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]. <u>Note:</u> 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). **<u>Thus, there are 2 parts to this question:</u>** <u>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</u> <u>2) Then the study should also have reported the age for each AED arm.</u> <u>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).</u> 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 assess ment 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) <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 <u>Note:</u> 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

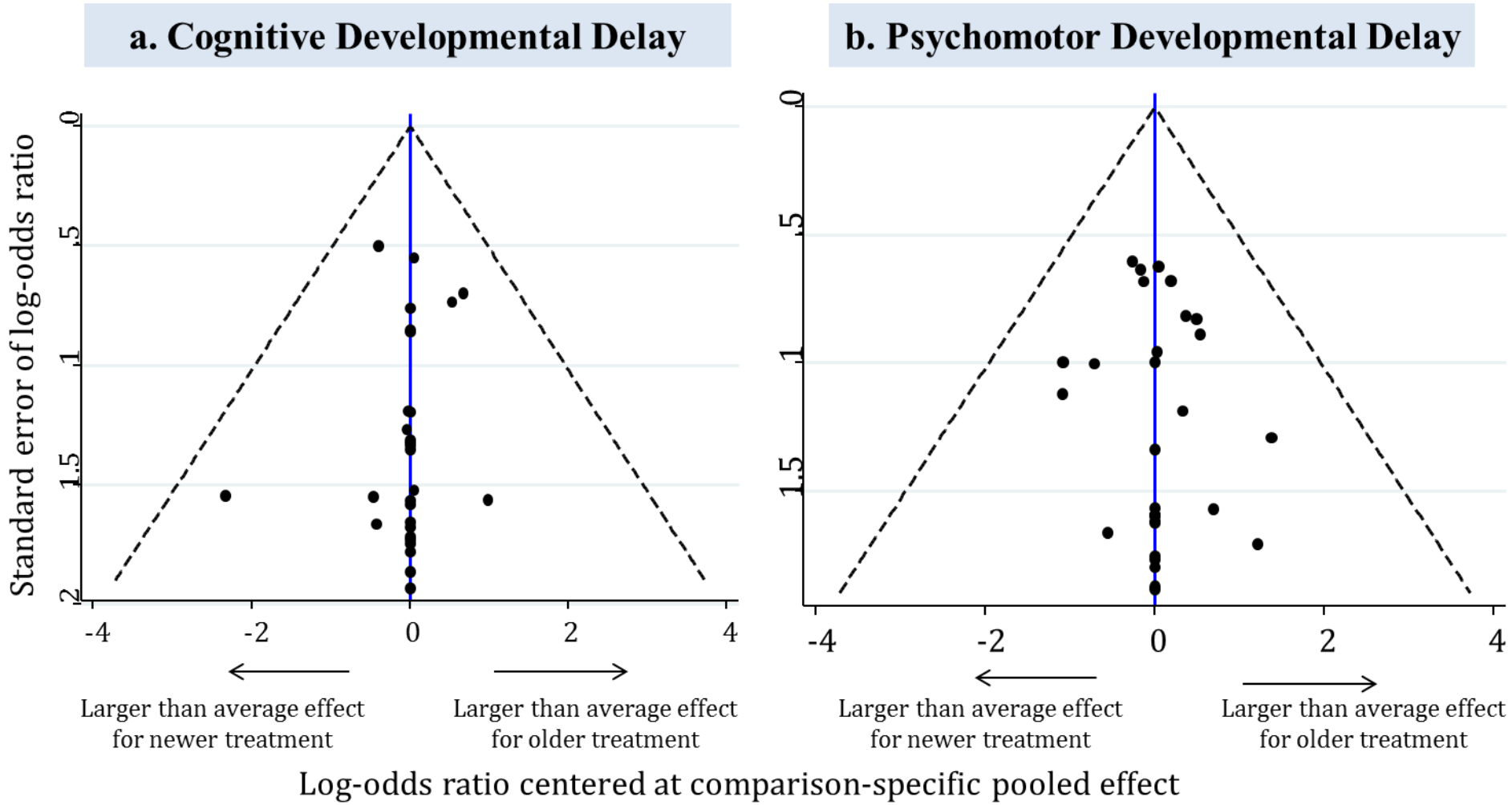
* Sample size used for analysis; ineligible treatment arms (i.e. treatment arms with excluded drugs or unspecified polytherapy) are not included in the count
† The mean age for children/follow-up period data were only collected for outcomes related to cognitive and/or psychomotor development
‡ Total 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*

Outcomes
Circles from outside in refer to :
1st: Cogn Dev Delay
2nd: Aut/Dyspr
3rd: Psyc Dev Delay
4th: Lang delay
5th: ADHD
White sectors including a '*' 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 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	$(5.25 - 4.5 \times 10^3)$	$(3.51 - 6.0 \times 10^3)$
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