BMJ Open

The Importance of Maternal Diet Quality During Pregnancy on Cognitive and Behavioural Outcomes in Children – A Systematic Review and Meta-Analysis

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-016777
Article Type:	Research
Date Submitted by the Author:	09-Mar-2017
Complete List of Authors:	Borge, Tiril; Nasjonalt folkehelseinstitutt, Department of Child Health Aase, Heidi; Nasjonalt folkehelseinstitutt, Department of Child Development Brantsaeter, Anne-Lise; Nasjonalt folkehelseinstitutt, Department of Environmental exposure and Epidemiology Biele, Guido; Nasjonalt folkehelseinstitutt, Department of Child Health
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Mental health, Nutrition and metabolism, Public health
Keywords:	EPIDEMIOLOGY, MENTAL HEALTH, NUTRITION & DIETETICS, Child & adolescent psychiatry < PSYCHIATRY, PUBLIC HEALTH

TITLE PAGE

Title: The Importance of Maternal Diet Quality During Pregnancy on Cognitive and Behavioural Outcomes in Children – A Systematic Review and Meta-Analysis

Corresponding author:
Tiril Cecilie Borge
Norwegian Institute of Public Health, Department of Child Health
P.O. Box 4404 Nydalen
0403 OSLO
Norway

Email address of the corresponding author: tibo@fhi.no
Tel: +4797613762

Co-authors (in correct author order):
Heidi Aase, Norwegian Institute of Public Health, Department of Child Development, Oslo, Norway
Anne Lise Brantsæter, Norwegian Institute of Public Health, Department of Environmental Exposure and Epidemiology, Oslo, Norway
Guido Biele, Norwegian Institute of Public Health, Department of Child Health, Oslo, Norway

Keywords: maternal diet quality, pregnancy, cognitive development, affective functioning, systematic, review, meta-analysis

Word count: 6165
Number of figures: 4
Number of tables: 3
Number of references: 94
Number of supplementary files for online only publication: 1

ABSTRACT

Objectives: This systematic review and meta-analysis provides a quantitative summary of the literature exploring the relationship between maternal diet quality during pregnancy and child cognitive and affective outcomes. We investigate whether there are indications for robust associations and aim to identify methodological strengths and challenges of the current research to provide suggestions of improvement for future research.

Design and participants: Relevant studies were identified through a systematic literature search in relevant databases. All studies investigating maternal diet quality during pregnancy in relation to child cognitive or affective functioning in children of elementary school age or younger were assessed for inclusion.

Results: 18 relevant studies, comprising 63861 participants were identified. The results indicated a small positive association between better maternal diet quality during pregnancy and child functioning. We observed publication bias and large heterogeneity between studies included in the affective domain, where type of diet classification accounted for about 35 % of this heterogeneity. Trim and fill analysis substantiated the presence of publication bias for studies in the affective domain and showed an adjusted effect size of Hedge's g=0.088 (p=0.0018) [unadjusted g=0.093(p=0.0012)]. We observed no publication bias in the cognitive domain, where results indicated a slightly larger effect size (g=0.14 (p<0.0001)) compared to that of the affective domain. The overall summary effect size was g=0.075 (p<0.0001) adjusted for publication bias [unadjusted g=0.112 (p<0.0001)]. Child diet was not systematically controlled for in the majority of the included studies.

Conclusion: The results indicated that a better maternal diet quality during pregnancy has a small positive association with child neurodevelopment, with more reliable results seen for cognitive development. These results warrant further research on the association between maternal diet quality during pregnancy and cognitive and affective aspects of child neurodevelopment, whereby it is crucial that future studies account for child diet in the analysis.

Strengths and limitations of this study

- This is the first article to summarize research into the association between maternal diet quality during pregnancy and child neurodevelopment
- Major strengths of this research are the use of meta-analytic methods for calculation of average effect sizes and investigation of publication bias
- This study highlights strengths and challenges of an emerging research field, thus building the foundation for improved future research
- A limitation is the relatively small number of relevant studies identified for inclusion in the meta-analysis
- Since this meta-analysis is based on observational studies, no strong causal interpretations about the association of maternal diet quality during pregnancy and child neurodevelopment can be made

INTRODUCTION

The importance of adequate nutrition during foetal life for long-term physical health is well documented.[1, 2] However, the relationship between maternal nutrition during pregnancy and child mental health is less established.[3] The prenatal environment is crucial in relation to cognitive development of the child, particularly during critical periods of brain development, which highlights the foetus' need for optimal nutrition.[4] There are documented detrimental effects of severe maternal malnutrition during pregnancy,[5] and severe deficiencies of certain micronutrients, like iron and iodine[6] on child neurodevelopment and general cognitive functions, as well as severe deficiencies of folate and choline on child neural tube defects,[7] but the impact of more subtle variations in maternal diet quality[1] on child neurodevelopment has received little attention until recently.

