.21 (0.00-1.00)	0.00 (0.00-0.00)
Valproate	7	50/160	0.21 (0.00-0.43)	0.35 (0.25-0.50)
Carbamazepine+Phenytoin	1	1/11	0.14 (0.00-0.79)	0.09 (0.09-0.09)
Carbamazepine+Phenobarbital+Valproate	1	2/3	0.07 (0.00-0.71)	0.67 (0.67-0.67)
Autism/Dyspraxia				
Control	2	1/180	0.91 (0.55-1.00)	0.00 (0.00-0.01)
Levetiracetam	1	0/11	0.73 (0.09-1.00)	0.00 (0.00-0.00)
Carbamazepine	5	8/518	0.64 (0.36-	0.03 (0.01-

Treatment	Frequency of treatment in network	Total # events/Total Sample size	Median SUCRA (95% CrI)	Total Group Risk Median (IQR)
			0.91)	0.06)
Carbamazepine+Lamotrigine	1	0/5	0.64 (0.00-1.00)	0.00 (0.00-0.00)
Clonazepam	1	3/269	0.64 (0.18-0.91)	0.01 (0.01-0.01)
Carbamazepine+Phenytoin	1	0/3	0.55 (0.00-1.00)	0.00 (0.00-0.00)
Phenytoin	1	0/5	0.55 (0.00-1.00)	0.00 (0.00-0.00)
Carbamazepine+Clonazepam	1	0/3	0.45 (0.00-1.00)	0.00 (0.00-0.00)
Lamotrigine	4	14/745	0.45 (0.18-0.82)	0.04 (0.01-0.08)
Oxcarbazepine	1	7/321	0.36 (0.09-0.82)	0.02 (0.02-0.02)
Valproate	5	21/485	0.27 (0.09-0.55)	0.05 (0.03-0.08)
Lamotrigine+Valproate	1	3/6	0.00 (0.00-0.27)	0.50 (0.50-0.50)
Neonatal Seizure				
Lamotrigine	1	3/42	NA	0.07 (NA)
Valproate	1	0/30	NA	0.00 (NA)
Psychomotor Developmental Delay				
Levetiracetam	1	0/11	0.94 (0.29-1.00)	0.00 (0.00-0.00)
Phenobarbital+Phenytoin	1	0/15	0.82 (0.12-1.00)	0.00 (0.00-0.00)
Carbamazepine+Phenytoin	1	0/11	0.76 (0.06-1.00)	0.00 (0.00-0.00)
Control	8	21/323	0.76 (0.53-0.94)	0.06 (0.03-0.10)
Phenobarbital	4	11/117	0.76 (0.47-0.94)	0.07 (0.02-0.17)
Carbamazepine	10	32/249	0.59 (0.35-0.82)	0.10 (0.06-0.16)
Carbamazepine+Phenobarbital	2	3/15	0.59 (0.18-0.94)	0.13 (0.00-0.25)
Lamotrigine	4	11/126	0.53 (0.24-0.82)	0.09 (0.06-0.12)
Clonazepam	1	4/27	0.47 (0.12-0.88)	0.15 (0.15-0.15)

Treatment	Frequency of treatment in network	Total # events/Total Sample size	Median SUCRA (95% CrI)	Total Group Risk Median (IQR)
Phenytoin+Valproate	1	0/5	0.47 (0.00-1.00)	0.00 (0.00-0.00)
Carbamazepine+Phenobarbital+Phenytoin	1	0/4	0.41 (0.00-1.00)	0.00 (0.00-0.00)
Clobazam	1	1/6	0.41 (0.00-0.94)	0.17 (0.17-0.17)
Phenytoin	3	10/83	0.41 (0.12-0.71)	0.04 (0.00-0.33)
Ethosuximide+Phenytoin	1	0/3	0.35 (0.00-1.00)	0.00 (0.00-0.00)
Topiramate	2	1/6	0.29 (0.00-0.88)	0.13 (0.00-0.25)
Valproate	7	36/137	0.24 (0.06-0.53)	0.28 (0.11-0.38)
Gabapentin	2	1/4	0.12 (0.00-0.76)	0.25 (0.00-0.50)
Carbamazepine+Phenobarbital+Valproate	1	2/3	0.06 (0.00-0.59)	0.67 (0.67-0.67)
Language Delay				
Control	2	17/209	0.75 (0.50-1.00)	0.13 (0.03-0.24)
Phenobarbital	1	10/41	0.75 (0.00-1.00)	0.24 (0.24-0.24)
Carbamazepine	4	17/117	0.50 (0.00-0.75)	0.15 (0.06-0.25)
Lamotrigine	3	6/59	0.50 (0.00-1.00)	0.00 (0.00-0.14)
Valproate	4	21/83	0.00 (0.00-0.50)	0.22 (0.12-0.35)
ADHD				
Phenytoin	2	2/41	1.00 (0.20-1.00)	0.05 (0.04-0.06)
Control	3	6/218	0.80 (0.20-1.00)	0.03 (0.00-0.05)
Phenobarbital	1	4/61	0.60 (0.00-1.00)	0.07 (0.07-0.07)
Lamotrigine	3	7/105	0.40 (0.00-0.80)	0.07 (0.00-0.13)
Carbamazepine	4	17/182	0.20 (0.00-0.60)	0.09 (0.04-0.13)
Valproate	4	12/143	0.20 (0.00-	0.08 (0.03-

Treatment	Frequency of treatment in network	Total # events/Total Sample size	Median SUCRA (95% CrI)	Total Group Risk Median (IQR)
			0.80)	0.16)
Social Impairment				
Carbamazepine	1	6/48	NA	0.13 (NA)
Control	1	37/276	NA	0.13 (NA)
Lamotrigine	1	9/71	NA	0.13 (NA)
Valproate	1	1/27	NA	0.04 (NA)

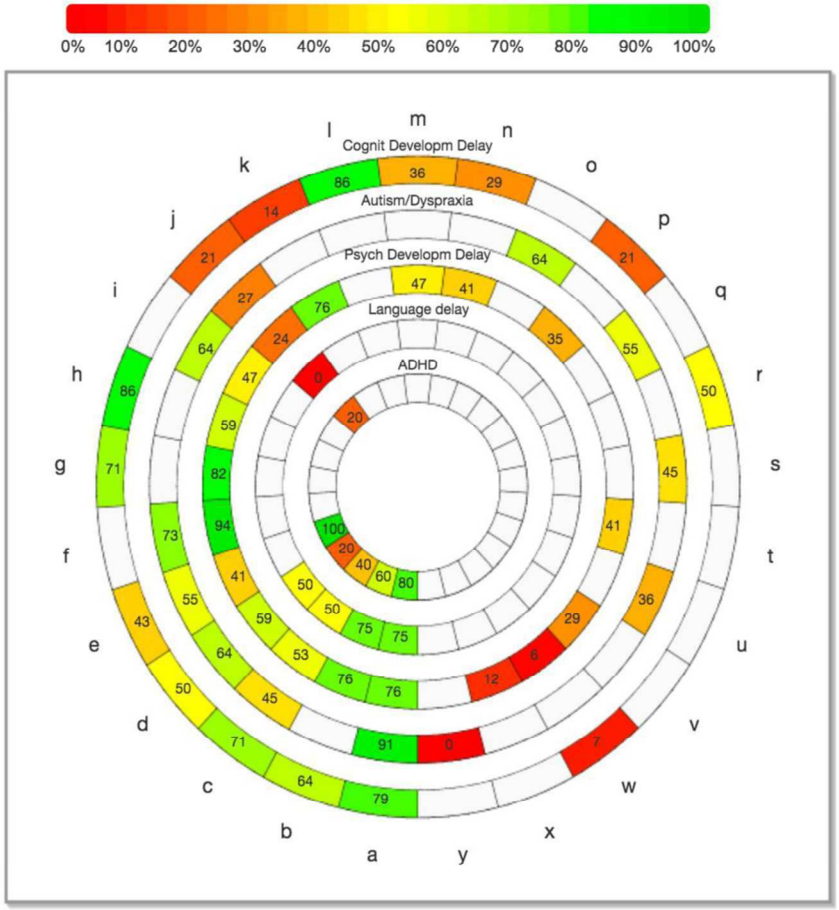
Abbreviations: ADHD - attention deficit hyperactivity disorder; CrI - Credible Interval; IQR - interquartile range; NA - Not applicable; SUCRA - surface under the cumulative ranking curve

Carbam = Carbamazepine; Clobaz = Clobazam; Clonaz = Clonazepam; Ethos = Ethosuximide; Gabap = Gabapentin; Lamot = Lamotrigine; Levet = Levetiracetam; Oxcar = Oxcarbazepine; Pheno = Phenobarbital; Pheny = Phenytoin; Pridmid = Primidone; Topir = Topiramate; Valpro = Valproate; Vigab = Viagabatrin

Appendix M. Rank-heat plot of cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcomes*

Treatments
a: Control
b: Pheno
c: Lamot
d: Carbam
e: Pheny
f: Levet
g: Pheno+Pheny
h: Carbam+Pheno
i: Clonaz
j: Valpro
k: Carbam+Pheny
l: Carbam+Levet
m: Pheny+Valpro
n: Carbam+Pheno+Pheny
o: Carbam+Lamot
p: Ethos+Pheny
q: Carbam+Pheny
r: Primid
s: Carbam+Clonaz
t: Clobaz
u: Oxcar
v: Topir
w: Carbam+Pheno+Valpro
x: Gabap
y: Lamot+Valpro

Outcomes
Circles from outside in refer to :
1st: Cognit Developm Delay
2nd: Autism/Dyspraxia
3rd: Psych Developm Delay
4th: Language delay
5th: ADHD
White sectors refer to treatments without data on the outcome within the circle.



Abbreviations: carbam - carbamazepine, clobaz - clobazam, clonaz - clonazepam, ethos - ethosuximide, gabap - gabapentin, lamot - lamotrigine, levet - levetiracetam, oxcar - oxcarbazepine, pheno - phenobarbital, pheny - phenytoin, primid - primidone, topir - topiramate, valpro - valproate, vigab - vigabatrin

*Rank-heat plot of cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcomes (5 circles) and 25 treatments (25 radii). Each sector is coloured according to the surface under the cumulative ranking curve value of the corresponding treatment and outcome using the transformation of three colours red (0%), yellow (50%), and green (100%).

Appendix N. Number of studies and treatments per outcome

Total studies	Range of study arms	# of treatments	# of patients	# of direct treatment comparisons	# of NMA treatment comparisons	Statistically significant NMA treatment effects	# of studies with zero events in all arms	# of studies with ineligible outcome definition*
Cognitive Developmental Delay								
10	(2,8)	14	748	53	105	6	0	5
Autism/Dyspraxia								
5	(4,6)	12	2551	34	66	8	0	4
Neonatal Seizure								
1	(2,2)	2	69	1	0	0	1	1
Psychomotor Developmental Delay								
11	(2,8)	18	1145	74	153	6	0	5
Language Delay								
5	(2,4)	5	509	7	10	1	0	3
ADHD								
4	(4,5)	6	750	14	15	0	0	0
Social Impairment								
1	(4,4)	4	422	1	0	0	0	0

Abbreviations: ADHD - Attention Deficit Hyperactivity Disorder; NMA - Network Meta-analysis
*See Appendix A for outcome definitions

BMJ Open

Comparative safety of anti-epileptic drugs for neurological development in children exposed during pregnancy and breastfeeding: a systematic review and network meta-analysis

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-017248.R1
Article Type:	Research
Date Submitted by the Author:	25-May-2017
Complete List of Authors:	Veroniki, Areti Angeliki; Li Ka Shing Knowledge Institute, St. Michael's Hospital Rios, Patricia; Li Ka Shing Knowledge Institute, St. Michael's Hospital Cogo, Elise; Li Ka Shing Knowledge Institute, St. Michael's Hospital Straus, Sharon; Li Ka Shing Knowledge Institute, St. Michael's Hospital; University of Toronto, Department of Medicine Finkelstein, Yaron; The Hospital for Sick Children; University of Toronto, Department of Paediatrics Kealey, M.; Li Ka Shing Knowledge Institute, St. Michael's Hospital Reynen, Emily; Li Ka Shing Knowledge Institute, St. Michael's Hospital Soobiah, Charlene; Li Ka Shing Knowledge Institute, St. Michael's Hospital; University of Toronto, Institute for Health Policy Management & Evaluation Thavorn, Kednapa; University of Ottawa, School of Epidemiology, Public Health and Preventive Medicine, Faculty of Medicine; The Ottawa Hospital Research Institute, Clinical Epidemiology Program Hutton, Brian; University of Ottawa, School of Epidemiology, Public Health and Preventive Medicine, Faculty of Medicine; Ottawa Hospital Research Institute, Center for Practice Changing Research Hemmelgarn, BR; University of Calgary, Departments of Medicine and Community Health Sciences Yazdi, Fatemeh; Li Ka Shing Knowledge Institute, St. Michael's Hospital D'Souza, Jennifer; Li Ka Shing Knowledge Institute, St. Michael's Hospital MacDonald, Heather; Li Ka Shing Knowledge Institute, St. Michael's Hospital Tricco, Andrea; Li Ka Shing Knowledge Institute, St. Michael's Hospital; University of Toronto, Epidemiology Division, Dalla Lana School of Public Health
Primary Subject Heading:	Neurology
Secondary Subject Heading:	Obstetrics and gynaecology
Keywords:	multiple treatment meta-analysis, knowledge synthesis, Epilepsy < NEUROLOGY, pregnancy, infants, developmental delay

Comparative safety of anti-epileptic drugs for neurological development in children exposed during pregnancy and breastfeeding: a systematic review and network meta-analysis

Areti Angeliki Veroniki, PhD, MSc[1] Email: VeronikiA@smh.ca

Patricia Rios, MSc[1] Email: RiosP@smh.ca

Elise Cogo, ND, MLIS[1] Email: CogoE@smh.ca

Sharon E. Straus, MD, MSc[1,2] Email: Sharon.straus@utoronto.ca

Yaron Finkelstein, MD[3,4,5] Email: Yaron.Finkelstein@sickkids.ca

Ryan Kealey, PhD[1] Email: ryan.kealey@utoronto.ca

Emily Reynen, MD, CM, PharmD[1] Email: ereynen@gmail.com

Charlene Soobiah, PhD (Cand.)[1,6] Email: SoobiahC@smh.ca

Kednapa Thavorn, PhD[7,8,9] Email: kthavorn@ohri.ca

Brian Hutton, PhD, MSc[7,10] Email: bhutton@ohri.ca

Brenda R. Hemmelgarn, MD, PhD[11] Email: Bhemmelg@ucalgary.ca

Fatemeh Yazdi, MSc[1] Email: SabaghYazdiF@smh.ca

Jennifer D'Souza, HBSc[1] Email: jennifer.dsouza@mail.utoronto.ca

Heather MacDonald, MSc[1] Email: hrmacdonald@gmail.com

Andrea C. Tricco, PhD, MSc[1,12,*] Email: TriccoA@smh.ca

AUTHOR DETAILS

[1] Li Ka Shing Knowledge Institute, St. Michael's Hospital, 209 Victoria Street, East Building, Toronto, Ontario, M5B 1W8, Canada

[2] Department of Medicine, University of Toronto, 27 King's College Circle, Toronto, Ontario, M5S 1A1, Canada

[3] The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario, M5G 1X8, Canada

[4] Department of Paediatrics, University of Toronto, 172 St. George Street, Toronto, Ontario, M5R 0A3, Canada

[5] Department of Pharmacology and Toxicology, University of Toronto, Medical Sciences Building, Room 4207, 1 King's College Circle, Toronto, Ontario, M5S 1A8, Canada

[6] Institute for Health Policy Management & Evaluation, University of Toronto, 4th Floor, 155 College Street, Toronto, Ontario, M5T 3M6, Canada

[7] School of Epidemiology, Public Health and Preventive Medicine, Faculty of Medicine, University of Ottawa, Roger-Guindon Building, 451 Smyth Road, Ottawa, Ontario, K1H 8M5, Canada

[8] Clinical Epidemiology Program, Ottawa Hospital Research Institute, The Ottawa Hospital, 501 Smyth Road, Ottawa, Ontario, K1H 8L6, Canada

[9] Institute of Clinical and Evaluative Sciences (ICES uOttawa), 1053 Carling Avenue, Ottawa, Ontario, K1Y 4E9, Canada

[10] Ottawa Hospital Research Institute, Center for Practice Changing Research, The Ottawa Hospital–General Campus, 501 Smyth Road, PO Box 201B, Ottawa, Ontario, K1H 8L6, Canada.

[11] Departments of Medicine and Community Health Sciences, University of Calgary, TRW Building, 3rd Floor, 3280 Hospital Drive NW, Calgary, Alberta, T2N 4Z6, Canada

[12] Epidemiology Division, Dalla Lana School of Public Health, University of Toronto, 6th Floor, 155 College Street, Toronto, Ontario, M5T 3M7, Canada

***Corresponding author**

Prof. Andrea C. Tricco, PhD

Scientist, Knowledge Translation Program,

Li Ka Shing Knowledge Institute, St. Michael's Hospital,

209 Victoria Street, East Building, Toronto, Ontario, M5B 1W8, Canada

Phone: 416-864-6060, Fax: 416-864-5805, Email: TriccoA@smh.ca

Keywords: multiple treatment meta-analysis, knowledge synthesis, epilepsy, pregnancy, infants, developmental delay.

Word count: abstract (300 words); main text (4,000 words); 2 tables; 3 figures; 3 additional files; 51 references

ABSTRACT

Objectives: Compare the safety of anti-epileptic drugs (AEDs) on neurodevelopment of infants/children exposed *in-utero* or during breastfeeding.

Design and Setting: Systematic review and Bayesian random-effects network meta-analysis (NMA). Medline, EMBASE, and the Cochrane Central Register of Controlled Trials were searched until April 27th, 2017. Screening, data abstraction, and quality appraisal were completed in duplicate by independent reviewers.

Participants: 29 cohort studies including 5,100 infants/children.

Interventions: Mono- and poly-therapy AEDs including first-generation (carbamazepine, clobazam, clonazepam, ethosuximide, phenobarbital, phenytoin, primidone, valproate) and newer-generation (gabapentin, lamotrigine, levetiracetam, oxcarbazepine, topiramate, vigabatrin) AEDs. Epileptic women who did not receive AEDs during pregnancy or breastfeeding served as the control group.

Primary and secondary Outcome measures: Cognitive developmental delay and autism/dyspraxia were primary outcomes. Attention deficit hyperactivity disorder, language delay, neonatal seizures, psychomotor developmental delay, and social impairment were secondary outcomes.

Results: The NMA on cognitive developmental delay (11 cohort studies, 933 children, 18 treatments) suggested among all AEDs only valproate was statistically significantly associated with more children experiencing cognitive developmental delay when compared with control (odds ratio (OR)=7.40, 95% credible interval (CrI): 3.00-18.46). The NMA on autism (5 cohort studies, 2,551 children, 12 treatments), suggested that oxcarbazepine (OR=13.51, CrI: 1.28-221.40), valproate (N=485, OR=17.29, 95% CrI: 2.40-217.60),

lamotrigine (OR=8.88, CrI: 1.28-112.00), and lamotrigine+valproate (OR=132.70, CrI: 7.41-3,851.00) were associated with significantly greater odds of developing autism compared with control. The NMA on Psychomotor developmental delay (11 cohort studies, 1,145 children, 18 treatments) found that valproate (OR=4.16, CrI: 2.04-8.75) and carbamazepine+phenobarbital+valproate (OR=19.12, CrI: 1.49-337.50) were associated with significantly greater odds of psychomotor delay compared with control.

Conclusions: Valproate alone or combined with another AED is associated with the greatest odds of adverse neurodevelopmental outcomes compared with control. Oxcarbazepine and lamotrigine were associated with increased occurrence of autism. Counselling is advised for women considering pregnancy to tailor the safest regimen.

Registration: PROSPERO database (CRD42014008925).

Keywords: multiple treatment meta-analysis, knowledge synthesis, epilepsy, pregnancy, infants, developmental delay.

ARTICLE SUMMARY

Strengths and limitations of this study

- 29 cohort studies involving 5,100 children of women who took AEDs were included in this systematic review. More evidence from long-term follow-up studies is required.
- This study was the first that compared and ranked the safety of AEDs, including comparative safety of treatments that have not been directly compared.

- Across all neurological outcomes and treatments compared with control, valproate alone or combined with another AED is associated with the greatest odds of adverse development.
- Oxcarbazepine and lamotrigine were associated with increased occurrence of autism.

INTRODUCTION

Anti-epileptic drugs (AEDs) are used by pregnant women for various conditions, such as epilepsy, pain syndromes, psychiatric disorders, and chronic migraine.[1] AED use during pregnancy is associated with risks to the fetus, as these drugs can cross the placenta or may be transferred to the infant through breastfeeding and may be associated with adverse neurodevelopment outcomes.[2-4] Two systematic reviews examined the association between AED exposure and neurodevelopment *in utero*, and reported that exposure to valproate was linked to significantly lower IQ scores and poorer overall neurodevelopmental outcomes in the children of women who used these medications.[5 6] No significant associations were found between neurodevelopment and exposure to other AEDs such as carbamazepine, lamotrigine, or phenytoin.[5-8] However, there is a lack of sufficiently powered studies to assess the impact of AEDs on neurodevelopment in children of women exposed to these agents, especially for newer generation drugs, thus highlighting the need for a systematic review.[9 10]

The aim of this study was to compare the safety of AEDs and assess their impact on neurodevelopment in infants and children exposed *in-utero* or during breastfeeding, employing a systematic review and network meta-analysis (NMA).

METHODS

The methods are briefly described here; details can be found in the published protocol (Additional File 1).[11] This study was registered with PROSPERO (CRD42014008925). We followed the ISPOR[12] guidelines for our NMA, and reported our findings using the PRISMA extension for NMA (Additional File 2).[13]

Eligibility criteria

All randomized clinical trials (RCTs), quasi-RCTs, and observational studies were eligible. Included studies assessed infants or children ≤12 years of age whose mothers consumed AEDs during pregnancy and/or while breastfeeding. Both mono- and poly-therapy AEDs were eligible, including first-generation (i.e., carbamazepine, clobazam, clonazepam, ethosuximide, phenobarbital, phenytoin, primidone, valproate) and newer-generation (i.e., marketed >1990: gabapentin, lamotrigine, levetiracetam, oxcarbazepine, topiramate, vigabatrin), with no restrictions on AED dosage. Placebo, no AED, other AEDs alone or in combination, were considered as comparators. Duplicate studies that used the same registry or population sample (i.e., companion studies) were used for supplementary information only. No language or other restrictions were imposed.

The primary neurological outcomes were cognitive developmental delay and autism/dyspraxia, and the secondary outcomes included attention deficit hyperactivity disorder (ADHD), language delay, neonatal seizures, psychomotor developmental delay, and social impairment. Table 2 shows the outcome measures and diagnostic scales used.

We initially intended to evaluate all safety outcomes in infants/ children exposed to AEDs *in-utero* or during breastfeeding in one publication, but given the breadth of evidence we

identified, we report results related to risk of major congenital malformations, birth, and prenatal outcomes in a companion paper.[14]

Information sources

An experienced librarian executed search strategies for MEDLINE, EMBASE, and the Cochrane Central Register of Controlled Trials up to March 18, 2014, and then updated the search in April 27th 2017. The search strategy for MEDLINE was peer-reviewed by another librarian using the PRESS checklist,[15] and is available in the protocol.[11] Additional studies were identified by scanning references and contacting authors. Unpublished studies were sought by searching clinical trial registries and conference abstracts.

Study selection and data collection

After a calibration exercise, titles/abstracts (level 1) and full-text papers (level 2) were screened by two reviewers independently. Upon completion of level 1, 6% of citations were discrepant between reviewer pairs, whereas at the conclusion of level 2, 16% of articles were discrepant. Conflicts were resolved through discussion or by a third reviewer. The same approach was used for data abstraction and appraisal of methodological quality. Three rounds of pilot testing were conducted prior to data abstraction to train reviewers and refine the data abstraction form. For studies published in the last 10 years, authors were contacted to request clarification or additional data.

Appraisal of methodological quality

Only observational studies were identified and included for analysis, and their methodological quality was appraised with the Newcastle-Ottawa Scale (NOS) (Additional File 3: Appendix A).[16] For each outcome with ≥10 studies, the comparison-adjusted funnel plot was used to assess small-study effects,[17] where the overall treatment effect for each

comparison was estimated under the fixed-effect meta-analysis model. All eligible medications were ordered from oldest to newest using their international market approval dates. Hence, the comparison-adjusted funnel plot additionally assesses the hypothesis that newer AEDs are favoured over older ones. To overcome some of the correlations induced by multi-arm studies, which may cause overestimation and mask funnel plot asymmetry, we plotted data points corresponding to the study-specific basic parameters (treatment comparisons with common comparator). In each study, we used the control group as the common comparator or if this was missing, we used the oldest treatment comparator against the remaining AEDs.

Synthesis of included studies

We used the odds ratio (OR) for each dichotomous outcome, and outcome data were pooled using hierarchical meta-analysis and NMA models and the Markov Chain Monte Carlo sampling method in a Bayesian framework. To account for anticipated methodological and clinical heterogeneity across studies, and to achieve the highest generalizability in the meta-analytical treatment effects, we applied a random-effects model.[18]

A NMA was applied for connected evidence networks and pre-specified treatment nodes.[19] We assessed the transitivity assumption for each outcome *a priori* using the effect modifiers: age, baseline risk, treatment indication, timing, and methodological quality. The mean of each continuous effect modifier and the mode of each categorical effect modifier for each pairwise comparison were presented in tables for each outcome.[20] The consistency assumption was evaluated for the entire network of each outcome using the random-effects design-by-treatment interaction model when multiple studies were available in

each network design or the fixed-effect design-by-treatment interaction model when a single study informed each network design.[21] If inconsistency was identified, further examination for local inconsistency in parts of the network was completed using the loop-specific method.[22 23] Common within-network between-study variance (τ^2) across treatment comparisons was assumed in the meta-analysis, NMA, and design-by-treatment interaction model, so that treatment comparisons including a single study can borrow strength from the remaining network. This assumption was clinically reasonable, as the treatments included were of the same nature. In the loop-specific approach, common within-loop τ^2 was assumed.

For cognitive developmental delay and autism/dyspraxia outcomes, network meta-regression analyses for maternal age and baseline risk (i.e., using the control group) were conducted, when ≥10 studies provided relevant information, assuming a common fixed coefficient across treatment comparisons for AEDs vs. control. Sensitivity analyses for cognitive developmental delay and autism/dyspraxia outcomes were performed for treatment indication of epilepsy, large study size (i.e., >300), maternal alcohol intake, maternal tobacco use, only first-generation AEDs, and methodological quality. The sensitivity analysis for methodological quality was restricted to studies with low risk of bias for the two items on the NOS where the greatest proportion of studies received a low-quality score: adequacy of follow-up of cohorts and comparability of cohorts. For autism/dyspraxia, a sensitivity analysis on maternal IQ/psychiatric history was additionally conducted. We measured the goodness of fit using the posterior mean of the residual deviance, the degree of τ^2, and the deviance information criterion (DIC). In a well-fitting model the posterior mean residual deviance should be close to the number of data

points.[24 25] A difference of 3 units in the DIC between a NMA and a network meta-regression model was considered important and the lowest value of the DIC corresponded to the model with the best fit.[24 25]

All analyses were conducted in OpenBUGS[26] assuming non-informative priors for all model parameters, and $\tau \sim N(0,1)$, $\tau > 0$. The first 10,000 iterations were discarded and then 100,000 simulations were run with thinning of 10 values. Convergence was checked by visual inspection of the evaluation of the mixing of two chains. The median and 95% CrI were calculated for each parameter value. The *network* command[27] was used to apply the design-by-treatment interaction model.

For NMA estimates, a 95% predictive interval (PrI) is also reported to capture the magnitude of τ^2 and present the interval within which the treatment effect of a future study is expected to lie.[28 29] The estimated safety of the included AEDs was ranked using the surface under the cumulative ranking (SUCRA) curve.[30] The larger the SUCRA for a treatment, the higher its safety rank among all the available treatment options. SUCRA values are presented along with 95% CrIs to capture the uncertainty in the parameter values.[31]

RESULTS

Literature search and included studies

Our literature search identified 5,707 titles and abstracts, which after the screening process yielded 681 articles potentially relevant for inclusion (Figure 1). After full-text review, 95 studies fulfilled eligibility criteria along with 17 studies identified through supplemental methods. Of the 112 total eligible studies in the complete review,[14] 29 articles with seven companion reports and two potentially overlapping registry studies included one or more relevant neurological outcomes (Additional File 3: Appendix B). Four of the studies included in this analysis were conference abstracts with usable data,[32-35] and four studies,[36-39] not captured in the original literature search, were identified through reference scanning. A table with the key excluded studies and a rationale for their exclusion is presented in Additional File 3: Appendix C.

Study and patient characteristics

We included 29 cohort studies (5,100 patients) published between 1989 and 2016 (Table 1; Additional File 3: Appendix D, E). The number of patients included in each study ranged from 23 to 2,011 (median 74.5). Most studies (76%) were published after 2000, 62% of the studies included fewer than 100 patients, and the 52% of the studies included a control group of pregnant/breastfeeding women with epilepsy who did not receive AEDs. The mean maternal age ranged from 24 to 34 years. About half of the studies (52%) were funded through government/public research funding.

Methodological quality results

Twenty-nine observational studies were appraised using the NOS (Additional File 3: Appendix F). Overall, the studies were of good methodological quality and were rated as

high quality across most items: 28 studies (97%) selected the non-exposed cohort from the same community as the exposed cohort, 26 (90%) included a representative or somewhat representative sample, 27 (93%) assessed outcomes independently, with blinding, or via a record linkage (e.g., identified through database records), and 23 (79%) ascertained exposure via secured records (e.g., database records) or structured interviews. The comparability of cohorts and adequacy of follow-up were the lowest scoring items across the studies with only 12 (41%) and 10 (34%) studies rated as high quality on these items. No evidence for small-study effects was identified by the visual inspection of the comparison-adjusted funnel plots (Additional File 3: Appendix G).

Statistical analysis results

No important concerns were raised regarding the violation of the transitivity assumption when maternal age, baseline risk, treatment indication, and timing were assessed (Additional File 3: Appendix H). However, the average methodological quality appraisal across treatment comparisons varied across treatment comparisons. The evaluation of the consistency assumption using the design-by-treatment interaction model suggested that there was no evidence of significant inconsistency across all outcomes (Additional File 3: Appendix H).

In the following sections, we present the significant NMA results by outcome for AEDs compared with control (i.e., no exposure to AEDs), while the SUCRA values from all outcomes are presented in Figure 2 and depicted in a rank-heat plot (http://rh.ktss.ca/)[40] in Additional File 3: Appendix I.

Cognitive developmental delay

The NMA for cognitive developmental delay (definitions in Table 1) included 11 cohort studies, 933 children, and examined 18 treatments (Figure 3a; Additional File 3: Appendix J; τ^2=0.12, 95% CrI: 0.00-1.15). One study included children exposed to AEDs both *in-utero* and through breastfeeding, and ten included children exposed to AEDs *in-utero*. Across all AEDs, only valproate was associated with significantly increased odds of cognitive developmental delay when compared with control (odds ratio (OR)=7.40, 95% credible interval (CrI): 3.00-18.46; Figure 2a; Additional File 3: Appendix H).

The same results were observed in a network meta-regression of baseline risk for offspring of women with epilepsy who were not exposed to AEDs (estimated regression coefficient on OR scale: 1.01, 95% CrI: 0.76-1.56; τ^2=0.16, 95% CrI: 0.00-1.24; residual deviance= 45.27, data points= 47, DIC= 80.17). Similarly, the sensitivity analyses restricted to: a) studies that only included women receiving AEDs to treat epilepsy (10 studies, 910 children, 17 treatments; τ^2=0.16, 95% CrI: 0.00-1.36), b) studies comparing only first-generation AEDs (6 studies, 480 children, 13 treatments; τ^2=0.28, 95% CrI: 0.00-2.97), c) studies that reported maternal alcohol or tobacco use (3 studies, 504 children, 7 treatments; τ^2=0.27, 95% CrI: 0.00-3.29), and d) studies with high methodological quality on NOS item 'comparability of cohorts' (3 studies, 366 children, 7 treatments; τ^2=0.38, 95% CrI: 0.00-4.14), were consistent with the NMA results (Additional File 3: Appendix K). The sensitivity analysis with studies of high methodological quality on the NOS item 'adequacy of follow-up' found no statistically significant results (4 studies, 283 patients, 12 treatments; τ^2=1.01, 95% CrI: 0.01-5.85; Additional File 3: Appendix K).

Autism/dyspraxia

The NMA on autism/dyspraxia (definitions in Table 1) included five cohort studies, 2,551 children exposed *in utero*, and examined 12 treatments (τ^2=0.16, 95% CrI: 0.00-1.95; Figure 3b; Additional File 3: Appendix H). Compared with control, only valproate (OR=17.29, 95% CrI: 2.40-217.60), oxcarbazepine (OR= 13.51, 95% CrI: 1.28-221.40), lamotrigine (OR= 8.88, 95% CrI: 1.28-112.00), and lamotrigine+valproate (OR=132.70, 95% CrI: 7.41-3851.00) were significantly associated with increased occurrence of autism/dyspraxia (Figure 2b).

Restricting the NMA to studies including only women with epilepsy as their treatment indication produced results that were generally in agreement with the NMA results, except that oxcarbazepine was no longer in the network (4 cohort studies, 540 children, 10 treatments; τ^2=0.31, 95% CrI: 0.00-304). Two cohort studies of 404 offspring of women with a history of tobacco use compared 4 treatments and found similar results except that oxcarbazepine and lamotrigine+valproate were no longer in the network (τ^2=0.39, 95% CrI: 0.00-4.47). The results were in agreement in sensitivity analyses including only higher methodological quality studies in the 'comparability of cohorts' item on the NOS (4 studies, 2,395 children, 12 treatments; τ^2=0.19, 95% CrI: 0.00-2.43) and the 'adequacy of follow-up of cohorts' (3 studies, 2244 children, 10 treatments; τ^2=0.23, 95% CrI: 0.00-2.88), except that lamotrigine was no longer statistically significant than control for the latter (Additional File 3: Appendix K).

Neonatal Seizure

One cohort study included 72 children who were exposed to AEDs *in-utero* as well as through breastfeeding reported on the incidence of neonatal seizures. The study compared

valproate against lamotrigine and found no significant difference in neonatal seizures between the two drugs (OR=0.18, 95% CI: 0.01-3.70).

Psychomotor developmental delay

The NMA on psychomotor developmental delay (definitions in Table 1) included 11 cohort studies, 1,145 children exposed *in utero*, and examined 18 treatments (τ^2=0.06, 95% CrI: 0.00-0.63; Figure 3c; Additional File 3: Appendices H, J). Valproate (OR=4.16, 95% CrI: 2.04-8.75) and carbamazepine+phenobarbital+valproate (OR=19.12, 95% CrI: 1.49-337.50) were significantly more harmful than control (Figure 2c).

Language delay

The NMA on language delay (definitions in Table 1) included five cohort studies, 509 children, and examined five treatments (τ^2=0.16, 95% CrI: 0.00-2.15; Figure 3d; Additional File 3: Appendices H, J). One study included children exposed to AEDs *in-utero* and through breastfeeding, and four included children exposed to AEDs *in-utero*. Compared with control, valproate was the only treatment significantly associated with increased odds of language delay (OR=7.95, 95% CrI: 1.50-49.13; Figure 2d).

Attention deficit hyperactivity disorder

The NMA on ADHD (definitions in Table 1) included five cohort studies, 816 children, and examined seven treatments (τ^2=0.11, 95% CrI: 0.00-1.29). One study included children exposed to AEDs *in-utero* and through breastfeeding, while four studies included children exposed to AEDs *in-utero*. None of the treatment comparisons reached statistical significance (Figure 3e; Figure 2e; Additional File 3: Appendices H, J).

Social Impairment

One cohort study included 422 children exposed to AEDs *in-utero* as well as through breastfeeding. The children were exposed to carbamazepine (n=48), lamotrigine (n=71), valproate (n=27) and control (n=278). No significant differences in social impairment were identified.[41]

DISCUSSION

Our results suggest that AEDs generally pose a risk for infants and children exposed *in-utero* or during breastfeeding. Valproate was significantly associated with more children experiencing autism/dyspraxia, language, cognitive and psychomotor developmental delays versus children who were not exposed to AEDs. Oxcarbazepine, lamotrigine and lamotrigine+valproate were associated with increased occurrence of autism/dyspraxia, whereas for the cognitive developmental delay and psychomotor developmental delay outcomes, children exposed to the combination of carbamazepine, phenobarbital, and valproate were at greater odds of harm than those who were not exposed to AEDs. However, these results should be interpreted with caution, as a number of factors (e.g., anticonvulsant dosing, severity of epilepsy, duration of exposure, serum concentrations of exposure, mother's IQ/education) that may all influence outcomes were not identified in these studies. Also, our subsequent analyses may be underpowered due to missing data (e.g., 17 of the 27 studies did not report maternal age, 23 of 27 studies did not report alcohol use, 22 of 27 studies did not report tobacco use, and 14 of 27 studies did not include control group).

NMA is a particularly useful tool for decision-makers because it allows the ranking of treatments for each outcome. However, the results of our SUCRA curves should be interpreted with caution, especially due to the small number of studies and children included in each NMA, which is also reflected in the high uncertainty around the SUCRA values (Figure 2).[31]

Our results are consistent with a longitudinal study of 311 children that found exposure to lamotrigine was associated with significantly higher IQ scores and verbal function at six

years of age compared to children exposed to valproate (Additional File 3: Appendix C).[7] As indicated in Additional File 3: Appendix C, we were unable to include this study because the outcome was reported as a continuous measure, where we focused on dichotomous outcomes to facilitate interpretation. Our results are supported by findings from a cohort study, which found that children exposed to levetiracetam were not at increased risk for delayed development compared to unexposed children (Additional File 3: Appendix C).[42] As indicated in Additional File 3: Appendix C, we were unable to include this study due to the same reason as above. A NMA of 195 RCTs (including 28,013 both male and female patients) showed that gabapentin and levetiracetam showed the best tolerability profile compared with other AEDs, whereas oxcarbazepine and topiramate had a higher withdrawal rate, and lamotrigine an intermediate withdrawal rate.[43]

Across all outcomes, valproate alone or combined with another AED (even with a newer-generation agent, e.g., lamotrigine) was associated with the greatest odds. Similarly, two previous systematic reviews that did not conduct a NMA found valproate was associated with significantly lower IQ scores and poorer overall neurodevelopmental outcomes when compared to an unexposed control group.[5 6] Also consistent with our results, a 2014 Cochrane review including 28 studies (10 of these studies were included in the meta-analyses; with a maximum number of five studies per meta-analysis) concluded that AED polytherapy led to poorer developmental outcomes and IQ compared to healthy controls, epileptic controls, and unspecified monotherapy.[5] This Cochrane review also concluded that insufficient data exist for newer AEDs. However, unlike our review, it included and analysed fewer studies, and did not differentiate between specific polytherapy regimens, and thus did not compare these regimens versus each other or specific monotherapy AEDs.

These risks must be balanced with the need to control seizure activity in pregnancy and thus informed decision-making by patients and clinicians is critical.

Strengths of our study include a comprehensive systematic review methodology that followed the Cochrane Handbook[44] and ISPOR[12] guidelines, and reported using the PRISMA extension for NMA.[13] To the best of our knowledge, our study was the first that compared and ranked the safety of AEDs. We evaluated the comparative safety of treatments that have not been directly compared head-to-head before. In addition, we calculated predictive intervals, which account for between-study variation and provide a predicted range for the treatment effect estimate, should a future study be conducted. On average, the predictive intervals suggested that our results are robust.

Our systematic review has a few limitations worth noting. First, due to the complexity of the data and the studies' underreporting, differences in drug dosages could not be accounted for, and it was assumed that different dosages of the same AED were equally effective. When a study reported multiple dosages for the same treatment, we combined the data for this treatment. This is common for cohort studies, which report on a number of different types of exposures amongst patients. Second, several polytherapies had high SUCRA estimates but very wide CrIs, which is due to the small number of studies included for each drug combination with underpowered sample sizes. Evidence suggests that ranking probabilities for a treatment of being the best may be biased toward the treatments with the smallest number of studies, which may have influenced our SUCRA results.[31 45] As such, the effect sizes need to be taken into account when considering the SUCRA values. Third, due to the absence of evidence from RCTs, our conclusions were based on evidence from observational studies only, and inherent biases because of

confounding and shortcomings of these studies may have impacted our findings. For example, the included studies often failed to report important confounding variables,[46] such as family history of autism, ADHD, and maternal IQ, severity of epilepsy making it impossible for us to control these variables through subgroup analysis and meta-regression. Recent research has explored methods to incorporate non-randomized with randomized evidence in a NMA and have highlighted the need to carefully explore the level of confidence in the non-randomized evidence.[47 48] The use of observational studies allows the assessment of the safety profile of AED treatments and offers the opportunity to evaluate effects in pregnancy.[49] Future large-scale observational studies are needed to allow the evaluation of rare adverse events that otherwise cannot be adequately evaluated in RCTs, especially during pregnancy. Fourth, although no intransitivity for most effect modifiers assessed was evident, there was an imbalance in the methodological study quality appraisal across treatment comparisons and most outcomes, which may impact our results. Unknown factors or factors that could not be assessed due to dearth of data may pose the risk of residual confounding bias, and hence risk the validity of the transitivity assumption. However, the assessment of consistency suggested no disagreement between the different sources of evidence in the network. Fifth, although the tendency towards small-study effects is greater with observational studies than with randomized trials,[50] the assessment of small-study effects using adjusted funnel plots suggested no evidence for their prevalence. Also, the majority of the included studies in this review compared multiple treatments inducing correlations in each funnel plot, which may mask asymmetry. Although we plotted data points corresponding to the study-specific basic parameters to reduce correlations, this issue may still exist. Sixth, we were unable to conduct subgroup

analysis by type of exposure (breastfeeding versus *in utero*) due to the small number of studies included in the NMA and due to the poor reporting; 22 studies did not report whether exposure was also in breastfeeding (additional to *in utero*). Hence, we included all studies in the analysis irrespective of the type of exposure.

More evidence from long-term follow-up studies is required to further delineate neurodevelopmental risks in children. Future studies should assess the genetic contribution from the biological father, maternal seizures during pregnancy, exposure through breastfeeding only, types of epilepsy, and maternal family history. Registries should aim to include a suitable control group and collect information on potential confounders, such as alcohol and tobacco use, allowing researchers to identify the safest agents for different patient-level covariates, and enhance decision-making for healthcare providers and patients. A critical evaluation of the validity of the control group is also necessary, in order to examine potential differences between the treated and the not treated populations. An individual patient data NMA would likely provide further clarity to the field, which allows the tailoring of management to specific patient characteristics.[51]

CONCLUSION

Across all outcomes and treatments compared with control, valproate alone or combined with another AED was associated with the greatest odds, whereas oxcarbazepine and lamotrigine were associated with increased occurrence of autism. Counselling is advised for women considering pregnancy to tailor the safest regimen.

LIST OF ABBREVIATIONS

AEDs: Anti-epileptic drugs; CrI: Credible interval; NMA: Network Meta-analysis; OR: Odds ratio; PrI: Predictive interval; SUCRA curve: Surface under the cumulative ranking curve

ADDITIONAL FILES

Additional File 1: Protocol

Additional File 2: PRISMA NMA Checklist

Additional File 3: Supplementary Online Content (Appendices A-O)

Appendix A. Newcastle-Ottawa Scale scoring guide

Appendix B. List of included studies

Appendix C. Key Excluded Studies

Appendix D. Table of Individual Study Characteristics

Appendix E. Table of Patient Characteristics

Appendix F. Methodological quality of observational studies – Newcastle Ottawa Scale

Appendix G. Comparison-adjusted funnel plots

Appendix H. Statistically significant network meta-analysis results along with meta-analysis results, transitivity, and inconsistency assessments

Appendix I. Rank-heat plot of cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcomes

Appendix J. Number of studies and treatments per outcome

Appendix K. Sensitivity and network meta-regression analyses - Anti-epileptic drugs compared with Control

FIGURE LEGENDS

Figure 1. Study flow diagram

Figure 2. Forest plots for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcome

Figure 3. Network diagrams for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcomes

Each treatment node is weighted according to the number of patients that have received the particular treatment, and each edge is weighted according to the number of studies comparing the treatments it connects.

