BMJ Open

Comparative safety of antiepileptic drugs for neurological development in children exposed during pregnancy and breast feeding: a systematic review and network meta-analysis

Areti Angeliki Veroniki,[1] Patricia Rios,[1] Elise Cogo,[1] Sharon E Straus,[1,2] Yaron Finkelstein,[3,4,5] Ryan Kealey,[1] Emily Reynen,[1] Charlene Soobiah,[1,6] Kednapa Thavorn,[7,8,9] Brian Hutton,[7,10] Brenda R Hemmelgarn,[11] Fatemeh Yazdi,[1] Jennifer D'Souza,[1] Heather MacDonald,[1] Andrea C Tricco[1,12]

For numbered affiliations see end of article.

Correspondence to
Dr Andrea C Tricco; triccoa@smh.ca

ABSTRACT

Objectives Compare the safety of antiepileptic drugs (AEDs) on neurodevelopment of infants/children exposed in utero or during breast feeding.

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 27 April 2017. Screening, data abstraction and quality appraisal were completed in duplicate by independent reviewers.

Participants 29 cohort studies including 5100 infants/children.

Interventions Monotherapy and polytherapy 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 breast feeding 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 that among all AEDs only valproate was statistically significantly associated with more children experiencing cognitive developmental delay compared with control (OR=7.40, 95% credible interval (CrI) 3.00 to 18.46). The NMA on autism (5 cohort studies, 2551 children, 12 treatments) suggested that oxcarbazepine (OR 13.51, CrI 1.28 to 221.40), valproate (OR 17.29, 95% CrI 2.40 to 217.60), lamotrigine (OR 8.88, CrI 1.28 to 112.00) and lamotrigine+valproate (OR 132.70, CrI 7.41 to 3851.00) were associated with significantly greater odds of developing autism compared with control. The NMA on psychomotor developmental delay (11 cohort studies, 1145 children, 18 treatments) found that valproate (OR 4.16, CrI 2.04 to 8.75) and

carbamazepine+phenobarbital+valproate (OR 19.12, CrI 1.49 to 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.

Trial registration number PROSPERO database (CRD42014008925).

Strengths and limitations of this study

► Twenty-nine cohort studies involving 5100 children of women who took antiepileptic drugs (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

Antiepileptic 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 breast feeding 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 breast feeding, employing a systematic review and network meta-analysis (NMA).

METHODS

The methods are briefly described here; details can be found in the published protocol (see online supplementary additional file 1).[11] This study was registered with PROSPERO (CRD42014008925). We followed the International Soceiety for Pharmacoepidemiology and Outcomes Research (ISPOR)[12] guidelines for our NMA and reported our findings using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) extension for NMA (see online supplementary additional file 2).[13]

Eligibility criteria

All randomised 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 breast feeding. Both monotherapy and polytherapy AEDs were eligible, including first-generation (ie, carbamazepine, clobazam, clonazepam, ethosuximide, phenobarbital, phenytoin, primidone, valproate) and newer-generation (ie, 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 (ie, 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 breast feeding in one publication, but

given the breadth of evidence 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 18 March 2014 and then updated the search on 27 April 2017. The search strategy for MEDLINE was peer-reviewed by another librarian using the Peer Review of Electronic Search Strategies (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. On 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) (see online supplementary 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 multiarm 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 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

Table 1 Outcome measures and diagnostic scales used in analysis

Cognitive developmental delay

Bayley Scales of Infant Development (children <42 months)	Score >2 SDs below the mean
Griffiths Scale of Infant Development (children >42 months)	Score >2 SDs below the mean
McCarthy Scales of Children's Abilities (children >30 months)	Score >1 SDs below the mean
Stanford-Binet IV Intelligence scale for children	IQ <80
Touwen's test	Above-average number of items rated abnormal in one or more domains
Wechsler Scale of Preschool and Primary Intelligence	IQ <90
Wechsler Intelligence Scale for Children—III	IQ <80; verbal IQ <69
Developmental assessment	Confirmed diagnosis by developmental paediatrician or paediatric 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 SDs from the test mean
Bayley Scales of Infant Development—Psychomotor Index	>2 SDs below the standardised 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 paediatrician); 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 SDs 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-month delay in age-appropriate language development

ADHD

Attention Problems and Hyperactivity Scales	Score >1 SDs 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 paediatrician or child psychiatrist