It has become increasingly recognized that investigating the impact of diet on most disease outcomes cannot be done solely by investigating single nutritional components separately. Considering that the human diet consists of at least 25 000 biologically active components, with only a fraction having been defined as nutrients, it is likely that the majority of the dietary constituents effect human health in an interdependent manner. Looking at overall diet quality, e.g. through dietary patterns, is believed to represent a valid and meaningful measure of overall nutrient intake[8] and is a promising approach when aiming to study diet related associations.[9]

To date, no meta-analysis has summarized research on maternal diet quality and child neurodevelopment. As the research interest for this topic is rapidly increasing, it is valuable to summarize the research to date on this topic using statistical procedures. The aim of this meta-analysis is to provide a quantitative summary of the existing literature exploring the relationship between maternal diet quality and child cognitive and affective outcomes. The goals are to investigate whether there are indications for robust associations, despite the limited amount of studies available, and to identify methodological strengths and challenges of the current research to provide suggestions of improvement for future research.

METHODS

As a scientific guideline for this manuscript, we followed the PRISMA statement.[10]

Defining exposure and outcome measures

Despite the increased interest in studying dietary patterns as a measure for diet quality, the current literature regarding the associations between maternal dietary patterns during pregnancy and child neurodevelopmental outcomes is sparse. A preliminary literature search resulted in only four articles with defined maternal dietary patterns as exposure relevant for inclusion into the meta-analysis. Consequently, in order to increase the basis for analysis, articles with dietary exposures believed to be good proxies for maternal diet quality were included. Based on the existing literature, fish intake,[11] Ω-6/Ω-3 fatty acid ratio,[12] saturated fat intake[13] and dietary fibre (reflecting intake of whole-grain foods, vegetables, fruits, legumes, and nuts)[14] were considered good proxies for maternal diet quality. In previous research, based on dietary data from large cohorts where dietary patterns have been identified with data driven methods, consumption of fish and fibre rich foods, as well as limited intake of saturated fats, have consistently been associated with a healthier dietary pattern, both in the general population,[15] and in pregnant women.[8, 16] Additionally, fibre rich foods, fat quality and fish are incorporated into

[1] When using the term "maternal diet quality" in this paper we are always referring to the maternal diet quality during pregnancy, unless otherwise stated

established healthy food indices, like the Mediterranean diet index[17] and the Healthy Eating Index (HEI-2010).[18] Considering the already limited amount of available relevant literature, few limitations were put on the possible outcome as long as it covered a child neurodevelopmental domain, like cognition (IQ and language) or affect (externalizing and internalizing difficulties).

Search criteria and strategies

An extensive search string was developed as to not exclude any relevant literature, and adapted to each database, including key words relating to maternal diet quality, child mental health, cognitive function (language, communication skills, IQ) neurodevelopmental disorders (ADHD, ODD, CD, ASD) and affective functioning. The following exclusion criteria were applied: children with very low birth weight; children older than elementary school age; and studies focusing on single micronutrients and/or supplements. The search string was developed by TCB in collaboration with a specialist librarian. For full search string, see supplementary table 1.