Abbreviations: carbam - carbamazepine, clobaz - clobazam, clonaz - clonazepam, ethos - ethosuximide, gabap - gabapentin, lamot - lamotrigine, levet - levetiracetam, oxcar - oxcarbazepine, pheno - phenobarbital, pheny - phenytoin, primid - primidone, topir - topiramate, valpro - valproate, vigab – vigabatrin

DECLARATIONS

CONTRIBUTORS

AAV analysed the data, interpreted the results, and drafted the manuscript. ACT and SES conceived and designed the study, helped obtain funding, interpreted the results, and helped write sections of the manuscript. PR and EC coordinated the review, screened citations and full-text articles, abstracted data, appraised quality, resolved discrepancies, contacted authors, and edited the manuscript. CS provided methodological support and screened citations and full-text articles and edited the manuscript. RK, ER, FY, JDS, KT, and HM screened citations and full-text articles, abstracted data, and/or appraised quality. BH, BRH and YF helped conceive the study and edited the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

We thank Dr. David Moher for providing his feedback on our protocol. We thank Dr. Laure Perrier for conducting the literature searches, Becky Skidmore for peer-reviewing the MEDLINE search, and Alissa Epworth for obtaining the full-text articles. We thank Alistair Scott, Wing Hui, and Geetha Sanmugalingham for screening some of the citations and/or abstracting some of the data for a few of the included studies, Misty Pratt and Mona Ghannad for helping scan reference lists, and Ana Guzman, Susan Le, and Inthuja Selvaratnam for contacting authors and formatting the manuscript.

FUNDING

This systematic review was funded by the Canadian Institutes for Health Research/Drug Safety and Effectiveness Network (CIHR/DSEN). AAV is funded by the Banting Postdoctoral Fellowship Program from the CIHR. SES is funded by a Tier 1 Canada Research Chair in

Knowledge Translation. BH is funded by a CIHR/DSEN New Investigator Award in Knowledge Synthesis. BRH receives funding from the Alberta Heritage Foundation for Medical Research. ACT is funded by a Tier 2 Canada Research Chair in Knowledge Synthesis. The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

COMPETING INTERESTS

None declared.

ETHICS APPROVAL

Not applicable.

PROVENANCE AND PEER REVIEW

Not commissioned; externally peer reviewed.

DATA SHARING STATEMENT

All datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

OPEN ACCESS

REFERENCES

1. Spina E, Perugi G. Antiepileptic drugs: indications other than epilepsy. *Epileptic Disord*. 2004;6(2):57-75.
2. Harden CL, Pennell PB, Koppel BS, et al. Management issues for women with epilepsy--focus on pregnancy (an evidence-based review): III. vitamin K, folic acid, blood levels, and breast-feeding: report of the quality standards subcommittee and therapeutics and technology assessment subcommittee of the American Academy of Neurology and the American Epilepsy Society. *Epilepsia*. 2009;50(5):1247-55.
3. Samren EB, van Duijn CM, Koch S, et al. Maternal use of antiepileptic drugs and the risk of major congenital malformations: a joint European prospective study of human teratogenesis associated with maternal epilepsy. *Epilepsia*. 1997;38(9):981-90.
4. Meador K, Reynolds MW, Crean S, Fahrbach K, Probst C. Pregnancy outcomes in women with epilepsy: a systematic review and meta-analysis of published pregnancy registries and cohorts. *Epilepsy Res*. 2008;81(1):1-13.
5. Bromley R, Weston J, Adab N, et al. Treatment for epilepsy in pregnancy: neurodevelopmental outcomes in the child. *The Cochrane database of systematic reviews*. 2014(10):Cd010236.
6. Banach R, Boskovic R, Einarson T, Koren G. Long-term developmental outcome of children of women with epilepsy, unexposed or exposed prenatally to antiepileptic drugs: a meta-analysis of cohort studies. *Drug Saf*. 2010;33(1):73-9.
7. Meador KJ, Baker GA, Browning N, et al. Fetal antiepileptic drug exposure and cognitive outcomes at age 6 years (NEAD study): a prospective observational study. *Lancet Neurol*. 2013;12(3):244-52.
8. Christensen J, Gronborg TK, Sorensen MJ, et al. Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. *JAMA*. 2013;309(16):1696-703.
9. Wlodarczyk BJ, Palacios AM, George TM, Finnell RH. Antiepileptic drugs and pregnancy outcomes. *Am J Med Genet A*. 2012;158a(8):2071-90.
10. Velez-Ruiz NJ, Meador KJ. Neurodevelopmental effects of fetal antiepileptic drug exposure. *Drug Saf*. 2015;38(3):271-8.
11. Tricco AC, Cogo E, Angeliki VA, et al. Comparative safety of anti-epileptic drugs among infants and children exposed in utero or during breastfeeding: protocol for a systematic review and network meta-analysis. *Systematic reviews*. 2014;3:68.
12. Jansen JP, Trikalinos T, Cappelleri JC, et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. *Value Health*. 2014;17(2):157-73.
13. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med*. 2015;162(11):777-84.
14. Veroniki AA, Cogo E, Rios P, et al. Comparative safety of anti-epileptic drugs during pregnancy: a systematic review and network meta-analysis of congenital malformations and prenatal outcomes. *BMC Med*. 2017;15(1):95.
15. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *J Clin Epidemiol*. 2016;75:40-6.
16. Wells G, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses2000. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.

17. Chaimani A, Higgins JP, Mavridis D, Spyridonos P, Salanti G. Graphical tools for network meta-analysis in STATA. *PLoS One*. 2013;8(10):e76654.
18. Furukawa TA, Guyatt GH, Griffith LE. Can we individualize the 'number needed to treat'? An empirical study of summary effect measures in meta-analyses. *Int J Epidemiol*. 2002;31(1):72-6.
19. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med*. 2004;23(20):3105-24.
20. Jansen JP, Naci H. Is network meta-analysis as valid as standard pairwise meta-analysis? It all depends on the distribution of effect modifiers. *BMC Med*. 2013;11:159.
21. White IR, Barrett JK, Jackson D, Higgins JP. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Res Synth Methods*. 2012;3(2):111-25.
22. Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ*. 2003;326(7387):472.
23. Veroniki AA, Vasiliadis HS, Higgins JP, Salanti G. Evaluation of inconsistency in networks of interventions. *Int J Epidemiol*. 2013;42(1):332-45.
24. Welton NJ, Sutton AJ, Cooper N, Abrams KR, Ades A. Evidence synthesis for decision making in healthcare. New York: Wiley; 2012.
25. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc Ser B Stat Methodol*. 2002;64(4):583-639.
26. Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: Evolution, critique and future directions. *Stat Med*. 2009;28(25):3049-67.
27. Palmer T, Sterne J. Meta-Analysis in Stata: An Updated Collection from the Stata Journal. White I, editor. Texas: Stata Press; 2016.
28. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ*. 2011;342:d549.
29. Higgins JP, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *J R Stat Soc Ser A Stat Soc*. 2009;172(1):137-59.
30. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol*. 2011;64(2):163-71.
31. Trinquart L, Attiche N, Bafeta A, Porcher R, Ravaud P. Uncertainty in Treatment Rankings: Reanalysis of Network Meta-analyses of Randomized Trials. *Ann Intern Med*. 2016;164(10):666-73.
32. Bromley R, Baxter N, Calderbank R, Mawer G, Clayton-Smith J, Baker G. A comprehensive review of the language abilities of children exposed to valproate or carbamazepine in utero. American Epilepsy Society; Texas2010.
33. Gogatishvili N, Ediberidze T, Lomidze G, Tatishvili N, Kasradze S. PO-0834 Long-term Developmental Outcome Of Children Prenatally Exposed To Antiepileptic Drugs. *Arch Dis Child*. 2014;99(Suppl 2):A526.
34. Gogatishvili N, Ediberidze T, Lomidze G, Tatishvili N, Kasradze S. Cognitive outcomes of children with fetal antiepileptic drug exposure at the age of 3-6 years-preliminary data. 1st Congress of the European Academy of Neurology; Berlin: European Journal of Neurology; 2015. p. 329.
35. Miskov S, Juraski RG, Fucic A, et al. Croatian Pregnant Women with Epilepsy and Effects of Antiepileptic Drugs Exposure in their Offspring - seven years of prospective surveillance. American Epilepsy Society; Texas2010.
36. Adab N, Kini U, Vinten J, et al. The longer term outcome of children born to mothers with epilepsy. *J Neurol Neurosurg Psychiatry*. 2004;75(11):1575-83.

37. Dean JCS, Hailey H, Moore SJ, Lloyd DJ, Turnpenny PD, Little J. Long term health and neurodevelopment in children exposed to antiepileptic drugs before birth. *J Med Genet*. 2002;39(4):251-9.
38. Gaily E. Development and growth in children of epileptic mothers: a prospective controlled study. Helsinki, Finland: University of Helsinki; 1990.
39. Katz JM, Pacia SV, Devinsky O. Current Management of Epilepsy and Pregnancy: Fetal Outcome, Congenital Malformations, and Developmental Delay. *Epilepsy Behav*. 2001;2(2):119-23.
40. Veroniki AA, Straus SE, Fyraridis A, Tricco AC. The rank-heat plot is a novel way to present the results from a network meta-analysis including multiple outcomes. *J Clin Epidemiol*. 2016.
41. Veiby G, Engelsen BA, Gilhus NE. Early child development and exposure to antiepileptic drugs prenatally and through breastfeeding: a prospective cohort study on children of women with epilepsy. *JAMA Neurol*. 2013;70(11):1367-74.
42. Shallcross R, Bromley RL, Irwin B, Bonnett LJ, Morrow J, Baker GA. Child development following in utero exposure: levetiracetam vs sodium valproate. *Neurology*. 2011;76(4):383-9.
43. Zaccara G, Giovannelli F, Giorgi FS, Franco V, Gasparini S, Benedetto U. Tolerability of new antiepileptic drugs: a network meta-analysis. *Eur J Clin Pharmacol*. 2017.
44. Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. 5.1.0 ed: The Cochrane Collaboration; 2009.
45. Kibret T, Richer D, Beyene J. Bias in identification of the best treatment in a Bayesian network meta-analysis for binary outcome: a simulation study. *Clin Epidemiol*. 2014;6:451-60.
46. Dalessio DJ. Seizure Disorders and Pregnancy. *N Engl J Med*. 1985;312(9):559-63.
47. Efthimiou O, Mavridis D, Debray TP, et al. Combining randomized and non-randomized evidence in network meta-analysis. *Stat Med*. 2017.
48. Schmitz S, Adams R, Walsh C. Incorporating data from various trial designs into a mixed treatment comparison model. *Stat Med*. 2013;32(17):2935-49.
49. Cameron C, Fireman B, Hutton B, et al. Network meta-analysis incorporating randomized controlled trials and non-randomized comparative cohort studies for assessing the safety and effectiveness of medical treatments: challenges and opportunities. *Systematic reviews*. 2015;4:147.
50. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet*. 1991;337(8746):867-72.
51. Veroniki AA, Straus SE, Soobiah C, Elliott MJ, Tricco AC. A scoping review of indirect comparison methods and applications using individual patient data. *BMC Med Res Methodol*. 2016;16(1):47.

666 **Table 1 Summary characteristics of included studies**

Table 1. Summary Characteristics of included studies		
Study/Patient Characteristic	**# of Studies (n=29)**	**% of Studies**
Year of publication		
1980-1989	1	3.45
1990-1999	6	20.69
2000-2009	5	17.24
2010-2015	17	58.62
Continent (of country of study conduct)		
Europe	20	68.97
North America	5	17.24
Asia	1	3.45
Australia	2	6.90
Trans-Continental	1	3.45
Study design		
Observational cohort	29	100.00
Case-control	0	0.00
Randomized clinical trial	0	0.00
Registry study		
Yes	11	37.93
No	18	62.07
Sample size		
0-99	18	62.07
100-299	9	31.03
300-499	1	3.45
500-699	0	0.00
700-999	0	0.00
1000+	1	3.45
Number of interventions		
2	4	13.79
3	5	17.24
4	8	27.59
5-7	8	27.59
8-10	2	6.90
11+	2	6.90
Outcomes[*,†]		
Cognitive Developmental Delay	12	58.62
Autism/Dyspraxia	5	17.24
Language Delay	5	17.24
ADHD	5	17.24

Table 1. Summary Characteristics of included studies

Study/Patient Characteristic	# of Studies (n=29)	% of Studies
Psychomotor Developmental Delay	11	37.93
Neonatal Seizures	2	6.90
Social Impairment	1	3.45
Funding		
Public	15	51.72
Private	0	0.00
Mixed public and private	4	13.79
NR/Unclear	10	34.48
Treatment indication		
Epilepsy	23	79.31
Mixed indications‡	0	0.00
Not reported	6	20.69
Epileptic control group§		
Yes	15	51.72
No/NR/NA	14	48.28
Mean maternal age		
24-26 y	2	6.90
27-29 y	5	17.24
30+ y	4	13.79
Not reported	18	62.07
AED exposure during pregnancy		
Reported as during 1st trimester	6	20.69
Reported as any time during pregnancy	6	20.69
Not reported	17	58.62
Alcohol use during pregnancy		
Yes	5	17.24
NR	24	82.76
Tobacco use during pregnancy		
Yes	7	24.14
NR	22	75.86

Abbreviations: ADHD - Attention Deficit Hyperactivity Disorder; AED - anti-epileptic drug(s); NA - Not applicable; NR - Not reported

* Values in this category do not match totals as some studies report more than one outcome

† Percentage of total number of included studies (n=29)

‡ Includes individuals taking AEDs for psychiatric disorders, migraine, and neuropathic/neurological pain

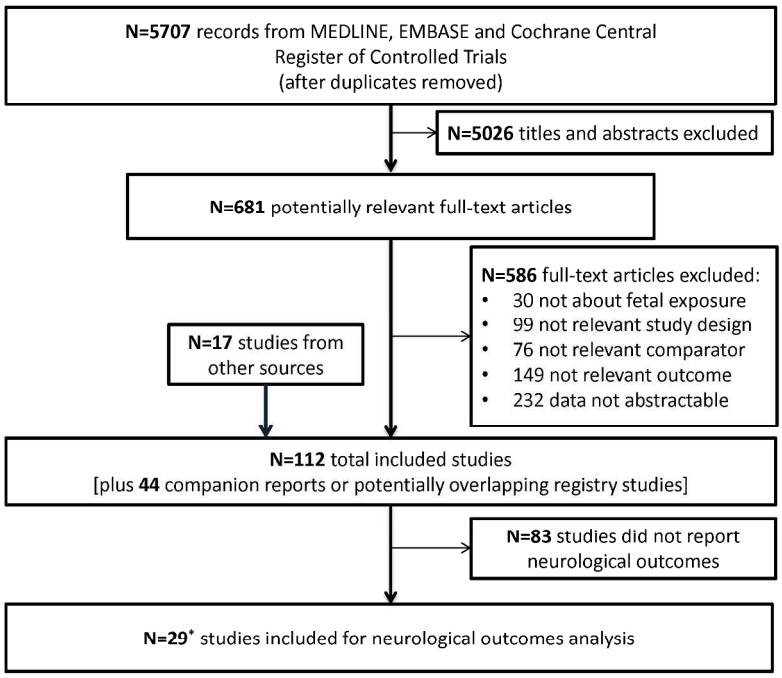
§ Consisted of women with Epilepsy who did not take AEDs during pregnancy

667 **Table 2 Outcome measures and diagnostic scales used in analysis**

Cognitive developmental delay	
Bayley Scales of Infant Development (children ≤42 mo.)	Score ≥2 standard deviations below the mean
Griffiths Scale of Infant Development (children >42 mo.)	Score ≥2 standard deviations below the mean
McCarthy Scales of Children's Abilities (children >30 mo.)	Score ≥1 standard deviations below the mean
Stanford-Binet IV Intelligence scale for children	Intelligence quotient ≤80
Touwen's Test	Above average number of items rated abnormal in one or more domains
Wechsler Scale of Preschool and Primary Intelligence	Intelligence quotient <90
Wechsler Intelligence Scale for Children - III	Intelligence quotient <80; verbal intelligence quotient <69
Developmental Assessment	Confirmed diagnosis by developmental pediatrician or pediatric neurologist
Autism/dyspraxia	
Developmental Assessment	Diagnosis confirmed by developmental specialists at 2 years of age
Medical Records	Confirmed diagnosis recorded in medical history; registry records (ICD-10 codes F84.0, F84.1, F84.5, F84.8, and F84.9)
Modified checklist for autism in toddlers	Scored positive for ≥2 out of 6 critical items OR ≥3 any items of the total scale
Psychomotor developmental delay	
Ages and Stages Questionnaire	>3 standard deviations from the test mean
Bayley Scales of Infant Development – Psychomotor Index	>2 standard deviations below the standardized mean for the test
Touwen's Test	Demonstrated dysfunctions in fine motor balance, fine motor functions, and coordination of extremities
Schedule of Growing Skills II	Scored as 'delayed' in ≥1 domain of the test

Developmental Assessment	Infant scored >2 negative items (administered by general practitioner or pediatrician); diagnosis of neuromotor deficit confirmed by a trained nurse practitioner; infant failing to sit by 10 months or walk by 18 months
Health/Medical Records	Diagnosis of psychomotor delay recorded in medical records
Language Delay	
Ages and Stages Questionnaire	>3 standard deviations from the test mean
Clinical Evaluation of Language Fundamentals – 4th Edition	Score <70 in core language domain; score <84 overall
Learning Accomplishment Profile	Below average performance in expressive speech (adjusted for age)
Comprehensive Language Assessment (Peabody Picture Vocabulary Test; Receptive Expressive Emergent Language Scale; Expressive One Word Picture Vocabulary Test, or Sequenced Inventory of Communication Development)	Scores/assessment indicate a >6 moth delay in age appropriate language development
ADHD	
Attention Problems and Hyperactivity Scales	Score >1 standard deviations from the test mean
Child Behaviour Checklist	≥6 positive items on checklist
Diagnostic and Statistical Manual – IV	≥5 positive items on checklist
Medical Records	Confirmed diagnosis in hospital/medical records made by a pediatrician or child psychiatrist
Neonatal Seizure	
Medical records	Record of seizures during 1st year; confirmation of neonatal seizure by electroencephalography or diagnosis
Social Impairment	
Developmental Assessment (Ages and Stages Questionnaire [6 and 18 months]; Child Behaviour Checklist [36 months])	Scores dichotomized into 'normal' or 'adverse' range based on pre-defined values used by scale, for scales without pre-defined values cut-off was set at a score >2 standard deviations outside the test mean

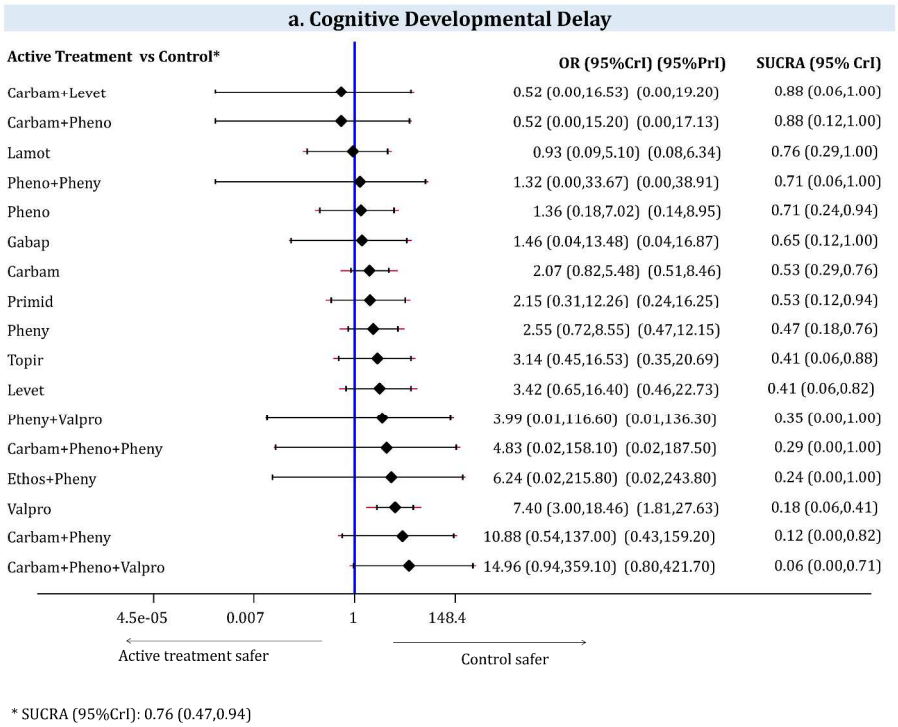
668



*29 publications reporting 30 included studies.

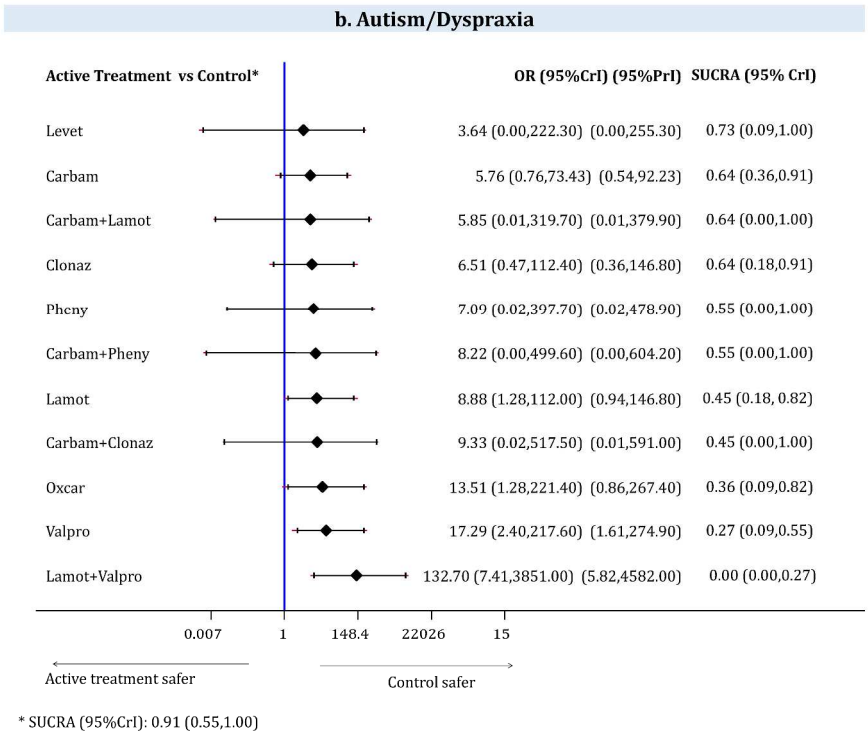
TITLE: Study flow diagram

254x190mm (300 x 300 DPI)



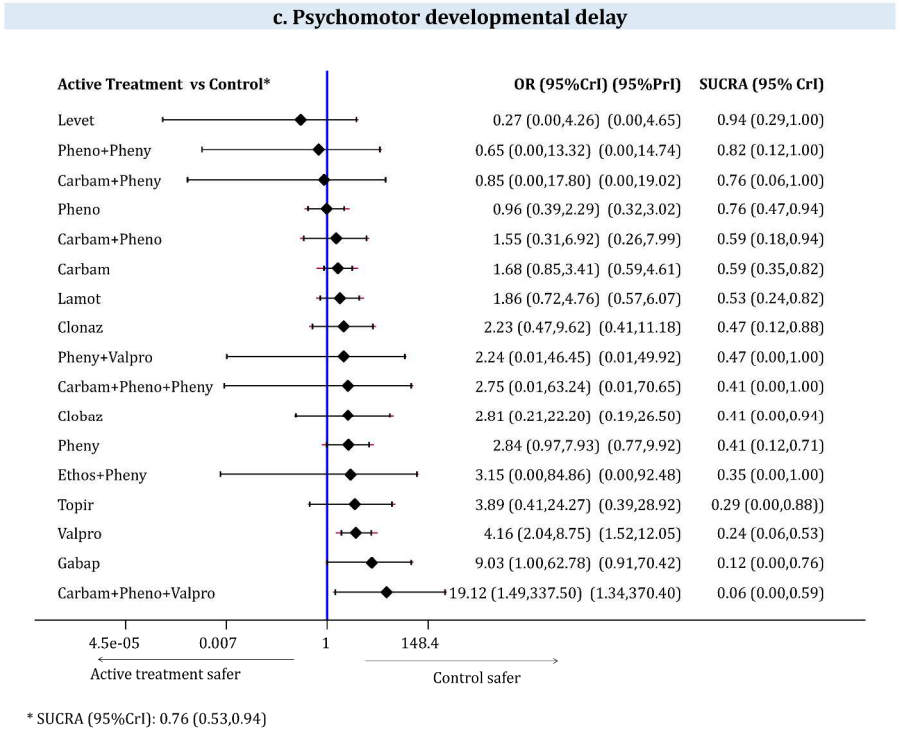
TITLE: Forest plots for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcome

533x400mm (300 x 300 DPI)



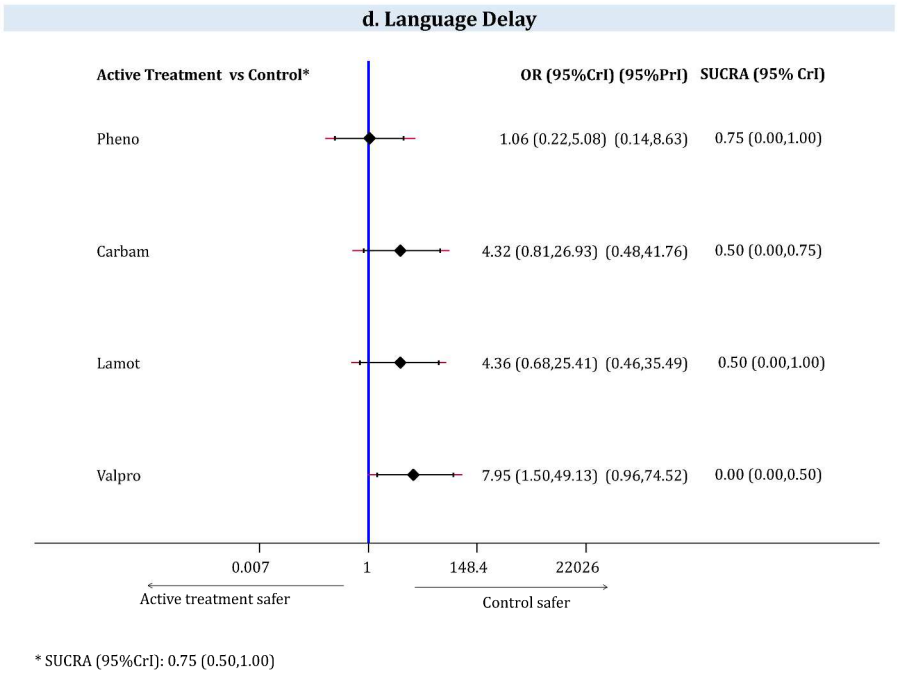
TITLE: Forest plots for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcome

533x400mm (300 x 300 DPI)



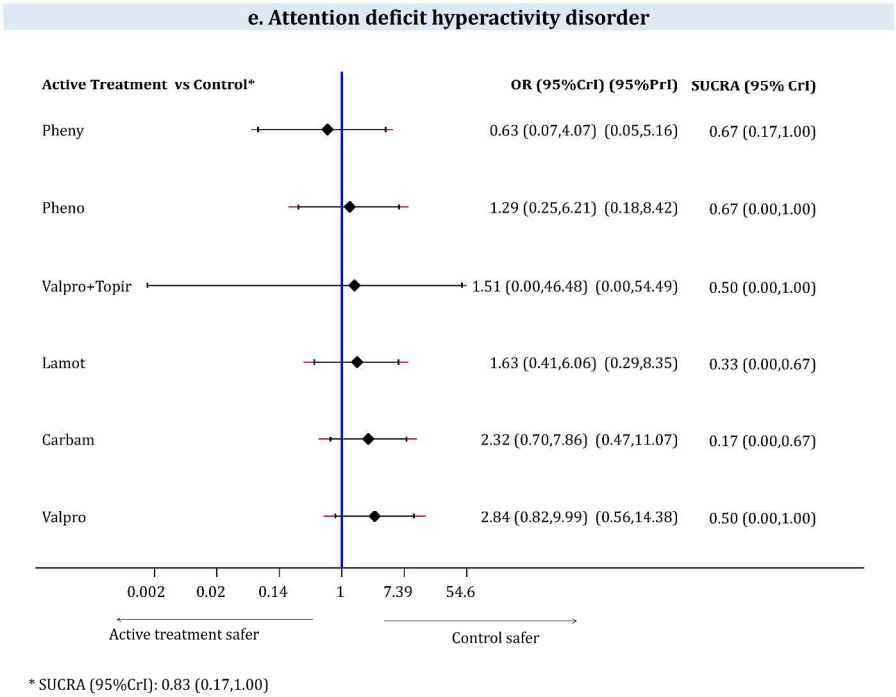
TITLE: Forest plots for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcome

533x400mm (300 x 300 DPI)



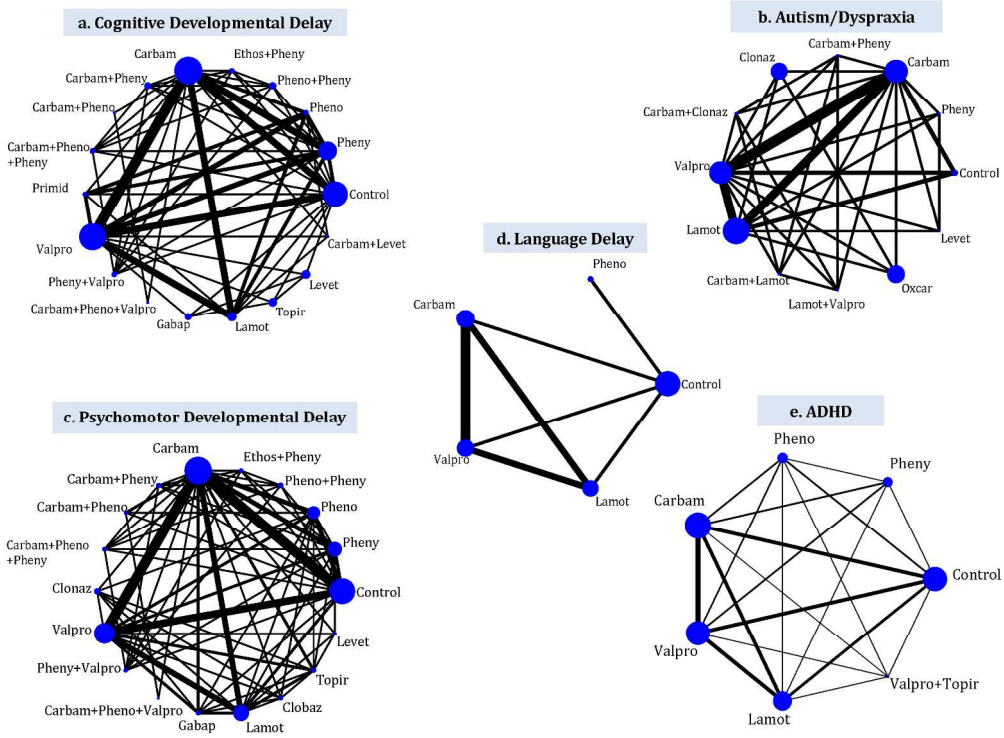
TITLE: Forest plots for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcome

533x400mm (300 x 300 DPI)



TITLE: Forest plots for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcome

533x400mm (300 x 300 DPI)



TITLE: Network diagrams for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcomes.

CAPTION: Each treatment node is weighted according to the number of patients that have received the particular treatment, and each edge is weighted according to the number of studies comparing the treatments it connects.

Abbreviations: carbam - carbamazepine, clobaz - clobazam, clonaz - clonazepam, ethos - ethosuximide, gabap - gabapentin, lamot - lamotrigine, levet - levetiracetam, oxcar - oxcarbazepine, pheno - phenobarbital, pheny - phenytoin, primid - primidone, topir - topiramate, valpro - valproate, vigab – vigabatrin

254x190mm (300 x 300 DPI)

Tricco et al. Systematic Reviews 2014, 3:68
http://www.systematicreviewsjournal.com/content/3/1/68

SYSTEMATIC REVIEWS

PROTOCOL Open Access

Comparative safety of anti-epileptic drugs among infants and children exposed *in utero* or during breastfeeding: protocol for a systematic review and network meta-analysis

Andrea C Tricco[1], Elise Cogo[1], Veroniki A Angeliki[1], Charlene Soobiah[1,2], Brian Hutton[3], Brenda R Hemmelgarn[4], David Moher[3], Yaron Finkelstein[5,6,7] and Sharon E Straus[1,8*]

Abstract

Background: Epilepsy affects about 1% of the general population. Anti-epileptic drugs (AEDs) prevent or terminate seizures in individuals with epilepsy. Pregnant women with epilepsy may continue taking AEDs. Many of these agents cross the placenta and increase the risk of major congenital malformations, early cognitive and developmental delays, and infant mortality. We aim to evaluate the comparative safety of AEDs approved for chronic use in Canada when administered to pregnant and breastfeeding women and the effects on their infants and children through a systematic review and network meta-analysis.

Methods: Studies examining the effects of AEDs administered to pregnant and breastfeeding women regardless of indication (e.g., epilepsy, migraine, pain, psychiatric disorders) on their infants and children will be included. We will include randomized clinical trials (RCTs), quasi-RCTs, non-RCTs, controlled before-after, interrupted time series, cohort, registry, and case-control studies. The main literature search will be executed in MEDLINE, EMBASE, and the Cochrane Central Register of Controlled Trials. We will seek unpublished literature through searches of trial protocol registries and conference abstracts. The literature search results screening, data abstraction, and risk of bias appraisal will be performed by two individuals, independently. Conflicts will be resolved through discussion. The risk of bias of experimental and quasi-experimental studies will be appraised using the Cochrane Effective Practice and Organization of Care Risk-of-Bias tool, methodological quality of observational studies will be appraised using the Newcastle-Ottawa Scale, and quality of reporting of safety outcomes will be conducted using the McMaster Quality Assessment Scale of Harms (McHarm) tool. If feasible and appropriate, we will conduct random effects meta-analysis. Network meta-analysis will be considered for outcomes that fulfill network meta-analysis assumptions.

The primary outcome is major congenital malformations (overall and by specific types), while secondary outcomes include fetal loss/miscarriage, minor congenital malformations (overall and by specific types), cognitive development, psychomotor development, small for gestational age, preterm delivery, and neonatal seizures.

(Continued on next page)

* Correspondence: sharon.straus@utoronto.ca
[1]Li Ka Shing Knowledge Institute, St. Michael's Hospital, 209 Victoria Street, East Building, Toronto, Ontario M5B 1 T8, Canada
[8]Department of Geriatric Medicine, University of Toronto, 172 St. George Street, Toronto, Ontario M5R 0A3, Canada
Full list of author information is available at the end of the article

(Continued from previous page)

Discussion: Our systematic review will address safety concerns regarding the use of AEDs during pregnancy and breastfeeding. Our results will be useful to healthcare providers, policy-makers, and women of childbearing age who are taking anti-epileptic medications.

Systematic review registration: PROSPERO CRD42014008925.

Keywords: Anti-epileptic drug, Breastfeeding, Comparative safety, Congenital malformation, Epilepsy, Fetus, Infant, Network meta-analysis, Pregnancy, Systematic review

Background

Epilepsy is the most common chronic neurological condition, affecting 0.6 to 1% of the population [1,2]. Individuals with uncontrolled epilepsy experience recurrent seizures, which can have psychosocial and physical consequences, including a compromised life expectancy [3,4]. The goal of anti-epileptic treatment is to improve quality of life and health outcomes by reducing the frequency of seizures [4].

Anti-epileptic medications decrease seizures by reducing excitation and enhancing inhibition of neurons [5-7]. Many of these medications target different channels, including calcium, sodium, and glutamate, and are broadly classified as first generation agents (e.g., phenobarbitone, phenytoin, carbamazepine, sodium valproate, ethosuximide) and second generation agents (e.g., lamotrigine, levetiracetam, topiramate, gabapentin, vigabatrin, oxcarbazepine, clobazam, clonazepam, zonisamide, lacosamide, rufinamide, primidone) [8]. Due to the broad and varied mechanisms of action, the indications for some of these medications also include pain syndromes, psychiatric disorders, and migraine headaches [8].

Many clinical practice guidelines recommend that women of childbearing age continue to take their anti-epileptic medications; however, medications with lower risk of teratogenic events are advised [9,10] since anti-epileptic drugs (AEDs) cross the placenta or transfer through breast milk, posing risks to the fetus and infant [9,11,12].

Some AEDs have been associated with increased risk of harm to the fetus and infants. For example, exposure to valproate has led to increased risk of major congenital malformations [10], cognitive delay, and minor congenital abnormalities [13-16]. Phenobarbital has been associated with minor congenital abnormalities and developmental delay [17,18]. Carbamazepine and lamotrigine have been associated with minor congenital abnormalities [19-22]. However, other than studies of the use of valproate, many studies have produced inconsistent findings regarding harm to the fetus and infant with use of other agents [23]. As such, our objective is to evaluate the comparative safety of AEDs for infants and children who were exposed *in utero* or during breastfeeding through a systematic review and network meta-analysis.

Methods/Design

Protocol

A systematic review protocol was developed and registered with the PROSPERO database (CRD42014008925, available at: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014008925). It was revised with feedback from the decision-makers who posed the query within Health Canada, healthcare practitioners, content experts, and research methodologists. The reporting of our systematic review protocol was guided by the Preferred Reporting Items for Systematic Reviews and Meta-analyses Protocols [24].

Eligibility criteria

We will include studies examining the effects of AEDs on infants and children who were exposed *in utero* or during breastfeeding. We will include experimental studies (randomized clinical trials [RCTs], quasi-RCTs, non-RCTs), quasi-experimental studies (controlled before and after studies, interrupted time series), and observational studies (cohort, case-control, registry studies) of pregnant women at any stage of pregnancy and breastfeeding women and their infants/children. The rationale for including other study designs in addition to RCTs is that there are ethical issues in conducting RCTs of AEDs in pregnancy, so RCT evidence might not exist for some or all of these drugs. Given that our review includes rare outcomes, including observational evidence is crucial. In contrast to efficacy evaluation, safety assessment usually requires very large sample sizes to be able to detect adverse events. Therefore, while RCTs have lower risk of bias, they usually do not have the statistical power needed to adequately evaluate uncommon/rare safety outcomes due to Type II (i.e., false negative) error [25]. Given that our review includes rare outcomes, including observational evidence is crucial [26]. Additionally, observational studies can often provide more generalizable evidence due to the strict participant inclusion criteria in most RCTs [27]. Real-world safety evidence that has external validity is important for the assessment of the possible risks of AEDs in pregnant and breastfeeding women.

The diagnosis of neurodevelopmental delay related to *in utero* exposure is made before adolescence, and

hence, we will limit inclusion to children up to 12 years of age. AEDs that are approved for chronic use in Canada will be included. Drugs that are only used acutely or those that are not currently approved for use in Canada will be excluded, as the focus of this review is on the Canadian setting [28-32]. However, most of the medications we will examine are available in other countries as well. The relevant 16 medications and their synonyms are listed in Additional file 1, and the excluded drugs are listed in Additional file 2. Studies of all combinations and doses of these medications are eligible for inclusion. Since we are only interested in exposures that occur *in utero* or during breastfeeding, studies examining AEDs administered directly to infants or children will be excluded. All indications for AEDs will be included such as epilepsy, migraine, pain, and psychiatric disorders.

In order to be included, studies must compare an anti-epileptic medication against another included anti-epileptic medication, placebo, a 'no intervention' control group, or combinations of two or more anti-epileptic medications. Only studies providing results for our outcomes of interest will be included. Our primary outcome is major congenital malformations (overall and by specific type, such as craniofacial defects and neural tube defects). Secondary outcomes include minor congenital malformations (overall and by specific type, such as epicanthal folds and microstomia), cognition (e.g., global cognitive functioning and specific cognitive domains such as attention), psychomotor development (e.g., autism, dyspraxia), small for gestational age, preterm delivery, neonatal seizures, and fetal loss/miscarriage. No other limitations will be imposed on the eligibility criteria, including published/unpublished material, language of dissemination, duration of follow-up, or year of publication. The draft eligibility criteria can be found in Additional file 3.

Information sources and literature search

Our main literature search will be executed in the MEDLINE database. The search terms were drafted by an experienced librarian and can be found in Additional file 4. The search was peer reviewed by another librarian using the Peer Review of Electronic Search Strategies checklist [33].

In addition to MEDLINE, we will also search the EMBASE and the Cochrane Central Register of Controlled Trials databases. We will follow the MEDLINE search strategy for these databases, and the search terms will be adjusted accordingly. The electronic database search will be supplemented by searching for unpublished literature [34]. This will be accomplished through exploring conference abstracts, clinical trial registries, and contacting manufacturers of AEDs. We will also scan the reference lists of included studies and previous reviews in the area [23,35,36].

Study selection process

The eligibility criteria screening form will be pilot-tested by the team and is presented in Additional file 3. We will calculate inter-rater reliability from the pilot-test and screening will only commence after high agreement (e.g., kappa statistic ≥60%) is observed [37]. Subsequently, two reviewers will screen each title/abstract and potentially relevant full-text articles from the literature search results, independently. Conflicts will be resolved through discussion. All screening will occur using our online screening software (synthesi.SR) [38].

Data items and data collection process

We will abstract data on the PICOS elements [39], including patient characteristics (e.g., age of the mother and infant/child, indication for anti-epileptic treatment, co-morbidities, concomitant medications), intervention details (e.g., type of anti-epileptic treatment, dose, route of administration, duration of treatment, timing [trimester] of treatment during pregnancy), comparator details (e.g., comparator agent, dose, route of administration), outcome results (e.g., major congenital abnormality, minor congenital abnormality, cognitive function, psychomotor development) at the longest duration of follow-up, and study characteristics (e.g., study design, country of conduct, year of conduct, sample size, setting). These characteristics will be abstracted using a data abstraction form created in Excel with an accompanying "cheat sheet" that will guide the reviewers with this process. The data abstraction form and cheat sheet will be pilot-tested and data abstraction will only commence when high agreement (e.g., kappa statistic ≥60%) [37] is observed. Each included study will be abstracted by two team members, independently, who will resolve disagreements through discussion.

Methodological quality/risk of bias appraisal

We will use various tools to assess the methodological quality/risk of bias of each of the studies that fulfill our eligibility criteria. This will be conducted by two reviewers, independently, and conflicts will be resolved through discussion. First, we will appraise the risk of bias of experimental and quasi-experimental studies using the Cochrane Effective Practice and Organization of Care Risk-of-Bias tool [40]. Second, we will assess the methodological quality of observational studies using the Newcastle-Ottawa Scale [41]. Third, the quality of reporting of harms will be appraised using the McMaster Quality Assessment Scale of Harms (McHarm) tool [42].

Synthesis of included studies

A narrative summary of study results will be presented along with evidence summary tables. When sufficient data are available, we will conduct random effects meta-

analysis to calculate pooled odds ratios for dichotomous data and pooled mean differences for continuous data [43,44]. Direct (pairwise) meta-analysis will be performed with RCTs alone in order to examine whether the data are consistent between direct and indirect evidence. If the large majority of included studies are observational, we will also conduct additional meta-analyses including observational studies alone. Analyses will be stratified by treatment indication (e.g., epilepsy, pain, etc.) to reduce clinical heterogeneity between different study populations whenever possible; for example, epilepsy itself in pregnant women is related to an increased baseline risk of certain neonatal adverse outcomes. Statistical, clinical, and methodological heterogeneity will be examined prior to conducting the meta-analysis. Funnel plots will be drawn for outcomes including at least 10 studies to explore asymmetry that might be explained by clinical, statistical, and methodological heterogeneity. The proportion of statistical heterogeneity will be examined using the I^2 measure [45] and the magnitude of statistical heterogeneity will be calculated using the restricted maximum likelihood [46]. Meta-regression will be conducted for clinically relevant subgroups or when extensive statistical heterogeneity is observed (e.g., $I^2 \geq 75\%$) [47]. This will allow the examination of the impact of important factors on our results, such as maternal age, dose, duration and timing (e.g., trimester) of antiepileptic treatment, co-morbidities, concomitant medications, risk of bias results, and sample size (due to Type II statistical power errors with rare adverse events). To ensure the meta-regression analysis is intuitive, the number of covariates examined will be less than 10% of the number of studies included in the meta-analysis for the particular outcome.