Neonatal seizure

Medical records	Record of seizures during first year; confirmation of neonatal seizure by electroencephalography or diagnosis

Continued

Table 1 Continued

Cognitive developmental delay

Social impairment

Developmental Assessment (Ages and Stages Questionnaire (6 and 18 months); Child Behaviour Checklist (36 months))	Scores dichotomised into 'normal' or 'adverse' range based on predefined values used by scale, for scales without predefined values cut-off was set at a score >2 SDs outside the test mean

ADHD, attention-deficit hyperactivity disorder, ICD-10, International Classification of Diseases, Tenth Revision; SD, standard deviation.

across studies, and to achieve the highest generalisability in the meta-analytical treatment effects, we applied a random-effects model.[18]

An NMA was applied for connected evidence networks and prespecified 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 (ie, using the control group) were conducted, when ≥10 studies provided relevant information, assuming a common fixed coefficient across treatment comparisons for AEDs versus control. Sensitivity analyses for cognitive developmental delay and autism/dyspraxia outcomes were performed for treatment indication of epilepsy, large study size (ie, >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 three units in the DIC between an 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% credible intervals (CrIs) 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 5707 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 (see online supplementary 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 the online supplementary appendix C.

Figure 1 Study flow diagram.
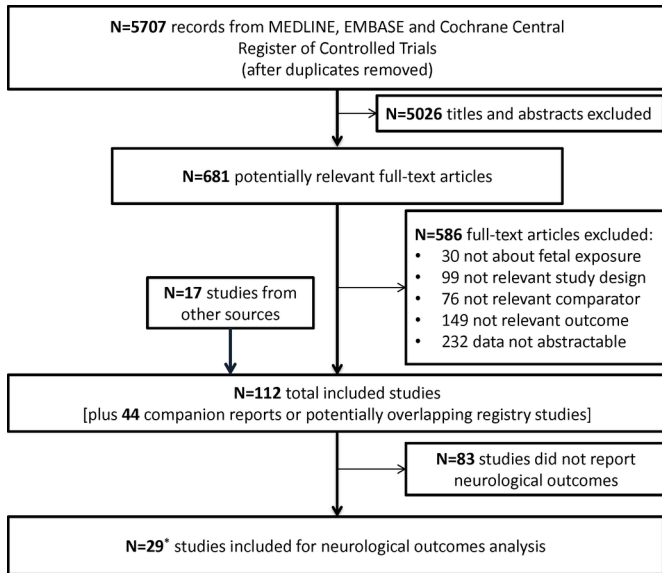
*29 publications reporting 30 included studies.

Study and patient characteristics

We included 29 cohort studies (5100 patients) published between 1989 and 2016 (table 2; see online supplementary appendix D,E). The number of patients included in each study ranged from 23 to 2011 (median 74.5). Most studies (76%) were published after 2000, 62% of the studies included <100 patients and 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 (see online supplementary 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 (eg, identified through database records) and 23 (79%) ascertained exposure via secured records (eg, 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 (see online supplementary appendix G).

Statistical analysis results

No important concerns were raised regarding the violation of the transitivity assumption when maternal

Table 2 Summary characteristics of included studies

Study/patient characteristic	Studies (n=29)	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
Randomised 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
Neonatal seizures	2	6.90
Social impairment	1	3.45
Funding		
Public	15	51.72
Private	0	0.00

Continued

Table 2 Continued

Study/patient characteristic	Studies (n=29)	Studies (%)
Mixed public and private	4	13.79
NR/unclear	10	34.48
Treatment indication		
Epilepsy	23	79.31
Mixed indications‡	0	0.00
NR	6	20.69
Epileptic control group§		
Yes	15	51.72
No/NR/NA	14	48.28
Mean maternal age (years)		
24–26	2	6.90
27–29	5	17.24
30+	4	13.79
NR	18	62.07
AED exposure during pregnancy		
Reported as during first trimester	5	17.24
Reported as any time during pregnancy	4	13.79
During pregnancy and breast feeding	5	17.24
NR	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

*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.
ADHD, attention-deficit hyperactivity disorder; AED, antiepileptic drug(s); NA, not applicable; NR, not reported.

age, baseline risk, treatment indication and timing were assessed (see online supplementary 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 (see online supplementary appendix H).