Data collection and extraction process

After identification of original articles through the initial search, excluding duplicates, TCB and ALB independently screened title and/or abstract of each study. TCB's and ALB's final list of eligible and possibly eligible studies were then crosschecked and read in full text by both. If both reviewers were unsure whether an article was eligible for inclusion, GB was consulted to assure coherence regarding the final selection of articles for inclusion. After identification of the eligible articles the relevant information from each study was extracted by TCB (e.g. year of publication, total number of participants, dietary exposure and outcome measures assessed, confounders controlled for, and reported effect sizes) in collaboration with GB and ALB. TCB and GB then assessed individual study quality with the Newcastle-Ottawa Scale (NOS) for assessing the quality of cohort studies in meta-analysis.[19]

Individual study quality assessment

For each eligible study included in the meta-analysis we performed an individual study quality assessment using the NOS.[19] The NOS provides an easy to use study quality checklist and is recognized by Cochrane.[20] The scoring system is based on the assessment of three aspects of a study; Selection (representativeness of cohort and exposure assessment); Comparability (ascertainment of confounding); and Outcome (assessment of outcome and follow-up). The scoring system categorizes studies as being of good, fair or poor methodological quality, whereby insufficiency in one of the domains results in a "poor" rating. While the NOS has been criticized for an overly general definition of quality criteria,[21] this generality allows for a wide application of the scale. Moreover, the intent of the scale is clear: A good rating of the Selection dimensions requires a representative sample and high quality measurement; a good rating of the Comparability dimension requires control of appropriate confounders; and a good rating of the Outcome dimensions requires a high quality measurement of outcomes and/or high follow up rates or correction for non-random drop out.

Analysis of reported effect sizes

To be able to compare the results of the studies, association measures reported in each individual study had to be transformed into a standardized effect size[2]. The effect size measure utilized for this meta-analysis was Hedges' g, which is a more conservative effect size measure compared to Cohen's d.[22]

[2] When referring to "effect size" in this paper we are not indicating causality – it is merely the statistical term of the reported outcome measures.

Effect sizes were calculated to reflect the association between better maternal diet quality and the different cognitive and affective outcomes, where a positive value indicate a better outcome, hence better language development or general cognitive functioning, or less affective problems.

The meta-analysis was conducted with the R statistical software (version 3.2.2.), using the Metafor package, version 1.9-8.[23, 24] Because the included studies were heterogeneous with regards to neurodevelopmental outcomes, choice of statistical procedures, and effect sizes, we used a random-effects model (REM) for analyses.[22] This approach models variance between the included studies, and assumes that observed differences in effect sizes are due to both sampling error and true effect size differences in the studies' background populations.[22]

A restricted maximum-likelihood method for estimation of heterogeneity was used to compute relevant Q-statistics, with corresponding I^2-statistics. The Q-statistics indicate whether there is statistically significant heterogeneity across the studies' effect sizes, whereas the I^2-statistics indicate the extent of heterogeneity. A significant Q-statistic indicates systematic (as opposed to random) variation of effect sizes between studies.

Possible moderators

If the REM analyses indicate presence of heterogeneity, moderator analyses can be used to investigate potential causes of this heterogeneity. We used meta-regressions where we added potential moderator variables individually to separate regression models in order to assess their effect on the association between exposure and outcome. The following factors were available for consideration as possible moderators: Publication year, diet category (type of diet classification - whether the exposure is defined as maternal dietary pattern or a proxy for maternal dietary pattern) and instrument category (measurement of outcome - questionnaire or neuropsychological test), as they are all factors which might moderate the association between exposure and outcome. The categorical factors were dichotomous.

Publication bias

Publication bias describes a situation in which the decision to publish research results depends on obtaining statistically significant results.[25] Indeed, studies reporting statistically significant results are more likely to be published than studies reporting results that are not statistically significant.[26] One visual meta-analytic tool traditionally used to investigate publication bias is the funnel plot.[27] To complement the potential subjectivity of visual inspection of funnel plots the Egger's regression test for funnel plot asymmetry[28] can be performed.

In the presence of publication bias, a "trim and fill" approach can be used to correct for it. The trim and fill method uses effect sizes and their standard error to generate a "complete" distribution of effect sizes that likely would have been reported without publication bias by adding imputed studies to the reported studies. If the average effect sizes calculated from published and "complete" effect sizes do not differ noticeably, one can have more confidence in the average effect size from a group of studies that appear afflicted by publication bias.[25] As there is some discussion in the literature regarding the optimal methods for adjustment for publication bias,[29] a meta-regression to adjust for publication bias using the standard error of effect sizes as a covariate should also be performed.

RESULTS

Study sample and selection

Ovid (Embase, Psychinfo, Medline), PubMed and ISI: WEB of science were searched on November 16th 2016 using the full search string. Additionally, we explicitly searched for relevant studies that employed a priori dietary quality indices (like the HEI-2010 and Mediterranean diet index).