We anticipate that many of these outcomes will be rare. To deal with studies reporting zero events in one treatment arm, 0.5 will be added to the numerator and 1 will be added to the denominator. We will exclude studies reporting zero events in all treatment arms for a particular outcome [48,49]. We also anticipate that we will encounter missing data in the included studies. We will contact the study authors for this data and if we are unable to receive the data, we will impute missing data (e.g., measures of variance) using established methods [50]. To ensure that our imputations do not bias our results, we will conduct a sensitivity analysis [51]. The meta-analysis and meta-regression will be analyzed in R using the *metafor* command [52].

A random-effects network meta-analysis will be conducted to make inferences regarding the comparative safety of the various AEDs [15], as well as rank their safety using rankograms and the surface under the cumulative ranking curve [53]. We will ensure the following factors are present prior to conducting network meta-analysis: i) transitivity (i.e., comparable distribution of effect modifiers across comparisons), which will be examined using boxplots or percentages to visually inspect potential effect modifiers of treatment effect [54]; ii) consistency between direct and indirect data, which will be examined locally (i.e., in certain paths of the network) using the loop-specific method [55,56] and the node-splitting method [57], and globally (i.e., evaluating the network as a whole), using the design-by-treatment interaction model [58]; and iii) we will quantify the amount of variability attributed to heterogeneity and inconsistency rather than sampling error, by calculating the I^2 [59]. We will estimate the amount of heterogeneity using the restricted maximum likelihood method and assuming common within-network heterogeneity. We will compare the magnitude of heterogeneity between consistency and inconsistency models, as well as between meta-regression and network meta-analysis models to determine how much heterogeneity will be explained by inconsistency or the explanatory variable, respectively. We will first use the design-by-treatment model for the evaluation of inconsistency in a network as a whole and then, if inconsistency is detected, we will employ the loop-specific and node-splitting methods to identify which piece of evidence is responsible for inconsistency. As mentioned above, analyses will be stratified by treatment indication when clinically appropriate. Important heterogeneity and inconsistency will be explored using network meta-regression using the same methods as described above, as necessary.

Prior to conducting the network meta-analysis, we will hold a team meeting to finalize which treatment nodes will be included in the analysis since we are unclear about the indications, dosages, patient populations, and outcomes reported in all of the studies. We will discuss issues, including conducting a class versus independent drug analysis, inclusion of drug routes of administration and dosages, as well as timing of drug administration. These decisions will be examined through a sensitivity analysis in which we will classify treatment nodes using a different classification to see how stable our results are. The network meta-analysis results will be presented as summary treatment effects for each pair of treatments. Network meta-analysis will be conducted in Stata with the *mvmeta* routine [60].

A sequential approach will be used for the network meta-analysis. We will first restrict our analysis to RCTs, which will be the primary analysis of interest. We will then include data from quasi-experimental studies, and finally, data from observational studies. This will provide an understanding of the contribution of each type of study design to our summary estimates, providing us with information on how these agents work above and beyond clinical trials.

Discussion

Epilepsy is the most common chronic neurological condition, affecting 0.6 to 1% of the population [1,2]. Given that approximately a third of patients receiving AEDs are of reproductive age and almost half of pregnancies are unplanned [61], the fetus may be exposed to these in the first trimester of pregnancy, including during the critical stage of embryogenesis [62].

The comparative safety of these agents is currently unknown and our results will be important for policy-makers, healthcare providers, and women of childbearing age. To ensure our results have wide dissemination and uptake, we will publish our results in open access journals, present our findings at scientific conferences, conduct dissemination meetings with key stakeholders (including policy-makers and healthcare providers), and produce policy briefs for Health Canada, the organization that posed this query.

Additional files

Additional file 1: List of relevant medications.
Additional file 2: Excluded drugs.
Additional file 3: Draft eligibility criteria.
Additional file 4: MEDLINE literature search.

Abbreviations

AEDs: Anti-epileptic drugs; RCTs: Randomized clinical trials.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

ACT conceived and designed the study, helped obtain funding for the study, and helped write the draft protocol. EC registered the protocol with the PROSPERO database and edited the draft protocol. AV helped write the draft protocol. CS edited the draft protocol. BH, BRH, DM, and YF provided input into the design, helped obtain funding for the study, and edited the draft protocol. SES conceived the study, designed the study, obtained the funding, and helped write the draft protocol. All authors read and approved the final protocol.

Acknowledgements

This systematic review was funded by the Canadian Institutes of Health Research/Drug Safety and Effectiveness Network (CIHR/DSEN). ACT and BH are funded by a CIHR/DSEN New Investigator Award in Knowledge Synthesis. BRH receives funding from the Alberta Heritage Foundation for Medical Research. DM is funded by a University of Ottawa Research Chair. SES is funded by a Tier 1 Canada Research Chair in Knowledge Translation.
We thank Laure Perrier for conducting the literature searches and Becky Skidmore for peer reviewing the MEDLINE search strategy. We also thank Dr. Joseph Beyene for providing feedback on our original proposal and Wing Hui and Judy Tran for formatting the paper.

Author details

[1]Li Ka Shing Knowledge Institute, St. Michael's Hospital, 209 Victoria Street, East Building, Toronto, Ontario M5B 1 T8, Canada. [2]Institute of Health Policy Management and Evaluation, University of Toronto, Health Sciences Building, 155 College Street, Suite 425, Toronto, Ontario M5T 3 M6, Canada. [3]Clinical Epidemiology Program, Centre for Practice-Changing Research, Ottawa Hospital Research Institute, The Ottawa Hospital – General Campus and University of Ottawa, 501 Smyth Road, Box 711, Ottawa, Ontario K1H 8 L6, Canada. [4]Departments of Medicine and Community Health Sciences, University of Calgary, TRW Building, 3rd Floor, 3280 Hospital Drive NW, Calgary, Alberta T2N 4Z6, Canada. [5]The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada. [6]Department of Pediatrics, University of Toronto, 172 St. George Street, Toronto, Ontario M5R 0A3, Canada. [7]Department of Pharmacology and Toxicology, University of Toronto, Medical Sciences Building, 1 King's College Circle, Room 4207, Toronto, Ontario M5S 1A8, Canada. [8]Department of Geriatric Medicine, University of Toronto, 172 St. George Street, Toronto, Ontario M5R 0A3, Canada.

Received: 9 April 2014 Accepted: 17 June 2014
Published: 25 June 2014

References

1. Hauser WA, Hesdorffer D: *Epilepsy, Frequency, Causes and Consequences.* New York: Demos Publications; 1990.
2. Wiebe S, Bellhouse DR, Fallahay C, Eliasziw M: **Burden of epilepsy: the Ontario Health Survey.** *Can J Neurol Sci* 1999, **26**(4):263–270.
3. Sperling MR: **The consequences of uncontrolled epilepsy.** *CNS Spectr* 2004, **9**(2):98–101. 106–109.
4. Jones MW: **Consequences of epilepsy: why do we treat seizures?** *Can J Neurol Sci* 1998, **25**(4):S24–S26.
5. Dickenson AH, Ghandehari J: **Anti-convulsants and anti-depressants.** *Handb Exp Pharmacol* 2007, **177:**145–177.
6. Stefani A, Spadoni F, Bernardi G: **Voltage-activated calcium channels: targets of antiepileptic drug therapy?** *Epilepsia* 1997, **38**(9):959–965.
7. Snutch TP, Reiner PB: **Ca^{2+} channels: diversity of form and function.** *Curr Opin Neurobiol* 1992, **2**(3):247–253.
8. Spina E, Perugi G: **Antiepileptic drugs: indications other than epilepsy.** *Epileptic Disord* 2004, **6**(2):57–75.
9. Harden CL, Pennell PB, Koppel BS, Hovinga CA, Gidal B, Meador KJ, Hopp J, Ting TY, Hauser WA, Thurman D, Kaplan PW, Robinson JN, French JA, Wiebe S, Wilner AN, Vazquez B, Holmes L, Krumholz A, Finnell R, Shafer PO, Le Guen CL, American Academy of Neurology; American Epilepsy Society: **Management issues for women with epilepsy–focus on pregnancy (an evidence-based review): III. Vitamin K, folic acid, blood levels, and breast-feeding: report of the quality standards subcommittee and therapeutics and technology assessment subcommittee of the American Academy of Neurology and the American Epilepsy Society.** *Epilepsia* 2009, **50**(5):1247–1255.
10. Harden CL, Meador KJ, Pennell PB, Hauser WA, Gronseth GS, French JA, Wiebe S, Thurman D, Koppel BS, Kaplan PW, Robinson JN, Hopp J, Ting TY, Gidal B, Hovinga CA, Wilner AN, Vazquez B, Holmes L, Krumholz A, Finnell R, Hirtz D, Le Guen C, American Academy of Neurology; American Epilepsy Society: **Management issues for women with epilepsy-Focus on pregnancy (an evidence-based review): II. Teratogenesis and perinatal outcomes: Report of the Quality Standards Subcommittee and Therapeutics and Technology Subcommittee of the American Academy of Neurology and the American Epilepsy Society.** *Epilepsia* 2009, **50**(5):1237–1246.
11. Samren EB, van Duijn CM, Koch S, Hiilesmaa VK, Klepel H, Bardy AH, Mannagetta GB, Deichl AW, Gaily E, Granstrom ML, Meinardi H, Grobbee DE, Hofman A, Janz D, Lindhout D: **Maternal use of antiepileptic drugs and the risk of major congenital malformations: a joint European prospective study of human teratogenesis associated with maternal epilepsy.** *Epilepsia* 1997, **38**(9):981–990.
12. Meador K, Reynolds MW, Crean S, Fahrbach K, Probst C: **Pregnancy outcomes in women with epilepsy: a systematic review and meta-analysis of published pregnancy registries and cohorts.** *Epilepsy Res* 2008, **81**(1):1–13.
13. Adab N, Jacoby A, Smith D, Chadwick D: **Additional educational needs in children born to mothers with epilepsy.** *J Neurol Neurosurg Psychiatry* 2001, **70**(1):15–21.
14. Adab N, Kini U, Vinten J, Ayres J, Baker G, Clayton-Smith J, Coyle H, Fryer A, Gorry J, Gregg J, Mawer G, Nicolaides P, Pickering L, Tunnicliffe L, Chadwick DW: **The longer term outcome of children born to mothers with epilepsy.** *J Neurol Neurosurg Psychiatry* 2004, **75**(11):1575–1583.
15. Gaily E, Kantola-Sorsa E, Hiilesmaa V, Isoaho M, Matila R, Kotila M, Nylund T, Bardy A, Kaaja E, Granstrom ML: **Normal intelligence in children with prenatal exposure to carbamazepine.** *Neurology* 2004, **62**(1):28–32.
16. Meador KJ, Baker GA, Browning N, Clayton-Smith J, Combs-Cantrell DT, Cohen M, Kalayjian LA, Kanner A, Liporace JD, Pennell PB, Privitera M, Loring DW, for the NEAD Study Group: **Cognitive function at 3 years of age after fetal exposure to antiepileptic drugs.** *N Engl J Med* 2009, **360**(16):1597–1605.

17. Holmes LB, Wyszynski DF, Lieberman E: **The AED (antiepileptic drug) pregnancy registry: a 6-year experience.** *Arch Neurol* 2004, **61**(5):673–678.
18. Reinisch JM, Sanders SA, Mortensen EL, Rubin DB: **In utero exposure to phenobarbital and intelligence deficits in adult men.** *JAMA* 1995, **274**(19):1518–1525.
19. Morrow J, Russell A, Guthrie E, Parsons L, Robertson I, Waddell R, Irwin B, McGivern RC, Morrison PJ, Craig J: **Malformation risks of antiepileptic drugs in pregnancy: a prospective study from the UK Epilepsy and Pregnancy Register.** *J Neurol Neurosurg Psychiatry* 2006, **77**(2):193–198.
20. Meador KJ, Baker GA, Finnell RH, Kalayjian LA, Liporace JD, Loring DW, Mawer G, Pennell PB, Smith JC, Wolff MC, NEAD Study Group: **In utero antiepileptic drug exposure: fetal death and malformations.** *Neurology* 2006, **67**(3):407–412.
21. Vajda FJ, Hitchcock A, Graham J, Solinas C, O'Brien TJ, Lander CM, Eadie MJ: **Foetal malformations and seizure control: 52 months data of the Australian Pregnancy Registry.** *Eur J Neurol* 2006, **13**(6):645–654.
22. Holmes LB, Baldwin EJ, Smith CR, Habecker E, Glassman L, Wong SL, Wyszynski DF: **Increased frequency of isolated cleft palate in infants exposed to lamotrigine during pregnancy.** *Neurology* 2008, **70**(22 Pt 2):2152–2158.
23. Meador KJ, Penovich P, Baker GA, Pennell PB, Bromfield E, Pack A, Liporace JD, Sam M, Kalayjian LA, Thurman DJ, Moore E, Loring DW, NEAD Study Group: **Antiepileptic drug use in women of childbearing age.** *Epilepsy Behav* 2009, **15**(3):339–343.
24. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L: **Reporting Guidelines for Systematic Review Protocols.** In *19th Cochrane Colloquium: 19–22 October 2011; Madrid, Spain.*
25. Ross CJ, Visscher H, Sistonen J, Brunham LR, Pussegoda K, Loo TT, Rieder MJ, Koren G, Carleton BC, Hayden MR, CPNDS Consortium: **The Canadian Pharmacogenomics Network for Drug Safety: a model for safety pharmacology.** *Thyroid* 2010, **20**(7):681–687.
26. Eypasch E, Lefering R, Kum CK, Troidl H: **Probability of adverse events that have not yet occurred: a statistical reminder.** *BMJ* 1995, **311**(7005):619–620.
27. Atkins D: **Creating and synthesizing evidence with decision makers in mind: integrating evidence from clinical trials and other study designs.** *Med Care* 2007, **45**(10 Supl 2):S16–S22.
28. *Health Canada: Drug Product Database.* http://www.hc-sc.gc.ca/dhp-mps/prodpharma/databasdon/index-eng.php.
29. *United States National Library of Medicine's ChemIDPlus Lite Database.* http://chem.sis.nlm.nih.gov/chemidplus/chemidlite.jsp.
30. *Canadian Pharmacists Association: E-CPS (Compendium of Pharmaceuticals and Specialties).* http://www.e-therapeutics.ca/home.whatsnew.action.
31. *Epilepsy Canada: Anticonvulsant Medications.* http://www.epilepsy.ca/en-CA/Diagnosis-and-Treatment/Anticonvulsant-Medications.html.
32. *Epilepsy Ontario: Anticonvulsant/Anti-Seizure Medication from A to Z.* http://epilepsyontario.org/anticonvulsantanti-seizure-medication-from-a-to-z/.
33. Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, Lefebvre C: **An evidence-based practice guideline for the peer review of electronic search strategies.** *J Clin Epidemiol* 2009, **62**(9):944–952.
34. *Canadian Agency for Drugs and Technologies in Health: Grey Matters: A Practical Search Tool for Evidence-Based Medicine.* http://www.cadth.ca/resources/grey-matters.
35. Adab N, Tudur SC, Vinten J, Williamson P, Winterbottom J: **Common antiepileptic drugs in pregnancy in women with epilepsy.** *Cochrane Database Syst Rev* 2004, **3**:CD004848.
36. Banach R, Boskovic R, Einarson T, Koren G: **Long-term developmental outcome of children of women with epilepsy, unexposed or exposed prenatally to antiepileptic drugs: a meta-analysis of cohort studies.** *Drug Saf* 2010, **33**(1):73–79.
37. Landis JR, Koch GG: **The measurement of observer agreement for categorical data.** *Biometrics* 1977, **33**(1):159–174.
38. *Synthesi.SR.* http://knowledgetranslation.ca/sysrev/login.php.
39. Stone PW: **Popping the (PICO) question in research and evidence-based practice.** *Appl Nurs Res* 2002, **15**(3):197–198.
40. *Cochrane Effective Practice and Organization of Care Group Draft Risk of Bias Tool.* http://epoc.cochrane.org/epoc-author-resources.
41. *The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-Analyses.* http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
42. Santaguida PL, Raina P, Ismaila A: *The Development of the McHarm Quality Assessment Scale for Adverse Events.* Hamilton, Ontario: McMaster University; 2008.
43. Raudenbush SW: **Analyzing Effect Sizes: Random Effects Models.** In *The Handbook of Research Synthesis and Meta-analysis.* 2nd edition. Edited by Cooper H, Hedges LV, Valentine JC. New York: Russell Sage Foundation; 2009:295–315.
44. Viechtbauer W: **Bias and efficiency of meta-analytic variance estimators in the random-effects model.** *J Educ Behav Stat* 2005, **30**(3):261–293.
45. Higgins JP, Thompson SG: **Quantifying heterogeneity in a meta-analysis.** *Stat Med* 2002, **21**(11):1539–1558.
46. Viechtbauer W: **Confidence intervals for the amount of heterogeneity in meta-analysis.** *Stat Med* 2007, **26**(1):37–52.
47. Higgins JPT, Green S: *Cochrane Handbook for Systematic Reviews of Interventions.* The Cochrane Collaboration; 2009. http://www.cochrane.org/handbook.
48. Sweeting MJ, Sutton AJ, Lambert PC: **What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data.** *Stat Med* 2004, **23**(9):1351–1375.
49. Bradburn MJ, Deeks JJ, Berlin JA, Russell Localio A: **Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events.** *Stat Med* 2007, **26**(1):53–77.
50. Littell JH, Corcoran J, Pillai V: *Systematic Reviews and Meta-Analysis.* New York: Oxford University Press; 2008.
51. Carpenter J, Rucker G, Schwarzer G: **Assessing the sensitivity of meta-analysis to selection bias: a multiple imputation approach.** *Biometrics* 2011, **67**(3):1066–1072.
52. *Conducting Meta-Analyses in R with the metafor Package.* http://www.jstatsoft.org/v36/i03/.
53. Salanti G, Ades AE, Ioannidis JP: **Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial.** *J Clin Epidemiol* 2011, **64**(2):163–171.
54. Salanti G: **Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool.** *Res Synth Methods* 2012, **3**(2):80–97.
55. Song F, Altman DG, Glenny AM, Deeks JJ: **Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses.** *BMJ* 2003, **326**(7387):472.
56. Veroniki AA, Vasiliadis HS, Higgins JP, Salanti G: **Evaluation of inconsistency in networks of interventions.** *Int J Epidemiol* 2013, **42**(1):332–345.
57. Dias S, Welton NJ, Caldwell DM, Ades AE: **Checking consistency in mixed treatment comparison meta-analysis.** *Stat Med* 2010, **29**(7–8):932–944.
58. White IR, Barrett JK, Jackson D, Higgins JPT: **Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression.** *Res Synth Methods* 2012, **3**(2):111–125.
59. Jackson D, Barrett JK, Stephen R, White IR, Higgins JPT: **A design-by-treatment interaction model for network meta-analysis with random inconsistency effects.** *Stat Med* 2014, In press.
60. White IR: **Multivariate random-effects meta-regression: updates to mvmeta.** *Stata J* 2011, **11**(2):255–270.
61. Centers for Disease Control and Prevention: *Unintended Pregnancy Prevention.* http://www.cdc.gov/reproductivehealth/unintendedpregnancy/.
62. Yerby MS: **Pregnancy, teratogenesis, and epilepsy.** *Neurol Clin* 1994, **12**(4):749–771.

doi:10.1186/2046-4053-3-68
Cite this article as: Tricco *et al.*: Comparative safety of anti-epileptic drugs among infants and children exposed *in utero* or during breastfeeding: protocol for a systematic review and network meta-analysis. *Systematic Reviews* 2014 **3**:68.

PRISMA NMA Checklist

Section/Topic	Item #	Checklist Item*	Reported on Page #
TITLE			
Title	1	Identify the report as a systematic review *incorporating a network meta-analysis (or related form of meta-analysis).*	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: **Background:** main objectives **Methods:** data sources; study eligibility criteria, participants, and interventions; study appraisal; and *synthesis methods, such as network meta-analysis.* **Results:** number of studies and participants identified; summary estimates with corresponding confidence/credible intervals; *treatment rankings may also be discussed. Authors may choose to summarize pairwise comparisons against a chosen treatment included in their analyses for brevity.* **Discussion/Conclusions:** limitations; conclusions and implications of findings. **Other:** primary source of funding; systematic review registration number with registry name.	4-5
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known, *including mention of why a network meta-analysis has been conducted.*	7
Objectives	4	Provide an explicit statement of questions being addressed, with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	7
METHODS			
Protocol and registration	5	Indicate whether a review protocol exists and if and where it can be accessed (e.g., Web address); and, if available, provide registration information, including registration number.	8

Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. *Clearly describe eligible treatments included in the treatment network, and note whether any have been clustered or merged into the same node (with justification).*	8
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	9
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Additional File 1
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	9
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	9
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	Additional File 1
Geometry of the network	**S1**	Describe methods used to explore the geometry of the treatment network under study and potential biases related to it. This should include how the evidence base has been graphically summarized for presentation, and what characteristics were compiled and used to describe the evidence base to readers.	10-12
Risk of bias within individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	9-10 (see also Appendix A)
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means). *Also describe the use of additional summary measures assessed, such as treatment rankings and surface under the cumulative ranking curve (SUCRA) values, as well as modified approaches used to present summary findings from meta-analyses.*	10-12

Planned methods of analysis	14	Describe the methods of handling data and combining results of studies for each network meta-analysis. This should include, but not be limited to: • *Handling of multi-arm trials;* • *Selection of variance structure;* • *Selection of prior distributions in Bayesian analyses; and* • *Assessment of model fit.*	10-12
Assessment of Inconsistency	**S2**	Describe the statistical methods used to evaluate the agreement of direct and indirect evidence in the treatment network(s) studied. Describe efforts taken to address its presence when found.	10-11
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	9-10
Additional analyses	16	Describe methods of additional analyses if done, indicating which were pre-specified. This may include, but not be limited to, the following: • Sensitivity or subgroup analyses; • Meta-regression analyses; • *Alternative formulations of the treatment network; and* • *Use of alternative prior distributions for Bayesian analyses (if applicable).*	11-12
RESULTS[†]			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	13 and Figure 1
Presentation of network structure	**S3**	Provide a network graph of the included studies to enable visualization of the geometry of the treatment network.	Figure 2
Summary of network geometry	**S4**	Provide a brief overview of characteristics of the treatment network. This may include commentary on the abundance of trials and randomized patients for the different interventions and pairwise comparisons in the network, gaps of evidence in the treatment network, and potential biases reflected by the network structure.	14-18
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Table 1, Appendices D and E

Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment.	Appendix F
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: 1) simple summary data for each intervention group, and 2) effect estimates and confidence intervals. *Modified approaches may be needed to deal with information from larger networks.*	N/A (data can be provided by the corresponding author)
Synthesis of results	21	Present results of each meta-analysis done, including confidence/credible intervals. *In larger networks, authors may focus on comparisons versus a particular comparator (e.g. placebo or standard care), with full findings presented in an appendix. League tables and forest plots may be considered to summarize pairwise comparisons.* If additional summary measures were explored (such as treatment rankings), these should also be presented.	15-18, Figure 3, Appendices H, I, J
Exploration for inconsistency	**S5**	Describe results from investigations of inconsistency. This may include such information as measures of model fit to compare consistency and inconsistency models, *P* values from statistical tests, or summary of inconsistency estimates from different parts of the treatment network.	14 (see also Appendix H)
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies for the evidence base being studied.	14 (see also Appendix G)
Results of additional analyses	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression analyses, *alternative network geometries studied, alternative choice of prior distributions for Bayesian analyses,* and so forth).	Appendix K
DISCUSSION			
Summary of evidence	24	Summarize the main findings, including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy-makers).	19-21
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review level (e.g., incomplete retrieval of identified research, reporting bias). *Comment on the validity of the assumptions, such as transitivity and consistency. Comment on any concerns regarding network geometry (e.g., avoidance of certain comparisons).*	21-23

Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	23
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. This should also include information regarding whether funding has been received from manufacturers of treatments in the network and/or whether some of the authors are content experts with professional conflicts of interest that could affect use of treatments in the network.	26-27

Abbreviations: PICOS - population, intervention, comparators, outcomes, study design
* Text in italics indicates wording specific to reporting of network meta-analyses that has been added to guidance from the PRISMA statement.
† Authors may wish to plan for use of appendices to present all relevant information in full detail for items in this section.

Supplementary Online Content

Appendix A. Newcastle-Ottawa Scale scoring guide

COHORT Studies

Excel Column	NOS* Answer Options**	NOS Coding Manual*
RefID	Enter the report's RefID.	
DA	Enter your initials.	
First author	Enter the first author's last name.	
Year of publication	Enter the year of the publication.	
SELECTION:		
1) Representative-ness of the exposed cohort	a) truly representative of the average pregnant woman taking AEDs in the community b) somewhat representative of the average pregnant woman taking AEDs in the community c) selected group of users e.g., nurses, volunteers d) no description of the derivation of the cohort	Item is assessing the representativeness of exposed individuals in the community, not the representativeness of the sample of women from some general population. For example, subjects derived from groups likely to contain middle class, better educated, health oriented women are likely to be representative of postmenopausal estrogen users while they are not representative of all women (e.g. members of a health maintenance organisation (HMO) will be a representative sample of estrogen users. While the HMO may have an under-representation of ethnic groups, the poor, and poorly educated, these excluded groups are not the predominant users of estrogen). <u>Note:</u> Truly representative (A) is a population-based cohort at the provincial or national levels (e.g., a sample from 2 cities is not enough). We need very 'broad' sample of the population. Somewhat representative (B) includes private clinics, hospital-based, or community-based.
2) Selection of the non-exposed cohort	a) drawn from the same community as the exposed	<u>Note:</u> In our review of mostly multi-arm studies, this question pertains to the

	cohort b) drawn from a different source c) no description of the derivation of the non-exposed cohort	study's comparator group(s) – including "active" controls (for example, a less teratogenic AED). Therefore, this will often be 'A' for our studies.
3) Ascertainment of exposure	a) secure record (e.g., surgical records) b) structured interview c) written self-report d) no description	Note: Option 'A' includes patient hospital records, prescription drug database, or hospital/clinic visits (e.g., patient is asked about "current" AED use during a visit with their doctor). Option 'B' includes a hospital/clinic visit, but the patients are asked to remember their AED use during pregnancy (e.g., retrospectively ascertained exposure). If a study used both medical records and interviews for everyone, select 'A'.
4) Demonstration that outcome of interest was not present at start of study	a) yes b) no	In the case of mortality studies, outcome of interest is still the presence of a disease/incident, rather than death. That is to say that a statement of 'no history of disease or incident' earns a star (i.e. option 'A'). Note: Since our review is on pregnant women, this question is 'A' for all. **Please email us if a study involves breastfeeding women.**
COMPARABILITY:		
1) Comparability of cohorts on the basis of the design or analysis	a) answer is BOTH B & C (i.e. study controls for age and one other important factor) b) study controls for age of the women c) study controls for any other important factor d) study does not control for any	Either exposed and non-exposed individuals must be matched in the design and/or confounders must be adjusted for in the analysis. Statements of no differences between groups or that differences were not statistically significant are not sufficient for establishing comparability. Note: If the relative risk for the exposure of interest is adjusted for the confounders listed, then the groups will be considered to be comparable on each variable used in the adjustment.

	important factor or it is not described	There may be multiple ratings for this item for different categories of exposure (e.g., ever vs. never, current vs. previous or never). [A maximum of 2 stars can be allotted in this category]. Note: The study should have initially matched the groups or presented adjusted odds ratios, AND in addition, since in our review we are analyzing each AED arm separately (instead of the whole exposed cohort), the study must also report the factor of interest **for 'each AED arm'** (or state that **'each AED arm'** is matched). **Thus, there are 2 parts to this question:** 1) The study should have matched/adjusted for age at whatever level of groups they were focused on (even if they aren't our abstracted AED arms); AND 2) Then the study should also have reported the age for each AED arm. If they haven't done both of these 2 things, it's a 'D' here (unless they happen to combine these by reporting adjusted ORs for each of our AED arms). For our review, this generally pertains to **the comparability of the MOTHERS.** The exception here is in studies of cognitive/psychomotor development disorders in children - when age of the children should be comparable. The "other important factors" here are any one of these: • history of congenital malformations (CMs), fetal losses, preterm deliveries or small babies. • family history of genetic problems or CMs.

		• alcohol use. • nutritional deficiencies (e.g., lack of folic acid). <u>Example:</u> - Option 'B' indicates that the study initially matched groups based on the women's age (or reported adjusted ORs) AND they report the mean women's age for EACH of our arms (e.g., for Tx1, Tx2, etc.).
OUTCOME:		
1) Assessment of outcome	a) independent OR blind assessment b) record linkage c) self-report d) no description	For some outcomes (e.g. fractured hip), reference to the medical record is sufficient to satisfy the requirement for confirmation of the fracture. This would not be adequate for vertebral fracture outcomes where reference to x-rays would be required. a) Independent or blind assessment stated in the paper, or confirmation of the outcome by reference to secure records (x-rays, medical records, etc.) b) Record linkage (e.g. identified through ICD codes on database records) c) Self-report (i.e. no reference to original medical records or x-rays to confirm the outcome) d) No description. <u>Note:</u> Blind (A) is if they tell us that the outcome assessors were blinded to exposures; or if the outcome is objective. For our purposes, we will focus on the primary outcome of interest of our systematic review, which is <u>major malformations (an objective outcome)</u>. So most of ours will be A, unless the study is only on a secondary outcome (e.g., cognitive development) and is based on the mother's self-report of their child (e.g., not a clinical examination).
2) Was follow-up	a) yes	An acceptable length of time should be decided before quality assessment

long enough for outcomes to occur	b) no	begins (e.g. 5 yrs. for exposure to breast implants) Note: For this component, focus only on the outcomes that are reported in the results. For our purposes, we will focus on the primary outcome of interest of our systematic review, which is major malformations. • For studies focusing on 'birth' outcomes (i.e. malformations, preterm, fetal losses, born small), the answer is 'A' if they follow the groups until birth. • For studies focusing on cognitive developmental disorders, an adequate follow-up period (i.e. child's age) is 4 years. • For studies focusing on psychomotor delays, an adequate follow-up period is the earliest point of detection of the disorder. • For studies focusing on neonatal seizures, an adequate follow-up period (i.e. infant's age) is 6 months.
3) Adequacy of follow up of cohorts	a) complete follow up - all subjects accounted for b) subjects lost to follow up unlikely to introduce bias - small number lost (see 'Note'), or description provided of those lost c) follow up rate is inadequate (see 'Note') and no description of those lost d) no statement	This item assesses the follow-up of the exposed and non-exposed cohorts to ensure that losses are not related to either the exposure or the outcome. Note: **Especially check ones that start their total sample size (or figure diagram) with only the ones who had "complete" data (or only those who they had "successfully" recruited), as these are often a 'D' (since they don't report on the ones NOT followed up).** • For a prospective study, ≥90% follow-up rate per year is adequate (e.g., 10% dropout or less for 1 year, 20% for 2 years of follow-up, etc.). This includes missing or incomplete data, etc. • For a retrospective cohort study, ≥80% follow-up rate is adequate; including the ones that they could NOT recruit or who would NOT participate. • For a survey/mail questionnaire, ≥75% response rate is adequate. (For

		a survey, a dropout rate is congruent to a survey response rate).

CASE-CONTROL Studies

Excel Column	NOS* Answer Options**	NOS Coding Manual*
RefID	Enter the report's RefID.	
DA	Enter your initials.	
First author	Enter the first author's last name.	
Year of publication	Enter the year of the publication.	
SELECTION:		
1) Is the case definition adequate?	a) yes, with independent validation b) yes, e.g., record linkage or based on self-reports c) no description	a) Requires some independent validation (e.g. >1 person/record/time/ process to extract information, or reference to primary record source such as x-rays or medical/hospital records) b) Record linkage (e.g. ICD codes in database) or self-report with no reference to primary record c) No description Note: This question is assessing the group of infants that have the outcome of interest (e.g., CMs) – i.e. the "cases" in a case-control study design.
2) Representative-ness of the cases	a) consecutive or obviously representative series of cases b) potential for selection biases, or not stated	a) All eligible cases with outcome of interest over a defined period of time, all cases in a defined catchment area, all cases in a defined hospital or clinic, group of hospitals, health maintenance organisation, or an appropriate sample of those cases (e.g. random sample) b) Not satisfying requirements in part (a), or not stated. Note: Option 'A' is a population-based sample.
3) Selection of controls	a) community controls b) hospital controls c) no description	This item assesses whether the control series used in the study is derived from the same population as the cases and essentially would have been cases had the outcome been present.

		a) Community controls (i.e. same community as cases and would be cases if had outcome) b) Hospital controls, within same community as cases (i.e. not another city) but derived from a hospitalised population c) No description Note: This question is assessing the group of infants that don't have the outcome (e.g., CMs) – i.e. the "controls" in a case-control study design. Community controls (A) includes a population-based sample.
4) Definition of controls	a) no history of disease (endpoint) b) no description of source	a) If cases are first occurrence of outcome, then it must explicitly state that controls have no history of this outcome. If cases have new (not necessarily first) occurrence of outcome, then controls with previous occurrences of outcome of interest should not be excluded. b) No mention of history of outcome Note: Since our review is on fetal effects, this question is 'A' for all studies. **Please email us if a study involves exposure during breastfeeding.**
COMPARABILITY:		
1) Comparability of cases and controls on the basis of the design or analysis	a) answer is BOTH B & C (i.e. study controls for age and one other important factor) b) study controls for age of the women c) study controls for any other important factor d) study does not control for any important factor or it is not described	Either cases and controls must be matched in the design and/or confounders must be adjusted for in the analysis. Statements of no differences between groups or that differences were not statistically significant are not sufficient for establishing comparability. Note: If the odds ratio for the exposure of interest is adjusted for the confounders listed, then the groups will be considered to be comparable on each variable used in the adjustment. There may be multiple ratings for this item for different categories of exposure (e.g. ever vs. never, current vs. previous or never). [A maximum

		of 2 stars can be allotted in this category]. Note: The study should have initially matched the groups, AND in addition, since in our review we are analyzing each AED arm separately (instead of the whole cases group), the study must also report the factor of interest **for 'each AED arm'** (or state that **'each AED arm'** is matched). For our review, this generally pertains to **the comparability of the MOTHERS of the cases and controls.** The exception here is in studies of cognitive/psychomotor development disorders in children - when age of the children should be comparable. The "other important factors" here are any one of these: • history of congenital malformations (CMs), fetal losses, preterm deliveries or small babies. • family history of genetic problems or CMs. • alcohol use. • nutritional deficiencies (e.g., lack of folic acid). For example, Option 'B' indicates that the study initially matched groups based on the women's age AND they report the mean women's age for EACH arm (e.g., for Tx1, Tx2, etc.).
EXPOSURE:		
1) Assessment of exposure	a) secure record (e.g., surgical records) b) structured interview where blind to case/control status c) interview not blinded to case/control status d) written self-report or medical	Note: Option 'A' includes patient hospital records, prescription drug database, or hospital/clinic visits (e.g., patient is asked about "current" AED use during a visit with their doctor). "Interview" here includes a hospital/clinic visit, but the patients are asked to remember their AED use during pregnancy (e.g., retrospectively

	record only e) no description	ascertained exposure).
2) Same method of ascertainment for cases and controls	a) yes b) no	Note: This question is asking whether the method of ascertainment of exposure was the same for 'cases' (with the outcome) and 'controls' (without the outcome; in this case-control study design).
3) Non-response rate	a) same rate for both groups b) non-respondents described c) rate different and no designation	Note: For our review, this pertains to either the infants or the mothers of the case and control groups. We're allowing 10% dropout per year for a prospective study – e.g., 10% for 1 year, 20% for 2 years of follow-up, etc. For a survey, we allow for a 75% response rate in order for it be adequate. For a survey, a dropout rate is congruent to a survey response rate.

*Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analyses. Available at: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

In the **"NOS Coding Manual" column, the first section for each item is copied straight from the NOS documentation while the lower portions in each item are our "Notes" tailored for the AED review.

Appendix B. List of included studies

A total of 29 cohort studies[1-29] with 9 companion reports[30-38] were included:

1. Adab N, Kini U, Vinten J, et al. The longer term outcome of children born to mothers with epilepsy. *J Neurol Neurosurg Psychiatry*. 2004;75(11):1575-83.
2. Arkilo D, Hanna J, Dickens D, et al. Pregnancy and neurodevelopmental outcomes with in-utero antiepileptic agent exposure. A pilot study. *Eur J Paediatr Neurol*. 2015;19(1):37-40.
3. Bromley R, Baxter N, Calderbank R, Mawer G, Clayton-Smith J, Baker G. A comprehensive review of the language abilities of children exposed to valproate or carbamazepine in utero. American Epilepsy Society; Texas2010.
4. Bromley RL, Mawer GE, Briggs M, et al. The prevalence of neurodevelopmental disorders in children prenatally exposed to antiepileptic drugs. *J Neurol Neurosurg Psychiatry*. 2013;84(6):637-43.
5. Bromley RL, Calderbank R, Cheyne CP, et al. Cognition in school-age children exposed to levetiracetam, topiramate, or sodium valproate. *Neurology*. 2016;87(18):1943-53.
6. Christensen J, Gronborg TK, Sorensen MJ, et al. Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. *JAMA*. 2013;309(16):1696-703.
7. Cohen MJ, Meador KJ, Browning N, et al. Fetal antiepileptic drug exposure: Adaptive and emotional/behavioral functioning at age 6years. *Epilepsy Behav*. 2013;29(2):308-15.
8. Cummings C, Stewart M, Stevenson M, Morrow J, Nelson J. Neurodevelopment of children exposed in utero to lamotrigine, sodium valproate and carbamazepine. *Arch Dis Child*. 2011;96(7):643-7.
9. Dean JCS, Hailey H, Moore SJ, Lloyd DJ, Turnpenny PD, Little J. Long term health and neurodevelopment in children exposed to antiepileptic drugs before birth. *J Med Genet*. 2002;39(4):251-9.
10. D'Souza SW, Robertson IG, Donnai D, Mawer G. Fetal phenytoin exposure, hypoplastic nails, and jitteriness. *Arch Dis Child*. 1991;66(3):320-4.
11. Eriksson K, Viinikainen K, Mönkkönen A, et al. Children exposed to valproate in utero—Population based evaluation of risks and confounding factors for long-term neurocognitive development. *Epilepsy Res*. 2005;65(3):189-200.
12. Gaily E. Development and growth in children of epileptic mothers: a prospective controlled study. Helsinki, Finland: University of Helsinki; 1990.
13. Gogatishvili N, Ediberidze T, Lomidze G, Tatishvili N, Kasradze S. PO-0834 Long-term Developmental Outcome Of Children Prenatally Exposed To Antiepileptic Drugs. *Arch Dis Child*. 2014;99(Suppl 2):A526.
14. Gogatishvili N, Ediberidze T, Lomidze G, Tatishvili N, Kasradze S. Cognitive outcomes of children with fetal antiepileptic drug exposure at the age of 3-6 years-preliminary data. 1st Congress of the European Academy of Neurology; Berlin: European Journal of Neurology; 2015. p. 329.
15. Hurault-Delarue C, Damase-Michel C, Finotto L, et al. Psychomotor developmental effects of prenatal exposure to psychotropic drugs: a study in EFEMERIS database. *Fundam Clin Pharmacol*. 2016;30(5):476-82.
16. Jones KL, Lacro RV, Johnson KA, Adams J. Pattern of malformations in the children of women treated with carbamazepine during pregnancy. *N Engl J Med*. 1989;320(25):1661-6.

17. Katz JM, Pacia SV, Devinsky O. Current Management of Epilepsy and Pregnancy: Fetal Outcome, Congenital Malformations, and Developmental Delay. *Epilepsy Behav*. 2001;2(2):119-23.
18. Koch S, Jager-Roman E, Losche G, Nau H, Rating D, Helge H. Antiepileptic drug treatment in pregnancy: drug side effects in the neonate and neurological outcome. *Acta Paediatr*. 1996;85(6):739-46.
19. Mawer G, Clayton-Smith J, Coyle H, Kini U. Outcome of pregnancy in women attending an outpatient epilepsy clinic: adverse features associated with higher doses of sodium valproate. *Seizure*. 2002;11(8):512-8.
20. Miskov S, Juraski RG, Fucic A, et al. Croatian Pregnant Women with Epilepsy and Effects of Antiepileptic Drugs Exposure in their Offspring - seven years of prospective surveillance. American Epilepsy Society; Texas2010.
21. Miskov S, Juraski RG, Mikula I, et al. The Croatian model of integrative prospective management of epilepsy and pregnancy. *Acta Clin Croat*. 2016;55(4):535-48.
22. Nadebaum C, Anderson VA, Vajda F, Reutens DC, Barton S, Wood AG. Language skills of school-aged children prenatally exposed to antiepileptic drugs. *Neurology*. 2011;76(8):719-26.
23. Rihtman T, Parush S, Ornoy A. Developmental outcomes at preschool age after fetal exposure to valproic acid and lamotrigine: cognitive, motor, sensory and behavioral function. *Reprod Toxicol*. 2013;41:115-25.
24. Scolnik D, Nulman I, Rovet J, et al. Neurodevelopment of children exposed in utero to phenytoin and carbamazepine monotherapy. *JAMA*. 1994;271(10):767-70.
25. Shankaran S, Woldt E, Nelson J, Bedard M, Delaney-Black V. Antenatal phenobarbital therapy and neonatal outcome. II: Neurodevelopmental outcome at 36 months. *Pediatrics*. 1996;97(5):649-52.
26. van der Pol MC, Hadders-Algra M, Huisjes HJ, Touwen BC. Antiepileptic medication in pregnancy: late effects on the children's central nervous system development. *Am J Obstet Gynecol*. 1991;164(1 Pt 1):121-8.
27. Veiby G, Engelsen BA, Gilhus NE. Early child development and exposure to antiepileptic drugs prenatally and through breastfeeding: a prospective cohort study on children of women with epilepsy. *JAMA Neurol*. 2013;70(11):1367-74.
28. Veiby G, Daltveit AK, Schjolberg S, et al. Exposure to antiepileptic drugs in utero and child development: a prospective population-based study. *Epilepsia*. 2013;54(8):1462-72.
29. Wood AG, Nadebaum C, Anderson V, et al. Prospective assessment of autism traits in children exposed to antiepileptic drugs during pregnancy. *Epilepsia*. 2015;56(7):1047-55.
30. Bromley RL, Mawer G, Clayton-Smith J, Baker GA. Autism spectrum disorders following in utero exposure to antiepileptic drugs. *Neurology*. 2008;71(23):1923-4.
31. Gaily EK, Granstrom ML, Hiilesmaa VK, Bardy AH. Head circumference in children of epileptic mothers: contributions of drug exposure and genetic background. *Epilepsy Res*. 1990;5(3):217-22.
32. Hiilesmaa V. A prospective study on maternal and fetal outcome in 139 women with epilepsy. Helsinki: University of Helsinki; 1982.
33. Hiilesmaa VK, Bardy A, Teramo K. Obstetric outcome in women with epilepsy. *Am J Obstet Gynecol*. 1985;152(5):499-504.
34. Rasalam AD, Hailey H, Williams JH, et al. Characteristics of fetal anticonvulsant syndrome associated autistic disorder. *Dev Med Child Neurol*. 2005;47(8):551-5.

35. Tomson T, Battino D, Bonizzoni E, et al. Antiepileptic drugs and intrauterine death: A prospective observational study from EURAP. *Neurology*. 2015;85(7):580-8.
36. Viinikainen K, Eriksson K, Monkkonen A, et al. The effects of valproate exposure in utero on behavior and the need for educational support in school-aged children. *Epilepsy Behav*. 2006;9(4):636-40.
37. Vinten J, Adab N, Kini U, Gorry J, Gregg J, Baker GA. Neuropsychological effects of exposure to anticonvulsant medication in utero. *Neurology*. 2005;64(6):949-54.
38. Vinten J, Bromley RL, Taylor J, Adab N, Kini U, Baker GA. The behavioral consequences of exposure to antiepileptic drugs in utero. *Epilepsy Behav*. 2009;14(1):197-201.