In the following sections, we present the significant NMA results by outcome for AEDs compared with control (ie, 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 the online supplementary appendix I.
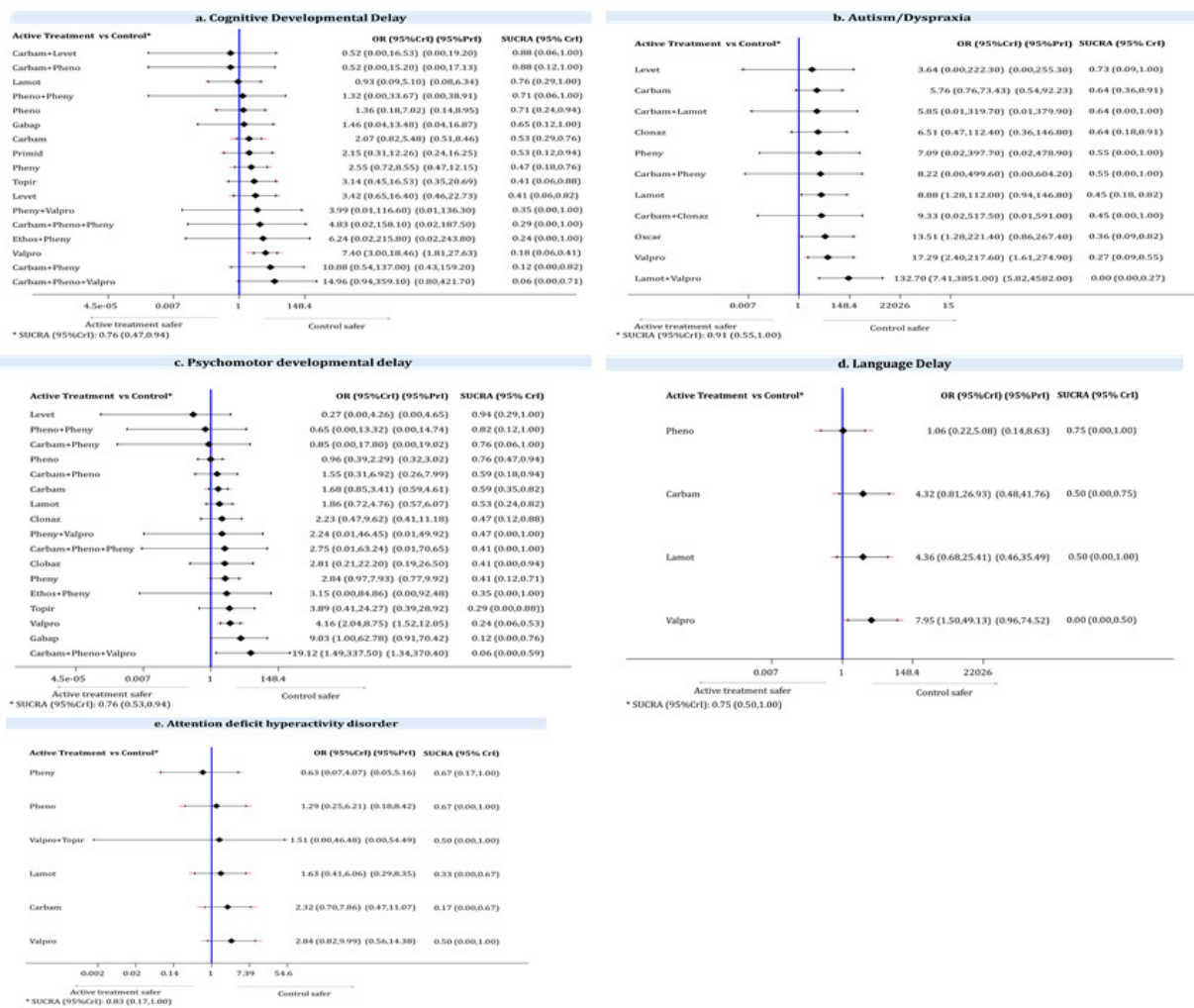
Cognitive developmental delay