A total of 18 studies fit the inclusion criteria,[30-47] the majority reporting several outcome measures. Four of the studies[30, 31, 34, 39] were based on subsamples of the ALSPAC cohort, and two studies[37, 47] were based on subsamples from the Project Viva cohort. However, these studies used different outcome measures at different time points and, considering the already limited amount of relevant studies, we included all six in the meta-analysis. The study selection process is visualized in Figure 1.

Figure 1 approximately here

All included studies were observational in nature and based on a prospective cohort design or case-control design, with baseline measures of maternal dietary intake during pregnancy and subsequent measurement of child cognitive or affective functioning, at one or more time points.

All studies collected information on maternal dietary intake with the use of a food frequency questionnaire (FFQ), either self-administered or by a trained interviewer, with some using validated FFQs; and/or a food-diary. Data obtained with these instruments was used the basis for the definition of dietary patterns, estimation of fish/seafood intake, fruit intake, saturated fat intake and estimation of Ω-6/Ω-3 fatty acid ratio based on intake.

Four studies used maternal dietary patterns as exposure variables,[30, 35, 39, 40] either defined by the use of principal component analysis (PCA) or confirmatory factor analysis (CFA). In three studies,[30, 35, 40] two distinct dietary patterns were identified; one "healthy" and one "unhealthy", while only an unhealthy dietary pattern was defined in the fourth.[39] The healthy dietary patterns were generally characterized by higher intakes of vegetables, fish, legumes, wholegrains and vegetable oils, while the unhealthy dietary patterns consisted of higher intakes of processed foods (fried foods, French fries, meats) confectionary foods (cakes, candy, sugary drinks), refined cereals, and salty snacks.

Eleven studies used maternal fish intake as exposure,[31-34, 36-38, 41, 42, 46, 47] where the studies categorized fish intake into groups based on meals/portions or grams eaten per day or week. The remaining three studies used Ω-6/Ω-3 fatty acid ratio,[43] saturated fat intake[45] and fruit intake[44] as their exposure variable. All studies were published in the period from 2004-2016, with study populations ranging from 48 to 23020 mother-child pairs (Table 1).

Table 1: Description of the studies and assessment of the exposure

Reference	Dietary data collection period	Country	Cohort name	Total n *	Dietary assessment	FFQ items	Exposure variable	Exposure variable scale
Barker et al, 2013[30]	1990-1992	UK	ALSPAC	6979	FFQ	103	"Healthy" and "unhealthy" patterns, grouped using CFA	Dietary patterns as continuous variable
Bernard et al, 2013[43]	2003-2005	France	Eden cohort	1335	FFQ	137	Estimated maternal intake of total n6 and total n3 (in g/d), then calculation of n6:n3 fatty acid ratio.	Fatty acid ratio as continuous variable
Bolduc et al, 2016[44]	2008-2012	Canada	CHILD Edmonton sub-cohort	688	FFQ	175	Total fruit intake estimated from FFQ	Maternal fruit intake as continuous predictor of neurodevelopment
Daniels et al, 2004[31]	1990-1992	UK	ALSPAC	7421	FFQ	103	Maternal fish intake as servings/week	Categorical variable; Four intake groups (of which one referent group)
Davidson et al, 2008[32]	2001-2002	Seychelles	Seychelles Child Development Study	229	FFQ and 4-day diet diary	NS	Maternal fish intake as g/day	Maternal fish intake as continuous variable
Gale et al, 2008[33]	1991-1992	UK	NS	217	FFQ	100	Maternal fish intake as times eaten/week	Categorical variable; Three intake groups (of which one referent group)
Gustafsson et al, 2016[45]	NS	USA	Ongoing longitudinal study	48	ASA24**	NS	Total and saturated fat intake	Fat intake as continuous variable
Hibbeln et al, 2007[34]	1990-1992	UK	ALSPAC	11875	FFQ	103	Maternal seafood intake as g/week	Categorical variable; Three intake groups (of which one referent group)
Jacka et al, 2013[35]	2002-2008	Norway	MoBa	23020	validated FFQ	255	"Healthy" and "unhealthy" data driven dietary pattern scores	Dietary patterns scores as continuous variables
Julvez et al, 2016[46]	2004-2008	Spain	INMA	1892	FFQ	101	Maternal seafood intake	Maternal seafood intake as Quintiles
Mendez et al, 2008[36]	1997-1998	Menorca	Prospective birth cohort	392	FFQ	42	Maternal fish intake	Categorical variable; Four intake groups (of which one referent group)
Oken, Radesky et al, 2008a[37]	1999-2002	USA	Project Viva	341	validated FFQ	>140	Maternal fish intake	Categorical variable; Three intake groups (of which one referent group)
Oken, Østerdal et al, 2008[38]	1997-2003	Denmark	Danish birth cohort	25446	Validated FFQ	>360	Maternal fish intake as servings(g)/week	Categorical variable; Three quintiles of intake (lowest, middle, highest)
Oken et al, 2016[47]	1999-2002	USA	Project Viva	1068	validated FFQ	>140	Maternal fish intake	Categorical variable; Three intake groups (of which one (no fish) referent group)
Pina-Camacho et al, 2015[39]	1990-1992	UK	ALSPAC	7814	FFQ	103	General unhealthy diet (second-order latent factor) generated with CFA	Unhealthy dietary pattern as continuous variable
Sagiv et al, 2012[42]	1993-1998	USA	New Bedford cohort	362	FFQ	NS	Maternal fish intake, expressed as total servings/week	Maternal fish intake as continuous variable
Steenweg-de Graaff et al, 2014[40]	2001-2006	Netherlands	Generation R	3104	Validated FFQ	293	"Healthy" and "unhealthy" data driven dietary pattern scores	Dietary patterns scores as continuous variables
Valent et al, 2013[41]	2007-2009	Italy	Set within project (PHIME).	606	FFQ (Adapted from a validated FFQ)	138	Maternal fish intake as servings/week	Maternal fish intake as continuous variable