Appendix C. Key excluded studies

Author, Year	Research Group	Title	Reason for Exclusion
Meador, 2009[39]	Neurodevelopmental Effects of Antiepileptic Drug (NEAD) Study Group	Cognitive Function at 3 Years of Age after Fetal Exposure to Antiepileptic Drugs	Outcomes only reported as continuous variables
Meador, 2010[40]	Neurodevelopmental Effects of Antiepileptic Drug (NEAD) Study Group	Effects of breastfeeding in children of women taking antiepileptic drugs	Outcomes only reported as continuous variables
Meador, 2011[41]	Neurodevelopmental Effects of Antiepileptic Drug (NEAD) Study Group	Foetal antiepileptic drug exposure and verbal versus non-verbal abilities at three years of age	Outcomes only reported as continuous variables
Meador, 2012[42]	Neurodevelopmental Effects of Antiepileptic Drug (NEAD) Study Group	Effects of fetal antiepileptic drug exposure: Outcomes at age 4.5 years	Outcomes only reported as continuous variables
Meador, 2013[43]	Neurodevelopmental Effects of Antiepileptic Drug (NEAD) Study Group	Fetal antiepileptic drug exposure and cognitive outcomes at age 6 years (NEAD study): a prospective observational study	Outcomes only reported as continuous variables
Shallcross, 2011[44]	Liverpool and Manchester Neurodevelopment Group and The UK Epilepsy and Pregnancy Register	Child development following in utero exposure: Levetiracetam vs. sodium valproate	Outcomes only reported as continuous variables
Shallcross, 2014[45]	Liverpool and Manchester	In utero exposure to levetiracetam vs. valproate: Development and language at 3 years of age	Outcomes only reported as continuous variables

	Neurodevelopment Group and The UK Epilepsy and Pregnancy Register		

References

39. Meador KJ, Baker GA, Browning N, et al. Cognitive function at 3 years of age after fetal exposure to antiepileptic drugs. *N Engl J Med*. 2009;360(16):1597-605.
40. Meador KJ, Baker GA, Browning N, et al. Effects of breastfeeding in children of women taking antiepileptic drugs. *Neurology*. 2010;75(22):1954-60.
41. Meador KJ, Baker GA, Browning N, et al. Foetal antiepileptic drug exposure and verbal versus non-verbal abilities at three years of age. *Brain*. 2011;134(Pt 2):396-404.
42. Meador KJ, Baker GA, Browning N, et al. Effects of fetal antiepileptic drug exposure: outcomes at age 4.5 years. *Neurology*. 2012;78(16):1207-14.
43. Meador KJ, Baker GA, Browning N, et al. Fetal antiepileptic drug exposure and cognitive outcomes at age 6 years (NEAD study): a prospective observational study. *Lancet Neurol*. 2013;12(3):244-52.
44. Shallcross R, Bromley RL, Irwin B, Bonnett LJ, Morrow J, Baker GA. Child development following in utero exposure: levetiracetam vs sodium valproate. *Neurology*. 2011;76(4):383-9.
45. Shallcross R, Bromley RL, Cheyne CP, et al. In utero exposure to levetiracetam vs valproate: development and language at 3 years of age. *Neurology*. 2014;82(3):213-21.

Appendix D. Table of Individual Study characteristics

Author, Year	Country of conduct	Registry or Setting	Study period	Interventions	Outcomes	Funding
Adab, 2004[*1] [CR: Vinten 2005[37] Vinten, 2009[38]]	UK	Mersey Regional Epilepsy Clinic; Epilepsy Clinic at the Manchester Royal Infirmary; Antenatal clinic at St Mary's Hospital, Manchester	2000-2001	Carbam, Control, Valpro	Cognitive Developmental Delay, Psychomotor Developmental Delay	NR
Arkilo, 2015[2]	USA	Minnesota Epilepsy Group	2006-2011	Carbam, Lamot, Levet, Pheny, Valpro	Autism/Dyspraxia, Psychomotor Developmental Delay	NR
Bromley, 2010[3]	UK	Liverpool and Manchester Neurodevelopment Group	NR	Carbam, Valpro	Language Delay	NR
Bromley, 2013[4] [CR: Bromley, 2008[30]]	UK	Liverpool and Manchester Neurodevelopment group	2000-2004	Carbam, Control, Lamot, Valpro	Autism/Dyspraxia, ADHD	mixed public & private
Bromley, 2016[5]†	UK	UK Epilepsy and Pregnancy Register	2004-2007	Control, Gabap, Levet, Topir, Valpro	Cognitive Developmental Delay	public
Christensen, 2013[6]†	Denmark	Danish Civil Registration System; Danish Prescription Register; Danish Psychiatric Central Register; Danish Birth Register; Danish	1996-2006	Carbam, Clonaz, Lamot, Oxcar, Valpro	Autism/Dyspraxia	public

		National Hospital Register				
Cohen, 2013[7]	USA;UK	Neurodevelopmental Effects of Antiepileptic Drugs Study Group	1999-2004	Carbam, Lamot, Pheny, Valpro,	ADHD	public
Cummings, 2011[8]† [CR: Tomson, 2015[35]]	Northern Ireland	UK Epilepsy and Pregnancy Register (Northern Ireland); Northern Ireland Child Health System	1996-2005	Carbam, Lamot, Valpro,	Cognitive Developmental Delay	public
Dean, 2002[9] [CR: Rasalam, 2005[34]]	Scotland	Aberdeen Maternity Hospital	1976-2000	Carbam, Carbam+Pheno, Carbam+Pheny, Carbam+Valpro, Control, Ethos, Pheno, Pheno+Pheny, Pheno+Valpro, Pheny, Primid, Valpro	Psychomotor Developmental Delay, ADHD	NR
D'Souza, 1991[10]	United Kingdom	St Mary's Hospital	1980-1982	Carbam, Control, Pheno, Pheny, Valpro	Cognitive Developmental Delay	public
Eriksson, 2005[11]† [CR: Viinikainen, 2006[36]]	Finland	Kuopio University Hospital	1989-2000	Carbam, Control, Valpro	Cognitive Developmental Delay, Psychomotor Developmental Delay	public

Gaily, 1990[12] [CR: Gaily, 1990[31]; Hiilesmaa, 1982[32]; Hiilesmaa, 1985[33]]	Finland	Helsinki University Central Hospital	1975-1979	Carbam, Carbam+Pheno+Pheny, Carbam+Pheny, Carbam+Valpro, Control, Ethos+Pheny, Pheno+Pheny, Pheny, Pheny+Primid, Pheny+Valpro	Cognitive Developmental Delay , Psychomotor Developmental Delay	mixed public & private
Gogatishvili, 2014[13]	Georgia	Georgian National AED-Pregnancy Registry	NR	Carbam, Lamot, Valpro	Cognitive Developmental Delay	public
Gogatishvili, 2015[14]	Georgia	Georgian National AED-Pregnancy Registry	NR	Carbam, Carbam+Levet, Lamot, Pheno, Valpro	Language Delay	public
Hurault-Delarue, 2012[15]	France	EFEMERIS database - Caisse Primaire d'Assurance Maladie of Haute-Garonne and Maternal and Infant Protection Service; Antenatal Diagnostic Centre	2004-2008	Carbam, Clobaz, Clonaz, Gabap, Lamot, Pheno, Topir, Valpro	Psychomotor Developmental Delay	NR
Jones, 1989[16]†	US	California Teratogen Registry	1979-1988	Carbam, Carbam+Pheno, Carbam+Pheno+Valpro, Carbam+Primid	Cognitive Developmental Delay , Psychomotor Developmental Delay	public
Katz, 2001[17]	USA	Mount Sinai Comprehensive Epilepsy Center	1990-2000	Carbam, Control, Lamot, Pheno, Pheny, Primid, Valpro	Cognitive Developmental Delay	NR

Koch, 1996[18]	Germany	NR	1976-1983	Pheno, Pheny, Primid, Valpro	Cognitive Developmental Delay	public
Mawer, 2002[19]	England	Manchester Royal Infirmary	1990-1999	Carbam, Lamot, Pheny, Valpro	Cognitive Developmental Delay	NR
Miskov, 2010[20]	Croatia	NR	2003-2010	Carbam, Control, Gabap, Lamot, Valpro	Psychomotor Developmental Delay, Neonatal Seizures	NR
Miskov, 2016[21]	Croatia	Sestre milosrdnice University Hospital Center	2003-2013	Carbam, Carbam+Lamot, Carbam+Pheno, Carbam+Pheny+Topir, Control, Clonaz+Valpro, Gabap, Lamot, Oxcar, Pheno, Pheny, Topir+Valpro, Valpro	Attention Deficit Hyperactivity Disorder	NR
Nadebaum, 2011[22]†	Australia	Australian Registry of Antiepileptic Drug Use in Pregnancy	2007-2009	Carbam, Lamot, Valpro	Language Delay	mixed public & private
Rihtman, 2013[23]	Israel	Israeli Teratogen Information Service	NR	Lamot, Valpro	Neonatal Seizure	mixed public & private
Scolnik, 1994[24]	Canada	Hospital for Sick Children - Motherisk Program; North York General Hospital; Toronto Hospital; Oshawa General Hospital	1987-1992	Carbam, Pheny	Cognitive Developmental Delay	public

Shankaran, 1996[25]	USA	Children's Hospital of Michigan	NR	Control, Pheno	Psychomotor Developmental Delay, Language Delay	public
Van der Pol, 1991[26]	Netherlands	Groningen University Hospital	1973-1981	Carbam, Carbam+Pheno, Control, Pheno	Psychomotor Developmental Delay	public
Veiby, 2013a[27]†	Norway	Norwegian Institute of Public Health- Mother and Child Cohort Study	1999-2009	Carbam, Control, Lamot, Valpro	Social Impairment	public
Veiby, 2013b[28]†	Norway	Medical Birth Registry of Norway	1999-2008	Carbam, Control, Lamot, Valpro	Psychomotor Developmental Delay, Autism/Dyspraxia, Language Delay, ADHD	public
Wood, 2015[29]†	Australia	Australian Registry of Antiepileptic Drug Use in Pregnancy	2007-2010	Carbam, Carbam+Clonaz, Carbam+Lamot, Carbam+Pheny, Lamot+Valpro, Valpro	Autism/Dyspraxia	public

Abbreviations: ADHD – Attention Deficit Hyperactivity Disorder; NR – Not Reported

Carbam = Carbamazepine; Clobaz = Clobazam; Clonaz = Clonazepam; Ethos = Ethosuximide; Gabap = Gabapentin; Lamot = Lamotrigine; Levet = Levetiracetam; Oxcar = Oxcarbazepine; Pheno = Phenobarbital; Pheny = Phenytoin; Primid = Primidone; Topir = Topiramate; Valpro = Valproate; Vigab = Viagabatrin

*Single publication reporting on two separate cohorts
†Registry Studies

Appendix E. Table of Patient characteristics

Author, Year	Indication	Sample Size*	Mean Age (Women)	Mean Age (Children)/ Follow-up period†	AED Exposure Timing	Maternal Alcohol Use n/N‡	Maternal Tobacco Use n/N‡
Adab, 2004a[1]§ [CR: Vinten 2005[37]; Vinten, 2009[38]]	Epilepsy	177	26.1	9-10.5	NR	24/279‡	68/249‡
Adab, 2004b[1]§ [CR: Vinten 2005[37]; Vinten, 2009[38]]	Epilepsy	81	26.1	3-3.33	NR	24/279‡	68/249‡
Arkilo, 2015[2]	Epilepsy	59	NR	NA	First trimester	NR	NR
Bromley, 2010[3]	NR	60	NR	6-7	Whole pregnancy	NR	NR
Bromley, 2013[4] [CR: Bromley, 2008[30]]	Epilepsy	156	28	6	NR	28/156	42/156
Bromley, 2016[5]	Epilepsy	185	NR	NR	NR	31/185	35/185
Christensen, 2013[6]	NR	2011	NR	NR	Whole pregnancy	NR	NR
Cohen, 2013[7]	Epilepsy	108	30	6	NR	12/192‡	NR
Cummings, 2011[8] [CR: Tomson, 2015[35]]	Epilepsy	142	NR	2-3	Whole pregnancy	32/108‡	19/108‡
Dean, 2002[9] [CR: Rasalam, 2005[34]]	Epilepsy	287	27	3.75-15.5	First trimester	NR	NR
D'Souza, 1991[10]	Epilepsy	42	26.5	2.5-3.5	Whole pregnancy	NR	NR
Eriksson, 2005[11] [CR: Viinikainen, 2006[36]]	Epilepsy	39	28.2	NR	NR	NR	NR

Gaily, 1990[12] [CR: Gaily, 1990[31]; Hiilesmaa, 1982[32]; Hiilesmaa, 1985[33]	Epilepsy	134	27.8	5.5	First trimester	NR	NR
Gogatishvili, 2014[13]	NR	39	NR	2 to 4	NR	NR	NR
Gogatishvili, 2015[14]	NR	23	NR	3 to 6	NR	NR	NR
Hurault-Delarue, 2012[15]	NR	109	NR	0.75	NR	NR	NR
Jones, 1989[16]	Epilepsy	63	NR	NR	Whole pregnancy	NR	NR
Katz, 2001[17]	Epilepsy	51	31	NR	NR	NR	NR
Koch, 1996[18]	Epilepsy	40	NR	6	First trimester	NR	NR
Mawer, 2002[19]	Epilepsy	52	NR	NR	NR	NR	NR
Miskov, 2010[20]	Epilepsy	55	NR	NR	NR	NR	NR
Miskov, 2016[21]	Epilepsy	74	34	NR	NR	NR	6/74
Nadebaum, 2011[22]	Epilepsy	66	31.6	7.4	First trimester	NR	5/66
Rihtman, 2013[23]	Epilepsy	72	NR	NR	Whole pregnancy	NR	NR
Scolnik, 1994[24]	Epilepsy	75	NR	1.5-3	1st trimester	NR	NR
Shankaran, 1996[25]	NR	96	NR	NR	NR	NR	NR
Van der Pol, 1991[26]	Epilepsy	57	NR	6-13	NR	NR	NR
Veiby, 2013a[27]	Epilepsy	422	NR	0.5	NR	NR	NR
Veiby, 2013b[28]	Epilepsy	248	28.9	3	NR	NR	68/726‡
Wood, 2015[29]	Epilepsy	77	NR	6-8	NR	NR	NR

Abbreviations: NA – Not applicable; NR – Not reported

* Sample size used for analysis; ineligible treatment arms (i.e. treatment arms with excluded drugs or unspecified polytherapy) are not included in the count
† The mean age for children/follow-up period data were only collected for outcomes related to cognitive and/or psychomotor development
‡ Total sample size is based on the number of women enrolled in the study; may differ from the sample size used for analysis
§ Single publication reporting on two separate cohorts

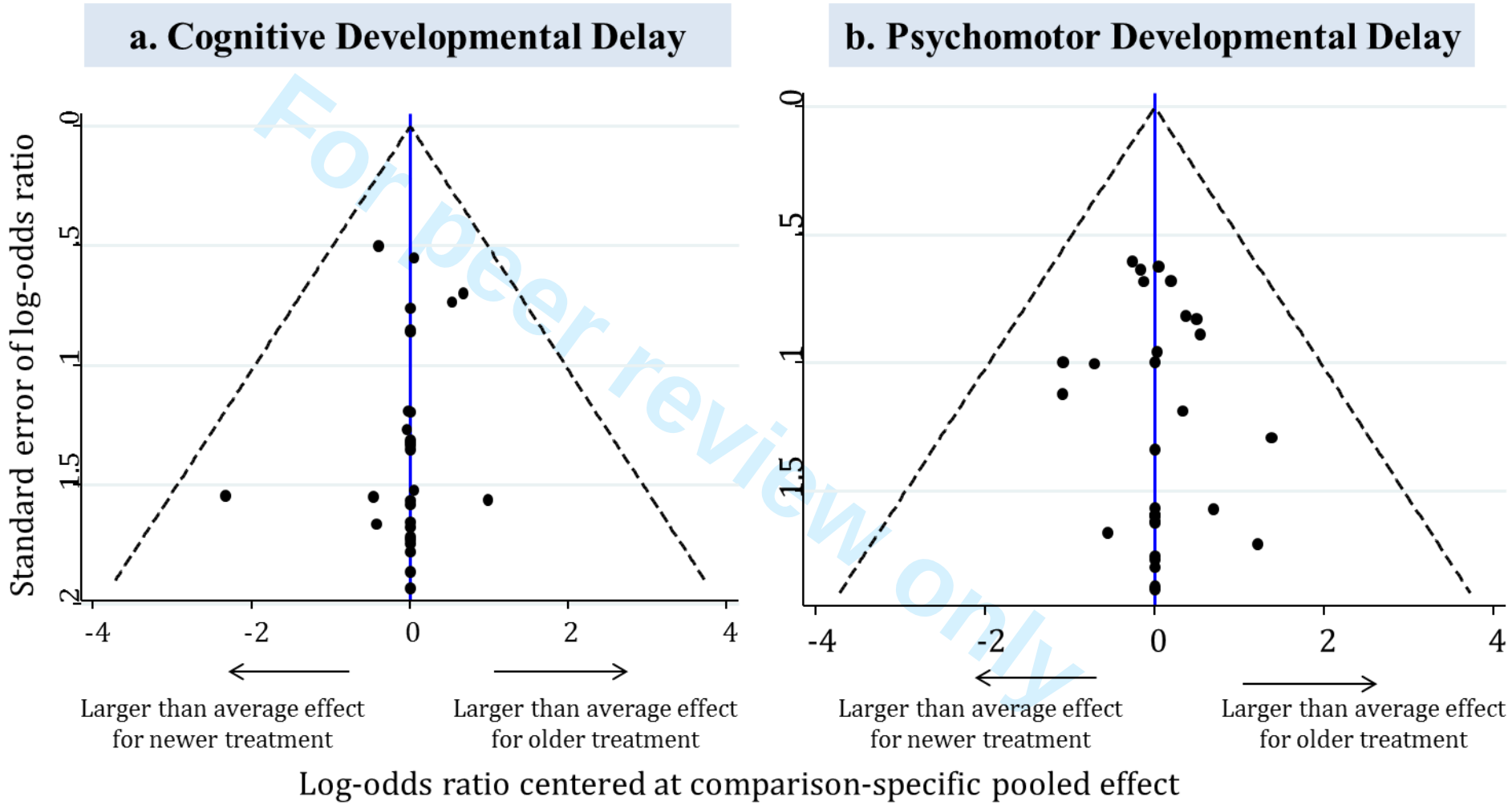
Appendix F. Methodological quality of observational studies – Newcastle Ottawa Scale results

First Author, Year	Representativeness of the exposed cohort	Selection of the non-exposed cohort	Ascertainment of exposure	Demonstration that outcome of interest was not present at start of study	Comparability of cohorts on the basis of the design or analysis	Assessment of outcome	Was follow-up long enough for outcomes to occur	Adequacy of follow up of cohorts
Adab, 2004[1]	B	A	A	A	C	A	A	C
Arkilo, 2015[2]	B	A	B	A	D	A	A	C
Bromley, 2010[3]	D	A	D	A	D	D	B	D
Bromley, 2013[4]	A	A	A	A	A	A	A	C
Bromley, 2016[5]	A	A	A	A	A	A	A	C
Christensen, 2013[6]	A	A	A	A	A	B	A	B
Cohen, 2013[7]	A	A	D	A	A	A	A	C
Cummings, 2011[8]	A	A	A	A	A	A	A	C
Dean, 2002[9]	B	A	A	A	D	A	A	C
D'Souza, 1991[10]	B	A	A	A	D	A	A	A
Eriksson, 2005[11]	B	A	A	A	B	A	A	D
Gaily, 1990[12]	B	A	A	A	D	A	A	A
Gogatishvili, 2014[13]	A	A	D	A	D	A	A	D
Gogatishvili, 2015[14]	A	A	D	A	D	A	A	D

Hurault-Delarue, 2012[15]	A	A	A	A	A	A	A	A
Jones, 1989[16]	A	A	B	A	D	A	A	B
Katz, 2001[17]	B	A	A	A	D	A	A	D
Koch, 1996[18]	B	A	B	A	D	A	A	C
Mawer, 2002[19]	B	A	A	A	D	A	A	B
Miskov, 2010[20]	D	A	D	A	D	D	A	D
Miskov, 2016[21]	C	A	A	A	D	A	A	D
Nadebaum, 2011[22]	A	A	A	A	A	A	A	B
Rihtman, 2013[23]	A	B	A	A	A	A	A	C
Scolnik, 1994[24]	B	A	A	A	D	A	A	A
Shankaran, 1996[25]	B	A	A	A	D	A	A	B
Van der Pol, 1991[26]	B	A	D	A	A	A	A	B
Veiby, 2013a[27]	A	A	A	A	A	A	A	D
Veiby, 2013b[28]	A	A	A	A	A	A	A	C
Wood, 2015[29]	A	A	A	A	D	A	A	C

Abbreviations: A – low risk; B – moderate risk; C – high risk; D – unclear risk

Appendix G. Comparison-adjusted funnel plots*



* Funnel plots have been produced only for outcomes with ≥10 studies. For multi-arm studies we plot data points from each study-specific basic parameter (treatment comparisons with a study-specific common comparator)

Appendix H. Statistically significant network meta-analysis results along with meta-analysis results, transitivity, and inconsistency assessments

Treatment Comparison	Number of Studies (Mean Baseline Risk)	Number of patients (Mean Age)	Treatment Indication	Timing	Comparability of cohorts	Adequacy of follow up of cohorts	MA Odds Ratio (95% CrI)	NMA Odds Ratio (95% CrI) (95% PrI)
Cognitive Developmental Delay (10 studies, 748 patients, 14 treatments)								
Lamot vs Valpro	4 (NA)	140 (31.00)	Epilepsy	NR	H	H	0.17 (0.02-0.87)	0.13 (0.01-0.57) (0.01-0.75)
Valpro vs Control	4 (0.06)	267 (28.80)	Epilepsy	1st trimester	H	H	8.15 (3.19-22.33)	7.40 (3.00-18.46) (1.81-27.63)
Valpro vs Carbam	6 (NA)	310 (27.80)	Epilepsy	NR	H	L	3.32 (1.56-7.04)	3.54 (1.69-7.26) (0.95-12.32)
Valpro vs Pheno	3 (NA)	36 (27.80)	Epilepsy	1st trimester	H	L	4.25 (0.82-34.07)	5.59 (1.21-35.07) (0.93-45.99)
Valpro vs Pheny	3 (NA)	58 (31.00)	Epilepsy	1st trimester	H	L	3.12 (0.75-14.12)	2.88 (1.04-8.49) (0.69-12.62)
Common between-study variance across treatment comparisons							0.13 (0.00-0.97)	0.12 (0.00-1.15) (NA)
Residual deviance: 44.72	Data points: 47	DIC: 78.7						
Evaluation of consistency using the design-by-treatment interaction model					Chi-square test: 14.15 Degrees of Freedom: 17		P- value: 0.66 Heterogeneity: 0	

Treatment Comparison	Number of Studies (Mean Baseline Risk)	Number of patients (Mean Age)	Treatment Indication	Timing	Comparability of cohorts	Adequacy of follow up of cohorts	MA Odds Ratio (95% CrI)	NMA Odds Ratio (95% CrI) (95% PrI)
Autism Dyspraxia (5 studies, 2551 patients, 12 treatments)								
Lamot vs Control	2 (0.00)	254 (27.75)	Epilepsy	1st trimester	H	H	13.77 (2.06-188.00)	8.88 (1.29-112.00) (0.94-146.80)
Lamot+Valpro vs Carbam	1 (NA)	40 (NR)	Epilepsy	NR	L	L	15.02 (2.04-171.90)	22.89 (2.58-219.00) (1.90-282.20)
Lamot+Valpro vs Clonaz	NA	NR	NR	NR	NR	NR	NA	20.21 (1.48-351.30) (1.15-455.00)
Lamot+Valpro vs Control	NA	NR	NR	NR	NR	NR	NA	132.70 ($7.41\text{-}3.9 \times 10^3$) ($5.82\text{-}4.6 \times 10^3$)
Lamot+Valpro vs Lamot	NA	NR	NR	NR	NR	NR	NA	14.61 (1.51-149.10) (1.14-196.80)
Oxcar vs Control	NA	NR	NR	NR	NR	NR	NA	13.51 (1.28-221.40) (0.86-267.40)
Valpro vs Carbam	5 (NA)	1003 (27.83)	Epilepsy	1st trimester	L	L	3.20 (1.20-8.68)	3.02 (1.09-8.40) (0.57-14.31)
Valpro vs Control	2 (0.00)	249 (27.75)	Epilepsy	1st trimester	H	H	9.19 (1.14-132.10)	17.29 (2.40-217.60) (1.61-274.90)
Common between-study variance across treatment comparisons							0.12 (0.00-1.37)	0.16 (0.00-1.95) (NA)
Residual deviance: 24	Data points: 24	DIC: 44						
Evaluation of consistency using the design-by-treatment interaction model			Chi-square test: 3.79 Degrees of Freedom: 5				P- value: 0.57 Heterogeneity: 0	

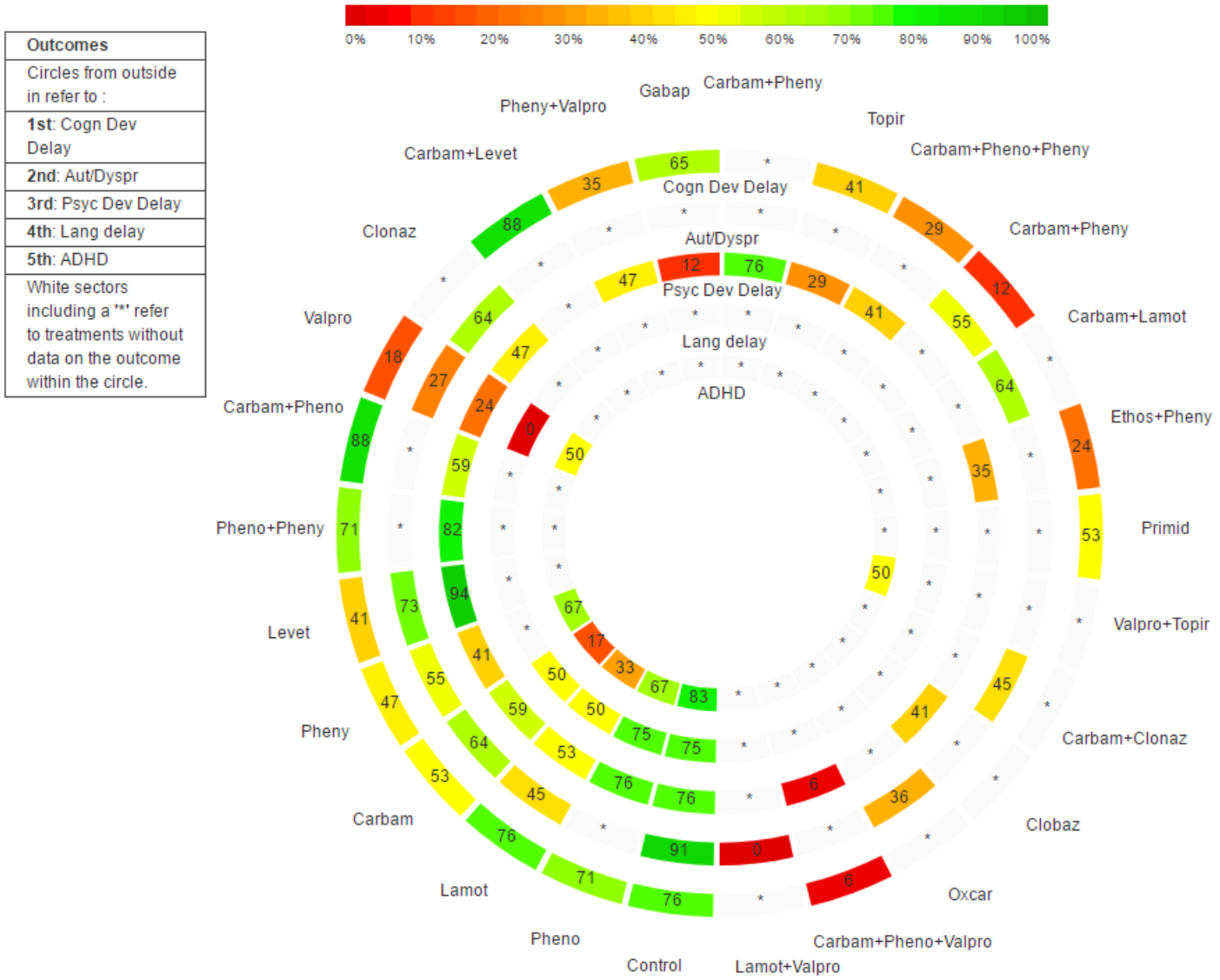
Treatment Comparison	Number of Studies (Mean Baseline Risk)	Number of patients (Mean Age)	Treatment Indication	Timing	Comparability of cohorts	Adequacy of follow up of cohorts	MA Odds Ratio (95% CrI)	NMA Odds Ratio (95% CrI) (95% PrI)
Psychomotor Developmental Delay (11 studies, 1145 patients, 18 treatments)								
Carbam+Pheno+Valpro vs Control	NA	NR	NR	NR	NR	NR	NA	19.12 (1.49-337.50) (1.34-370.40)
Carbam+Pheno+Valpro vs Pheno	NA	NR	NR	NR	NR	NR	NA	19.86 (1.38-393.60) (1.26-423.30)
Levet vs Carbam+Pheno+Valpro	NA	NR	NR	NR	NR	NR	NA	0.01 (0.00-0.58) (0.00-0.62)
Valpro vs Carbam	7 (NA)	331 (27.80)	Epilepsy	1st trimester	H	H	2.72 (1.39-5.67)	2.45 (1.27-4.88) (0.95-6.77)
Valpro vs Control	5 (0.07)	331 (28.38)	Epilepsy	1st trimester	H	H	3.53 (1.60-8.64)	4.16 (2.04-8.75) (1.52-12.05)
Valpro vs Pheno	2 (NA)	141 (NR)	Epilepsy	1st trimester	H	H	3.68 (1.17-12.30)	4.32 (1.72-11.20) (1.34-14.51)
Common between-study variance across treatment comparisons							0.05 (0.00-0.49)	0.06 (0.00-0.63) (NA)
Residual deviance: 45 Data points: 51 DIC: 78								
Evaluation of consistency using the design-by-treatment interaction model				Chi-square test: 13.46 Degrees of Freedom: 21			P- value: 0.89 Heterogeneity: 0	

Treatment Comparison	Number of Studies (Mean Baseline Risk)	Number of patients (Mean Age)	Treatment Indication	Timing	Comparability of cohorts	Adequacy of follow up of cohorts	MA Odds Ratio (95% CrI)	NMA Odds Ratio (95% CrI) (95% PrI)
Language Delay (5 studies, 509 patients, 5 treatments)								
Valpro vs Control	1 (0.03)	173 (28.90)	Epilepsy	NR	L	H	6.96 (1.14-37.03)	7.95 (1.50-49.13) (0.96-74.52)
Common between-study variance across treatment comparisons							0.15 (0.00-1.85)	0.16 (0.00-2.15) (NA)
Residual deviance: 12	Data points: 14	DIC: 23						
Evaluation of consistency using the design-by-treatment interaction model				Chi-square test: 2.33 Degrees of Freedom: 3			P- value: 0.50 Heterogeneity: 0	
ADHD (4 studies, 750 patients, 6 treatments)								
No statistically significant results								
Residual deviance: 12	Data points: 17	DIC: 22						

Abbreviations: ADHD - Attention Deficit Hyperactivity Disorder; CrI - Credible Interval; DIC - Deviance Information Criterion; H- high risk of bias; L - low risk of bias; MA - Meta-analysis; NA - Not applicable; NMA - Network Meta-analysis; NR- Not Reported; PrI - Predictive Interval

Carbam = Carbamazepine; Clobaz = Clobazam; Clonaz = Clonazepam; Ethos = Ethosuximide; Gabap = Gabapentin; Lamot = Lamotrigine; Levet = Levetiracetam; Oxcar = Oxcarbazepine; Pheno = Phenobarbital; Pheny = Phenytoin; Pridmid = Primidone; Topir = Topiramate; Valpro = Valproate; Vigab = Viagabatrin

Appendix I. Rank-heat plot of cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcomes*



Abbreviations: carbam - carbamazepine, clobaz - clobazam, clonaz - clonazepam, ethos - ethosuximide, gabap - gabapentin, lamot - lamotrigine, levet - levetiracetam, oxcar - oxcarbazepine, pheno - phenobarbital, pheny - phenytoin, primid - primidone, topir - topiramate, valpro - valproate, vigab - vigabatrin

*Rank-heat plot of cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcomes (5 circles) and 25 treatments (25 radii). Each sector is coloured according to the surface under the cumulative ranking curve value of the corresponding treatment and outcome using the transformation of three colours red (0%), yellow (50%), and green (100%).

Appendix J. Number of studies and treatments per outcome

Total studies	Range of study arms	# of treatments	# of patients	# of direct treatment comparisons	# of NMA treatment comparisons	Statistically significant NMA treatment effects	# of studies with zero events in all arms	# of studies with ineligible outcome definition*
Cognitive Developmental Delay								
11	(2,8)	18	933	62	153	5	1	5
Autism/Dyspraxia								
5	(4,6)	12	2551	34	66	8	0	4
Neonatal Seizure								
1	(2,2)	2	69	1	0	0	1	1
Psychomotor Developmental Delay								
11	(2,8)	18	1145	74	153	6	0	5
Language Delay								
5	(2,4)	5	509	7	10	1	0	3
ADHD								
5	(4,6)	7	816	20	21	0	0	0
Social Impairment								
1	(4,4)	4	422	1	0	0	0	0

Abbreviations: ADHD - Attention Deficit Hyperactivity Disorder; NMA - Network Meta-analysis

Appendix K. Sensitivity and network meta-regression analyses - Anti-epileptic drugs compared with Control

Treatment Comparison	NMA Odds Ratio	95% CrI	95% PrI
Cognitive Developmental Delay – Sensitivity Analysis - Epilepsy only (10 studies, 910 patients, 17 treatments)			
Carbamazepine vs Control	2.08	(0.79 - 5.82)	(0.47 - 9.34)
Carbamazepine+Phenobarbital vs Control	0.62	(0.00 - 15.31)	(0.00 - 19.29)
Carbamazepine+Phenobarbital+Phenytoin vs Control	4.75	(0.01 - 164.80)	(0.01 - 192.50)
Carbamazepine+Phenobarbital+Valproate vs Control	15.00	(1.00 - 367.10)	(0.82 - 426.90)
Carbamazepine+Phenytoin vs Control	9.84	(0.60 - 136.30)	(0.49 - 164.50)
Ethosuximide+Phenytoin vs Control	6.53	(0.02 - 216.00)	(0.02 - 251.30)
Gabapentin vs Control	1.43	(0.05 - 14.28)	(0.04 - 18.20)
Lamotrigine vs Control	0.79	(0.05 - 5.12)	(0.05 - 6.66)
Levetiracetam vs Control	3.46	(0.65 - 17.14)	(0.47 - 23.57)
Phenobarbital vs Control	0.55	(0.01 - 5.38)	(0.01 - 6.85)
Phenobarbital+Phenytoin vs Control	1.28	(0.00 - 36.18)	(0.00 - 44.03)
Phenytoin vs Control	2.47	(0.65 - 8.25)	(0.41 - 12.47)
Phenytoin+Valproate vs Control	3.68	(0.01 - 121.00)	(0.01 - 135.00)
Primidone vs Control	1.97	(0.25 - 12.16)	(0.19 - 16.25)
Topiramate vs Control	3.06	(0.42 - 17.51)	(0.32 - 23.57)
Valproate vs Control	7.48	(2.99 - 19.04)	(1.67 - 31.21)
Common within-network between-study variance	0.16	(0.00 - 1.36)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 12.98 Degrees of Freedom: 14	P-value: 0.53 Heterogeneity: 0.00
Cognitive Developmental Delay - Sensitivity Analysis - First generation AEDs only (6 studies, 480 patients, 13 treatments)			
Carbamazepine vs Control	1.68	(0.37 - 7.82)	(0.19 - 14.98)
Carbamazepine+Phenytoin vs Control	8.98	(0.36 - 169.90)	(0.26 - 243.60)
Carbamazepine+Phenobarbital vs Control	0.46	(0.00 - 21.02)	(0.00 - 28.01)
Carbamazepine+Phenobarbital+Phenytoin vs Control	4.12	(0.01 - 180.10)	(0.00 - 236.30)
Carbamazepine+Phenobarbital+Valproate vs Control	12.84	(0.50 - 435.70)	(0.35 - 604.30)
Ethosuximide+Phenytoin vs Control	5.65	(0.01 - 219.00)	(0.01 - 291.50)

Treatment Comparison	NMA Odds Ratio	95% CrI	95% PrI
Phenobarbital vs Control	0.64	(0.00 - 26.02)	(0.00 - 35.36)
Phenobarbital+Phenytoin vs Control	1.06	(0.00 - 37.64)	(0.00 - 50.85)
Phenytoin vs Control	2.08	(0.26 - 12.50)	(0.13 - 22.02)
Phenytoin+Valproate vs Control	3.14	(0.00 - 135.80)	(0.00 - 178.90)
Primidone vs Control	3.30	(0.18 - 43.76)	(0.12 - 68.72)
Valproate vs Control	13.22	(3.20 - 64.06)	(1.50 - 128.40)
Common within-network between-study variance	0.27	(0.00 - 2.97)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 3.31 Degrees of Freedom: 3	P-value: 0.35 Heterogeneity: 0.00
Cognitive Developmental Delay - Sensitivity Analysis - Maternal Alcohol or Tobacco use (3 studies, 504 patients, 7 treatments)			
Carbamazepine vs Control	1.97	(0.40 - 10.01)	(0.19 - 21.27)
Gabapentin vs Control	1.47	(0.04 - 19.01)	(0.02 - 27.11)
Lamotrigine vs Control	0.41	(0.00 - 10.09)	(0.00 - 13.61)
Levetiracetam vs Control	3.55	(0.43 - 24.13)	(0.23 - 42.39)
Topiramate vs Control	3.17	(0.30 - 24.07)	(0.18 - 44.87)
Valproate vs Control	7.79	(1.84 - 29.60)	(0.84 - 62.77)
Common within-network between-study variance	0.27	(0.00 - 3.29)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 2.69 Degrees of Freedom: 2	P-value: 0.26 Heterogeneity: NA
Cognitive Developmental Delay - Sensitivity Analysis - Low Risk of Bias: "Adequacy of follow-up" (4 studies, 283 patients, 12 treatments)			
Carbamazepine vs Control	2.68	(0.05 - 2.9 x 10^3)	(0.03 - 4.3 x 10^3)
Carbamazepine+Phenobarbital vs Control	0.67	(0.00 - 2.2 x 10^3)	(0.00 - 2.9 x 10^3)
Carbamazepine+Phenobarbital+Phenytoin vs Control	5.23	(0.01 - 7.2 x 10^3)	(0.00 - 1.1 x 10^4)
Carbamazepine+Phenobarbital+Valproate vs Control	22.18	(0.10 - 4.8 x 10^4)	(0.06 - 7.7 x 10^4)
Carbamazepine+Phenytoin vs Control	11.45	(0.13 - 1.2 x 10^4)	(0.07 - 1.8 x 10^4)
Ethosuximide+Phenytoin vs Control	6.45	(0.01 - 8.3 x 10^3)	(0.00 - 1.4 x 10^4)
Lamotrigine vs Control	0.52	(0.00 - 1.2 x 10^3)	(0.00 - 1.9 x 10^3)
Phenobarbital+Phenytoin vs Control	1.33	(0.00 - 1.8 x 10^3)	(0.00 - 2.7 x 10^3)
Phenytoin vs Control	1.67	(0.03 - 1.8 x 10^3)	(0.01 - 2.5 x 10^3)

Treatment Comparison	NMA Odds Ratio	95% CrI	95% PrI
Phenytoin+Valproate vs Control	3.94	(0.00 - 6.7 x 10^3)	(0.00 - 8.8 x10^3)
Valproate vs Control	5.9	(0.06 - 9.7 x 10^3)	(0.03 - 1.5 x 10^4)
Common within-network between-study variance	1.01	(0.01 - 5.85)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 5.07 Degrees of Freedom: 2	P-value: 0.08 Heterogeneity: 0.00
Cognitive Developmental Delay - Sensitivity Analysis - Low Risk of Bias: "Comparability of cohorts" (3 studies, 366 patients, 7 treatments)			
Carbamazepine vs Control	1.46	(0.11 - 19.59)	(0.06 - 38.10)
Gabapentin vs Control	1.19	(0.03 - 22.80)	(0.02 - 39.35)
Lamotrigine vs Control	0.27	(0.00 - 11.80)	(0.00 - 19.37)
Levetiracetam vs Control	2.90	(0.30 - 32.81)	(0.15 - 62.97)
Topiramate vs Control	2.55	(0.22 - 29.21)	(0.11 - 64.23)
Valproate vs Control	5.79	(1.05 - 47.35)	(0.47 - 102.90)
Common within-network between-study variance	0.38	(0.00 - 4.14)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 1.47 Degrees of Freedom: 2	P-value: 0.48 Heterogeneity: NA
Cognitive Developmental Delay – Network Meta-regression Analysis (11 studies, 933 patients, 18 treatments)			
Carbamazepine vs Control	1.99	(0.64 - 6.18)	(0.40 - 9.77)
Carbamazepine+Levetiracetam vs Control	0.54	(0.00 - 16.36)	(0.00 - 19.87)
Carbamazepine+Phenobarbital vs Control	0.50	(0.00 - 16.10)	(0.00 - 19.36)
Carbamazepine+Phenobarbital+Phenytoin vs Control	4.36	(0.01 - 171.20)	(0.01 - 194.60)
Carbamazepine+Phenobarbital+Valproate vs Control	14.58	(0.90 - 413.20)	(0.74 - 488.90)
Carbamazepine+Phenytoin vs Control	9.44	(0.50 - 130.50)	(0.39 - 162.40)
Ethosuximide+Phenytoin vs Control	5.77	(0.01 - 234.70)	(0.01 - 268.10)
Gabapentin vs Control	1.37	(0.04 - 15.51)	(0.03 - 19.10)
Lamotrigine vs Control	0.87	(0.07 - 5.14)	(0.06 - 6.76)
Levetiracetam vs Control	3.43	(0.57 - 18.78)	(0.42 - 24.85)
Phenobarbital vs Control	1.16	(0.13 - 8.59)	(0.10 - 11.43)
Phenobarbital+Phenytoin vs Control	1.34	(0.00 - 39.21)	(0.00 - 49.39)

Treatment Comparison	NMA Odds Ratio	95% CrI	95% PrI
Phenytoin vs Control	2.43	(0.55 - 9.14)	(0.36 - 13.45)
Phenytoin+Valproate vs Control	3.58	(0.01 - 134.20)	(0.01 - 161.70)
Primidone vs Control	2.03	(0.21 - 16.49)	(0.16 - 21.39)
Topiramate vs Control	2.93	(0.41 - 16.34)	(0.31 - 22.91)
Valproate vs Control	7.03	(2.26 - 20.02)	(1.41 - 30.92)
Common within-network between-study variance	0.16	(0.00 - 1.27)	
Regression Coefficient	1.01	(0.76 - 1.56)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 14.15 Degrees of Freedom: 17	P-value: 0.66 Heterogeneity: 0.00
Autism/Dyspraxia - Sensitivity Analysis - Large cohort (>300 patients) - (1 study, 2,551 patients, 5 treatments)**			
Clonazepam vs Carbamazepine	1.08	(0.24 - 4.85)	-
Lamotrigine vs Carbamazepine	1.20	(0.36 - 4.00)	-
Oxcarbazepine vs Carbamazepine	2.13	(0.62 - 7.35)	-
Valproate vs Carbamazepine	3.05	(0.97 - 9.52)	-
Common within-network between-study variance	NA	NA	
Evaluation of inconsistency using the design-by-treatment interaction model		NA	NA
Autism/Dyspraxia - Sensitivity Analysis - Epilepsy only (4 studies, 540 patients, 10 treatments)			
Carbamazepine vs Control	5.20	(0.54 - 90.53)	(0.33 - 133.00)
Carbamazepine+Clonazepam vs Control	7.90	(0.01 - 653.30)	(0.01 - 881.00)
Carbamazepine+Lamotrigine vs Control	4.25	(0.01 - 333.60)	(0.01 - 446.90)
Carbamazepine+Phenytoin vs Control	9.03	(0.01 - 666.30)	(0.01 - 893.00)
Lamotrigine vs Control	10.24	(1.25 - 171.40)	(0.67 - 248.50)
Lamotrigine+Valproate vs Control	120.20	(5.25 - 4.5 x 10^3)	(3.51 - 6.0 x 10^3)
Levetiracetam vs Control	3.52	(0.00 - 272.20)	(0.00 - 364.30)
Phenytoin vs Control	8.10	(0.01 - 577.50)	(0.01 - 754.60)
Valproate vs Control	14.41	(1.66 - 252.10)	(0.88 - 378.00)
Common within-network between-study variance	0.31	(0.00 - 3.04)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 2.9 Degrees of Freedom: 3	P-value: 0.41 Heterogeneity: 0.00