The NMA for cognitive developmental delay (definitions in table 1) included 11 cohort studies, 933 children and examined 18 treatments (figure 3A; see online supplementary appendix J; τ^2=0.12, 95% CrI 0.00 to 1.15). One study included children exposed to AEDs both in utero and through breast feeding, and 10 included children exposed to AEDs in utero. Across all AEDs, only valproate was associated with significantly increased odds of cognitive developmental delay compared with control (OR 7.40, 95% CrI 3.00 to 18.46; figure 2A; see online supplementary 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 to 1.56; τ^2=0.16, 95% CrI 0.00 to 1.24; residual deviance=45.27, data points=47, DIC=80.17). Similarly, the sensitivity analyses restricted to (1) studies that only included women receiving AEDs to treat epilepsy (10 studies, 910 children, 17 treatments; τ^2=0.16, 95% CrI 0.00 to 1.36), (2) studies comparing only first-generation AEDs (6 studies, 480 children, 13 treatments; τ^2=0.28, 95% CrI 0.00 to 2.97), (3) studies that reported maternal alcohol or tobacco use (3 studies, 504 children, 7 treatments; τ^2=0.27, 95% CrI 0.00 to 3.29) and (4) 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 to 4.14) were consistent with the NMA results (see online supplementary 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 to 5.85; see online supplementary appendix K).

Autism/dyspraxia

The NMA on autism/dyspraxia (definitions in table 1) included 5 cohort studies, 2551 children exposed in utero and examined 12 treatments (τ^2=0.16, 95% CrI 0.00 to 1.95; figure 3B; see online supplementary appendix H). Compared with control, only valproate (OR 17.29, 95% CrI 2.40 to 217.60), oxcarbazepine (OR 13.51, 95% CrI 1.28 to 221.40), lamotrigine (OR 8.88, 95% CrI 1.28 to 112.00) and lamotrigine+valproate (OR 132.70, 95% CrI 7.41 to 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 to 304). Two cohort studies of 404 offspring of women with a history of tobacco use

Figure 2 Forest plots for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay and attention-deficit hyperactivity disorder outcome. carbam, carbamazepine; ethos, ethosuximide; gabap, gabapentin; lamot, lamotrigine; levet, levetiracetam; pheno, phenobarbital; pheny, phenytoin; PrI, predictive interval; primid, primidone; SUCRA, surface under the cumulative ranking; topir, topiramate; valpro, valproate.

compared four 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 to 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, 2395 children, 12 treatments; τ^2=0.19, 95% CrI 0.00 to 2.43) and the 'adequacy of follow-up of cohorts' (3 studies, 2244 children, 10 treatments; τ^2=0.23, 95% CrI 0.00 to 2.88), except that lamotrigine was no longer statistically significant than control for the latter (see online supplementary appendix K).

Neonatal seizure
One cohort study included 72 children who were exposed to AEDs in utero as well as through breast feeding 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 to 3.70).

Psychomotor developmental delay
The NMA on psychomotor developmental delay (definitions in table 1) included 11 cohort studies, 1145 children exposed in utero and examined 18 treatments (τ^2=0.06, 95% CrI 0.00 to 0.63; figure 3C; see online supplementary appendices H,J). Valproate (OR 4.16, 95% CrI 2.04 to 8.75) and carbamazepine+phenobarbital+valproate (OR 19.12, 95% CrI 1.49 to 337.50) were significantly more harmful than control (figure 2C).

Language delay
The NMA on language delay (definitions in table 1) included 5 cohort studies, 509 children and examined 5 treatments (τ^2=0.16, 95% CrI 0.00 to 2.15; figure 3D; see online supplementary appendices H,J). One study included children exposed to AEDs in utero and through breast feeding, 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 to 49.13; figure 2D).
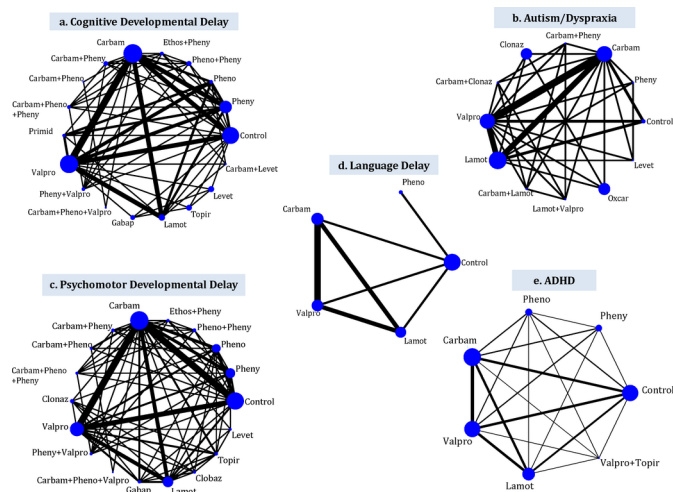
Figure 3 Network diagrams for cognitive developmental delay, autism/dyspraxia, psychomotor developmental delay, language delay and attention-deficit hyperactivity disorder outcomes. Each treatment node is weighted according to the number of patients that have received the particular treatment, and each edge is weighted according to the number of studies comparing the treatments it connects. 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.