* differs from total n used in analysis (due to missing data, excluded participants, twin births etc.) **Automated Self-Administered 24-Hour (ASA24®) Dietary Assessment Tool

Outcome measures

Table 2 summarizes the wide range of different neuropsychological instruments that were used across the studies to assess cognitive and behavioural functions. A total of 18 original instruments were used in addition to one self-developed instrument, comprising both questionnaires and neuropsychological tests.

Table 2: Overview of outcomes assessment methods

Reference	Cognitive outcome assessment*†	Affective outcome assessment*†	Child age at assessment
Barker et al, 2013[30]	WISC-III[n]		8 years
Bernard et al, 2013[43]	MCDI[q], ASQ[q]		2 and 3 years
Bolduc et al, 2016[44]	BSID-III (cognitive subscale)[n]		1 year
Daniels et al, 2004[31]	MCDI[q], DDST[n]		1.25 and 1.5 years
Davidson et al, 2008[32]	BSID-II (Psychomotor Developmental Index)[n]		2.5 years
Gale et al, 2008[33]	WASI[n]	SDQ[q]	9 years
Gustafsson et al, 2016[45]		IBQ-R[q]	4 months
Hibbeln et al, 2007[34]	WISC-III[n],	SDQ[q]	7 years
Jacka et al, 2013[35]		CBCL[q]	1.5 years
Julvez et al, 2016[46]	BSID (mental and psychomotor developmental index)[n], MCSA[n]		14 months (BSID) & 5 years (MCSA)
Mendez et al, 2008[36]	MCSA[n]		4 years
Oken, Radesky et al, 2008[37]	PPVT[n], WRAVMA[n]		3 years
Oken, Østerdal et al, 2008[38]	Self-developed instrument (9 q's regarding developmental milestones)[q]		1.5 years
Oken et al, 2016[47]	WRAML[n], KBIT-II [n]		7.7 years**
Pina-Camacho et al, 2015[39]		CITS[q]	2 years
Sagiv et al, 2012[42]	WISC-III[n]	CRS-T[q]	8 years
Steenweg-de Graaff et al, 2014[40]		CBCL[n]	3 years
Valent et al, 2013[41]		BSID-III (socio emotional subscale)[n]	1.5 years

* Outcomes are administered either as a questionnaire (q) or neuropsychological test (n)