Treatment Comparison	NMA Odds Ratio	95% CrI	95% PrI
Autism/Dyspraxia - Sensitivity Analysis - Maternal Tobacco Use (4 studies, 540 patients, 10 treatments)			
Carbamazepine vs Control	2.51	(0.05 - 154.30)	(0.04 - 254.50)
Lamotrigine vs Control	24.84	(2.14 - 1.2 x 10^3)	(1.23 - 2.2 x 10^3)
Valproate vs Control	33.40	(2.60 -1.7 x 10^3)	(1.45 - 2.9 x 10^3)
Common within-network between-study variance	0.39	(0.00 - 4.47)	
Evaluation of inconsistency using the design-by-treatment interaction model		NA - all closed loops are formed from a multi-arm study	
Autism/Dyspraxia - Sensitivity Analysis - Maternal Alcohol Use (1 study, 156 patients, 4 treatments)			
Carbamazepine vs Control	Excluded due to zero events	-	-
Lamotrigine vs Control	4.65	(0.21 - 100.00)	-
Valproate vs Control	7.75	(0.42 - 142.86)	-
Common within-network between-study variance	*1.91*	(0.36 - 10.13)	
Evaluation of inconsistency using the design-by-treatment interaction model		NA	NA
Autism/Dyspraxia - Sensitivity Analysis - Low Risk of Bias: "Adequacy of Follow-up" (3 studies, 2,244 patients, 10 treatments)			
Carbamazepine vs Control	3.97	(0.17 - 2.4 x 10^3)	(0.11 - 3.0 x 10^3)
Carbamazepine+Clonazepam vs Control	7.48	(0.01 - 7.8 x 10^3)	(0.01 - 9.0 x 10^3)
Carbamazepine+Lamotrigine vs Control	4.47	(0.00 - 5.0 x 10^3)	(0.00 - 5.7 x 10^3)
Carbamazepine+Phenytoin vs Control	7.23	(0.01 - 6.6 x 10^3)	(0.01 - 8.2 x 10^3)
Clonazepam vs Control	4.88	(0.12 - 3.2 x 10^3)	(0.09 - 3.8 x 10^3)
Lamotrigine vs Control	6.55	(0.30 - 4.4 x 10^3)	(0.21 - 4.7 x 10^3)
Lamotrigine+Valproate vs Control	113.50	(2.33 - 7.8 x 10^4)	(1.62 - 8.9 x 10^4)
Oxcarbazepine vs Control	10.23	(0.36 - 6.8 x 10^3)	(0.26 - 7.5 x 10^3)
Valproate vs Control	13.97	(0.68 - 8.4 x 10^3)	(0.47 - 1.0 x 10^4)
Common within-network between-study variance	0.23	(0.00 - 2.88)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 2.17 Degrees of Freedom: 3	P-value: 0.54 Heterogeneity: 0.00
Autism/Dyspraxia - Sensitivity Analysis - Low Risk of Bias: "Comparability of Cohorts" (4 studies, 2,395 patients, 12 treatments)			
Carbamazepine vs Control	9.55	(0.90 - 246.20)	(0.61 - 329.40)

Treatment Comparison	NMA Odds Ratio	95% CrI	95% PrI
Carbamazepine+Clonazepam vs Control	13.58	(0.01 - 1.3 x 10^3)	(0.01 - 1.6 x 10^3)
Carbamazepine+Lamotrigine vs Control	7.11	(0.01 - 614.20)	(0.01 - 717.60)
Carbamazepine+Phenytoin vs Control	10.97	(0.01 - 1.1 x 10^3)	(0.01 - 1.4 x 10^3)
Clonazepam vs Control	8.33	(0.45 - 263.10)	(0.33 - 353.70)
Lamotrigine vs Control	10.98	(1.07 - 283.50)	(0.71 - 358.20)
Lamotrigine+Valproate vs Control	194.10	(8.06 - 8.4 x 10^3)	(6.28 - 1.0 x 10^4)
Levetiracetam vs Control	4.25	(0.00 - 390.90)	(0.00 - 485.30)
Oxcarbazepine vs Control	17.60	(1.22 - 552.20)	(0.86 - 727.40)
Phenytoin vs Control	9.76	(0.01 - 861.60)	(0.01 - 1.0 x 10^3)
Valproate vs Control	21.06	(1.86 - 525.40)	(1.25 - 681.90)
Common within-network between-study variance	0.19	(0.00 - 2.43)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 3.36 Degrees of Freedom: 5	P-value: 0.64 Heterogeneity: 0.00
Autism/Dyspraxia - Sensitivity Analysis - Maternal IQ (1 study, 77 patients, 6 treatments)**			
Carbamazepine+Clonazepam vs Carbamazepine	1.86	(0.07 - 47.62)	-
Carbamazepine+Lamotrigine vs Carbamazepine	1.18	(0.05 - 27.78)	-
Carbamazepine+Phenytoin vs Carbamazepine	1.86	(0.07 - 47.62)	-
Lamotrigine+Valproate vs Carbamazepine	15.87	(1.87 - 142.86)	-
Valproate vs Carbamazepine	1.33	(0.18 - 10.20)	-
Common within-network between-study variance	NA	NA	
Evaluation of inconsistency using the design-by-treatment interaction model		NA	NA

Abbreviations: NMA – Network Meta-analysis; OR – odds ratio; CrI – Credible Interval; PrI – Predictive Interval
** Network did not include a control arm, comparison with Carbamazepine is reported instead

BMJ Open

Comparative safety of anti-epileptic drugs for neurological development in children exposed during pregnancy and breastfeeding: a systematic review and network meta-analysis

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-017248.R2
Article Type:	Research
Date Submitted by the Author:	01-Jun-2017
Complete List of Authors:	Veroniki, Areti Angeliki; Li Ka Shing Knowledge Institute, St. Michael's Hospital Rios, Patricia; Li Ka Shing Knowledge Institute, St. Michael's Hospital Cogo, Elise; Li Ka Shing Knowledge Institute, St. Michael's Hospital Straus, Sharon; Li Ka Shing Knowledge Institute, St. Michael's Hospital; University of Toronto, Department of Medicine Finkelstein, Yaron; The Hospital for Sick Children; University of Toronto, Department of Paediatrics Kealey, M.; Li Ka Shing Knowledge Institute, St. Michael's Hospital Reynen, Emily; Li Ka Shing Knowledge Institute, St. Michael's Hospital Soobiah, Charlene; Li Ka Shing Knowledge Institute, St. Michael's Hospital; University of Toronto, Institute for Health Policy Management & Evaluation Thavorn, Kednapa; University of Ottawa, School of Epidemiology, Public Health and Preventive Medicine, Faculty of Medicine; The Ottawa Hospital Research Institute, Clinical Epidemiology Program Hutton, Brian; University of Ottawa, School of Epidemiology, Public Health and Preventive Medicine, Faculty of Medicine; Ottawa Hospital Research Institute, Center for Practice Changing Research Hemmelgarn, BR; University of Calgary, Departments of Medicine and Community Health Sciences Yazdi, Fatemeh; Li Ka Shing Knowledge Institute, St. Michael's Hospital D'Souza, Jennifer; Li Ka Shing Knowledge Institute, St. Michael's Hospital MacDonald, Heather; Li Ka Shing Knowledge Institute, St. Michael's Hospital Tricco, Andrea; Li Ka Shing Knowledge Institute, St. Michael's Hospital; University of Toronto, Epidemiology Division, Dalla Lana School of Public Health
Primary Subject Heading:	Neurology
Secondary Subject Heading:	Obstetrics and gynaecology
Keywords:	multiple treatment meta-analysis, knowledge synthesis, Epilepsy < NEUROLOGY, pregnancy, infants, developmental delay

Comparative safety of anti-epileptic drugs for neurological development in children exposed during pregnancy and breastfeeding: a systematic review and network meta-analysis

Areti Angeliki Veroniki, PhD, MSc[1] Email: VeronikiA@smh.ca

Patricia Rios, MSc[1] Email: RiosP@smh.ca

Elise Cogo, ND, MLIS[1] Email: CogoE@smh.ca

Sharon E. Straus, MD, MSc[1,2] Email: Sharon.straus@utoronto.ca

Yaron Finkelstein, MD[3,4,5] Email: Yaron.Finkelstein@sickkids.ca

Ryan Kealey, PhD[1] Email: ryan.kealey@utoronto.ca

Emily Reynen, MD, CM, PharmD[1] Email: ereynen@gmail.com

Charlene Soobiah, PhD (Cand.)[1,6] Email: SoobiahC@smh.ca

Kednapa Thavorn, PhD[7,8,9] Email: kthavorn@ohri.ca

Brian Hutton, PhD, MSc[7,10] Email: bhutton@ohri.ca

Brenda R. Hemmelgarn, MD, PhD[11] Email: Bhemmelg@ucalgary.ca

Fatemeh Yazdi, MSc[1] Email: SabaghYazdiF@smh.ca

Jennifer D'Souza, HBSc[1] Email: jennifer.dsouza@mail.utoronto.ca

Heather MacDonald, MSc[1] Email: hrmacdonald@gmail.com

Andrea C. Tricco, PhD, MSc[1,12,*] Email: TriccoA@smh.ca

AUTHOR DETAILS

[1] Li Ka Shing Knowledge Institute, St. Michael's Hospital, 209 Victoria Street, East Building, Toronto, Ontario, M5B 1W8, Canada

[2] Department of Medicine, University of Toronto, 27 King's College Circle, Toronto, Ontario, M5S 1A1, Canada

[3] The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario, M5G 1X8, Canada

[4] Department of Paediatrics, University of Toronto, 172 St. George Street, Toronto, Ontario, M5R 0A3, Canada

[5] Department of Pharmacology and Toxicology, University of Toronto, Medical Sciences Building, Room 4207, 1 King's College Circle, Toronto, Ontario, M5S 1A8, Canada

[6] Institute for Health Policy Management & Evaluation, University of Toronto, 4th Floor, 155 College Street, Toronto, Ontario, M5T 3M6, Canada

[7] School of Epidemiology, Public Health and Preventive Medicine, Faculty of Medicine, University of Ottawa, Roger-Guindon Building, 451 Smyth Road, Ottawa, Ontario, K1H 8M5, Canada

[8] Clinical Epidemiology Program, Ottawa Hospital Research Institute, The Ottawa Hospital, 501 Smyth Road, Ottawa, Ontario, K1H 8L6, Canada

[9] Institute of Clinical and Evaluative Sciences (ICES uOttawa), 1053 Carling Avenue, Ottawa, Ontario, K1Y 4E9, Canada

[10] Ottawa Hospital Research Institute, Center for Practice Changing Research, The Ottawa Hospital–General Campus, 501 Smyth Road, PO Box 201B, Ottawa, Ontario, K1H 8L6, Canada.

[11] Departments of Medicine and Community Health Sciences, University of Calgary, TRW Building, 3rd Floor, 3280 Hospital Drive NW, Calgary, Alberta, T2N 4Z6, Canada

[12] Epidemiology Division, Dalla Lana School of Public Health, University of Toronto, 6th Floor, 155 College Street, Toronto, Ontario, M5T 3M7, Canada

***Corresponding author**

Prof. Andrea C. Tricco, PhD

Scientist, Knowledge Translation Program,

Li Ka Shing Knowledge Institute, St. Michael's Hospital,

209 Victoria Street, East Building, Toronto, Ontario, M5B 1W8, Canada

Phone: 416-864-6060, Fax: 416-864-5805, Email: TriccoA@smh.ca

Keywords: multiple treatment meta-analysis, knowledge synthesis, epilepsy, pregnancy, infants, developmental delay.

Word count: abstract (300 words); main text (4,000 words); 2 tables; 3 figures; 3 additional files; 51 references

ABSTRACT

Objectives: Compare the safety of anti-epileptic drugs (AEDs) on neurodevelopment of infants/children exposed *in-utero* or during breastfeeding.

Design and Setting: Systematic review and Bayesian random-effects network meta-analysis (NMA). Medline, EMBASE, and the Cochrane Central Register of Controlled Trials were searched until April 27th, 2017. Screening, data abstraction, and quality appraisal were completed in duplicate by independent reviewers.

Participants: 29 cohort studies including 5,100 infants/children.

Interventions: Mono- and poly-therapy AEDs including first-generation (carbamazepine, clobazam, clonazepam, ethosuximide, phenobarbital, phenytoin, primidone, valproate) and newer-generation (gabapentin, lamotrigine, levetiracetam, oxcarbazepine, topiramate, vigabatrin) AEDs. Epileptic women who did not receive AEDs during pregnancy or breastfeeding served as the control group.

Primary and secondary Outcome measures: Cognitive developmental delay and autism/dyspraxia were primary outcomes. Attention deficit hyperactivity disorder, language delay, neonatal seizures, psychomotor developmental delay, and social impairment were secondary outcomes.

Results: The NMA on cognitive developmental delay (11 cohort studies, 933 children, 18 treatments) suggested among all AEDs only valproate was statistically significantly associated with more children experiencing cognitive developmental delay when compared with control (odds ratio (OR)=7.40, 95% credible interval (CrI): 3.00-18.46). The NMA on autism (5 cohort studies, 2,551 children, 12 treatments), suggested that oxcarbazepine (OR=13.51, CrI: 1.28-221.40), valproate (N=485, OR=17.29, 95% CrI: 2.40-217.60),

lamotrigine (OR=8.88, CrI: 1.28-112.00), and lamotrigine+valproate (OR=132.70, CrI: 7.41-3,851.00) were associated with significantly greater odds of developing autism compared with control. The NMA on Psychomotor developmental delay (11 cohort studies, 1,145 children, 18 treatments) found that valproate (OR=4.16, CrI: 2.04-8.75) and carbamazepine+phenobarbital+valproate (OR=19.12, CrI: 1.49-337.50) were associated with significantly greater odds of psychomotor delay compared with control.

Conclusions: Valproate alone or combined with another AED is associated with the greatest odds of adverse neurodevelopmental outcomes compared with control. Oxcarbazepine and lamotrigine were associated with increased occurrence of autism. Counselling is advised for women considering pregnancy to tailor the safest regimen.

Registration: PROSPERO database (CRD42014008925).

Keywords: multiple treatment meta-analysis, knowledge synthesis, epilepsy, pregnancy, infants, developmental delay.

ARTICLE SUMMARY

Strengths and limitations of this study

- 29 cohort studies involving 5,100 children of women who took AEDs were included in this systematic review. More evidence from long-term follow-up studies is required.
- This study was the first that compared and ranked the safety of AEDs, including comparative safety of treatments that have not been directly compared.

- Across all neurological outcomes and treatments compared with control, valproate alone or combined with another AED is associated with the greatest odds of adverse development.
- Oxcarbazepine and lamotrigine were associated with increased occurrence of autism.

INTRODUCTION

Anti-epileptic drugs (AEDs) are used by pregnant women for various conditions, such as epilepsy, pain syndromes, psychiatric disorders, and chronic migraine.[1] AED use during pregnancy is associated with risks to the fetus, as these drugs can cross the placenta or may be transferred to the infant through breastfeeding and may be associated with adverse neurodevelopment outcomes.[2-4] Two systematic reviews examined the association between AED exposure and neurodevelopment *in utero*, and reported that exposure to valproate was linked to significantly lower IQ scores and poorer overall neurodevelopmental outcomes in the children of women who used these medications.[5 6] No significant associations were found between neurodevelopment and exposure to other AEDs such as carbamazepine, lamotrigine, or phenytoin.[5-8] However, there is a lack of sufficiently powered studies to assess the impact of AEDs on neurodevelopment in children of women exposed to these agents, especially for newer generation drugs, thus highlighting the need for a systematic review.[9 10]

The aim of this study was to compare the safety of AEDs and assess their impact on neurodevelopment in infants and children exposed *in-utero* or during breastfeeding, employing a systematic review and network meta-analysis (NMA).

METHODS

The methods are briefly described here; details can be found in the published protocol (Additional File 1).[11] This study was registered with PROSPERO (CRD42014008925). We followed the ISPOR[12] guidelines for our NMA, and reported our findings using the PRISMA extension for NMA (Additional File 2).[13]

Eligibility criteria

All randomized clinical trials (RCTs), quasi-RCTs, and observational studies were eligible. Included studies assessed infants or children ≤12 years of age whose mothers consumed AEDs during pregnancy and/or while breastfeeding. Both mono- and poly-therapy AEDs were eligible, including first-generation (i.e., carbamazepine, clobazam, clonazepam, ethosuximide, phenobarbital, phenytoin, primidone, valproate) and newer-generation (i.e., marketed >1990: gabapentin, lamotrigine, levetiracetam, oxcarbazepine, topiramate, vigabatrin), with no restrictions on AED dosage. Placebo, no AED, other AEDs alone or in combination, were considered as comparators. Duplicate studies that used the same registry or population sample (i.e., companion studies) were used for supplementary information only. No language or other restrictions were imposed.

The primary neurological outcomes were cognitive developmental delay and autism/dyspraxia, and the secondary outcomes included attention deficit hyperactivity disorder (ADHD), language delay, neonatal seizures, psychomotor developmental delay, and social impairment. Table 1 shows the outcome measures and diagnostic scales used. We initially intended to evaluate all safety outcomes in infants/ children exposed to AEDs *in-utero* or during breastfeeding in one publication, but given the breadth of evidence we

identified, we report results related to risk of major congenital malformations, birth, and prenatal outcomes in a companion paper.[14]

Information sources

An experienced librarian executed search strategies for MEDLINE, EMBASE, and the Cochrane Central Register of Controlled Trials up to March 18, 2014, and then updated the search in April 27th 2017. The search strategy for MEDLINE was peer-reviewed by another librarian using the PRESS checklist,[15] and is available in the protocol.[11] Additional studies were identified by scanning references and contacting authors. Unpublished studies were sought by searching clinical trial registries and conference abstracts.

Study selection and data collection

After a calibration exercise, titles/abstracts (level 1) and full-text papers (level 2) were screened by two reviewers independently. Upon completion of level 1, 6% of citations were discrepant between reviewer pairs, whereas at the conclusion of level 2, 16% of articles were discrepant. Conflicts were resolved through discussion or by a third reviewer. The same approach was used for data abstraction and appraisal of methodological quality. Three rounds of pilot testing were conducted prior to data abstraction to train reviewers and refine the data abstraction form. For studies published in the last 10 years, authors were contacted to request clarification or additional data.

Appraisal of methodological quality

Only observational studies were identified and included for analysis, and their methodological quality was appraised with the Newcastle-Ottawa Scale (NOS) (Additional File 3: Appendix A).[16] For each outcome with ≥10 studies, the comparison-adjusted funnel plot was used to assess small-study effects,[17] where the overall treatment effect for each

comparison was estimated under the fixed-effect meta-analysis model. All eligible medications were ordered from oldest to newest using their international market approval dates. Hence, the comparison-adjusted funnel plot additionally assesses the hypothesis that newer AEDs are favoured over older ones. To overcome some of the correlations induced by multi-arm studies, which may cause overestimation and mask funnel plot asymmetry, we plotted data points corresponding to the study-specific basic parameters (treatment comparisons with common comparator). In each study, we used the control group as the common comparator or if this was missing, we used the oldest treatment comparator against the remaining AEDs.

Synthesis of included studies

We used the odds ratio (OR) for each dichotomous outcome, and outcome data were pooled using hierarchical meta-analysis and NMA models and the Markov Chain Monte Carlo sampling method in a Bayesian framework. To account for anticipated methodological and clinical heterogeneity across studies, and to achieve the highest generalizability in the meta-analytical treatment effects, we applied a random-effects model.[18]

A NMA was applied for connected evidence networks and pre-specified treatment nodes.[19] We assessed the transitivity assumption for each outcome *a priori* using the effect modifiers: age, baseline risk, treatment indication, timing, and methodological quality. The mean of each continuous effect modifier and the mode of each categorical effect modifier for each pairwise comparison were presented in tables for each outcome.[20] The consistency assumption was evaluated for the entire network of each outcome using the random-effects design-by-treatment interaction model when multiple studies were available in

each network design or the fixed-effect design-by-treatment interaction model when a single study informed each network design.[21] If inconsistency was identified, further examination for local inconsistency in parts of the network was completed using the loop-specific method.[22 23] Common within-network between-study variance (τ^2) across treatment comparisons was assumed in the meta-analysis, NMA, and design-by-treatment interaction model, so that treatment comparisons including a single study can borrow strength from the remaining network. This assumption was clinically reasonable, as the treatments included were of the same nature. In the loop-specific approach, common within-loop τ^2 was assumed.

For cognitive developmental delay and autism/dyspraxia outcomes, network meta-regression analyses for maternal age and baseline risk (i.e., using the control group) were conducted, when ≥10 studies provided relevant information, assuming a common fixed coefficient across treatment comparisons for AEDs vs. control. Sensitivity analyses for cognitive developmental delay and autism/dyspraxia outcomes were performed for treatment indication of epilepsy, large study size (i.e., >300), maternal alcohol intake, maternal tobacco use, only first-generation AEDs, and methodological quality. The sensitivity analysis for methodological quality was restricted to studies with low risk of bias for the two items on the NOS where the greatest proportion of studies received a low-quality score: adequacy of follow-up of cohorts and comparability of cohorts. For autism/dyspraxia, a sensitivity analysis on maternal IQ/psychiatric history was additionally conducted. We measured the goodness of fit using the posterior mean of the residual deviance, the degree of τ^2, and the deviance information criterion (DIC). In a well-fitting model the posterior mean residual deviance should be close to the number of data

points.[24 25] A difference of 3 units in the DIC between a NMA and a network meta-regression model was considered important and the lowest value of the DIC corresponded to the model with the best fit.[24 25]

All analyses were conducted in OpenBUGS[26] assuming non-informative priors for all model parameters, and $\tau \sim N(0,1)$, $\tau > 0$. The first 10,000 iterations were discarded and then 100,000 simulations were run with thinning of 10 values. Convergence was checked by visual inspection of the evaluation of the mixing of two chains. The median and 95% CrI were calculated for each parameter value. The *network* command[27] was used to apply the design-by-treatment interaction model.

For NMA estimates, a 95% predictive interval (PrI) is also reported to capture the magnitude of τ^2 and present the interval within which the treatment effect of a future study is expected to lie.[28 29] The estimated safety of the included AEDs was ranked using the surface under the cumulative ranking (SUCRA) curve.[30] The larger the SUCRA for a treatment, the higher its safety rank among all the available treatment options. SUCRA values are presented along with 95% CrIs to capture the uncertainty in the parameter values.[31]

RESULTS

Literature search and included studies

Our literature search identified 5,707 titles and abstracts, which after the screening process yielded 681 articles potentially relevant for inclusion (Figure 1). After full-text review, 95 studies fulfilled eligibility criteria along with 17 studies identified through supplemental methods. Of the 112 total eligible studies in the complete review,[14] 29 articles with seven companion reports and two potentially overlapping registry studies included one or more relevant neurological outcomes (Additional File 3: Appendix B). Four of the studies included in this analysis were conference abstracts with usable data,[32-35] and four studies,[36-39] not captured in the original literature search, were identified through reference scanning. A table with the key excluded studies and a rationale for their exclusion is presented in Additional File 3: Appendix C.

Study and patient characteristics

We included 29 cohort studies (5,100 patients) published between 1989 and 2016 (Table 2; Additional File 3: Appendix D, E). The number of patients included in each study ranged from 23 to 2,011 (median 74.5). Most studies (76%) were published after 2000, 62% of the studies included fewer than 100 patients, and the 52% of the studies included a control group of pregnant/breastfeeding women with epilepsy who did not receive AEDs. The mean maternal age ranged from 24 to 34 years. About half of the studies (52%) were funded through government/public research funding.

Methodological quality results

Twenty-nine observational studies were appraised using the NOS (Additional File 3: Appendix F). Overall, the studies were of good methodological quality and were rated as

high quality across most items: 28 studies (97%) selected the non-exposed cohort from the same community as the exposed cohort, 26 (90%) included a representative or somewhat representative sample, 27 (93%) assessed outcomes independently, with blinding, or via a record linkage (e.g., identified through database records), and 23 (79%) ascertained exposure via secured records (e.g., database records) or structured interviews. The comparability of cohorts and adequacy of follow-up were the lowest scoring items across the studies with only 12 (41%) and 10 (34%) studies rated as high quality on these items. No evidence for small-study effects was identified by the visual inspection of the comparison-adjusted funnel plots (Additional File 3: Appendix G).

Statistical analysis results

No important concerns were raised regarding the violation of the transitivity assumption when maternal age, baseline risk, treatment indication, and timing were assessed (Additional File 3: Appendix H). However, the average methodological quality appraisal across treatment comparisons varied across treatment comparisons. The evaluation of the consistency assumption using the design-by-treatment interaction model suggested that there was no evidence of significant inconsistency across all outcomes (Additional File 3: Appendix H).

In the following sections, we present the significant NMA results by outcome for AEDs compared with control (i.e., no exposure to AEDs), while the SUCRA values from all outcomes are presented in Figure 2 and depicted in a rank-heat plot (http://rh.ktss.ca/)[40] in Additional File 3: Appendix I.

Cognitive developmental delay

The NMA for cognitive developmental delay (definitions in Table 1) included 11 cohort studies, 933 children, and examined 18 treatments (Figure 3a; Additional File 3: Appendix J; τ^2=0.12, 95% CrI: 0.00-1.15). One study included children exposed to AEDs both *in-utero* and through breastfeeding, and ten included children exposed to AEDs *in-utero*. Across all AEDs, only valproate was associated with significantly increased odds of cognitive developmental delay when compared with control (odds ratio (OR)=7.40, 95% credible interval (CrI): 3.00-18.46; Figure 2a; Additional File 3: Appendix H).

The same results were observed in a network meta-regression of baseline risk for offspring of women with epilepsy who were not exposed to AEDs (estimated regression coefficient on OR scale: 1.01, 95% CrI: 0.76-1.56; τ^2=0.16, 95% CrI: 0.00-1.24; residual deviance= 45.27, data points= 47, DIC= 80.17). Similarly, the sensitivity analyses restricted to: a) studies that only included women receiving AEDs to treat epilepsy (10 studies, 910 children, 17 treatments; τ^2=0.16, 95% CrI: 0.00-1.36), b) studies comparing only first-generation AEDs (6 studies, 480 children, 13 treatments; τ^2=0.28, 95% CrI: 0.00-2.97), c) studies that reported maternal alcohol or tobacco use (3 studies, 504 children, 7 treatments; τ^2=0.27, 95% CrI: 0.00-3.29), and d) studies with high methodological quality on NOS item 'comparability of cohorts' (3 studies, 366 children, 7 treatments; τ^2=0.38, 95% CrI: 0.00-4.14), were consistent with the NMA results (Additional File 3: Appendix K). The sensitivity analysis with studies of high methodological quality on the NOS item 'adequacy of follow-up' found no statistically significant results (4 studies, 283 patients, 12 treatments; τ^2=1.01, 95% CrI: 0.01-5.85; Additional File 3: Appendix K).

Autism/dyspraxia

The NMA on autism/dyspraxia (definitions in Table 1) included five cohort studies, 2,551 children exposed *in utero*, and examined 12 treatments (τ^2=0.16, 95% CrI: 0.00-1.95; Figure 3b; Additional File 3: Appendix H). Compared with control, only valproate (OR=17.29, 95% CrI: 2.40-217.60), oxcarbazepine (OR= 13.51, 95% CrI: 1.28-221.40), lamotrigine (OR= 8.88, 95% CrI: 1.28-112.00), and lamotrigine+valproate (OR=132.70, 95% CrI: 7.41-3851.00) were significantly associated with increased occurrence of autism/dyspraxia (Figure 2b).

Restricting the NMA to studies including only women with epilepsy as their treatment indication produced results that were generally in agreement with the NMA results, except that oxcarbazepine was no longer in the network (4 cohort studies, 540 children, 10 treatments; τ^2=0.31, 95% CrI: 0.00-304). Two cohort studies of 404 offspring of women with a history of tobacco use compared 4 treatments and found similar results except that oxcarbazepine and lamotrigine+valproate were no longer in the network (τ^2=0.39, 95% CrI: 0.00-4.47). The results were in agreement in sensitivity analyses including only higher methodological quality studies in the 'comparability of cohorts' item on the NOS (4 studies, 2,395 children, 12 treatments; τ^2=0.19, 95% CrI: 0.00-2.43) and the 'adequacy of follow-up of cohorts' (3 studies, 2244 children, 10 treatments; τ^2=0.23, 95% CrI: 0.00-2.88), except that lamotrigine was no longer statistically significant than control for the latter (Additional File 3: Appendix K).

Neonatal Seizure

One cohort study included 72 children who were exposed to AEDs *in-utero* as well as through breastfeeding reported on the incidence of neonatal seizures. The study compared

valproate against lamotrigine and found no significant difference in neonatal seizures between the two drugs (OR=0.18, 95% CI: 0.01-3.70).

Psychomotor developmental delay

The NMA on psychomotor developmental delay (definitions in Table 1) included 11 cohort studies, 1,145 children exposed *in utero*, and examined 18 treatments (τ^2=0.06, 95% CrI: 0.00-0.63; Figure 3c; Additional File 3: Appendices H, J). Valproate (OR=4.16, 95% CrI: 2.04-8.75) and carbamazepine+phenobarbital+valproate (OR=19.12, 95% CrI: 1.49-337.50) were significantly more harmful than control (Figure 2c).

Language delay

The NMA on language delay (definitions in Table 1) included five cohort studies, 509 children, and examined five treatments (τ^2=0.16, 95% CrI: 0.00-2.15; Figure 3d; Additional File 3: Appendices H, J). One study included children exposed to AEDs *in-utero* and through breastfeeding, and four included children exposed to AEDs *in-utero*. Compared with control, valproate was the only treatment significantly associated with increased odds of language delay (OR=7.95, 95% CrI: 1.50-49.13; Figure 2d).

Attention deficit hyperactivity disorder

The NMA on ADHD (definitions in Table 1) included five cohort studies, 816 children, and examined seven treatments (τ^2=0.11, 95% CrI: 0.00-1.29). One study included children exposed to AEDs *in-utero* and through breastfeeding, while four studies included children exposed to AEDs *in-utero*. None of the treatment comparisons reached statistical significance (Figure 3e; Figure 2e; Additional File 3: Appendices H, J).

Social Impairment

One cohort study included 422 children exposed to AEDs *in-utero* as well as through breastfeeding. The children were exposed to carbamazepine (n=48), lamotrigine (n=71), valproate (n=27) and control (n=278). No significant differences in social impairment were identified.[41]

DISCUSSION

Our results suggest that AEDs generally pose a risk for infants and children exposed *in-utero* or during breastfeeding. Valproate was significantly associated with more children experiencing autism/dyspraxia, language, cognitive and psychomotor developmental delays versus children who were not exposed to AEDs. Oxcarbazepine, lamotrigine and lamotrigine+valproate were associated with increased occurrence of autism/dyspraxia, whereas for the cognitive developmental delay and psychomotor developmental delay outcomes, children exposed to the combination of carbamazepine, phenobarbital, and valproate were at greater odds of harm than those who were not exposed to AEDs. However, these results should be interpreted with caution, as a number of factors (e.g., anticonvulsant dosing, severity of epilepsy, duration of exposure, serum concentrations of exposure, mother's IQ/education) that may all influence outcomes were not identified in these studies. Also, our subsequent analyses may be underpowered due to missing data (e.g., 17 of the 27 studies did not report maternal age, 23 of 27 studies did not report alcohol use, 22 of 27 studies did not report tobacco use, and 14 of 27 studies did not include control group).

NMA is a particularly useful tool for decision-makers because it allows the ranking of treatments for each outcome. However, the results of our SUCRA curves should be interpreted with caution, especially due to the small number of studies and children included in each NMA, which is also reflected in the high uncertainty around the SUCRA values (Figure 2).[31]

Our results are consistent with a longitudinal study of 311 children that found exposure to lamotrigine was associated with significantly higher IQ scores and verbal function at six

years of age compared to children exposed to valproate (Additional File 3: Appendix C).[7] As indicated in Additional File 3: Appendix C, we were unable to include this study because the outcome was reported as a continuous measure, where we focused on dichotomous outcomes to facilitate interpretation. Our results are supported by findings from a cohort study, which found that children exposed to levetiracetam were not at increased risk for delayed development compared to unexposed children (Additional File 3: Appendix C).[42] As indicated in Additional File 3: Appendix C, we were unable to include this study due to the same reason as above. A NMA of 195 RCTs (including 28,013 both male and female patients) showed that gabapentin and levetiracetam showed the best tolerability profile compared with other AEDs, whereas oxcarbazepine and topiramate had a higher withdrawal rate, and lamotrigine an intermediate withdrawal rate.[43]

Across all outcomes, valproate alone or combined with another AED (even with a newer-generation agent, e.g., lamotrigine) was associated with the greatest odds. Similarly, two previous systematic reviews that did not conduct a NMA found valproate was associated with significantly lower IQ scores and poorer overall neurodevelopmental outcomes when compared to an unexposed control group.[5 6] Also consistent with our results, a 2014 Cochrane review including 28 studies (10 of these studies were included in the meta-analyses; with a maximum number of five studies per meta-analysis) concluded that AED polytherapy led to poorer developmental outcomes and IQ compared to healthy controls, epileptic controls, and unspecified monotherapy.[5] This Cochrane review also concluded that insufficient data exist for newer AEDs. However, unlike our review, it included and analysed fewer studies, and did not differentiate between specific polytherapy regimens, and thus did not compare these regimens versus each other or specific monotherapy AEDs.

These risks must be balanced with the need to control seizure activity in pregnancy and thus informed decision-making by patients and clinicians is critical.

Strengths of our study include a comprehensive systematic review methodology that followed the Cochrane Handbook[44] and ISPOR[12] guidelines, and reported using the PRISMA extension for NMA.[13] To the best of our knowledge, our study was the first that compared and ranked the safety of AEDs. We evaluated the comparative safety of treatments that have not been directly compared head-to-head before. In addition, we calculated predictive intervals, which account for between-study variation and provide a predicted range for the treatment effect estimate, should a future study be conducted. On average, the predictive intervals suggested that our results are robust.

Our systematic review has a few limitations worth noting. First, due to the complexity of the data and the studies' underreporting, differences in drug dosages could not be accounted for, and it was assumed that different dosages of the same AED were equally effective. When a study reported multiple dosages for the same treatment, we combined the data for this treatment. This is common for cohort studies, which report on a number of different types of exposures amongst patients. Second, several polytherapies had high SUCRA estimates but very wide CrIs, which is due to the small number of studies included for each drug combination with underpowered sample sizes. Evidence suggests that ranking probabilities for a treatment of being the best may be biased toward the treatments with the smallest number of studies, which may have influenced our SUCRA results.[31 45] As such, the effect sizes need to be taken into account when considering the SUCRA values. Third, due to the absence of evidence from RCTs, our conclusions were based on evidence from observational studies only, and inherent biases because of

confounding and shortcomings of these studies may have impacted our findings. For example, the included studies often failed to report important treatment effect modifiers,[46] such as family history of autism, ADHD, and maternal IQ, severity of epilepsy making it impossible for us to explore their impact through subgroup analysis and meta-regression. Recent research has explored methods to incorporate non-randomized with randomized evidence in a NMA and have highlighted the need to carefully explore the level of confidence in the non-randomized evidence.[47 48] The use of observational studies allows the assessment of the safety profile of AED treatments and offers the opportunity to evaluate effects in pregnancy.[49] Future large-scale observational studies are needed to allow the evaluation of rare adverse events that otherwise cannot be adequately evaluated in RCTs, especially during pregnancy. Fourth, although no intransitivity for most effect modifiers assessed was evident, there was an imbalance in the methodological study quality appraisal across treatment comparisons and most outcomes, which may impact our results. Unknown factors or factors that could not be assessed due to dearth of data may pose the risk of residual confounding bias, and hence risk the validity of the transitivity assumption. However, the assessment of consistency suggested no disagreement between the different sources of evidence in the network. Fifth, although the tendency towards small-study effects is greater with observational studies than with randomized trials,[50] the assessment of small-study effects using adjusted funnel plots suggested no evidence for their prevalence. Also, the majority of the included studies in this review compared multiple treatments inducing correlations in each funnel plot, which may mask asymmetry. Although we plotted data points corresponding to the study-specific basic parameters to reduce correlations, this issue may still exist. Sixth, we were unable to conduct subgroup

analysis by type of exposure (breastfeeding versus *in utero*) due to the small number of studies included in the NMA and due to the poor reporting; 22 studies did not report whether exposure was also in breastfeeding (additional to *in utero*). Hence, we included all studies in the analysis irrespective of the type of exposure.

More evidence from long-term follow-up studies is required to further delineate neurodevelopmental risks in children. Future studies should assess the genetic contribution from the biological father, maternal seizures during pregnancy, exposure through breastfeeding only, types of epilepsy, and maternal family history. Registries should aim to include a suitable control group and collect information on potential confounders, such as alcohol and tobacco use, allowing researchers to identify the safest agents for different patient-level covariates, and enhance decision-making for healthcare providers and patients. A critical evaluation of the validity of the control group is also necessary, in order to examine potential differences between the treated and the not treated populations. An individual patient data NMA would likely provide further clarity to the field, which allows the tailoring of management to specific patient characteristics.[51]

CONCLUSION

Across all outcomes and treatments compared with control, valproate alone or combined with another AED was associated with the greatest odds, whereas oxcarbazepine and lamotrigine were associated with increased occurrence of autism. Counselling is advised for women considering pregnancy to tailor the safest regimen.

LIST OF ABBREVIATIONS

AEDs: Anti-epileptic drugs; CrI: Credible interval; NMA: Network Meta-analysis; OR: Odds ratio; PrI: Predictive interval; SUCRA curve: Surface under the cumulative ranking curve

ADDITIONAL FILES

Additional File 1: Protocol

Additional File 2: PRISMA NMA Checklist

Additional File 3: Supplementary Online Content (Appendices A-O)

Appendix A. Newcastle-Ottawa Scale scoring guide

Appendix B. List of included studies

Appendix C. Key Excluded Studies

Appendix D. Table of Individual Study Characteristics

Appendix E. Table of Patient Characteristics

Appendix F. Methodological quality of observational studies – Newcastle Ottawa Scale

Appendix G. Comparison-adjusted funnel plots

Appendix H. Statistically significant network meta-analysis results along with meta-analysis results, transitivity, and inconsistency assessments

Appendix I. Rank-heat plot of cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcomes

Appendix J. Number of studies and treatments per outcome

Appendix K. Sensitivity and network meta-regression analyses - Anti-epileptic drugs compared with Control

FIGURE LEGENDS

Figure 1. Study flow diagram

Figure 2. Forest plots for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcome

Figure 3. Network diagrams for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcomes

Each treatment node is weighted according to the number of patients that have received the particular treatment, and each edge is weighted according to the number of studies comparing the treatments it connects.

Abbreviations: carbam - carbamazepine, clobaz - clobazam, clonaz - clonazepam, ethos - ethosuximide, gabap - gabapentin, lamot - lamotrigine, levet - levetiracetam, oxcar - oxcarbazepine, pheno - phenobarbital, pheny - phenytoin, primid - primidone, topir - topiramate, valpro - valproate, vigab – vigabatrin

DECLARATIONS

CONTRIBUTORS

AAV analysed the data, interpreted the results, and drafted the manuscript. ACT and SES conceived and designed the study, helped obtain funding, interpreted the results, and helped write sections of the manuscript. PR and EC coordinated the review, screened citations and full-text articles, abstracted data, appraised quality, resolved discrepancies, contacted authors, and edited the manuscript. CS provided methodological support and screened citations and full-text articles and edited the manuscript. RK, ER, FY, JDS, KT, and HM screened citations and full-text articles, abstracted data, and/or appraised quality. BH, BRH and YF helped conceive the study and edited the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

We thank Dr. David Moher for providing his feedback on our protocol. We thank Dr. Laure Perrier for conducting the literature searches, Becky Skidmore for peer-reviewing the MEDLINE search, and Alissa Epworth for obtaining the full-text articles. We thank Alistair Scott, Wing Hui, and Geetha Sanmugalingham for screening some of the citations and/or abstracting some of the data for a few of the included studies, Misty Pratt and Mona Ghannad for helping scan reference lists, and Ana Guzman, Susan Le, and Inthuja Selvaratnam for contacting authors and formatting the manuscript.

FUNDING

This systematic review was funded by the Canadian Institutes for Health Research/Drug Safety and Effectiveness Network (CIHR/DSEN). AAV is funded by the Banting Postdoctoral Fellowship Program from the CIHR. SES is funded by a Tier 1 Canada Research Chair in

Knowledge Translation. BH is funded by a CIHR/DSEN New Investigator Award in Knowledge Synthesis. BRH receives funding from the Alberta Heritage Foundation for Medical Research. ACT is funded by a Tier 2 Canada Research Chair in Knowledge Synthesis. The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

COMPETING INTERESTS

None declared.

ETHICS APPROVAL

Not applicable.

PROVENANCE AND PEER REVIEW

Not commissioned; externally peer reviewed.

DATA SHARING STATEMENT

All datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

OPEN ACCESS

REFERENCES

1. Spina E, Perugi G. Antiepileptic drugs: indications other than epilepsy. *Epileptic Disord*. 2004;6(2):57-75.
2. Harden CL, Pennell PB, Koppel BS, et al. Management issues for women with epilepsy--focus on pregnancy (an evidence-based review): III. vitamin K, folic acid, blood levels, and breast-feeding: report of the quality standards subcommittee and therapeutics and technology assessment subcommittee of the American Academy of Neurology and the American Epilepsy Society. *Epilepsia*. 2009;50(5):1247-55.
3. Samren EB, van Duijn CM, Koch S, et al. Maternal use of antiepileptic drugs and the risk of major congenital malformations: a joint European prospective study of human teratogenesis associated with maternal epilepsy. *Epilepsia*. 1997;38(9):981-90.
4. Meador K, Reynolds MW, Crean S, Fahrbach K, Probst C. Pregnancy outcomes in women with epilepsy: a systematic review and meta-analysis of published pregnancy registries and cohorts. *Epilepsy Res*. 2008;81(1):1-13.
5. Bromley R, Weston J, Adab N, et al. Treatment for epilepsy in pregnancy: neurodevelopmental outcomes in the child. *The Cochrane database of systematic reviews*. 2014(10):Cd010236.
6. Banach R, Boskovic R, Einarson T, Koren G. Long-term developmental outcome of children of women with epilepsy, unexposed or exposed prenatally to antiepileptic drugs: a meta-analysis of cohort studies. *Drug Saf*. 2010;33(1):73-9.
7. Meador KJ, Baker GA, Browning N, et al. Fetal antiepileptic drug exposure and cognitive outcomes at age 6 years (NEAD study): a prospective observational study. *Lancet Neurol*. 2013;12(3):244-52.
8. Christensen J, Gronborg TK, Sorensen MJ, et al. Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. *JAMA*. 2013;309(16):1696-703.
9. Wlodarczyk BJ, Palacios AM, George TM, Finnell RH. Antiepileptic drugs and pregnancy outcomes. *Am J Med Genet A*. 2012;158a(8):2071-90.
10. Velez-Ruiz NJ, Meador KJ. Neurodevelopmental effects of fetal antiepileptic drug exposure. *Drug Saf*. 2015;38(3):271-8.
11. Tricco AC, Cogo E, Angeliki VA, et al. Comparative safety of anti-epileptic drugs among infants and children exposed in utero or during breastfeeding: protocol for a systematic review and network meta-analysis. *Systematic reviews*. 2014;3:68.
12. Jansen JP, Trikalinos T, Cappelleri JC, et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. *Value Health*. 2014;17(2):157-73.
13. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med*. 2015;162(11):777-84.
14. Veroniki AA, Cogo E, Rios P, et al. Comparative safety of anti-epileptic drugs during pregnancy: a systematic review and network meta-analysis of congenital malformations and prenatal outcomes. *BMC Med*. 2017;15(1):95.
15. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *J Clin Epidemiol*. 2016;75:40-6.
16. Wells G, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses2000. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.