Attention-deficit hyperactivity disorder (ADHD)

The NMA on ADHD (definitions in table 1) included 5 cohort studies, 816 children and examined 7 treatments (τ^2=0.11, 95% CrI 0.00 to 1.29). One study included children exposed to AEDs in utero and through breast feeding, while four studies included children exposed to AEDs in utero. None of the treatment comparisons reached statistical significance (figures 3E and 2E; supplementary appendices H,J).

Social impairment

One cohort study included 422 children exposed to AEDs in utero as well as through breast feeding. 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 breast feeding. 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 (eg, 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 (eg, 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 6 years of age compared with children exposed to valproate (see online supplementary appendix C).[7] As indicated in the online supplementary 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 with unexposed children (see online supplementary appendix C).[42] As indicated in the online supplementary appendix C, we were unable to include this study due to the same reason as above. An 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, eg, lamotrigine) was associated with the greatest odds. Similarly, two previous systematic reviews that did not conduct an NMA found valproate was associated with significantly lower IQ scores and poorer overall neurodevelopmental outcomes compared with 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 5 studies per meta-analysis) concluded that AED polytherapy led to poorer developmental outcomes and IQ compared with 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' under-reporting, 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 among 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 towards 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-randomised with randomised evidence in an NMA and have highlighted the need to carefully explore the level of confidence in the non-randomised 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 a 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 randomised 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 (breast feeding vs 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 breast feeding (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 breast feeding 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 non-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.

Author affiliations
[1]Li Ka Shing Knowledge Institute,St. Michael's Hospital, Toronto, Canada
[2]Department of Medicine, University of Toronto, 27 King's College Circle, Toronto, Canada
[3]The Hospital for Sick Children,555 University Avenue, Toronto, Canada
[4]Department of Paediatrics, University of Toronto, Toronto, Canada
[5]Department of Pharmacology and Toxicology, University of Toronto, Medical Sciences Building, Toronto, Canada
[6]Institute for Health Policy Management & Evaluation,University of Toronto, Toronto, Canada
[7]School of Epidemiology,Public Health and Preventive Medicine, University of Ottawa, Ottawa, Canada
[8]Clinical Epidemiology Program,Ottawa Hospital Research Institute, The Ottawa Hospital, Ottawa, Canada

(b)

⁹Institute of Clinical and Evaluative Sciences (ICES uOttawa), Ottawa, Canada
¹⁰Ottawa Hospital Research Institute, Center for Practice Changing Research, Ottawa, Canada
¹¹Departments of Medicine and Community Health Sciences, University of Calgary, Calgary, Canada
¹²Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, Canada