**median age in years of children at outcome assessment

†Full name of instruments in alphabetical order: ASQ: Ages and stages questionnaire, BSID: Bayley Scales of Infant Development, CAST: Childhood Asperger Syndrome Test, CBCL: Child behaviour checklist, CITS: Carey infant temperament scale, CRS-T: Conners rating scale – teacher, DDST: Denver Developmental Screening Test, IBQ-R: Revised infant behaviour questionnaire, KBIT-II: Kaufman Brief Intelligence Test 2nd edition, MCDI: MacArthur Communicative Development Inventory, MCSA: McCarthy Scales of Children's Abilities, PDI: Psychomotor Developmental Index, PPVT: Peabody Picture Vocabulary Test, SDQ: Strength and Difficulties Questionnaire, WASI: Wechsler Abbreviated Scale of Intelligence, WISC: Wechsler Intelligence Scale for Children, WRAML: Wide Range Assessment of Memory and Learning, WRAVMA: Wide Range Assessment of Visual Motor Abilities

Confounders

Overall, the studies controlled for a number of different factors, depicted in supplementary table 2. Studies varied greatly in which confounders they included in the analysis, with SES being the only confounder considered by all studies.

Individual study quality assessment

Each study was evaluated with the NOS checklist, please see supplementary table 3 for individual study scoring information. Of the 18 included studies, 9 were rated as of "fair" quality and 9 as of "poor" quality. No study received the rating "good" because none of the studies that used high quality

measurements also adequately dealt with self-selection into studies and selective dropout of participants. All "poor" ratings were due to insufficiencies in the "Outcome" dimension of the NOS in studies that measured outcomes through self-reports (independent blind assessments or record linkage is preferred by the NOS) and that additionally did not account for selective dropout between exposure and outcome assessments.

Computation of Effect Sizes

None of the included studies reported Hedges' g as their effect size. The compute.es package[48] was used to calculate Hedges' g for the studies reporting the following: 1) For odds ratio (OR) Hedges' g was calculated using the "lores" function (based on log of OR and its corresponding variance).; 2) for p-values (with information of group sample sizes), Hedges' g was calculated using the "pes" function; 3) for correlation coefficient (r), Hedges' g was calculated using the "res" function.

Some studies did not report all the required information to calculate Hedges' g in the compute.es package. For the studies reporting OR, where group sample sizes were not reported, the method for converting OR to Cohens' d proposed by Chinn[49] was used. For mean difference in standardized test score, Hedges' g was calculated using standard deviation (SD) and mean difference. For regression coefficient (β), Hedges' g was calculated using p-value, standard error (SE) and z-statistics. For Cohens' d, Hedges' g was calculated using the formula proposed by Lakens.[50] For studies lacking p-value, confidence intervals (CIs), standard deviations or standard errors, the required inferential statistics were calculated in advance of the final effect size estimation by using appropriate formulas.[51-53]

Summarizing effect sizes

After Hedges' g had been calculated, the effect sizes were further summarized as many studies reported several effect sizes for the same exposure-outcome combination. Preferably, only one effect size per study should be retained,[22] however this was considered inapplicable, due to the large variations in neurodevelopmental outcomes assessed in the different studies. After careful consideration, four outcome dimensions (externalizing, internalizing, socio-emotional, cognitive) covering the affective and general cognitive domains were chosen. Selection of the outcome measures into each respective domain was based on 1) a thorough review of the properties of each instrument with regards to what area of development the instrument is aimed at measuring based on the manual for each instrument, and 2) research indicating that language, cognition, and executive functions are more strongly correlated with each other than with affective functioning.[54, 55]

Additionally, we applied the following rules to reduce the number of effect sizes per study, aiming to obtain only one effect size per outcome dimension per study:

1. For fish/seafood intake, if more than two intake groups had been defined, we only included the group which best corresponded to what is considered a healthy diet by the national health authorities, which is 2-3 servings per week for total fish intake, where about half should be fatty fish[11]
2. If studies reported statistics for a total score as well as sub scales of an outcome measure, only the effect size for the total score was included in the analyses
3. If the effect sizes were based on sample stratification (e.g. by breastfeeding duration), measures were collapsed. Collapsed effect sizes were calculated as weighted means of Hedges' g over comparisons, whereby weights depended on the n of each comparison
4. If OR was calculated for both high and low test score in the original article, only the effect sizes corresponding to the high test score was retained
5. If effect sizes were reported for all types of fish as well as oily fish - only all types of fish were included in the final analysis, as to make the exposure definition as homogenous as possible