17. Chaimani A, Higgins JP, Mavridis D, Spyridonos P, Salanti G. Graphical tools for network meta-analysis in STATA. *PLoS One*. 2013;8(10):e76654.
18. Furukawa TA, Guyatt GH, Griffith LE. Can we individualize the 'number needed to treat'? An empirical study of summary effect measures in meta-analyses. *Int J Epidemiol*. 2002;31(1):72-6.
19. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med*. 2004;23(20):3105-24.
20. Jansen JP, Naci H. Is network meta-analysis as valid as standard pairwise meta-analysis? It all depends on the distribution of effect modifiers. *BMC Med*. 2013;11:159.
21. White IR, Barrett JK, Jackson D, Higgins JP. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Res Synth Methods*. 2012;3(2):111-25.
22. Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ*. 2003;326(7387):472.
23. Veroniki AA, Vasiliadis HS, Higgins JP, Salanti G. Evaluation of inconsistency in networks of interventions. *Int J Epidemiol*. 2013;42(1):332-45.
24. Welton NJ, Sutton AJ, Cooper N, Abrams KR, Ades A. Evidence synthesis for decision making in healthcare. New York: Wiley; 2012.
25. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc Ser B Stat Methodol*. 2002;64(4):583-639.
26. Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: Evolution, critique and future directions. *Stat Med*. 2009;28(25):3049-67.
27. Palmer T, Sterne J. Meta-Analysis in Stata: An Updated Collection from the Stata Journal. White I, editor. Texas: Stata Press; 2016.
28. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ*. 2011;342:d549.
29. Higgins JP, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *J R Stat Soc Ser A Stat Soc*. 2009;172(1):137-59.
30. Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol*. 2011;64(2):163-71.
31. Trinquart L, Attiche N, Bafeta A, Porcher R, Ravaud P. Uncertainty in Treatment Rankings: Reanalysis of Network Meta-analyses of Randomized Trials. *Ann Intern Med*. 2016;164(10):666-73.
32. Bromley R, Baxter N, Calderbank R, Mawer G, Clayton-Smith J, Baker G. A comprehensive review of the language abilities of children exposed to valproate or carbamazepine in utero. American Epilepsy Society; Texas2010.
33. Gogatishvili N, Ediberidze T, Lomidze G, Tatishvili N, Kasradze S. PO-0834 Long-term Developmental Outcome Of Children Prenatally Exposed To Antiepileptic Drugs. *Arch Dis Child*. 2014;99(Suppl 2):A526.
34. Gogatishvili N, Ediberidze T, Lomidze G, Tatishvili N, Kasradze S. Cognitive outcomes of children with fetal antiepileptic drug exposure at the age of 3-6 years-preliminary data. 1st Congress of the European Academy of Neurology; Berlin: European Journal of Neurology; 2015. p. 329.
35. Miskov S, Juraski RG, Fucic A, et al. Croatian Pregnant Women with Epilepsy and Effects of Antiepileptic Drugs Exposure in their Offspring - seven years of prospective surveillance. American Epilepsy Society; Texas2010.
36. Adab N, Kini U, Vinten J, et al. The longer term outcome of children born to mothers with epilepsy. *J Neurol Neurosurg Psychiatry*. 2004;75(11):1575-83.

37. Dean JCS, Hailey H, Moore SJ, Lloyd DJ, Turnpenny PD, Little J. Long term health and neurodevelopment in children exposed to antiepileptic drugs before birth. *J Med Genet*. 2002;39(4):251-9.
38. Gaily E. Development and growth in children of epileptic mothers: a prospective controlled study. Helsinki, Finland: University of Helsinki; 1990.
39. Katz JM, Pacia SV, Devinsky O. Current Management of Epilepsy and Pregnancy: Fetal Outcome, Congenital Malformations, and Developmental Delay. *Epilepsy Behav*. 2001;2(2):119-23.
40. Veroniki AA, Straus SE, Fyraridis A, Tricco AC. The rank-heat plot is a novel way to present the results from a network meta-analysis including multiple outcomes. *J Clin Epidemiol*. 2016.
41. Veiby G, Engelsen BA, Gilhus NE. Early child development and exposure to antiepileptic drugs prenatally and through breastfeeding: a prospective cohort study on children of women with epilepsy. *JAMA Neurol*. 2013;70(11):1367-74.
42. Shallcross R, Bromley RL, Irwin B, Bonnett LJ, Morrow J, Baker GA. Child development following in utero exposure: levetiracetam vs sodium valproate. *Neurology*. 2011;76(4):383-9.
43. Zaccara G, Giovannelli F, Giorgi FS, Franco V, Gasparini S, Benedetto U. Tolerability of new antiepileptic drugs: a network meta-analysis. *Eur J Clin Pharmacol*. 2017.
44. Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions. 5.1.0 ed: The Cochrane Collaboration; 2009.
45. Kibret T, Richer D, Beyene J. Bias in identification of the best treatment in a Bayesian network meta-analysis for binary outcome: a simulation study. *Clin Epidemiol*. 2014;6:451-60.
46. Dalessio DJ. Seizure Disorders and Pregnancy. *N Engl J Med*. 1985;312(9):559-63.
47. Efthimiou O, Mavridis D, Debray TP, et al. Combining randomized and non-randomized evidence in network meta-analysis. *Stat Med*. 2017.
48. Schmitz S, Adams R, Walsh C. Incorporating data from various trial designs into a mixed treatment comparison model. *Stat Med*. 2013;32(17):2935-49.
49. Cameron C, Fireman B, Hutton B, et al. Network meta-analysis incorporating randomized controlled trials and non-randomized comparative cohort studies for assessing the safety and effectiveness of medical treatments: challenges and opportunities. *Systematic reviews*. 2015;4:147.
50. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet*. 1991;337(8746):867-72.
51. Veroniki AA, Straus SE, Soobiah C, Elliott MJ, Tricco AC. A scoping review of indirect comparison methods and applications using individual patient data. *BMC Med Res Methodol*. 2016;16(1):47.

666 **Table 1. Outcome measures and diagnostic scales used in analysis**

Cognitive developmental delay	
Bayley Scales of Infant Development (children ≤42 mo.)	Score ≥2 standard deviations below the mean
Griffiths Scale of Infant Development (children >42 mo.)	Score ≥2 standard deviations below the mean
McCarthy Scales of Children's Abilities (children >30 mo.)	Score ≥1 standard deviations below the mean
Stanford-Binet IV Intelligence scale for children	Intelligence quotient ≤80
Touwen's Test	Above average number of items rated abnormal in one or more domains
Wechsler Scale of Preschool and Primary Intelligence	Intelligence quotient <90
Wechsler Intelligence Scale for Children - III	Intelligence quotient <80; verbal intelligence quotient <69
Developmental Assessment	Confirmed diagnosis by developmental pediatrician or pediatric neurologist
Autism/dyspraxia	
Developmental Assessment	Diagnosis confirmed by developmental specialists at 2 years of age
Medical Records	Confirmed diagnosis recorded in medical history; registry records (ICD-10 codes F84.0, F84.1, F84.5, F84.8, and F84.9)
Modified checklist for autism in toddlers	Scored positive for ≥2 out of 6 critical items OR ≥3 any items of the total scale
Psychomotor developmental delay	
Ages and Stages Questionnaire	>3 standard deviations from the test mean
Bayley Scales of Infant Development – Psychomotor Index	>2 standard deviations below the standardized mean for the test
Touwen's Test	Demonstrated dysfunctions in fine motor balance, fine motor functions, and coordination of extremities
Schedule of Growing Skills II	Scored as 'delayed' in ≥1 domain of the test

Developmental Assessment	Infant scored >2 negative items (administered by general practitioner or pediatrician); diagnosis of neuromotor deficit confirmed by a trained nurse practitioner; infant failing to sit by 10 months or walk by 18 months
Health/Medical Records	Diagnosis of psychomotor delay recorded in medical records
Language Delay	
Ages and Stages Questionnaire	>3 standard deviations from the test mean
Clinical Evaluation of Language Fundamentals – 4th Edition	Score <70 in core language domain; score <84 overall
Learning Accomplishment Profile	Below average performance in expressive speech (adjusted for age)
Comprehensive Language Assessment (Peabody Picture Vocabulary Test; Receptive Expressive Emergent Language Scale; Expressive One Word Picture Vocabulary Test, or Sequenced Inventory of Communication Development)	Scores/assessment indicate a >6 moth delay in age appropriate language development
ADHD	
Attention Problems and Hyperactivity Scales	Score >1 standard deviations from the test mean
Child Behaviour Checklist	≥6 positive items on checklist
Diagnostic and Statistical Manual – IV	≥5 positive items on checklist
Medical Records	Confirmed diagnosis in hospital/medical records made by a pediatrician or child psychiatrist
Neonatal Seizure	
Medical records	Record of seizures during 1st year; confirmation of neonatal seizure by electroencephalography or diagnosis
Social Impairment	
Developmental Assessment (Ages and Stages Questionnaire [6 and 18 months]; Child Behaviour Checklist [36 months])	Scores dichotomized into 'normal' or 'adverse' range based on pre-defined values used by scale, for scales without pre-defined values cut-off was set at a score >2 standard deviations outside the test mean

667

668 **Table 2. Summary characteristics of included studies**

Study/Patient Characteristic	# of Studies (n=29)	% of Studies
Year of publication		
1980-1989	1	3.45
1990-1999	6	20.69
2000-2009	5	17.24
2010-2015	17	58.62
Continent (of country of study conduct)		
Europe	20	68.97
North America	5	17.24
Asia	1	3.45
Australia	2	6.90
Trans-Continental	1	3.45
Study design		
Observational cohort	29	100.00
Case-control	0	0.00
Randomized clinical trial	0	0.00
Registry study		
Yes	11	37.93
No	18	62.07
Sample size		
0-99	18	62.07
100-299	9	31.03
300-499	1	3.45
500-699	0	0.00
700-999	0	0.00
1000+	1	3.45
Number of interventions		
2	4	13.79
3	5	17.24
4	8	27.59
5-7	8	27.59
8-10	2	6.90
11+	2	6.90
Outcomes[*,†]		
Cognitive Developmental Delay	12	58.62
Autism/Dyspraxia	5	17.24
Language Delay	5	17.24
ADHD	5	17.24
Psychomotor Developmental Delay	11	37.93

Study/Patient Characteristic	# of Studies (n=29)	% of Studies
Neonatal Seizures	2	6.90
Social Impairment	1	3.45
Funding		
Public	15	51.72
Private	0	0.00
Mixed public and private	4	13.79
NR/Unclear	10	34.48
Treatment indication		
Epilepsy	23	79.31
Mixed indications‡	0	0.00
Not reported	6	20.69
Epileptic control group§		
Yes	15	51.72
No/NR/NA	14	48.28
Mean maternal age		
24-26 y	2	6.90
27-29 y	5	17.24
30+ y	4	13.79
Not reported	18	62.07
AED exposure during pregnancy		
Reported as during 1st trimester	5	17.24
Reported as any time during pregnancy	4	13.79
During pregnancy and breastfeeding	5	17.24
Not reported	15	51.72
Alcohol use during pregnancy		
Yes	5	17.24
NR	24	82.76
Tobacco use during pregnancy		
Yes	7	24.14
NR	22	75.86

Abbreviations: ADHD - Attention Deficit Hyperactivity Disorder; AED - anti-epileptic drug(s); NA - Not applicable; NR - Not reported

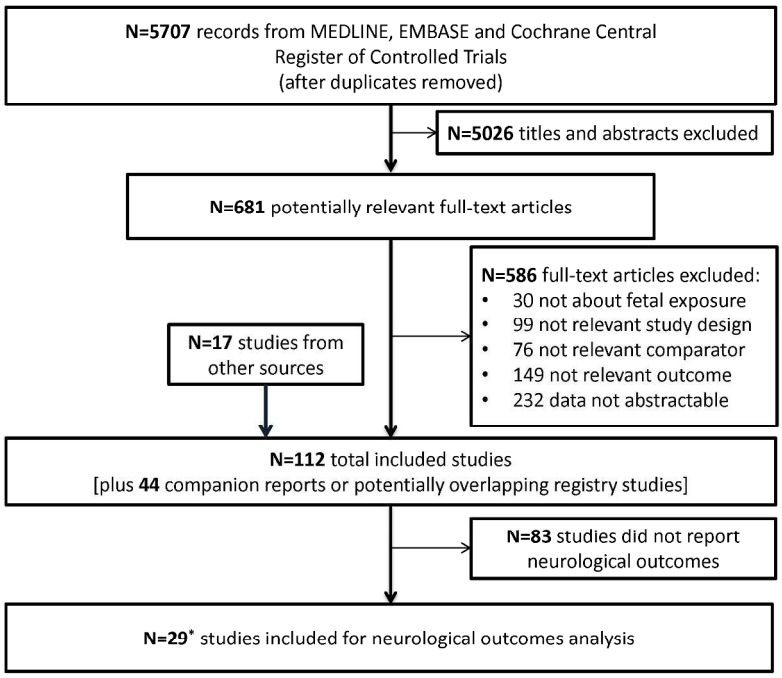
* Values in this category do not match totals as some studies report more than one outcome

† Percentage of total number of included studies (n=29)

‡ Includes individuals taking AEDs for psychiatric disorders, migraine, and neuropathic/neurological pain

§ Consisted of women with Epilepsy who did not take AEDs during pregnancy

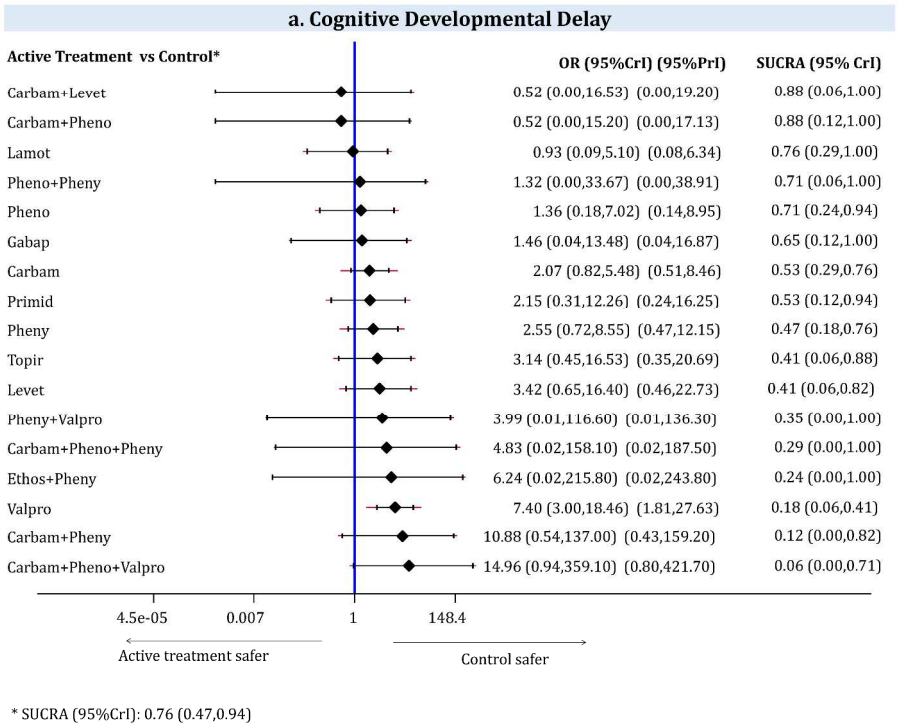
669



*29 publications reporting 30 included studies.

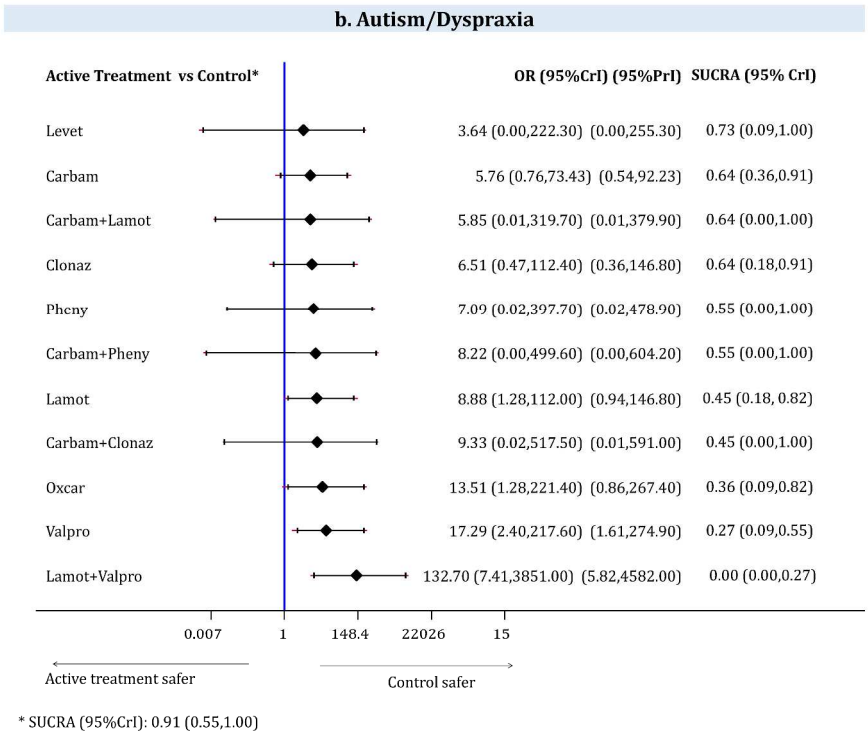
TITLE: Study flow diagram

254x190mm (300 x 300 DPI)



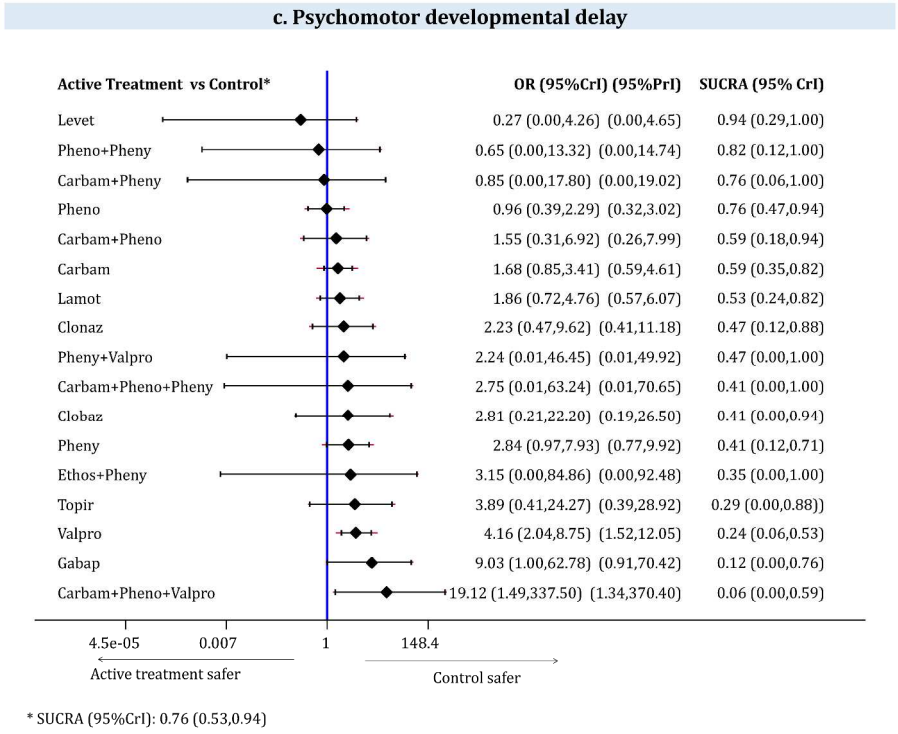
TITLE: Forest plots for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcome

533x400mm (300 x 300 DPI)



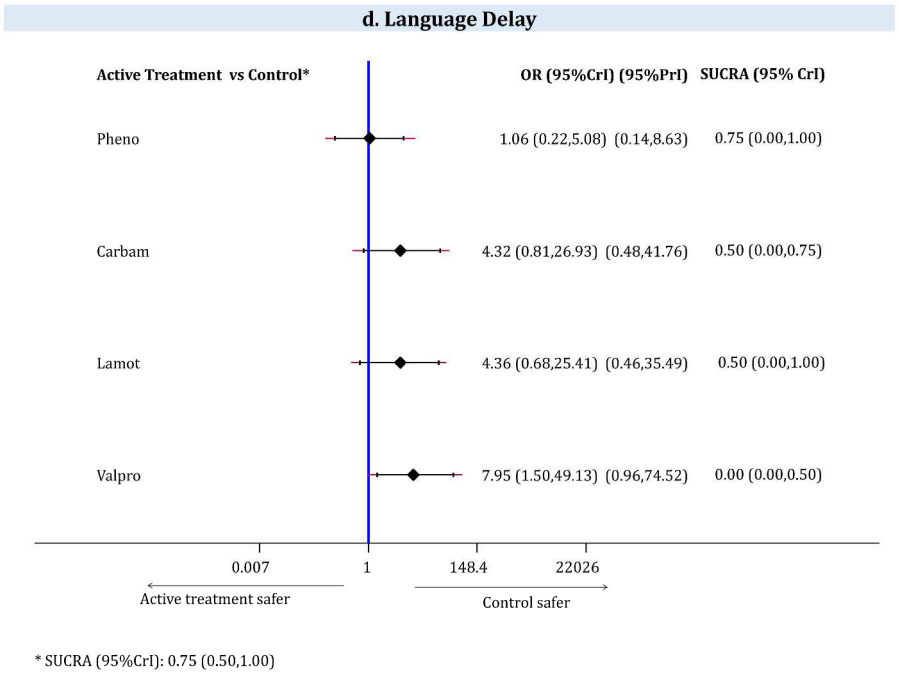
TITLE: Forest plots for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcome

533x400mm (300 x 300 DPI)



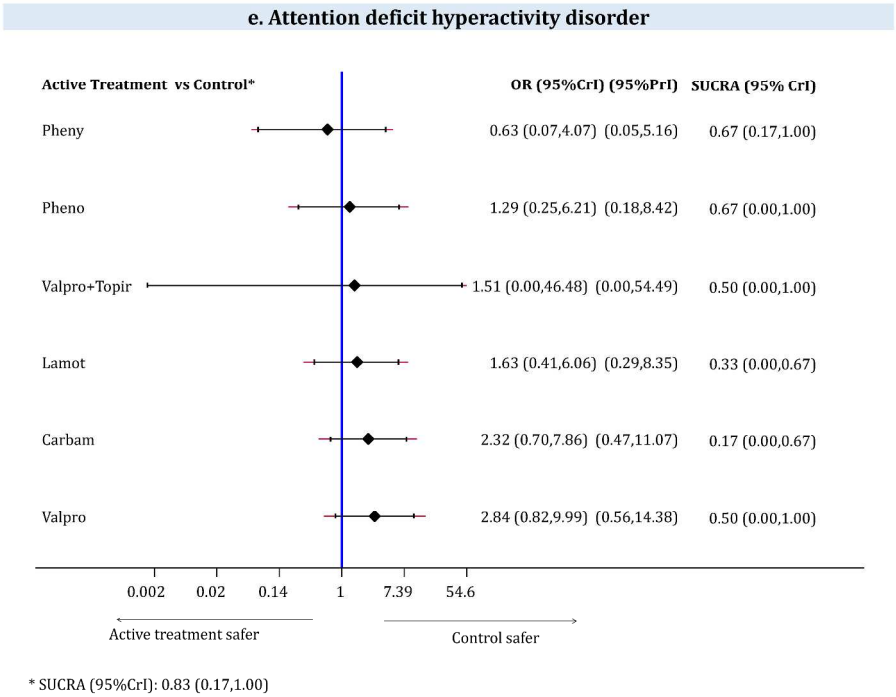
TITLE: Forest plots for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcome

533x400mm (300 x 300 DPI)



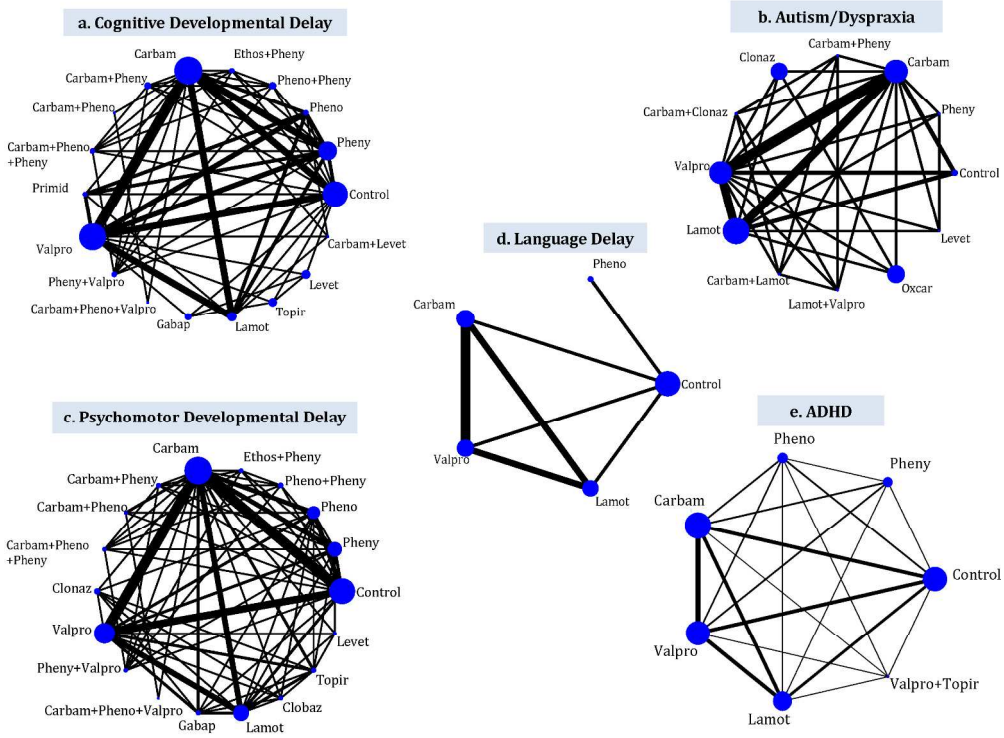
TITLE: Forest plots for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcome

533x400mm (300 x 300 DPI)



TITLE: Forest plots for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcome

533x400mm (300 x 300 DPI)



TITLE: Network diagrams for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcomes.

CAPTION: Each treatment node is weighted according to the number of patients that have received the particular treatment, and each edge is weighted according to the number of studies comparing the treatments it connects.

Abbreviations: carbam - carbamazepine, clobaz - clobazam, clonaz - clonazepam, ethos - ethosuximide, gabap - gabapentin, lamot - lamotrigine, levet - levetiracetam, oxcar - oxcarbazepine, pheno - phenobarbital, pheny - phenytoin, primid - primidone, topir - topiramate, valpro - valproate, vigab – vigabatrin

254x190mm (300 x 300 DPI)

Tricco et al. Systematic Reviews 2014, 3:68
http://www.systematicreviewsjournal.com/content/3/1/68

SYSTEMATIC REVIEWS

PROTOCOL Open Access

Comparative safety of anti-epileptic drugs among infants and children exposed *in utero* or during breastfeeding: protocol for a systematic review and network meta-analysis

Andrea C Tricco[1], Elise Cogo[1], Veroniki A Angeliki[1], Charlene Soobiah[1,2], Brian Hutton[3], Brenda R Hemmelgarn[4], David Moher[3], Yaron Finkelstein[5,6,7] and Sharon E Straus[1,8*]

Abstract

Background: Epilepsy affects about 1% of the general population. Anti-epileptic drugs (AEDs) prevent or terminate seizures in individuals with epilepsy. Pregnant women with epilepsy may continue taking AEDs. Many of these agents cross the placenta and increase the risk of major congenital malformations, early cognitive and developmental delays, and infant mortality. We aim to evaluate the comparative safety of AEDs approved for chronic use in Canada when administered to pregnant and breastfeeding women and the effects on their infants and children through a systematic review and network meta-analysis.

Methods: Studies examining the effects of AEDs administered to pregnant and breastfeeding women regardless of indication (e.g., epilepsy, migraine, pain, psychiatric disorders) on their infants and children will be included. We will include randomized clinical trials (RCTs), quasi-RCTs, non-RCTs, controlled before-after, interrupted time series, cohort, registry, and case-control studies. The main literature search will be executed in MEDLINE, EMBASE, and the Cochrane Central Register of Controlled Trials. We will seek unpublished literature through searches of trial protocol registries and conference abstracts. The literature search results screening, data abstraction, and risk of bias appraisal will be performed by two individuals, independently. Conflicts will be resolved through discussion. The risk of bias of experimental and quasi-experimental studies will be appraised using the Cochrane Effective Practice and Organization of Care Risk-of-Bias tool, methodological quality of observational studies will be appraised using the Newcastle-Ottawa Scale, and quality of reporting of safety outcomes will be conducted using the McMaster Quality Assessment Scale of Harms (McHarm) tool. If feasible and appropriate, we will conduct random effects meta-analysis. Network meta-analysis will be considered for outcomes that fulfill network meta-analysis assumptions.

The primary outcome is major congenital malformations (overall and by specific types), while secondary outcomes include fetal loss/miscarriage, minor congenital malformations (overall and by specific types), cognitive development, psychomotor development, small for gestational age, preterm delivery, and neonatal seizures.

(Continued on next page)

* Correspondence: sharon.straus@utoronto.ca
[1]Li Ka Shing Knowledge Institute, St. Michael's Hospital, 209 Victoria Street, East Building, Toronto, Ontario M5B 1 T8, Canada
[8]Department of Geriatric Medicine, University of Toronto, 172 St. George Street, Toronto, Ontario M5R 0A3, Canada
Full list of author information is available at the end of the article

(Continued from previous page)

Discussion: Our systematic review will address safety concerns regarding the use of AEDs during pregnancy and breastfeeding. Our results will be useful to healthcare providers, policy-makers, and women of childbearing age who are taking anti-epileptic medications.

Systematic review registration: PROSPERO CRD42014008925.

Keywords: Anti-epileptic drug, Breastfeeding, Comparative safety, Congenital malformation, Epilepsy, Fetus, Infant, Network meta-analysis, Pregnancy, Systematic review

Background

Epilepsy is the most common chronic neurological condition, affecting 0.6 to 1% of the population [1,2]. Individuals with uncontrolled epilepsy experience recurrent seizures, which can have psychosocial and physical consequences, including a compromised life expectancy [3,4]. The goal of anti-epileptic treatment is to improve quality of life and health outcomes by reducing the frequency of seizures [4].

Anti-epileptic medications decrease seizures by reducing excitation and enhancing inhibition of neurons [5-7]. Many of these medications target different channels, including calcium, sodium, and glutamate, and are broadly classified as first generation agents (e.g., phenobarbitone, phenytoin, carbamazepine, sodium valproate, ethosuximide) and second generation agents (e.g., lamotrigine, levetiracetam, topiramate, gabapentin, vigabatrin, oxcarbazepine, clobazam, clonazepam, zonisamide, lacosamide, rufinamide, primidone) [8]. Due to the broad and varied mechanisms of action, the indications for some of these medications also include pain syndromes, psychiatric disorders, and migraine headaches [8].

Many clinical practice guidelines recommend that women of childbearing age continue to take their anti-epileptic medications; however, medications with lower risk of teratogenic events are advised [9,10] since anti-epileptic drugs (AEDs) cross the placenta or transfer through breast milk, posing risks to the fetus and infant [9,11,12].

Some AEDs have been associated with increased risk of harm to the fetus and infants. For example, exposure to valproate has led to increased risk of major congenital malformations [10], cognitive delay, and minor congenital abnormalities [13-16]. Phenobarbital has been associated with minor congenital abnormalities and developmental delay [17,18]. Carbamazepine and lamotrigine have been associated with minor congenital abnormalities [19-22]. However, other than studies of the use of valproate, many studies have produced inconsistent findings regarding harm to the fetus and infant with use of other agents [23]. As such, our objective is to evaluate the comparative safety of AEDs for infants and children who were exposed *in utero* or during breastfeeding through a systematic review and network meta-analysis.

Methods/Design

Protocol

A systematic review protocol was developed and registered with the PROSPERO database (CRD42014008925, available at: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014008925). It was revised with feedback from the decision-makers who posed the query within Health Canada, healthcare practitioners, content experts, and research methodologists. The reporting of our systematic review protocol was guided by the Preferred Reporting Items for Systematic Reviews and Meta-analyses Protocols [24].

Eligibility criteria

We will include studies examining the effects of AEDs on infants and children who were exposed *in utero* or during breastfeeding. We will include experimental studies (randomized clinical trials [RCTs], quasi-RCTs, non-RCTs), quasi-experimental studies (controlled before and after studies, interrupted time series), and observational studies (cohort, case-control, registry studies) of pregnant women at any stage of pregnancy and breastfeeding women and their infants/children. The rationale for including other study designs in addition to RCTs is that there are ethical issues in conducting RCTs of AEDs in pregnancy, so RCT evidence might not exist for some or all of these drugs. Given that our review includes rare outcomes, including observational evidence is crucial. In contrast to efficacy evaluation, safety assessment usually requires very large sample sizes to be able to detect adverse events. Therefore, while RCTs have lower risk of bias, they usually do not have the statistical power needed to adequately evaluate uncommon/rare safety outcomes due to Type II (i.e., false negative) error [25]. Given that our review includes rare outcomes, including observational evidence is crucial [26]. Additionally, observational studies can often provide more generalizable evidence due to the strict participant inclusion criteria in most RCTs [27]. Real-world safety evidence that has external validity is important for the assessment of the possible risks of AEDs in pregnant and breastfeeding women.

The diagnosis of neurodevelopmental delay related to *in utero* exposure is made before adolescence, and

hence, we will limit inclusion to children up to 12 years of age. AEDs that are approved for chronic use in Canada will be included. Drugs that are only used acutely or those that are not currently approved for use in Canada will be excluded, as the focus of this review is on the Canadian setting [28-32]. However, most of the medications we will examine are available in other countries as well. The relevant 16 medications and their synonyms are listed in Additional file 1, and the excluded drugs are listed in Additional file 2. Studies of all combinations and doses of these medications are eligible for inclusion. Since we are only interested in exposures that occur *in utero* or during breastfeeding, studies examining AEDs administered directly to infants or children will be excluded. All indications for AEDs will be included such as epilepsy, migraine, pain, and psychiatric disorders.

In order to be included, studies must compare an anti-epileptic medication against another included anti-epileptic medication, placebo, a 'no intervention' control group, or combinations of two or more anti-epileptic medications. Only studies providing results for our outcomes of interest will be included. Our primary outcome is major congenital malformations (overall and by specific type, such as craniofacial defects and neural tube defects). Secondary outcomes include minor congenital malformations (overall and by specific type, such as epicanthal folds and microstomia), cognition (e.g., global cognitive functioning and specific cognitive domains such as attention), psychomotor development (e.g., autism, dyspraxia), small for gestational age, preterm delivery, neonatal seizures, and fetal loss/miscarriage. No other limitations will be imposed on the eligibility criteria, including published/unpublished material, language of dissemination, duration of follow-up, or year of publication. The draft eligibility criteria can be found in Additional file 3.

Information sources and literature search

Our main literature search will be executed in the MEDLINE database. The search terms were drafted by an experienced librarian and can be found in Additional file 4. The search was peer reviewed by another librarian using the Peer Review of Electronic Search Strategies checklist [33].

In addition to MEDLINE, we will also search the EMBASE and the Cochrane Central Register of Controlled Trials databases. We will follow the MEDLINE search strategy for these databases, and the search terms will be adjusted accordingly. The electronic database search will be supplemented by searching for unpublished literature [34]. This will be accomplished through exploring conference abstracts, clinical trial registries, and contacting manufacturers of AEDs. We will also scan the reference lists of included studies and previous reviews in the area [23,35,36].

Study selection process

The eligibility criteria screening form will be pilot-tested by the team and is presented in Additional file 3. We will calculate inter-rater reliability from the pilot-test and screening will only commence after high agreement (e.g., kappa statistic ≥60%) is observed [37]. Subsequently, two reviewers will screen each title/abstract and potentially relevant full-text articles from the literature search results, independently. Conflicts will be resolved through discussion. All screening will occur using our online screening software (synthesi.SR) [38].

Data items and data collection process

We will abstract data on the PICOS elements [39], including patient characteristics (e.g., age of the mother and infant/child, indication for anti-epileptic treatment, co-morbidities, concomitant medications), intervention details (e.g., type of anti-epileptic treatment, dose, route of administration, duration of treatment, timing [trimester] of treatment during pregnancy), comparator details (e.g., comparator agent, dose, route of administration), outcome results (e.g., major congenital abnormality, minor congenital abnormality, cognitive function, psychomotor development) at the longest duration of follow-up, and study characteristics (e.g., study design, country of conduct, year of conduct, sample size, setting). These characteristics will be abstracted using a data abstraction form created in Excel with an accompanying "cheat sheet" that will guide the reviewers with this process. The data abstraction form and cheat sheet will be pilot-tested and data abstraction will only commence when high agreement (e.g., kappa statistic ≥60%) [37] is observed. Each included study will be abstracted by two team members, independently, who will resolve disagreements through discussion.

Methodological quality/risk of bias appraisal

We will use various tools to assess the methodological quality/risk of bias of each of the studies that fulfill our eligibility criteria. This will be conducted by two reviewers, independently, and conflicts will be resolved through discussion. First, we will appraise the risk of bias of experimental and quasi-experimental studies using the Cochrane Effective Practice and Organization of Care Risk-of-Bias tool [40]. Second, we will assess the methodological quality of observational studies using the Newcastle-Ottawa Scale [41]. Third, the quality of reporting of harms will be appraised using the McMaster Quality Assessment Scale of Harms (McHarm) tool [42].

Synthesis of included studies

A narrative summary of study results will be presented along with evidence summary tables. When sufficient data are available, we will conduct random effects meta-

analysis to calculate pooled odds ratios for dichotomous data and pooled mean differences for continuous data [43,44]. Direct (pairwise) meta-analysis will be performed with RCTs alone in order to examine whether the data are consistent between direct and indirect evidence. If the large majority of included studies are observational, we will also conduct additional meta-analyses including observational studies alone. Analyses will be stratified by treatment indication (e.g., epilepsy, pain, etc.) to reduce clinical heterogeneity between different study populations whenever possible; for example, epilepsy itself in pregnant women is related to an increased baseline risk of certain neonatal adverse outcomes. Statistical, clinical, and methodological heterogeneity will be examined prior to conducting the meta-analysis. Funnel plots will be drawn for outcomes including at least 10 studies to explore asymmetry that might be explained by clinical, statistical, and methodological heterogeneity. The proportion of statistical heterogeneity will be examined using the I^2 measure [45] and the magnitude of statistical heterogeneity will be calculated using the restricted maximum likelihood [46]. Meta-regression will be conducted for clinically relevant subgroups or when extensive statistical heterogeneity is observed (e.g., $I^2 \geq 75\%$) [47]. This will allow the examination of the impact of important factors on our results, such as maternal age, dose, duration and timing (e.g., trimester) of antiepileptic treatment, co-morbidities, concomitant medications, risk of bias results, and sample size (due to Type II statistical power errors with rare adverse events). To ensure the meta-regression analysis is intuitive, the number of covariates examined will be less than 10% of the number of studies included in the meta-analysis for the particular outcome.

We anticipate that many of these outcomes will be rare. To deal with studies reporting zero events in one treatment arm, 0.5 will be added to the numerator and 1 will be added to the denominator. We will exclude studies reporting zero events in all treatment arms for a particular outcome [48,49]. We also anticipate that we will encounter missing data in the included studies. We will contact the study authors for this data and if we are unable to receive the data, we will impute missing data (e.g., measures of variance) using established methods [50]. To ensure that our imputations do not bias our results, we will conduct a sensitivity analysis [51]. The meta-analysis and meta-regression will be analyzed in R using the *metafor* command [52].

A random-effects network meta-analysis will be conducted to make inferences regarding the comparative safety of the various AEDs [15], as well as rank their safety using rankograms and the surface under the cumulative ranking curve [53]. We will ensure the following factors are present prior to conducting network meta-analysis: i) transitivity (i.e., comparable distribution of effect modifiers across comparisons), which will be examined using boxplots or percentages to visually inspect potential effect modifiers of treatment effect [54]; ii) consistency between direct and indirect data, which will be examined locally (i.e., in certain paths of the network) using the loop-specific method [55,56] and the node-splitting method [57], and globally (i.e., evaluating the network as a whole), using the design-by-treatment interaction model [58]; and iii) we will quantify the amount of variability attributed to heterogeneity and inconsistency rather than sampling error, by calculating the I^2 [59]. We will estimate the amount of heterogeneity using the restricted maximum likelihood method and assuming common within-network heterogeneity. We will compare the magnitude of heterogeneity between consistency and inconsistency models, as well as between meta-regression and network meta-analysis models to determine how much heterogeneity will be explained by inconsistency or the explanatory variable, respectively. We will first use the design-by-treatment model for the evaluation of inconsistency in a network as a whole and then, if inconsistency is detected, we will employ the loop-specific and node-splitting methods to identify which piece of evidence is responsible for inconsistency. As mentioned above, analyses will be stratified by treatment indication when clinically appropriate. Important heterogeneity and inconsistency will be explored using network meta-regression using the same methods as described above, as necessary.

Prior to conducting the network meta-analysis, we will hold a team meeting to finalize which treatment nodes will be included in the analysis since we are unclear about the indications, dosages, patient populations, and outcomes reported in all of the studies. We will discuss issues, including conducting a class versus independent drug analysis, inclusion of drug routes of administration and dosages, as well as timing of drug administration. These decisions will be examined through a sensitivity analysis in which we will classify treatment nodes using a different classification to see how stable our results are. The network meta-analysis results will be presented as summary treatment effects for each pair of treatments. Network meta-analysis will be conducted in Stata with the *mvmeta* routine [60].

A sequential approach will be used for the network meta-analysis. We will first restrict our analysis to RCTs, which will be the primary analysis of interest. We will then include data from quasi-experimental studies, and finally, data from observational studies. This will provide an understanding of the contribution of each type of study design to our summary estimates, providing us with information on how these agents work above and beyond clinical trials.

Discussion

Epilepsy is the most common chronic neurological condition, affecting 0.6 to 1% of the population [1,2]. Given that approximately a third of patients receiving AEDs are of reproductive age and almost half of pregnancies are unplanned [61], the fetus may be exposed to these in the first trimester of pregnancy, including during the critical stage of embryogenesis [62].

The comparative safety of these agents is currently unknown and our results will be important for policy-makers, healthcare providers, and women of childbearing age. To ensure our results have wide dissemination and uptake, we will publish our results in open access journals, present our findings at scientific conferences, conduct dissemination meetings with key stakeholders (including policy-makers and healthcare providers), and produce policy briefs for Health Canada, the organization that posed this query.

Additional files

Additional file 1: List of relevant medications.
Additional file 2: Excluded drugs.
Additional file 3: Draft eligibility criteria.
Additional file 4: MEDLINE literature search.

Abbreviations

AEDs: Anti-epileptic drugs; RCTs: Randomized clinical trials.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

ACT conceived and designed the study, helped obtain funding for the study, and helped write the draft protocol. EC registered the protocol with the PROSPERO database and edited the draft protocol. AV helped write the draft protocol. CS edited the draft protocol. BH, BRH, DM, and YF provided input into the design, helped obtain funding for the study, and edited the draft protocol. SES conceived the study, designed the study, obtained the funding, and helped write the draft protocol. All authors read and approved the final protocol.

Acknowledgements

This systematic review was funded by the Canadian Institutes of Health Research/Drug Safety and Effectiveness Network (CIHR/DSEN). ACT and BH are funded by a CIHR/DSEN New Investigator Award in Knowledge Synthesis. BRH receives funding from the Alberta Heritage Foundation for Medical Research. DM is funded by a University of Ottawa Research Chair. SES is funded by a Tier 1 Canada Research Chair in Knowledge Translation.
We thank Laure Perrier for conducting the literature searches and Becky Skidmore for peer reviewing the MEDLINE search strategy. We also thank Dr. Joseph Beyene for providing feedback on our original proposal and Wing Hui and Judy Tran for formatting the paper.

Author details

[1]Li Ka Shing Knowledge Institute, St. Michael's Hospital, 209 Victoria Street, East Building, Toronto, Ontario M5B 1 T8, Canada. [2]Institute of Health Policy Management and Evaluation, University of Toronto, Health Sciences Building, 155 College Street, Suite 425, Toronto, Ontario M5T 3 M6, Canada. [3]Clinical Epidemiology Program, Centre for Practice-Changing Research, Ottawa Hospital Research Institute, The Ottawa Hospital – General Campus and University of Ottawa, 501 Smyth Road, Box 711, Ottawa, Ontario K1H 8 L6, Canada. [4]Departments of Medicine and Community Health Sciences, University of Calgary, TRW Building, 3rd Floor, 3280 Hospital Drive NW, Calgary, Alberta T2N 4Z6, Canada. [5]The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada. [6]Department of Pediatrics, University of Toronto, 172 St. George Street, Toronto, Ontario M5R 0A3, Canada. [7]Department of Pharmacology and Toxicology, University of Toronto, Medical Sciences Building, 1 King's College Circle, Room 4207, Toronto, Ontario M5S 1A8, Canada. [8]Department of Geriatric Medicine, University of Toronto, 172 St. George Street, Toronto, Ontario M5R 0A3, Canada.