REFERENCES

1. Spina E, Perugi G. Antiepileptic drugs: indications other than epilepsy. *Epileptic Disord* 2004;6:57–75.
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;50:1247–55.
3. Samrén 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;38:981–90.
4. Meador K, Reynolds MW, Crean S, *et al*. Pregnancy outcomes in women with epilepsy: a systematic review and meta-analysis of published pregnancy registries and cohorts. *Epilepsy Res* 2008;81:1–13.
5. Bromley R, Weston J, Adab N, *et al*. Treatment for epilepsy in pregnancy: neurodevelopmental outcomes in the child. *Cochrane Database Syst Rev* 2014;10:Cd010236.
6. Banach R, Boskovic R, Einarson T, *et al*. 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;33:73–9.
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;12:244–52.
8. Christensen J, Grønborg TK, Sørensen MJ, *et al*. Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. *JAMA* 2013;309:1696–703.
9. Wlodarczyk BJ, Palacios AM, George TM, *et al*. Antiepileptic drugs and pregnancy outcomes. *Am J Med Genet A* 2012;158A:2071–90.
10. Velez-Ruiz NJ, Meador KJ. Neurodevelopmental effects of fetal antiepileptic drug exposure. *Drug Saf* 2015;38:271–8.
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. *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;17:157–73.
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;162:777–84.
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;15:95.
15. McGowan J, Sampson M, Salzwedel DM, *et al*. 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-analyses 2000. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
17. Chaimani A, Higgins JP, Mavridis D, *et al*. Graphical tools for network meta-analysis in STATA. *PLoS One* 2013;8: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;31:72–6.
19. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004;23:3105–24.
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, *et al*. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Res Synth Methods* 2012;3:111–25.
22. Song F, Altman DG, Glenny AM, *et al*. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* 2003;326:472.
23. Veroniki AA, Vasiliadis HS, Higgins JP, *et al*. Evaluation of inconsistency in networks of interventions. *Int J Epidemiol* 2013;42:332–45.
24. Welton NJ, Sutton AJ, Cooper N, *et al*. *Evidence synthesis for decision making in healthcare*. New York: Wiley, 2012.
25. Spiegelhalter DJ, Best NG, Carlin BP, *et al*. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B* 2002;64:583–639.
26. Lunn D, Spiegelhalter D, Thomas A, *et al*. The BUGS project: Evolution, critique and future directions. *Stat Med* 2009;28:3049–67.
27. Palmer T, Sterne J. White I, ed. *Meta-Analysis in Stata: An Updated Collection from the Stata Journal*. 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;172:137–59.
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;64:163–71.
31. Trinquart L, Attiche N, Bafeta A, *et al*. Uncertainty in Treatment Rankings: Reanalysis of Network Meta-analyses of Randomized Trials. *Ann Intern Med* 2016;164:666–73.

32. Bromley R, Baxter N, Calderbank R, *et al*. *A comprehensive review of the language abilities of children exposed to valproate or carbamazepine in utero*. Texas: American Epilepsy Society, 2010.
33. Gogatishvili N, Ediberidze T, Lomidze G, *et al*. PO-0834 Long-term Developmental Outcome Of Children Prenatally Exposed To Antiepileptic Drugs. *Arch Dis Child* 2014;99(Suppl 2):A526.3–A526.
34. Gogatishvili N, Ediberidze T, Lomidze G, *et al*. Cognitive outcomes of children with fetal antiepileptic drug exposure at the age of 3-6 years-preliminary data. *1st 624 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*. Texas: American Epilepsy Society, 2010.
36. Adab N, *et al*. The longer term outcome of children born to mothers with epilepsy. *Journal of Neurology, Neurosurgery & Psychiatry* 2004;75:1575–83.
37. Dean JC, Hailey H, Moore SJ, *et al*. Long term health and neurodevelopment in children exposed to antiepileptic drugs before birth. *J Med Genet* 2002;39:251–9.
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;2:119–23.
40. Veroniki AA, Straus SE, Fyraridis A, *et al*. The rank-heat plot is a novel way to present the results from a network meta-analysis including multiple outcomes. *J Clin Epidemiol* 2016;76:193–9.
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;70:1367–74.
42. Shallcross R, Bromley RL, Irwin B, *et al*. Child development following in utero exposure: levetiracetam vs sodium valproate. *Neurology* 2011;76:383–9.
43. Zaccara G, Giovannelli F, Giorgi FS, *et al*. Tolerability of new antiepileptic drugs: a network meta-analysis. *Eur J Clin Pharmacol* 2017;73:811–7.
44. Higgins JPT, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. 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 Overseas Ed* 1985;312:559–63.
47. Efthimiou O, Mavridis D, Debray TP, *et al*. Combining randomized and non-randomized evidence in network meta-analysis. *Stat Med* 2017;36:1210–26.
48. Schmitz S, Adams R, Walsh C. Incorporating data from various trial designs into a mixed treatment comparison model. *Stat Med* 2013;32:2935–49.
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. *Syst Rev* 2015;4:147.
50. Easterbrook PJ, Berlin JA, Gopalan R, *et al*. Publication bias in clinical research. *Lancet* 1991;337:867–72.
51. Veroniki AA, Straus SE, Soobiah C, *et al*. A scoping review of indirect comparison methods and applications using individual patient data. *BMC Med Res Methodol* 2016;16:47.