Received: 9 April 2014 Accepted: 17 June 2014
Published: 25 June 2014

References

1. Hauser WA, Hesdorffer D: *Epilepsy, Frequency, Causes and Consequences.* New York: Demos Publications; 1990.
2. Wiebe S, Bellhouse DR, Fallahay C, Eliasziw M: **Burden of epilepsy: the Ontario Health Survey.** *Can J Neurol Sci* 1999, **26**(4):263–270.
3. Sperling MR: **The consequences of uncontrolled epilepsy.** *CNS Spectr* 2004, **9**(2):98–101. 106–109.
4. Jones MW: **Consequences of epilepsy: why do we treat seizures?** *Can J Neurol Sci* 1998, **25**(4):S24–S26.
5. Dickenson AH, Ghandehari J: **Anti-convulsants and anti-depressants.** *Handb Exp Pharmacol* 2007, **177:**145–177.
6. Stefani A, Spadoni F, Bernardi G: **Voltage-activated calcium channels: targets of antiepileptic drug therapy?** *Epilepsia* 1997, **38**(9):959–965.
7. Snutch TP, Reiner PB: **Ca^{2+} channels: diversity of form and function.** *Curr Opin Neurobiol* 1992, **2**(3):247–253.
8. Spina E, Perugi G: **Antiepileptic drugs: indications other than epilepsy.** *Epileptic Disord* 2004, **6**(2):57–75.
9. Harden CL, Pennell PB, Koppel BS, Hovinga CA, Gidal B, Meador KJ, Hopp J, Ting TY, Hauser WA, Thurman D, Kaplan PW, Robinson JN, French JA, Wiebe S, Wilner AN, Vazquez B, Holmes L, Krumholz A, Finnell R, Shafer PO, Le Guen CL, American Academy of Neurology; American Epilepsy Society: **Management issues for women with epilepsy–focus on pregnancy (an evidence-based review): III. Vitamin K, folic acid, blood levels, and breast-feeding: report of the quality standards subcommittee and therapeutics and technology assessment subcommittee of the American Academy of Neurology and the American Epilepsy Society.** *Epilepsia* 2009, **50**(5):1247–1255.
10. Harden CL, Meador KJ, Pennell PB, Hauser WA, Gronseth GS, French JA, Wiebe S, Thurman D, Koppel BS, Kaplan PW, Robinson JN, Hopp J, Ting TY, Gidal B, Hovinga CA, Wilner AN, Vazquez B, Holmes L, Krumholz A, Finnell R, Hirtz D, Le Guen C, American Academy of Neurology; American Epilepsy Society: **Management issues for women with epilepsy-Focus on pregnancy (an evidence-based review): II. Teratogenesis and perinatal outcomes: Report of the Quality Standards Subcommittee and Therapeutics and Technology Subcommittee of the American Academy of Neurology and the American Epilepsy Society.** *Epilepsia* 2009, **50**(5):1237–1246.
11. Samren EB, van Duijn CM, Koch S, Hiilesmaa VK, Klepel H, Bardy AH, Mannagetta GB, Deichl AW, Gaily E, Granstrom ML, Meinardi H, Grobbee DE, Hofman A, Janz D, Lindhout D: **Maternal use of antiepileptic drugs and the risk of major congenital malformations: a joint European prospective study of human teratogenesis associated with maternal epilepsy.** *Epilepsia* 1997, **38**(9):981–990.
12. Meador K, Reynolds MW, Crean S, Fahrbach K, Probst C: **Pregnancy outcomes in women with epilepsy: a systematic review and meta-analysis of published pregnancy registries and cohorts.** *Epilepsy Res* 2008, **81**(1):1–13.
13. Adab N, Jacoby A, Smith D, Chadwick D: **Additional educational needs in children born to mothers with epilepsy.** *J Neurol Neurosurg Psychiatry* 2001, **70**(1):15–21.
14. Adab N, Kini U, Vinten J, Ayres J, Baker G, Clayton-Smith J, Coyle H, Fryer A, Gorry J, Gregg J, Mawer G, Nicolaides P, Pickering L, Tunnicliffe L, Chadwick DW: **The longer term outcome of children born to mothers with epilepsy.** *J Neurol Neurosurg Psychiatry* 2004, **75**(11):1575–1583.
15. Gaily E, Kantola-Sorsa E, Hiilesmaa V, Isoaho M, Matila R, Kotila M, Nylund T, Bardy A, Kaaja E, Granstrom ML: **Normal intelligence in children with prenatal exposure to carbamazepine.** *Neurology* 2004, **62**(1):28–32.
16. Meador KJ, Baker GA, Browning N, Clayton-Smith J, Combs-Cantrell DT, Cohen M, Kalayjian LA, Kanner A, Liporace JD, Pennell PB, Privitera M, Loring DW, for the NEAD Study Group: **Cognitive function at 3 years of age after fetal exposure to antiepileptic drugs.** *N Engl J Med* 2009, **360**(16):1597–1605.

17. Holmes LB, Wyszynski DF, Lieberman E: **The AED (antiepileptic drug) pregnancy registry: a 6-year experience.** *Arch Neurol* 2004, **61**(5):673–678.
18. Reinisch JM, Sanders SA, Mortensen EL, Rubin DB: **In utero exposure to phenobarbital and intelligence deficits in adult men.** *JAMA* 1995, **274**(19):1518–1525.
19. Morrow J, Russell A, Guthrie E, Parsons L, Robertson I, Waddell R, Irwin B, McGivern RC, Morrison PJ, Craig J: **Malformation risks of antiepileptic drugs in pregnancy: a prospective study from the UK Epilepsy and Pregnancy Register.** *J Neurol Neurosurg Psychiatry* 2006, **77**(2):193–198.
20. Meador KJ, Baker GA, Finnell RH, Kalayjian LA, Liporace JD, Loring DW, Mawer G, Pennell PB, Smith JC, Wolff MC, NEAD Study Group: **In utero antiepileptic drug exposure: fetal death and malformations.** *Neurology* 2006, **67**(3):407–412.
21. Vajda FJ, Hitchcock A, Graham J, Solinas C, O'Brien TJ, Lander CM, Eadie MJ: **Foetal malformations and seizure control: 52 months data of the Australian Pregnancy Registry.** *Eur J Neurol* 2006, **13**(6):645–654.
22. Holmes LB, Baldwin EJ, Smith CR, Habecker E, Glassman L, Wong SL, Wyszynski DF: **Increased frequency of isolated cleft palate in infants exposed to lamotrigine during pregnancy.** *Neurology* 2008, **70**(22 Pt 2):2152–2158.
23. Meador KJ, Penovich P, Baker GA, Pennell PB, Bromfield E, Pack A, Liporace JD, Sam M, Kalayjian LA, Thurman DJ, Moore E, Loring DW, NEAD Study Group: **Antiepileptic drug use in women of childbearing age.** *Epilepsy Behav* 2009, **15**(3):339–343.
24. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L: **Reporting Guidelines for Systematic Review Protocols.** In *19th Cochrane Colloquium: 19–22 October 2011; Madrid, Spain.*
25. Ross CJ, Visscher H, Sistonen J, Brunham LR, Pussegoda K, Loo TT, Rieder MJ, Koren G, Carleton BC, Hayden MR, CPNDS Consortium: **The Canadian Pharmacogenomics Network for Drug Safety: a model for safety pharmacology.** *Thyroid* 2010, **20**(7):681–687.
26. Eypasch E, Lefering R, Kum CK, Troidl H: **Probability of adverse events that have not yet occurred: a statistical reminder.** *BMJ* 1995, **311**(7005):619–620.
27. Atkins D: **Creating and synthesizing evidence with decision makers in mind: integrating evidence from clinical trials and other study designs.** *Med Care* 2007, **45**(10 Supl 2):S16–S22.
28. *Health Canada: Drug Product Database.* http://www.hc-sc.gc.ca/dhp-mps/prodpharma/databasdon/index-eng.php.
29. *United States National Library of Medicine's ChemIDPlus Lite Database.* http://chem.sis.nlm.nih.gov/chemidplus/chemidlite.jsp.
30. *Canadian Pharmacists Association: E-CPS (Compendium of Pharmaceuticals and Specialties).* http://www.e-therapeutics.ca/home.whatsnew.action.
31. *Epilepsy Canada: Anticonvulsant Medications.* http://www.epilepsy.ca/en-CA/Diagnosis-and-Treatment/Anticonvulsant-Medications.html.
32. *Epilepsy Ontario: Anticonvulsant/Anti-Seizure Medication from A to Z.* http://epilepsyontario.org/anticonvulsantanti-seizure-medication-from-a-to-z/.
33. Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, Lefebvre C: **An evidence-based practice guideline for the peer review of electronic search strategies.** *J Clin Epidemiol* 2009, **62**(9):944–952.
34. *Canadian Agency for Drugs and Technologies in Health: Grey Matters: A Practical Search Tool for Evidence-Based Medicine.* http://www.cadth.ca/resources/grey-matters.
35. Adab N, Tudur SC, Vinten J, Williamson P, Winterbottom J: **Common antiepileptic drugs in pregnancy in women with epilepsy.** *Cochrane Database Syst Rev* 2004, **3**:CD004848.
36. Banach R, Boskovic R, Einarson T, Koren G: **Long-term developmental outcome of children of women with epilepsy, unexposed or exposed prenatally to antiepileptic drugs: a meta-analysis of cohort studies.** *Drug Saf* 2010, **33**(1):73–79.
37. Landis JR, Koch GG: **The measurement of observer agreement for categorical data.** *Biometrics* 1977, **33**(1):159–174.
38. *Synthesi.SR.* http://knowledgetranslation.ca/sysrev/login.php.
39. Stone PW: **Popping the (PICO) question in research and evidence-based practice.** *Appl Nurs Res* 2002, **15**(3):197–198.
40. *Cochrane Effective Practice and Organization of Care Group Draft Risk of Bias Tool.* http://epoc.cochrane.org/epoc-author-resources.
41. *The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-Analyses.* http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
42. Santaguida PL, Raina P, Ismaila A: *The Development of the McHarm Quality Assessment Scale for Adverse Events.* Hamilton, Ontario: McMaster University; 2008.
43. Raudenbush SW: **Analyzing Effect Sizes: Random Effects Models.** In *The Handbook of Research Synthesis and Meta-analysis.* 2nd edition. Edited by Cooper H, Hedges LV, Valentine JC. New York: Russell Sage Foundation; 2009:295–315.
44. Viechtbauer W: **Bias and efficiency of meta-analytic variance estimators in the random-effects model.** *J Educ Behav Stat* 2005, **30**(3):261–293.
45. Higgins JP, Thompson SG: **Quantifying heterogeneity in a meta-analysis.** *Stat Med* 2002, **21**(11):1539–1558.
46. Viechtbauer W: **Confidence intervals for the amount of heterogeneity in meta-analysis.** *Stat Med* 2007, **26**(1):37–52.
47. Higgins JPT, Green S: *Cochrane Handbook for Systematic Reviews of Interventions.* The Cochrane Collaboration; 2009. http://www.cochrane.org/handbook.
48. Sweeting MJ, Sutton AJ, Lambert PC: **What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data.** *Stat Med* 2004, **23**(9):1351–1375.
49. Bradburn MJ, Deeks JJ, Berlin JA, Russell Localio A: **Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events.** *Stat Med* 2007, **26**(1):53–77.
50. Littell JH, Corcoran J, Pillai V: *Systematic Reviews and Meta-Analysis.* New York: Oxford University Press; 2008.
51. Carpenter J, Rucker G, Schwarzer G: **Assessing the sensitivity of meta-analysis to selection bias: a multiple imputation approach.** *Biometrics* 2011, **67**(3):1066–1072.
52. *Conducting Meta-Analyses in R with the metafor Package.* http://www.jstatsoft.org/v36/i03/.
53. Salanti G, Ades AE, Ioannidis JP: **Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial.** *J Clin Epidemiol* 2011, **64**(2):163–171.
54. Salanti G: **Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool.** *Res Synth Methods* 2012, **3**(2):80–97.
55. Song F, Altman DG, Glenny AM, Deeks JJ: **Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses.** *BMJ* 2003, **326**(7387):472.
56. Veroniki AA, Vasiliadis HS, Higgins JP, Salanti G: **Evaluation of inconsistency in networks of interventions.** *Int J Epidemiol* 2013, **42**(1):332–345.
57. Dias S, Welton NJ, Caldwell DM, Ades AE: **Checking consistency in mixed treatment comparison meta-analysis.** *Stat Med* 2010, **29**(7–8):932–944.
58. White IR, Barrett JK, Jackson D, Higgins JPT: **Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression.** *Res Synth Methods* 2012, **3**(2):111–125.
59. Jackson D, Barrett JK, Stephen R, White IR, Higgins JPT: **A design-by-treatment interaction model for network meta-analysis with random inconsistency effects.** *Stat Med* 2014, In press.
60. White IR: **Multivariate random-effects meta-regression: updates to mvmeta.** *Stata J* 2011, **11**(2):255–270.
61. Centers for Disease Control and Prevention: *Unintended Pregnancy Prevention.* http://www.cdc.gov/reproductivehealth/unintendedpregnancy/.
62. Yerby MS: **Pregnancy, teratogenesis, and epilepsy.** *Neurol Clin* 1994, **12**(4):749–771.

doi:10.1186/2046-4053-3-68
Cite this article as: Tricco *et al.*: Comparative safety of anti-epileptic drugs among infants and children exposed *in utero* or during breastfeeding: protocol for a systematic review and network meta-analysis. *Systematic Reviews* 2014 **3**:68.

PRISMA NMA Checklist

Section/Topic	Item #	Checklist Item*	Reported on Page #
TITLE			
Title	1	Identify the report as a systematic review *incorporating a network meta-analysis (or related form of meta-analysis).*	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: **Background:** main objectives **Methods:** data sources; study eligibility criteria, participants, and interventions; study appraisal; and *synthesis methods, such as network meta-analysis.* **Results:** number of studies and participants identified; summary estimates with corresponding confidence/credible intervals; *treatment rankings may also be discussed. Authors may choose to summarize pairwise comparisons against a chosen treatment included in their analyses for brevity.* **Discussion/Conclusions:** limitations; conclusions and implications of findings. **Other:** primary source of funding; systematic review registration number with registry name.	4-5
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known, *including mention of why a network meta-analysis has been conducted.*	7
Objectives	4	Provide an explicit statement of questions being addressed, with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	7
METHODS			
Protocol and registration	5	Indicate whether a review protocol exists and if and where it can be accessed (e.g., Web address); and, if available, provide registration information, including registration number.	8

Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale. *Clearly describe eligible treatments included in the treatment network, and note whether any have been clustered or merged into the same node (with justification).*	8
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	9
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Additional File 1
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	9
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	9
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	Additional File 1
Geometry of the network	**S1**	Describe methods used to explore the geometry of the treatment network under study and potential biases related to it. This should include how the evidence base has been graphically summarized for presentation, and what characteristics were compiled and used to describe the evidence base to readers.	10-12
Risk of bias within individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	9-10 (see also Appendix A)
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means). *Also describe the use of additional summary measures assessed, such as treatment rankings and surface under the cumulative ranking curve (SUCRA) values, as well as modified approaches used to present summary findings from meta-analyses.*	10-12

Planned methods of analysis	14	Describe the methods of handling data and combining results of studies for each network meta-analysis. This should include, but not be limited to: • *Handling of multi-arm trials;* • *Selection of variance structure;* • *Selection of prior distributions in Bayesian analyses; and* • *Assessment of model fit.*	10-12
Assessment of Inconsistency	**S2**	Describe the statistical methods used to evaluate the agreement of direct and indirect evidence in the treatment network(s) studied. Describe efforts taken to address its presence when found.	10-11
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	9-10
Additional analyses	16	Describe methods of additional analyses if done, indicating which were pre-specified. This may include, but not be limited to, the following: • Sensitivity or subgroup analyses; • Meta-regression analyses; • *Alternative formulations of the treatment network; and* • *Use of alternative prior distributions for Bayesian analyses (if applicable).*	11-12
RESULTS[†]			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	13 and Figure 1
Presentation of network structure	**S3**	Provide a network graph of the included studies to enable visualization of the geometry of the treatment network.	Figure 2
Summary of network geometry	**S4**	Provide a brief overview of characteristics of the treatment network. This may include commentary on the abundance of trials and randomized patients for the different interventions and pairwise comparisons in the network, gaps of evidence in the treatment network, and potential biases reflected by the network structure.	14-18
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Table 1, Appendices D and E

Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment.	Appendix F
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: 1) simple summary data for each intervention group, and 2) effect estimates and confidence intervals. *Modified approaches may be needed to deal with information from larger networks.*	N/A (data can be provided by the corresponding author)
Synthesis of results	21	Present results of each meta-analysis done, including confidence/credible intervals. *In larger networks, authors may focus on comparisons versus a particular comparator (e.g. placebo or standard care), with full findings presented in an appendix. League tables and forest plots may be considered to summarize pairwise comparisons.* If additional summary measures were explored (such as treatment rankings), these should also be presented.	15-18, Figure 3, Appendices H, I, J
Exploration for inconsistency	**S5**	Describe results from investigations of inconsistency. This may include such information as measures of model fit to compare consistency and inconsistency models, *P* values from statistical tests, or summary of inconsistency estimates from different parts of the treatment network.	14 (see also Appendix H)
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies for the evidence base being studied.	14 (see also Appendix G)
Results of additional analyses	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression analyses, *alternative network geometries studied, alternative choice of prior distributions for Bayesian analyses,* and so forth).	Appendix K
DISCUSSION			
Summary of evidence	24	Summarize the main findings, including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy-makers).	19-21
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review level (e.g., incomplete retrieval of identified research, reporting bias). *Comment on the validity of the assumptions, such as transitivity and consistency. Comment on any concerns regarding network geometry (e.g., avoidance of certain comparisons).*	21-23

Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	23
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review. This should also include information regarding whether funding has been received from manufacturers of treatments in the network and/or whether some of the authors are content experts with professional conflicts of interest that could affect use of treatments in the network.	26-27

Abbreviations: PICOS - population, intervention, comparators, outcomes, study design
* Text in italics indicates wording specific to reporting of network meta-analyses that has been added to guidance from the PRISMA statement.
† Authors may wish to plan for use of appendices to present all relevant information in full detail for items in this section.

Supplementary Online Content

Appendix A. Newcastle-Ottawa Scale scoring guide

COHORT Studies

Excel Column	NOS* Answer Options**	NOS Coding Manual*
RefID	Enter the report's RefID.	
DA	Enter your initials.	
First author	Enter the first author's last name.	
Year of publication	Enter the year of the publication.	
SELECTION:		
1) Representative-ness of the exposed cohort	a) truly representative of the average pregnant woman taking AEDs in the community b) somewhat representative of the average pregnant woman taking AEDs in the community c) selected group of users e.g., nurses, volunteers d) no description of the derivation of the cohort	Item is assessing the representativeness of exposed individuals in the community, not the representativeness of the sample of women from some general population. For example, subjects derived from groups likely to contain middle class, better educated, health oriented women are likely to be representative of postmenopausal estrogen users while they are not representative of all women (e.g. members of a health maintenance organisation (HMO) will be a representative sample of estrogen users. While the HMO may have an under-representation of ethnic groups, the poor, and poorly educated, these excluded groups are not the predominant users of estrogen). <u>Note:</u> Truly representative (A) is a population-based cohort at the provincial or national levels (e.g., a sample from 2 cities is not enough). We need very 'broad' sample of the population. Somewhat representative (B) includes private clinics, hospital-based, or

		community-based.
2) Selection of the non-exposed cohort	a) drawn from the same community as the exposed cohort b) drawn from a different source c) no description of the derivation of the non-exposed cohort	Note: In our review of mostly multi-arm studies, this question pertains to the study's comparator group(s) – including "active" controls (for example, a less teratogenic AED). Therefore, this will often be 'A' for our studies.
3) Ascertainment of exposure	a) secure record (e.g., surgical records) b) structured interview c) written self-report d) no description	Note: Option 'A' includes patient hospital records, prescription drug database, or hospital/clinic visits (e.g., patient is asked about "current" AED use during a visit with their doctor). Option 'B' includes a hospital/clinic visit, but the patients are asked to remember their AED use during pregnancy (e.g., retrospectively ascertained exposure). If a study used both medical records and interviews for everyone, select 'A'.
4) Demonstration that outcome of interest was not present at start of study	a) yes b) no	In the case of mortality studies, outcome of interest is still the presence of a disease/incident, rather than death. That is to say that a statement of 'no history of disease or incident' earns a star (i.e. option 'A'). Note: Since our review is on pregnant women, this question is 'A' for all. **Please email us if a study involves breastfeeding women.**
COMPARABILITY:		
1) Comparability of cohorts on the basis of the design or analysis	a) answer is BOTH B & C (i.e. study controls for age and one other important factor) b) study controls for age of the women	Either exposed and non-exposed individuals must be matched in the design and/or confounders must be adjusted for in the analysis. Statements of no differences between groups or that differences were not statistically significant are not sufficient for establishing comparability.

	c) study controls for any other important factor d) study does not control for any important factor or it is not described	Note: If the relative risk for the exposure of interest is adjusted for the confounders listed, then the groups will be considered to be comparable on each variable used in the adjustment. There may be multiple ratings for this item for different categories of exposure (e.g., ever vs. never, current vs. previous or never). [A maximum of 2 stars can be allotted in this category]. Note: The study should have initially matched the groups or presented adjusted odds ratios, AND in addition, since in our review we are analyzing each AED arm separately (instead of the whole exposed cohort), the study must also report the factor of interest **for 'each AED arm'** (or state that **'each AED arm'** is matched). **Thus, there are 2 parts to this question:** 1) The study should have matched/adjusted for age at whatever level of groups they were focused on (even if they aren't our abstracted AED arms); AND 2) Then the study should also have reported the age for each AED arm. If they haven't done both of these 2 things, it's a 'D' here (unless they happen to combine these by reporting adjusted ORs for each of our AED arms). For our review, this generally pertains to **the comparability of the MOTHERS.** The exception here is in studies of cognitive/psychomotor development disorders in children - when age of the children should be comparable. The "other important factors" here are any one of these:

		• history of congenital malformations (CMs), fetal losses, preterm deliveries or small babies. • family history of genetic problems or CMs. • alcohol use. • nutritional deficiencies (e.g., lack of folic acid). <u>Example:</u> - Option 'B' indicates that the study initially matched groups based on the women's age (or reported adjusted ORs) AND they report the mean women's age for EACH of our arms (e.g., for Tx1, Tx2, etc.).
OUTCOME:		
1) Assessment of outcome	a) independent OR blind assessment b) record linkage c) self-report d) no description	For some outcomes (e.g. fractured hip), reference to the medical record is sufficient to satisfy the requirement for confirmation of the fracture. This would not be adequate for vertebral fracture outcomes where reference to x-rays would be required. a) Independent or blind assessment stated in the paper, or confirmation of the outcome by reference to secure records (x-rays, medical records, etc.) b) Record linkage (e.g. identified through ICD codes on database records) c) Self-report (i.e. no reference to original medical records or x-rays to confirm the outcome) d) No description. <u>Note:</u> Blind (A) is if they tell us that the outcome assessors were blinded to exposures; or if the outcome is objective. For our purposes, we will focus on the primary outcome of interest of our systematic review, which is <u>major malformations (an objective outcome)</u>.

		So most of ours will be A, unless the study is only on a secondary outcome (e.g., cognitive development) and is based on the mother's self-report of their child (e.g., not a clinical examination).
2) Was follow-up long enough for outcomes to occur	a) yes b) no	An acceptable length of time should be decided before quality assessment begins (e.g. 5 yrs. for exposure to breast implants) <u>Note:</u> For this component, focus only on the outcomes that are reported in the results. For our purposes, we will focus on the primary outcome of interest of our systematic review, which is <u>major malformations</u>. • For studies focusing on 'birth' outcomes (i.e. malformations, preterm, fetal losses, born small), the answer is 'A' if they follow the groups until birth. • For studies focusing on cognitive developmental disorders, an adequate follow-up period (i.e. child's age) is 4 years. • For studies focusing on psychomotor delays, an adequate follow-up period is the earliest point of detection of the disorder. • For studies focusing on neonatal seizures, an adequate follow-up period (i.e. infant's age) is 6 months.
3) Adequacy of follow up of cohorts	a) complete follow up - all subjects accounted for b) subjects lost to follow up unlikely to introduce bias - small number lost (see 'Note'), or description provided of those lost c) follow up rate is inadequate (see 'Note') and no description of those lost d) no statement	This item assesses the follow-up of the exposed and non-exposed cohorts to ensure that losses are not related to either the exposure or the outcome. <u>Note:</u> **<u>Especially check ones that start their total sample size (or figure diagram) with only the ones who had "complete" data (or only those who they had "successfully" recruited), as these are often a 'D' (since they don't report on the ones NOT followed up).</u>** • For a prospective study, ≥90% follow-up rate per year is adequate (e.g., 10% dropout or less for 1 year, 20% for 2 years of follow-up, etc.). This includes missing or incomplete data, etc.

		• For a retrospective cohort study, ≥80% follow-up rate is adequate; including the ones that they could NOT recruit or who would NOT participate. • For a survey/mail questionnaire, ≥75% response rate is adequate. (For a survey, a dropout rate is congruent to a survey response rate).

CASE-CONTROL Studies

Excel Column	NOS* Answer Options**	NOS Coding Manual*
RefID	Enter the report's RefID.	
DA	Enter your initials.	
First author	Enter the first author's last name.	
Year of publication	Enter the year of the publication.	
SELECTION:		
1) Is the case definition adequate?	a) yes, with independent validation b) yes, e.g., record linkage or based on self-reports c) no description	a) Requires some independent validation (e.g. >1 person/record/time/ process to extract information, or reference to primary record source such as x-rays or medical/hospital records) b) Record linkage (e.g. ICD codes in database) or self-report with no reference to primary record c) No description Note: This question is assessing the group of infants that have the outcome of interest (e.g., CMs) – i.e. the "cases" in a case-control study design.
2) Representative-ness of the cases	a) consecutive or obviously representative series of cases b) potential for selection biases, or not stated	a) All eligible cases with outcome of interest over a defined period of time, all cases in a defined catchment area, all cases in a defined hospital or clinic, group of hospitals, health maintenance organisation, or an appropriate sample of those cases (e.g. random sample) b) Not satisfying requirements in part (a), or not stated.

		Note: Option 'A' is a population-based sample.
3) Selection of controls	a) community controls b) hospital controls c) no description	This item assesses whether the control series used in the study is derived from the same population as the cases and essentially would have been cases had the outcome been present. a) Community controls (i.e. same community as cases and would be cases if had outcome) b) Hospital controls, within same community as cases (i.e. not another city) but derived from a hospitalised population c) No description Note: This question is assessing the group of infants that don't have the outcome (e.g., CMs) – i.e. the "controls" in a case-control study design. Community controls (A) includes a population-based sample.
4) Definition of controls	a) no history of disease (endpoint) b) no description of source	a) If cases are first occurrence of outcome, then it must explicitly state that controls have no history of this outcome. If cases have new (not necessarily first) occurrence of outcome, then controls with previous occurrences of outcome of interest should not be excluded. b) No mention of history of outcome Note: Since our review is on fetal effects, this question is 'A' for all studies. **Please email us if a study involves exposure during breastfeeding.**
COMPARABILITY:		
1) Comparability of cases and controls on the basis of the design	a) answer is BOTH B & C (i.e. study controls for age and one other important factor) b) study controls for age of the	Either cases and controls must be matched in the design and/or confounders must be adjusted for in the analysis. Statements of no differences between groups or that differences were not statistically significant are not sufficient for establishing comparability.

or analysis	women c) study controls for any other important factor d) study does not control for any important factor or it is not described	Note: If the odds ratio for the exposure of interest is adjusted for the confounders listed, then the groups will be considered to be comparable on each variable used in the adjustment. There may be multiple ratings for this item for different categories of exposure (e.g. ever vs. never, current vs. previous or never). [A maximum of 2 stars can be allotted in this category]. <u>Note:</u> The study should have initially matched the groups, AND in addition, since in our review we are analyzing each AED arm separately (instead of the whole cases group), the study must also report the factor of interest **for 'each AED arm'** (or state that **'each AED arm'** is matched). For our review, this generally pertains to **the comparability of the MOTHERS of the cases and controls.** The exception here is in studies of cognitive/psychomotor development disorders in children - when age of the children should be comparable. The "other important factors" here are any one of these: • history of congenital malformations (CMs), fetal losses, preterm deliveries or small babies. • family history of genetic problems or CMs. • alcohol use. • nutritional deficiencies (e.g., lack of folic acid). For example, Option 'B' indicates that the study initially matched groups based on the women's age AND they report the mean women's age for EACH arm (e.g., for Tx1, Tx2, etc.).
EXPOSURE:		

1) Assessment of exposure	a) secure record (e.g., surgical records) b) structured interview where blind to case/control status c) interview not blinded to case/control status d) written self-report or medical record only e) no description	Note: Option 'A' includes patient hospital records, prescription drug database, or hospital/clinic visits (e.g., patient is asked about "current" AED use during a visit with their doctor). "Interview" here includes a hospital/clinic visit, but the patients are asked to remember their AED use during pregnancy (e.g., retrospectively ascertained exposure).
2) Same method of ascertainment for cases and controls	a) yes b) no	Note: This question is asking whether the method of ascertainment of exposure was the same for 'cases' (with the outcome) and 'controls' (without the outcome; in this case-control study design).
3) Non-response rate	a) same rate for both groups b) non-respondents described c) rate different and no designation	Note: For our review, this pertains to either the infants or the mothers of the case and control groups. We're allowing 10% dropout per year for a prospective study – e.g., 10% for 1 year, 20% for 2 years of follow-up, etc. For a survey, we allow for a 75% response rate in order for it be adequate. For a survey, a dropout rate is congruent to a survey response rate.

*Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analyses. Available at: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

In the **"NOS Coding Manual" column, the first section for each item is copied straight from the NOS documentation while the lower portions in each item are our "Notes" tailored for the AED review.

Appendix B. List of included studies

A total of 29 cohort studies[1-29] with 9 companion reports[30-38] were included

1. Adab N, Kini U, Vinten J, et al. The longer term outcome of children born to mothers with epilepsy. *J Neurol Neurosurg Psychiatry*. 2004;75(11):1575-83.
2. Arkilo D, Hanna J, Dickens D, et al. Pregnancy and neurodevelopmental outcomes with in-utero antiepileptic agent exposure. A pilot study. *Eur J Paediatr Neurol*. 2015;19(1):37-40.
3. Bromley R, Baxter N, Calderbank R, Mawer G, Clayton-Smith J, Baker G. A comprehensive review of the language abilities of children exposed to valproate or carbamazepine in utero. American Epilepsy Society; Texas2010.
4. Bromley RL, Calderbank R, Cheyne CP, et al. Cognition in school-age children exposed to levetiracetam, topiramate, or sodium valproate. *Neurology*. 2016;87(18):1943-53.
5. Bromley RL, Mawer GE, Briggs M, et al. The prevalence of neurodevelopmental disorders in children prenatally exposed to antiepileptic drugs. *J Neurol Neurosurg Psychiatry*. 2013;84(6):637-43.
6. Christensen J, Gronborg TK, Sorensen MJ, et al. Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. *JAMA*. 2013;309(16):1696-703.
7. Cohen MJ, Meador KJ, Browning N, et al. Fetal antiepileptic drug exposure: Adaptive and emotional/behavioral functioning at age 6years. *Epilepsy Behav*. 2013;29(2):308-15.
8. Cummings C, Stewart M, Stevenson M, Morrow J, Nelson J. Neurodevelopment of children exposed in utero to lamotrigine, sodium valproate and carbamazepine. *Arch Dis Child*. 2011;96(7):643-7.
9. Dean JCS, Hailey H, Moore SJ, Lloyd DJ, Turnpenny PD, Little J. Long term health and neurodevelopment in children exposed to antiepileptic drugs before birth. *J Med Genet*. 2002;39(4):251-9.
10. D'Souza SW, Robertson IG, Donnai D, Mawer G. Fetal phenytoin exposure, hypoplastic nails, and jitteriness. *Arch Dis Child*. 1991;66(3):320-4.
11. Eriksson K, Viinikainen K, Mönkkönen A, et al. Children exposed to valproate in utero—Population based evaluation of risks and confounding factors for long-term neurocognitive development. *Epilepsy Res*. 2005;65(3):189-200.
12. Gaily E. Development and growth in children of epileptic mothers: a prospective controlled study. Helsinki, Finland: University of Helsinki; 1990.
13. Gogatishvili N, Ediberidze T, Lomidze G, Tatishvili N, Kasradze S. PO-0834 Long-term Developmental Outcome Of Children Prenatally Exposed To Antiepileptic Drugs. *Arch Dis Child*. 2014;99(Suppl 2):A526.
14. Gogatishvili N, Ediberidze T, Lomidze G, Tatishvili N, Kasradze S. Cognitive outcomes of children with fetal antiepileptic drug exposure at the age of 3-6 years-preliminary data. 1st Congress of the European Academy of Neurology; Berlin: European Journal of Neurology; 2015. p. 329.
15. Hurault-Delarue C, Damase-Michel C, Finotto L, et al. Psychomotor developmental effects of prenatal exposure to psychotropic drugs: a study in EFEMERIS database. *Fundam Clin Pharmacol*. 2016;30(5):476-82.
16. Jones KL, Lacro RV, Johnson KA, Adams J. Pattern of malformations in the children of women treated with carbamazepine during pregnancy. *N Engl J Med*. 1989;320(25):1661-6.

17. Katz JM, Pacia SV, Devinsky O. Current Management of Epilepsy and Pregnancy: Fetal Outcome, Congenital Malformations, and Developmental Delay. *Epilepsy Behav*. 2001;2(2):119-23.
18. Koch S, Jager-Roman E, Losche G, Nau H, Rating D, Helge H. Antiepileptic drug treatment in pregnancy: drug side effects in the neonate and neurological outcome. *Acta Paediatr*. 1996;85(6):739-46.
19. Mawer G, Clayton-Smith J, Coyle H, Kini U. Outcome of pregnancy in women attending an outpatient epilepsy clinic: adverse features associated with higher doses of sodium valproate. *Seizure*. 2002;11(8):512-8.
20. Miskov S, Juraski RG, Fucic A, et al. Croatian Pregnant Women with Epilepsy and Effects of Antiepileptic Drugs Exposure in their Offspring - seven years of prospective surveillance. American Epilepsy Society; Texas2010.
21. Miskov S, Juraski RG, Mikula I, et al. The Croatian model of integrative prospective management of epilepsy and pregnancy. *Acta Clin Croat*. 2016;55(4):535-48.
22. Nadebaum C, Anderson VA, Vajda F, Reutens DC, Barton S, Wood AG. Language skills of school-aged children prenatally exposed to antiepileptic drugs. *Neurology*. 2011;76(8):719-26.
23. Rihtman T, Parush S, Ornoy A. Developmental outcomes at preschool age after fetal exposure to valproic acid and lamotrigine: cognitive, motor, sensory and behavioral function. *Reprod Toxicol*. 2013;41:115-25.
24. Scolnik D, Nulman I, Rovet J, et al. Neurodevelopment of children exposed in utero to phenytoin and carbamazepine monotherapy. *JAMA*. 1994;271(10):767-70.
25. Shankaran S, Woldt E, Nelson J, Bedard M, Delaney-Black V. Antenatal phenobarbital therapy and neonatal outcome. II: Neurodevelopmental outcome at 36 months. *Pediatrics*. 1996;97(5):649-52.
26. van der Pol MC, Hadders-Algra M, Huisjes HJ, Touwen BC. Antiepileptic medication in pregnancy: late effects on the children's central nervous system development. *Am J Obstet Gynecol*. 1991;164(1 Pt 1):121-8.
27. Veiby G, Daltveit AK, Schjolberg S, et al. Exposure to antiepileptic drugs in utero and child development: a prospective population-based study. *Epilepsia*. 2013;54(8):1462-72.
28. Veiby G, Engelsen BA, Gilhus NE. Early child development and exposure to antiepileptic drugs prenatally and through breastfeeding: a prospective cohort study on children of women with epilepsy. *JAMA Neurol*. 2013;70(11):1367-74.
29. Wood AG, Nadebaum C, Anderson V, et al. Prospective assessment of autism traits in children exposed to antiepileptic drugs during pregnancy. *Epilepsia*. 2015;56(7):1047-55.
30. Bromley RL, Mawer G, Clayton-Smith J, Baker GA. Autism spectrum disorders following in utero exposure to antiepileptic drugs. *Neurology*. 2008;71(23):1923-4.
31. Gaily EK, Granstrom ML, Hiilesmaa VK, Bardy AH. Head circumference in children of epileptic mothers: contributions of drug exposure and genetic background. *Epilepsy Res*. 1990;5(3):217-22.
32. Hiilesmaa V. A prospective study on maternal and fetal outcome in 139 women with epilepsy. Helsinki: University of Helsinki; 1982.
33. Hiilesmaa VK, Bardy A, Teramo K. Obstetric outcome in women with epilepsy. *Am J Obstet Gynecol*. 1985;152(5):499-504.
34. Rasalam AD, Hailey H, Williams JH, et al. Characteristics of fetal anticonvulsant syndrome associated autistic disorder. *Dev Med Child Neurol*. 2005;47(8):551-5.

35. Tomson T, Battino D, Bonizzoni E, et al. Antiepileptic drugs and intrauterine death: A prospective observational study from EURAP. *Neurology*. 2015;85(7):580-8.
36. Viinikainen K, Eriksson K, Monkkonen A, et al. The effects of valproate exposure in utero on behavior and the need for educational support in school-aged children. *Epilepsy Behav*. 2006;9(4):636-40.
37. Vinten J, Adab N, Kini U, Gorry J, Gregg J, Baker GA. Neuropsychological effects of exposure to anticonvulsant medication in utero. *Neurology*. 2005;64(6):949-54.
38. Vinten J, Bromley RL, Taylor J, Adab N, Kini U, Baker GA. The behavioral consequences of exposure to antiepileptic drugs in utero. *Epilepsy Behav*. 2009;14(1):197-201.

Appendix C. Key excluded studies

Author, Year	Research Group	Title	Reason for Exclusion
Meador, 2009[39]	Neurodevelopmental Effects of Antiepileptic Drug (NEAD) Study Group	Cognitive Function at 3 Years of Age after Fetal Exposure to Antiepileptic Drugs	Outcomes only reported as continuous variables
Meador, 2010[40]	Neurodevelopmental Effects of Antiepileptic Drug (NEAD) Study Group	Effects of breastfeeding in children of women taking antiepileptic drugs	Outcomes only reported as continuous variables
Meador, 2011[41]	Neurodevelopmental Effects of Antiepileptic Drug (NEAD) Study Group	Foetal antiepileptic drug exposure and verbal versus non-verbal abilities at three years of age	Outcomes only reported as continuous variables
Meador, 2012[42]	Neurodevelopmental Effects of Antiepileptic Drug (NEAD) Study Group	Effects of fetal antiepileptic drug exposure: Outcomes at age 4.5 years	Outcomes only reported as continuous variables
Meador, 2013[43]	Neurodevelopmental Effects of Antiepileptic Drug (NEAD) Study Group	Fetal antiepileptic drug exposure and cognitive outcomes at age 6 years (NEAD study): a prospective observational study	Outcomes only reported as continuous variables
Shallcross, 2011[44]	Liverpool and Manchester Neurodevelopment Group and The UK Epilepsy and Pregnancy Register	Child development following in utero exposure: Levetiracetam vs. sodium valproate	Outcomes only reported as continuous variables
Shallcross, 2014[45]	Liverpool and Manchester	In utero exposure to levetiracetam vs. valproate: Development and language at 3 years of age	Outcomes only reported as continuous variables

	Neurodevelopment Group and The UK Epilepsy and Pregnancy Register		

References

39. Meador KJ, Baker GA, Browning N, et al. Cognitive function at 3 years of age after fetal exposure to antiepileptic drugs. *N Engl J Med*. 2009;360(16):1597-605.
40. Meador KJ, Baker GA, Browning N, et al. Effects of breastfeeding in children of women taking antiepileptic drugs. *Neurology*. 2010;75(22):1954-60.
41. Meador KJ, Baker GA, Browning N, et al. Foetal antiepileptic drug exposure and verbal versus non-verbal abilities at three years of age. *Brain*. 2011;134(Pt 2):396-404.
42. Meador KJ, Baker GA, Browning N, et al. Effects of fetal antiepileptic drug exposure: outcomes at age 4.5 years. *Neurology*. 2012;78(16):1207-14.
43. Meador KJ, Baker GA, Browning N, et al. Fetal antiepileptic drug exposure and cognitive outcomes at age 6 years (NEAD study): a prospective observational study. *Lancet Neurol*. 2013;12(3):244-52.
44. Shallcross R, Bromley RL, Irwin B, Bonnett LJ, Morrow J, Baker GA. Child development following in utero exposure: levetiracetam vs sodium valproate. *Neurology*. 2011;76(4):383-9.
45. Shallcross R, Bromley RL, Cheyne CP, et al. In utero exposure to levetiracetam vs valproate: development and language at 3 years of age. *Neurology*. 2014;82(3):213-21.
46. Cohen MJ, Meador KJ, Browning N, et al. Fetal antiepileptic drug exposure: Adaptive and emotional/behavioral functioning at age 6 years. *Epilepsy Behav*. 2013;29(2):308-15.

Appendix D. Table of Individual Study characteristics

Author, Year	Country of conduct	Registry or Setting	Study period	Interventions	Outcomes	Funding
Adab, 2004[*1] [CR: Vinten 2005[37]Vinten, 2009[38]]	UK	Mersey Regional Epilepsy Clinic; Epilepsy Clinic at the Manchester Royal Infirmary; Antenatal clinic at St Mary's Hospital, Manchester	2000-2001	Carbam, Control, Valpro	Cognitive Developmental Delay, Psychomotor Developmental Delay	NR
Arkilo, 2015[2]	USA	Minnesota Epilepsy Group	2006-2011	Carbam, Lamot, Levet, Pheny, Valpro	Autism/Dyspraxia, Psychomotor Developmental Delay	NR
Bromley, 2010[3]	UK	Liverpool and Manchester Neurodevelopment Group	NR	Carbam, Valpro	Language Delay	NR
Bromley, 2013[5] [CR: Bromley, 2008[30]]	UK	Liverpool and Manchester Neurodevelopment group	2000-2004	Carbam, Control, Lamot, Valpro	Autism/Dyspraxia, ADHD	mixed public & private
Bromley, 2016[4]†	UK	UK Epilepsy and Pregnancy Register	2004-2007	Control, Gabap, Levet, Topir, Valpro	Cognitive Developmental Delay	public
Christensen, 2013[6]†	Denmark	Danish Civil Registration System; Danish Prescription Register; Danish Psychiatric Central Register; Danish	1996-2006	Carbam, Clonaz, Lamot, Oxcar, Valpro	Autism/Dyspraxia	public

		Birth Register; Danish National Hospital Register				
Cohen, 2013[46]	USA;UK	Neurodevelopmental Effects of Antiepileptic Drugs Study Group	1999-2004	Carbam, Lamot, Pheny, Valpro,	ADHD	public
Cummings, 2011[8]† [CR: Tomson, 2015[35]]	Northern Ireland	UK Epilepsy and Pregnancy Register (Northern Ireland); Northern Ireland Child Health System	1996-2005	Carbam, Lamot, Valpro,	Cognitive Developmental Delay	public
Dean, 2002[9] [CR: Rasalam, 2005[34]]	Scotland	Aberdeen Maternity Hospital	1976-2000	Carbam, Carbam+Pheno, Carbam+Pheny, Carbam+Valpro, Control, Ethos, Pheno, Pheno+Pheny, Pheno+Valpro, Pheny, Primid, Valpro	Psychomotor Developmental Delay, ADHD	NR
D'Souza, 1991[10]	United Kingdom	St Mary's Hospital	1980-1982	Carbam, Control, Pheno, Pheny, Valpro	Cognitive Developmental Delay	public
Eriksson, 2005[11]† [CR: Viinikainen, 2006[36]]	Finland	Kuopio University Hospital	1989-2000	Carbam, Control, Valpro	Cognitive Developmental Delay, Psychomotor Developmental Delay	public

Gaily, 1990[12] [CR: Gaily, 1990[31]; Hiilesmaa, 1982[32]; Hiilesmaa, 1985[33]]	Finland	Helsinki University Central Hospital	1975-1979	Carbam, Carbam+Pheno+Pheny, Carbam+Pheny, Carbam+Valpro, Control, Ethos+Pheny, Pheno+Pheny, Pheny, Pheny+Primid, Pheny+Valpro	Cognitive Developmental Delay , Psychomotor Developmental Delay	mixed public & private
Gogatishvili, 2014[13]	Georgia	Georgian National AED-Pregnancy Registry	NR	Carbam, Lamot, Valpro	Cognitive Developmental Delay	public
Gogatishvili, 2015[14]	Georgia	Georgian National AED-Pregnancy Registry	NR	Carbam, Carbam+Levet, Lamot, Pheno, Valpro	Language Delay	public
Hurault-Delarue, 2012[15]	France	EFEMERIS database - Caisse Primaire d'Assurance Maladie of Haute-Garonne and Maternal and Infant Protection Service; Antenatal Diagnostic Centre	2004-2008	Carbam, Clobaz, Clonaz, Gabap, Lamot, Pheno, Topir, Valpro	Psychomotor Developmental Delay	NR
Jones, 1989[16]†	US	California Teratogen Registry	1979-1988	Carbam, Carbam+Pheno, Carbam+Pheno+Valpro, Carbam+Primid	Cognitive Developmental Delay , Psychomotor Developmental Delay	public
Katz, 2001[17]	USA	Mount Sinai Comprehensive Epilepsy Center	1990-2000	Carbam, Control, Lamot, Pheno, Pheny, Primid, Valpro	Cognitive Developmental Delay	NR

Koch, 1996[18]	Germany	NR	1976-1983	Pheno, Pheny, Primid, Valpro	Cognitive Developmental Delay	public
Mawer, 2002[19]	England	Manchester Royal Infirmary	1990-1999	Carbam, Lamot, Pheny, Valpro	Cognitive Developmental Delay	NR
Miskov, 2010[20]	Croatia	NR	2003-2010	Carbam, Control, Gabap, Lamot, Valpro	Psychomotor Developmental Delay, Neonatal Seizures	NR
Miskov, 2016[21]	Croatia	Sestre milosrdnice University Hospital Center	2003-2013	Carbam, Carbam+Lamot, Carbam+Pheno, Carbam+Pheny+Topir, Control, Clonaz+Valpro, Gabap, Lamot, Oxcar, Pheno, Pheny, Topir+Valpro, Valpro	Attention Deficit Hyperactivity Disorder	NR
Nadebaum, 2011[22]†	Australia	Australian Registry of Antiepileptic Drug Use in Pregnancy	2007-2009	Carbam, Lamot, Valpro	Language Delay	mixed public & private
Rihtman, 2013[23]	Israel	Israeli Teratogen Information Service	NR	Lamot, Valpro	Neonatal Seizure	mixed public & private
Scolnik, 1994[24]	Canada	Hospital for Sick Children - Motherisk Program; North York General Hospital; Toronto Hospital; Oshawa General Hospital	1987-1992	Carbam, Pheny	Cognitive Developmental Delay	public

Shankaran, 1996[25]	USA	Children's Hospital of Michigan	NR	Control, Pheno	Psychomotor Developmental Delay, Language Delay	public
Van der Pol, 1991[26]	Netherlands	Groningen University Hospital	1973-1981	Carbam, Carbam+Pheno, Control, Pheno	Psychomotor Developmental Delay	public
Veiby, 2013a[27]†	Norway	Norwegian Institute of Public Health- Mother and Child Cohort Study	1999-2009	Carbam, Control, Lamot, Valpro	Social Impairment	public
Veiby, 2013b[28]†	Norway	Medical Birth Registry of Norway	1999-2008	Carbam, Control, Lamot, Valpro	Psychomotor Developmental Delay, Autism/Dyspraxia, Language Delay, ADHD	public
Wood, 2015[29]†	Australia	Australian Registry of Antiepileptic Drug Use in Pregnancy	2007-2010	Carbam, Carbam+Clonaz, Carbam+Lamot, Carbam+Pheny, Lamot+Valpro, Valpro	Autism/Dyspraxia	public

Abbreviations: ADHD – Attention Deficit Hyperactivity Disorder; NR – Not Reported

Carbam = Carbamazepine; Clobaz = Clobazam; Clonaz = Clonazepam; Ethos = Ethosuximide; Gabap = Gabapentin; Lamot = Lamotrigine; Levet = Levetiracetam; Oxcar = Oxcarbazepine; Pheno = Phenobarbital; Pheny = Phenytoin; Primid = Primidone; Topir = Topiramate; Valpro = Valproate; Vigab = Viagabatrin

*Single publication reporting on two separate cohorts
†Registry Studies

Appendix E. Table of Patient characteristics

Author, Year	Indication	Sample Size*	Mean Age (Women)	Mean Age (Children)/ Follow-up period†	AED Exposure Timing	Maternal Alcohol Use n/N‡	Maternal Tobacco Use n/N‡
Adab, 2004a[1]§ [CR: Vinten 2005[37]; Vinten, 2009[38]]	Epilepsy	177	26.1	9-10.5	NR	24/279‡	68/249‡
Adab, 2004b[1]§ [CR: Vinten 2005[37]; Vinten, 2009[38]]	Epilepsy	81	26.1	3-3.33	NR	24/279‡	68/249‡
Arkilo, 2015[2]	Epilepsy	59	NR	NA	First trimester	NR	NR
Bromley, 2010[3]	NR	60	NR	6-7	Whole pregnancy	NR	NR
Bromley, 2013[5] [CR: Bromley, 2008[30]]	Epilepsy	156	28	6	NR	28/156	42/156
Bromley, 2016[4]	Epilepsy	185	NR	NR	NR	31/185	35/185
Christensen, 2013[6]	NR	2011	NR	NR	Whole pregnancy	NR	NR
Cohen, 2013[46]	Epilepsy	108	30	6	During pregnancy and breastfeeding	12/192‡	NR
Cummings, 2011[8] [CR: Tomson, 2015[35]]	Epilepsy	142	NR	2-3	During pregnancy and breastfeeding	32/108‡	19/108‡
Dean, 2002[9] [CR: Rasalam, 2005[34]]	Epilepsy	287	27	3.75-15.5	First trimester	NR	NR
D'Souza, 1991[10]	Epilepsy	42	26.5	2.5-3.5	Whole pregnancy	NR	NR
Eriksson, 2005[11] [CR: Viinikainen, 2006[36]]	Epilepsy	39	28.2	NR	NR	NR	NR

Gaily, 1990[12] [CR: Gaily, 1990[31]; Hiilesmaa, 1982[32]; Hiilesmaa, 1985[33]	Epilepsy	134	27.8	5.5	First trimester	NR	NR
Gogatishvili, 2014[13]	NR	39	NR	2 to 4	NR	NR	NR
Gogatishvili, 2015[14]	NR	23	NR	3 to 6	NR	NR	NR
Hurault-Delarue, 2012[15]	NR	109	NR	0.75	NR	NR	NR
Jones, 1989[16]	Epilepsy	63	NR	NR	Whole pregnancy	NR	NR
Katz, 2001[17]	Epilepsy	51	31	NR	NR	NR	NR
Koch, 1996[18]	Epilepsy	40	NR	6	First trimester	NR	NR
Mawer, 2002[19]	Epilepsy	52	NR	NR	NR	NR	NR
Miskov, 2010[20]	Epilepsy	55	NR	NR	NR	NR	NR
Miskov, 2016[21]	Epilepsy	74	34	NR	NR	NR	6/74
Nadebaum, 2011[22]	Epilepsy	66	31.6	7.4	During pregnancy and breastfeeding	NR	5/66
Rihtman, 2013[23]	Epilepsy	72	NR	NR	During pregnancy and breastfeeding	NR	NR
Scolnik, 1994[24]	Epilepsy	75	NR	1.5-3	1st trimester	NR	NR
Shankaran, 1996[25]	NR	96	NR	NR	NR	NR	NR
Van der Pol, 1991[26]	Epilepsy	57	NR	6-13	NR	NR	NR
Veiby, 2013a[27]	Epilepsy	422	NR	0.5	During pregnancy and breastfeeding	NR	NR
Veiby, 2013b[28]	Epilepsy	248	28.9	3	NR	NR	68/726‡
Wood, 2015[29]	Epilepsy	77	NR	6-8	NR	NR	NR

Abbreviations: NA – Not applicable; NR – Not reported

* Sample size used for analysis; ineligible treatment arms (i.e. treatment arms with excluded drugs or unspecified polytherapy) are not included in the count
† The mean age for children/follow-up period data were only collected for outcomes related to cognitive and/or psychomotor development
‡ Total sample size is based on the number of women enrolled in the study; may differ from the sample size used for analysis
§ Single publication reporting on two separate cohorts

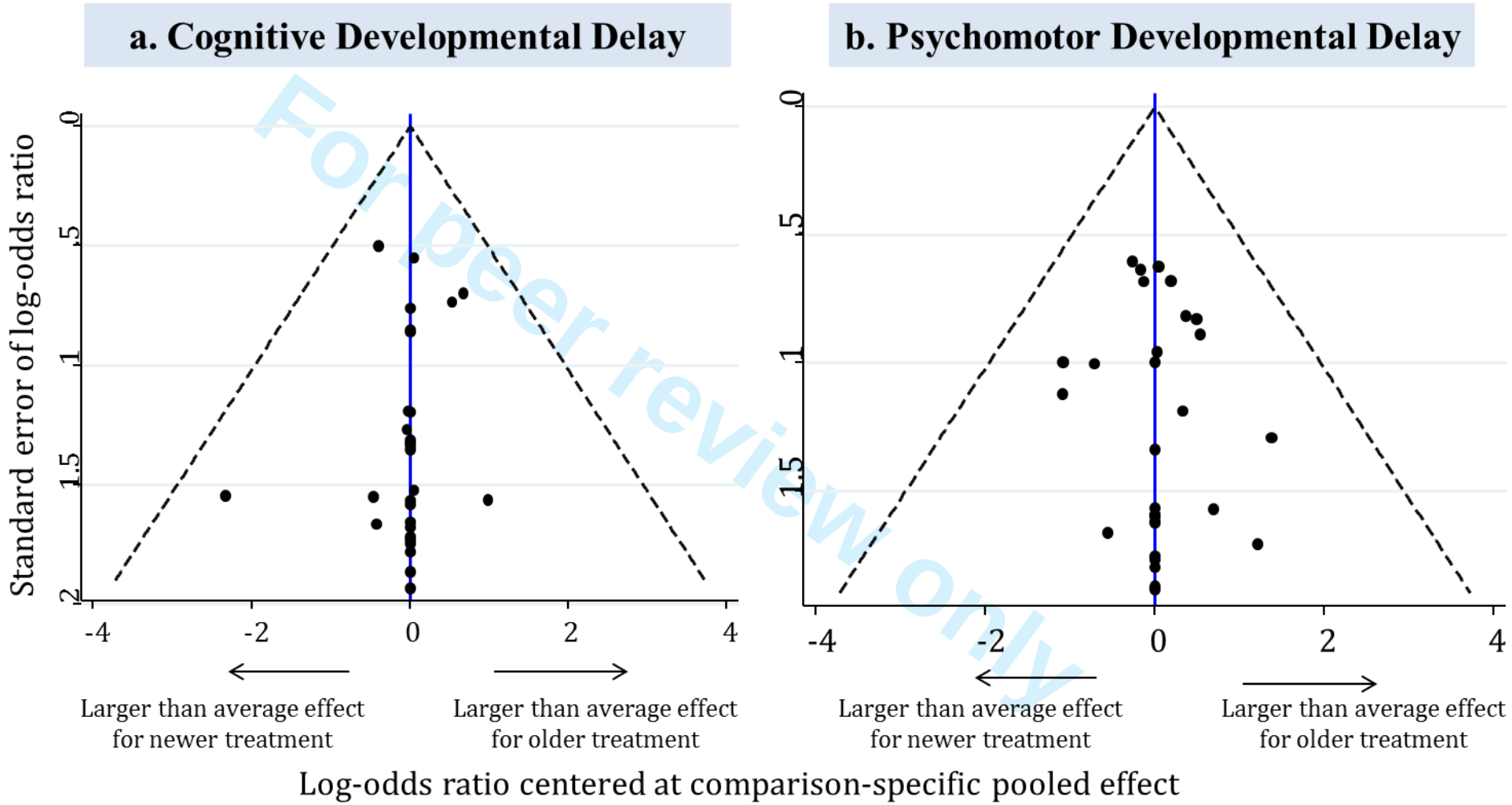
Appendix F. Methodological quality of observational studies – Newcastle Ottawa Scale results

First Author, Year	Representativeness of the exposed cohort	Selection of the non-exposed cohort	Ascertainment of exposure	Demonstration that outcome of interest was not present at start of study	Comparability of cohorts on the basis of the design or analysis	Assessment of outcome	Was follow-up long enough for outcomes to occur	Adequacy of follow up of cohorts
Adab, 2004[1]	B	A	A	A	C	A	A	C
Arkilo, 2015[2]	B	A	B	A	D	A	A	C
Bromley, 2010[3]	D	A	D	A	D	D	B	D
Bromley, 2013[5]	A	A	A	A	A	A	A	C
Bromley, 2016[4]	A	A	A	A	A	A	A	C
Christensen, 2013[6]	A	A	A	A	A	B	A	B
Cohen, 2013[46]	A	A	D	A	A	A	A	C
Cummings, 2011[8]	A	A	A	A	A	A	A	C
Dean, 2002[9]	B	A	A	A	D	A	A	C
D'Souza, 1991[10]	B	A	A	A	D	A	A	A
Eriksson, 2005[11]	B	A	A	A	B	A	A	D
Gaily, 1990[12]	B	A	A	A	D	A	A	A
Gogatishvili, 2014[13]	A	A	D	A	D	A	A	D
Gogatishvili, 2015[14]	A	A	D	A	D	A	A	D

Hurault-Delarue, 2012[15]	A	A	A	A	A	A	A	A
Jones, 1989[16]	A	A	B	A	D	A	A	B
Katz, 2001[17]	B	A	A	A	D	A	A	D
Koch, 1996[18]	B	A	B	A	D	A	A	C
Mawer, 2002[19]	B	A	A	A	D	A	A	B
Miskov, 2010[20]	D	A	D	A	D	D	A	D
Miskov, 2016[21]	C	A	A	A	D	A	A	D
Nadebaum, 2011[22]	A	A	A	A	A	A	A	B
Rihtman, 2013[23]	A	B	A	A	A	A	A	C
Scolnik, 1994[24]	B	A	A	A	D	A	A	A
Shankaran, 1996[25]	B	A	A	A	D	A	A	B
Van der Pol, 1991[26]	B	A	D	A	A	A	A	B
Veiby, 2013a[27]	A	A	A	A	A	A	A	D
Veiby, 2013b[28]	A	A	A	A	A	A	A	C
Wood, 2015[29]	A	A	A	A	D	A	A	C

Abbreviations: A – low risk; B – moderate risk; C – high risk; D – unclear risk

Appendix G. Comparison-adjusted funnel plots*



* Funnel plots have been produced only for outcomes with ≥10 studies. For multi-arm studies we plot data points from each study-specific basic parameter (treatment comparisons with a study-specific common comparator)

Appendix H. Statistically significant network meta-analysis results along with meta-analysis results, transitivity, and inconsistency assessments

Treatment Comparison	Number of Studies (Mean Baseline Risk)	Number of patients (Mean Age)	Treatment Indication	Timing	Comparability of cohorts	Adequacy of follow up of cohorts	MA Odds Ratio (95% CrI)	NMA Odds Ratio (95% CrI) (95% PrI)
Cognitive Developmental Delay (10 studies, 748 patients, 14 treatments)								
Lamot vs Valpro	4 (NA)	140 (31.00)	Epilepsy	NR	H	H	0.17 (0.02-0.87)	0.13 (0.01-0.57) (0.01-0.75)
Valpro vs Control	4 (0.06)	267 (28.80)	Epilepsy	1st trimester	H	H	8.15 (3.19-22.33)	7.40 (3.00-18.46) (1.81-27.63)
Valpro vs Carbam	6 (NA)	310 (27.80)	Epilepsy	NR	H	L	3.32 (1.56-7.04)	3.54 (1.69-7.26) (0.95-12.32)
Valpro vs Pheno	3 (NA)	36 (27.80)	Epilepsy	1st trimester	H	L	4.25 (0.82-34.07)	5.59 (1.21-35.07) (0.93-45.99)
Valpro vs Pheny	3 (NA)	58 (31.00)	Epilepsy	1st trimester	H	L	3.12 (0.75-14.12)	2.88 (1.04-8.49) (0.69-12.62)
Common between-study variance across treatment comparisons							0.13 (0.00-0.97)	0.12 (0.00-1.15) (NA)
Residual deviance: 44.72 Data points: 47 DIC: 78.7								
Evaluation of consistency using the design-by-treatment interaction model					Chi-square test: 14.15 Degrees of Freedom: 17		P- value: 0.66 Heterogeneity: 0	

Treatment Comparison	Number of Studies (Mean Baseline Risk)	Number of patients (Mean Age)	Treatment Indication	Timing	Comparability of cohorts	Adequacy of follow up of cohorts	MA Odds Ratio (95% CrI)	NMA Odds Ratio (95% CrI) (95% PrI)
Autism Dyspraxia (5 studies, 2551 patients, 12 treatments)								
Lamot vs Control	2 (0.00)	254 (27.75)	Epilepsy	1st trimester	H	H	13.77 (2.06-188.00)	8.88 (1.29-112.00) (0.94-146.80)
Lamot+Valpro vs Carbam	1 (NA)	40 (NR)	Epilepsy	NR	L	L	15.02 (2.04-171.90)	22.89 (2.58-219.00) (1.90-282.20)
Lamot+Valpro vs Clonaz	NA	NR	NR	NR	NR	NR	NA	20.21 (1.48-351.30) (1.15-455.00)
Lamot+Valpro vs Control	NA	NR	NR	NR	NR	NR	NA	132.70 (7.41-3.9 x 10^3) (5.82-4.6 x 10^3)
Lamot+Valpro vs Lamot	NA	NR	NR	NR	NR	NR	NA	14.61 (1.51-149.10) (1.14-196.80)
Oxcar vs Control	NA	NR	NR	NR	NR	NR	NA	13.51 (1.28-221.40) (0.86-267.40)
Valpro vs Carbam	5 (NA)	1003 (27.83)	Epilepsy	1st trimester	L	L	3.20 (1.20-8.68)	3.02 (1.09-8.40) (0.57-14.31)
Valpro vs Control	2 (0.00)	249 (27.75)	Epilepsy	1st trimester	H	H	9.19 (1.14-132.10)	17.29 (2.40-217.60) (1.61-274.90)
Common between-study variance across treatment comparisons							0.12 (0.00-1.37)	0.16 (0.00-1.95) (NA)
Residual deviance: 24 Data points: 24 DIC: 44								
Evaluation of consistency using the design-by-treatment interaction model				Chi-square test: 3.79 Degrees of Freedom: 5			P- value: 0.57 Heterogeneity: 0	

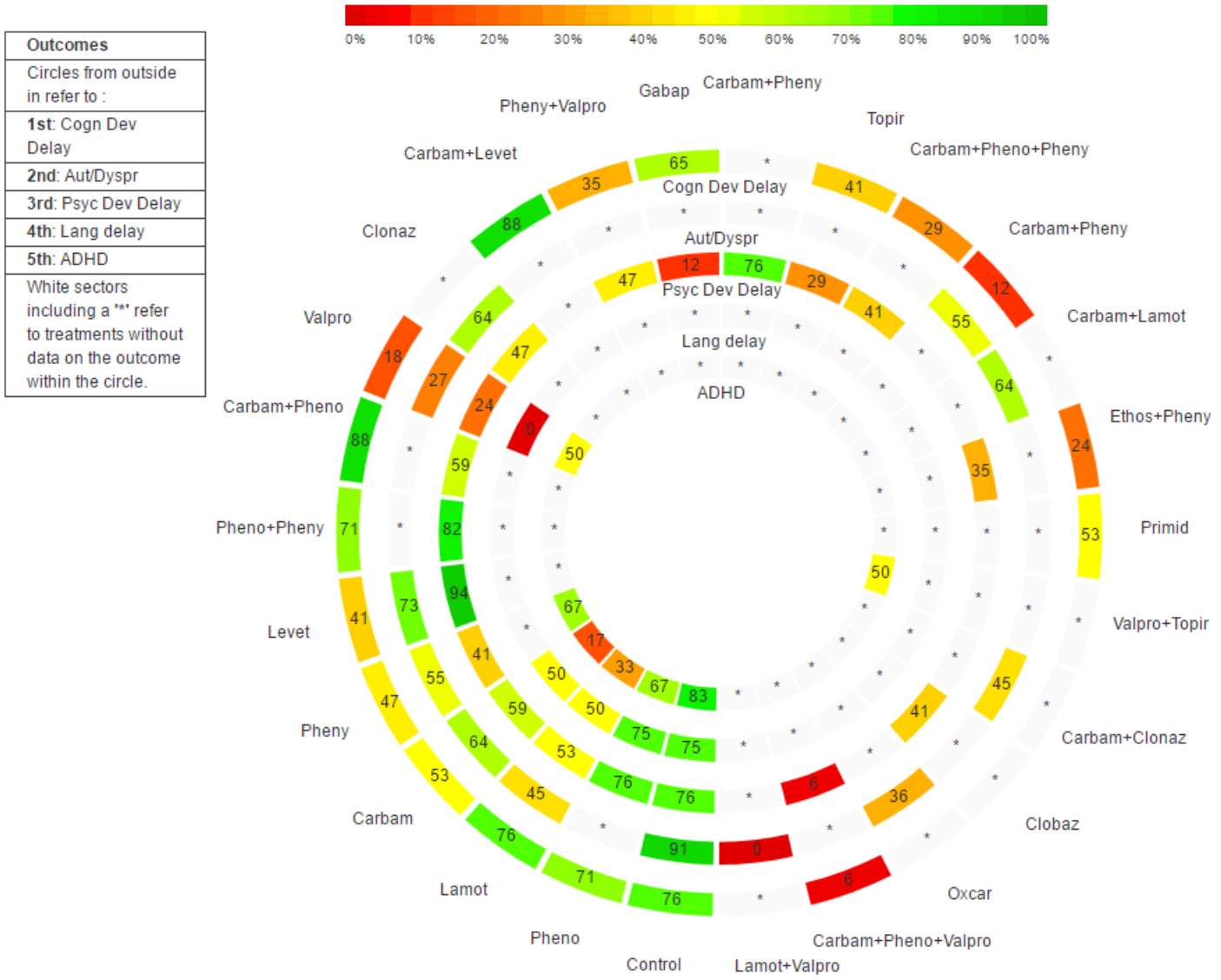
Treatment Comparison	Number of Studies (Mean Baseline Risk)	Number of patients (Mean Age)	Treatment Indication	Timing	Comparability of cohorts	Adequacy of follow up of cohorts	MA Odds Ratio (95% CrI)	NMA Odds Ratio (95% CrI) (95% PrI)
Psychomotor Developmental Delay (11 studies, 1145 patients, 18 treatments)								
Carbam+Pheno+Valpro vs Control	NA	NR	NR	NR	NR	NR	NA	19.12 (1.49-337.50) (1.34-370.40)
Carbam+Pheno+Valpro vs Pheno	NA	NR	NR	NR	NR	NR	NA	19.86 (1.38-393.60) (1.26-423.30)
Levet vs Carbam+Pheno+Valpro	NA	NR	NR	NR	NR	NR	NA	0.01 (0.00-0.58) (0.00-0.62)
Valpro vs Carbam	7 (NA)	331 (27.80)	Epilepsy	1st trimester	H	H	2.72 (1.39-5.67)	2.45 (1.27-4.88) (0.95-6.77)
Valpro vs Control	5 (0.07)	331 (28.38)	Epilepsy	1st trimester	H	H	3.53 (1.60-8.64)	4.16 (2.04-8.75) (1.52-12.05)
Valpro vs Pheno	2 (NA)	141 (NR)	Epilepsy	1st trimester	H	H	3.68 (1.17-12.30)	4.32 (1.72-11.20) (1.34-14.51)
Common between-study variance across treatment comparisons							0.05 (0.00-0.49)	0.06 (0.00-0.63) (NA)
Residual deviance: 45 Data points: 51 DIC: 78								
Evaluation of consistency using the design-by-treatment interaction model				Chi-square test: 13.46 Degrees of Freedom: 21			P- value: 0.89 Heterogeneity: 0	

Treatment Comparison	Number of Studies (Mean Baseline Risk)	Number of patients (Mean Age)	Treatment Indication	Timing	Comparability of cohorts	Adequacy of follow up of cohorts	MA Odds Ratio (95% CrI)	NMA Odds Ratio (95% CrI) (95% PrI)
Language Delay (5 studies, 509 patients, 5 treatments)								
Valpro vs Control	1 (0.03)	173 (28.90)	Epilepsy	NR	L	H	6.96 (1.14-37.03)	7.95 (1.50-49.13) (0.96-74.52)
Common between-study variance across treatment comparisons							0.15 (0.00-1.85)	0.16 (0.00-2.15) (NA)
Residual deviance: 12	Data points: 14	DIC: 23						
Evaluation of consistency using the design-by-treatment interaction model				Chi-square test: 2.33 Degrees of Freedom: 3			P- value: 0.50 Heterogeneity: 0	
ADHD (4 studies, 750 patients, 6 treatments)								
No statistically significant results								
Residual deviance: 12	Data points: 17	DIC: 22						

Abbreviations: ADHD - Attention Deficit Hyperactivity Disorder; CrI - Credible Interval; DIC - Deviance Information Criterion; H- high risk of bias; L - low risk of bias; MA - Meta-analysis; NA - Not applicable; NMA - Network Meta-analysis; NR- Not Reported; PrI - Predictive Interval

Carbam = Carbamazepine; Clobaz = Clobazam; Clonaz = Clonazepam; Ethos = Ethosuximide; Gabap = Gabapentin; Lamot = Lamotrigine; Levet = Levetiracetam; Oxcar = Oxcarbazepine; Pheno = Phenobarbital; Pheny = Phenytoin; Pridmid = Primidone; Topir = Topiramate; Valpro = Valproate; Vigab = Viagabatrin

Appendix I. Rank-heat plot of cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcomes*



Abbreviations: carbam - carbamazepine, clobaz - clobazam, clonaz - clonazepam, ethos - ethosuximide, gabap - gabapentin, lamot - lamotrigine, levet - levetiracetam, oxcar - oxcarbazepine, pheno - phenobarbital, pheny - phenytoin, primid - primidone, topir - topiramate, valpro - valproate, vigab - vigabatrin

*Rank-heat plot of cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay, and attention deficit hyperactivity disorder outcomes (5 circles) and 25 treatments (25 radii). Each sector is coloured according to the surface under the cumulative ranking curve value of the corresponding treatment and outcome using the transformation of three colours red (0%), yellow (50%), and green (100%).

Appendix J. Number of studies and treatments per outcome

Total studies	Range of study arms	# of treatments	# of patients	# of direct treatment comparisons	# of NMA treatment comparisons	Statistically significant NMA treatment effects	# of studies with zero events in all arms	# of studies with ineligible outcome definition*
Cognitive Developmental Delay								
11	(2,8)	18	933	62	153	5	1	5
Autism/Dyspraxia								
5	(4,6)	12	2551	34	66	8	0	4
Neonatal Seizure								
1	(2,2)	2	69	1	0	0	1	1
Psychomotor Developmental Delay								
11	(2,8)	18	1145	74	153	6	0	5
Language Delay								
5	(2,4)	5	509	7	10	1	0	3
ADHD								
5	(4,6)	7	816	20	21	0	0	0
Social Impairment								
1	(4,4)	4	422	1	0	0	0	0

Abbreviations: ADHD - Attention Deficit Hyperactivity Disorder; NMA - Network Meta-analysis

Appendix K. Sensitivity and network meta-regression analyses - Anti-epileptic drugs compared with Control

Treatment Comparison	NMA Odds Ratio	95% CrI	95% PrI
Cognitive Developmental Delay – Sensitivity Analysis - Epilepsy only (10 studies, 910 patients, 17 treatments)			
Carbamazepine vs Control	2.08	(0.79 - 5.82)	(0.47 - 9.34)
Carbamazepine+Phenobarbital vs Control	0.62	(0.00 - 15.31)	(0.00 - 19.29)
Carbamazepine+Phenobarbital+Phenytoin vs Control	4.75	(0.01 - 164.80)	(0.01 - 192.50)
Carbamazepine+Phenobarbital+Valproate vs Control	15.00	(1.00 - 367.10)	(0.82 - 426.90)
Carbamazepine+Phenytoin vs Control	9.84	(0.60 - 136.30)	(0.49 - 164.50)
Ethosuximide+Phenytoin vs Control	6.53	(0.02 - 216.00)	(0.02 - 251.30)
Gabapentin vs Control	1.43	(0.05 - 14.28)	(0.04 - 18.20)
Lamotrigine vs Control	0.79	(0.05 - 5.12)	(0.05 - 6.66)
Levetiracetam vs Control	3.46	(0.65 - 17.14)	(0.47 - 23.57)
Phenobarbital vs Control	0.55	(0.01 - 5.38)	(0.01 - 6.85)
Phenobarbital+Phenytoin vs Control	1.28	(0.00 - 36.18)	(0.00 - 44.03)
Phenytoin vs Control	2.47	(0.65 - 8.25)	(0.41 - 12.47)
Phenytoin+Valproate vs Control	3.68	(0.01 - 121.00)	(0.01 - 135.00)
Primidone vs Control	1.97	(0.25 - 12.16)	(0.19 - 16.25)
Topiramate vs Control	3.06	(0.42 - 17.51)	(0.32 - 23.57)
Valproate vs Control	7.48	(2.99 - 19.04)	(1.67 - 31.21)
Common within-network between-study variance	0.16	(0.00 - 1.36)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 12.98 Degrees of Freedom: 14	P-value: 0.53 Heterogeneity: 0.00
Cognitive Developmental Delay - Sensitivity Analysis - First generation AEDs only (6 studies, 480 patients, 13 treatments)			
Carbamazepine vs Control	1.68	(0.37 - 7.82)	(0.19 - 14.98)
Carbamazepine+Phenytoin vs Control	8.98	(0.36 - 169.90)	(0.26 - 243.60)
Carbamazepine+Phenobarbital vs Control	0.46	(0.00 - 21.02)	(0.00 - 28.01)
Carbamazepine+Phenobarbital+Phenytoin vs Control	4.12	(0.01 - 180.10)	(0.00 - 236.30)
Carbamazepine+Phenobarbital+Valproate vs Control	12.84	(0.50 - 435.70)	(0.35 - 604.30)
Ethosuximide+Phenytoin vs Control	5.65	(0.01 - 219.00)	(0.01 - 291.50)

Treatment Comparison	NMA Odds Ratio	95% CrI	95% PrI
Phenobarbital vs Control	0.64	(0.00 - 26.02)	(0.00 - 35.36)
Phenobarbital+Phenytoin vs Control	1.06	(0.00 - 37.64)	(0.00 - 50.85)
Phenytoin vs Control	2.08	(0.26 - 12.50)	(0.13 - 22.02)
Phenytoin+Valproate vs Control	3.14	(0.00 - 135.80)	(0.00 - 178.90)
Primidone vs Control	3.30	(0.18 - 43.76)	(0.12 - 68.72)
Valproate vs Control	13.22	(3.20 - 64.06)	(1.50 - 128.40)
Common within-network between-study variance	0.27	(0.00 - 2.97)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 3.31 Degrees of Freedom: 3	P-value: 0.35 Heterogeneity: 0.00
Cognitive Developmental Delay - Sensitivity Analysis - Maternal Alcohol or Tobacco use (3 studies, 504 patients, 7 treatments)			
Carbamazepine vs Control	1.97	(0.40 - 10.01)	(0.19 - 21.27)
Gabapentin vs Control	1.47	(0.04 - 19.01)	(0.02 - 27.11)
Lamotrigine vs Control	0.41	(0.00 - 10.09)	(0.00 - 13.61)
Levetiracetam vs Control	3.55	(0.43 - 24.13)	(0.23 - 42.39)
Topiramate vs Control	3.17	(0.30 - 24.07)	(0.18 - 44.87)
Valproate vs Control	7.79	(1.84 - 29.60)	(0.84 - 62.77)
Common within-network between-study variance	0.27	(0.00 - 3.29)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 2.69 Degrees of Freedom: 2	P-value: 0.26 Heterogeneity: NA
Cognitive Developmental Delay - Sensitivity Analysis - Low Risk of Bias: "Adequacy of follow-up" (4 studies, 283 patients, 12 treatments)			
Carbamazepine vs Control	2.68	(0.05 - 2.9 x 10^3)	(0.03 - 4.3 x 10^3)
Carbamazepine+Phenobarbital vs Control	0.67	(0.00 - 2.2 x 10^3)	(0.00 - 2.9 x 10^3)
Carbamazepine+Phenobarbital+Phenytoin vs Control	5.23	(0.01 - 7.2 x 10^3)	(0.00 - 1.1 x 10^4)
Carbamazepine+Phenobarbital+Valproate vs Control	22.18	(0.10 - 4.8 x 10^4)	(0.06 - 7.7 x 10^4)
Carbamazepine+Phenytoin vs Control	11.45	(0.13 - 1.2 x 10^4)	(0.07 - 1.8 x 10^4)
Ethosuximide+Phenytoin vs Control	6.45	(0.01 - 8.3 x 10^3)	(0.00 - 1.4 x 10^4)
Lamotrigine vs Control	0.52	(0.00 - 1.2 x 10^3)	(0.00 - 1.9 x 10^3)
Phenobarbital+Phenytoin vs Control	1.33	(0.00 - 1.8 x 10^3)	(0.00 - 2.7 x 10^3)
Phenytoin vs Control	1.67	(0.03 - 1.8 x 10^3)	(0.01 - 2.5 x 10^3)

Treatment Comparison	NMA Odds Ratio	95% CrI	95% PrI
Phenytoin+Valproate vs Control	3.94	(0.00 - 6.7 x 10^3)	(0.00 - 8.8 x10^3)
Valproate vs Control	5.9	(0.06 - 9.7 x 10^3)	(0.03 - 1.5 x 10^4)
Common within-network between-study variance	1.01	(0.01 - 5.85)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 5.07 Degrees of Freedom: 2	P-value: 0.08 Heterogeneity: 0.00
Cognitive Developmental Delay - Sensitivity Analysis - Low Risk of Bias: "Comparability of cohorts" (3 studies, 366 patients, 7 treatments)			
Carbamazepine vs Control	1.46	(0.11 - 19.59)	(0.06 - 38.10)
Gabapentin vs Control	1.19	(0.03 - 22.80)	(0.02 - 39.35)
Lamotrigine vs Control	0.27	(0.00 - 11.80)	(0.00 - 19.37)
Levetiracetam vs Control	2.90	(0.30 - 32.81)	(0.15 - 62.97)
Topiramate vs Control	2.55	(0.22 - 29.21)	(0.11 - 64.23)
Valproate vs Control	5.79	(1.05 - 47.35)	(0.47 - 102.90)
Common within-network between-study variance	0.38	(0.00 - 4.14)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 1.47 Degrees of Freedom: 2	P-value: 0.48 Heterogeneity: NA
Cognitive Developmental Delay – Network Meta-regression Analysis (11 studies, 933 patients, 18 treatments)			
Carbamazepine vs Control	1.99	(0.64 - 6.18)	(0.40 - 9.77)
Carbamazepine+Levetiracetam vs Control	0.54	(0.00 - 16.36)	(0.00 - 19.87)
Carbamazepine+Phenobarbital vs Control	0.50	(0.00 - 16.10)	(0.00 - 19.36)
Carbamazepine+Phenobarbital+Phenytoin vs Control	4.36	(0.01 - 171.20)	(0.01 - 194.60)
Carbamazepine+Phenobarbital+Valproate vs Control	14.58	(0.90 - 413.20)	(0.74 - 488.90)
Carbamazepine+Phenytoin vs Control	9.44	(0.50 - 130.50)	(0.39 - 162.40)
Ethosuximide+Phenytoin vs Control	5.77	(0.01 - 234.70)	(0.01 - 268.10)
Gabapentin vs Control	1.37	(0.04 - 15.51)	(0.03 - 19.10)
Lamotrigine vs Control	0.87	(0.07 - 5.14)	(0.06 - 6.76)
Levetiracetam vs Control	3.43	(0.57 - 18.78)	(0.42 - 24.85)
Phenobarbital vs Control	1.16	(0.13 - 8.59)	(0.10 - 11.43)
Phenobarbital+Phenytoin vs Control	1.34	(0.00 - 39.21)	(0.00 - 49.39)

Treatment Comparison	NMA Odds Ratio	95% CrI	95% PrI
Phenytoin vs Control	2.43	(0.55 - 9.14)	(0.36 - 13.45)
Phenytoin+Valproate vs Control	3.58	(0.01 - 134.20)	(0.01 - 161.70)
Primidone vs Control	2.03	(0.21 - 16.49)	(0.16 - 21.39)
Topiramate vs Control	2.93	(0.41 - 16.34)	(0.31 - 22.91)
Valproate vs Control	7.03	(2.26 - 20.02)	(1.41 - 30.92)
Common within-network between-study variance	0.16	(0.00 - 1.27)	
Regression Coefficient	1.01	(0.76 - 1.56)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 14.15 Degrees of Freedom: 17	P-value: 0.66 Heterogeneity: 0.00
Autism/Dyspraxia - Sensitivity Analysis - Large cohort (>300 patients) - (1 study, 2,551 patients, 5 treatments)**			
Clonazepam vs Carbamazepine	1.08	(0.24 - 4.85)	-
Lamotrigine vs Carbamazepine	1.20	(0.36 - 4.00)	-
Oxcarbazepine vs Carbamazepine	2.13	(0.62 - 7.35)	-
Valproate vs Carbamazepine	3.05	(0.97 - 9.52)	-
Common within-network between-study variance	NA	NA	
Evaluation of inconsistency using the design-by-treatment interaction model		NA	NA
Autism/Dyspraxia - Sensitivity Analysis - Epilepsy only (4 studies, 540 patients, 10 treatments)			
Carbamazepine vs Control	5.20	(0.54 - 90.53)	(0.33 - 133.00)
Carbamazepine+Clonazepam vs Control	7.90	(0.01 - 653.30)	(0.01 - 881.00)
Carbamazepine+Lamotrigine vs Control	4.25	(0.01 - 333.60)	(0.01 - 446.90)
Carbamazepine+Phenytoin vs Control	9.03	(0.01 - 666.30)	(0.01 - 893.00)
Lamotrigine vs Control	10.24	(1.25 - 171.40)	(0.67 - 248.50)
Lamotrigine+Valproate vs Control	120.20	(5.25 - 4.5 x 10^3)	(3.51 - 6.0 x 10^3)
Levetiracetam vs Control	3.52	(0.00 - 272.20)	(0.00 - 364.30)
Phenytoin vs Control	8.10	(0.01 - 577.50)	(0.01 - 754.60)
Valproate vs Control	14.41	(1.66 - 252.10)	(0.88 - 378.00)
Common within-network between-study variance	0.31	(0.00 - 3.04)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 2.9 Degrees of Freedom: 3	P-value: 0.41 Heterogeneity: 0.00

Treatment Comparison	NMA Odds Ratio	95% CrI	95% PrI
Autism/Dyspraxia - Sensitivity Analysis - Maternal Tobacco Use (4 studies, 540 patients, 10 treatments)			
Carbamazepine vs Control	2.51	(0.05 - 154.30)	(0.04 - 254.50)
Lamotrigine vs Control	24.84	(2.14 - 1.2 x 10^3)	(1.23 - 2.2 x 10^3)
Valproate vs Control	33.40	(2.60 -1.7 x 10^3)	(1.45 - 2.9 x 10^3)
Common within-network between-study variance	0.39	(0.00 - 4.47)	
Evaluation of inconsistency using the design-by-treatment interaction model		NA - all closed loops are formed from a multi-arm study	
Autism/Dyspraxia - Sensitivity Analysis - Maternal Alcohol Use (1 study, 156 patients, 4 treatments)			
Carbamazepine vs Control	Excluded due to zero events	-	-
Lamotrigine vs Control	4.65	(0.21 - 100.00)	-
Valproate vs Control	7.75	(0.42 - 142.86)	-
Common within-network between-study variance	*1.91*	(0.36 - 10.13)	
Evaluation of inconsistency using the design-by-treatment interaction model		NA	NA
Autism/Dyspraxia - Sensitivity Analysis - Low Risk of Bias: "Adequacy of Follow-up" (3 studies, 2,244 patients, 10 treatments)			
Carbamazepine vs Control	3.97	(0.17 - 2.4 x 10^3)	(0.11 - 3.0 x 10^3)
Carbamazepine+Clonazepam vs Control	7.48	(0.01 - 7.8 x 10^3)	(0.01 - 9.0 x 10^3)
Carbamazepine+Lamotrigine vs Control	4.47	(0.00 - 5.0 x 10^3)	(0.00 - 5.7 x 10^3)
Carbamazepine+Phenytoin vs Control	7.23	(0.01 - 6.6 x 10^3)	(0.01 - 8.2 x 10^3)
Clonazepam vs Control	4.88	(0.12 - 3.2 x 10^3)	(0.09 - 3.8 x 10^3)
Lamotrigine vs Control	6.55	(0.30 - 4.4 x 10^3)	(0.21 - 4.7 x 10^3)
Lamotrigine+Valproate vs Control	113.50	(2.33 - 7.8 x 10^4)	(1.62 - 8.9 x 10^4)
Oxcarbazepine vs Control	10.23	(0.36 - 6.8 x 10^3)	(0.26 - 7.5 x 10^3)
Valproate vs Control	13.97	(0.68 - 8.4 x 10^3)	(0.47 - 1.0 x 10^4)
Common within-network between-study variance	0.23	(0.00 - 2.88)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 2.17 Degrees of Freedom: 3	P-value: 0.54 Heterogeneity: 0.00

Treatment Comparison	NMA Odds Ratio	95% CrI	95% PrI
Autism/Dyspraxia - Sensitivity Analysis - Low Risk of Bias: "Comparability of Cohorts" (4 studies, 2,395 patients, 12 treatments)			
Carbamazepine vs Control	9.55	(0.90 - 246.20)	(0.61 - 329.40)
Carbamazepine+Clonazepam vs Control	13.58	(0.01 - 1.3 x 10^3)	(0.01 - 1.6 x 10^3)
Carbamazepine+Lamotrigine vs Control	7.11	(0.01 - 614.20)	(0.01 - 717.60)
Carbamazepine+Phenytoin vs Control	10.97	(0.01 - 1.1 x 10^3)	(0.01 - 1.4 x 10^3)
Clonazepam vs Control	8.33	(0.45 - 263.10)	(0.33 - 353.70)
Lamotrigine vs Control	10.98	(1.07 - 283.50)	(0.71 - 358.20)
Lamotrigine+Valproate vs Control	194.10	(8.06 - 8.4 x 10^3)	(6.28 - 1.0 x 10^4)
Levetiracetam vs Control	4.25	(0.00 - 390.90)	(0.00 - 485.30)
Oxcarbazepine vs Control	17.60	(1.22 - 552.20)	(0.86 - 727.40)
Phenytoin vs Control	9.76	(0.01 - 861.60)	(0.01 - 1.0 x 10^3)
Valproate vs Control	21.06	(1.86 - 525.40)	(1.25 - 681.90)
Common within-network between-study variance	0.19	(0.00 - 2.43)	
Evaluation of inconsistency using the design-by-treatment interaction model		Chi-square test: 3.36 Degrees of Freedom: 5	P-value: 0.64 Heterogeneity: 0.00
Autism/Dyspraxia - Sensitivity Analysis - Maternal IQ (1 study, 77 patients, 6 treatments)**			
Carbamazepine+Clonazepam vs Carbamazepine	1.86	(0.07 - 47.62)	-
Carbamazepine+Lamotrigine vs Carbamazepine	1.18	(0.05 - 27.78)	-
Carbamazepine+Phenytoin vs Carbamazepine	1.86	(0.07 - 47.62)	-
Lamotrigine+Valproate vs Carbamazepine	15.87	(1.87 - 142.86)	-
Valproate vs Carbamazepine	1.33	(0.18 - 10.20)	-
Common within-network between-study variance	NA	NA	
Evaluation of inconsistency using the design-by-treatment interaction model		NA	NA

Abbreviations: NMA – Network Meta-analysis; OR – odds ratio; CrI – Credible Interval; PrI – Predictive Interval
** Network did not include a control arm, comparison with Carbamazepine is reported instead