6. If studies reported associations with both an unhealthy and a healthy dietary pattern, the effect size for the unhealthy dietary pattern was reversed and then averaged with the healthy dietary pattern
7. For studies reporting several eligible effect sizes for an outcome dimension, the effect size based on the most valid measurement instrument were used (e.g. effect size from a neuropsychological test vs questionnaire or a validated questionnaire vs a not validated questionnaire)
8. For studies that report more than one effect size relevant for inclusion in one domain based on the same type of instrument, the average effect size across those reported were included after Hedges g had been calculated
9. For reported effect sizes where corresponding outcome dimensions were unclear (e.g. based on inadequate reporting of instrument properties) and not resolved with discussion among the reviewers, these effect sizes were excluded. This occurred for three separate instruments, each from different studies[43, 44, 46]

Application of these rules resulted in a total of 26 separate effect sizes. Only the fully adjusted effect sizes from each study were chosen. We initially considered including the corresponding unadjusted effect sizes for each study, however only four studies provided this information,[33, 35, 41, 42] and the studies that reported minimally adjusted results adjusted for different variables.

The 18 studies included in the final meta-analysis comprised a total of 63861[3] participants and 26 separate effect sizes divided into four different cognitive or affective dimensions. These effect sizes with corresponding confidence intervals (CI) are depicted in the forest plot in Figure 2[4]. The size of each square reflects the precision of the effect size estimate by means of the weight that is assigned to each respective study when the summary effect size is computed. A larger square equals larger weight assigned to that study.

For studies reporting more than one effect size, the appropriate statistical techniques were applied so that this was accounted for when calculating the summary effect size. When calculating the summary effect size in the two domains separately, a study that reported effect sizes for one cognitive and one affective outcome would have its originally calculated standard errors. When calculating the summary effect size across domains, a re-calculated standard error was used (the N across the two effect sizes were averaged and then divided by two (((N1 +N2)/2)/2), before calculating the adjusted variance and subsequently standard error. For the studies with more than two reported effect sizes across domains, the denominators were equivalent to the total number of effect sizes included per study). Hence, one study could be allocated different weights, depending on the analysis. In summary, whenever multiple effect sizes from one study entered the analysis, those were weighted correctly by adjusting the standard errors of the effect sizes accordingly.

Figure 2 approximately here

A REM was fit for each of the four dimensions, as well as for the two overall domains. Table 4 provides a summary of the REM for the cognitive and different affective dimensions, including results for the summary effect size and test for heterogeneity (Q- and I^2-statistics). The effect sizes are typically

[3] As more than one effect size per study was incorporated into the meta-analysis, with sometimes differing n between neurocognitive domains, the largest n for each study was used as a basis for total n.
[4] The data in Figure 2, 3 and 4 is based on adjusted standard errors to account for studies contributing with multiple outcomes as described in the "summarizing effect sizes" section

larger for the cognitive domain (uncorrected; g=0.14), compared to the affective domain (uncorrected; g=0.093). The average uncorrected effect size across both domains is g=0.112.

Table 4: Random Effects Model statistics, with test for overall summary effect size (Hedges g) and test for heterogeneity, including publication bias

Original REM	n studies	n participants	Hedges g	SE	z	p-value	df	Q	Qp	I^2 (%)
Affective*		38219†	0.093	0.029	3.2	0.0012	12	67	<0.0001	77
Externalizing	6	37517	0.113	0.033	3.4	0.0008	5	33	<0.0001	82
Internalizing	5	29437	0.018	0.036	0.6	0.5219	4	14	0.0065	68
Socio-emotional	2	3752	0.196	0.030	5.7	<0.0001	1	0.06	0.806	0
Cognitive	13	29269	0.140	0.017	8.2	<0.0001	12	12	0.445	27
Summary effect size	26	63861†	0.112	0.019	5.9	<0.0001	25	102	<0.0001	69
With "Trim and fill"										
Affective*	14††	38219†	0.088	0.028	3.1	0.0018	13	68	<0.0001	76
Cognitive	13	29269	0.140	0.017	8.2	<0.0001	12	12	0.445	27
Summary effect size	35††	63861†	0.075	0.019	3.9	<0.0001	34	150	<0.0001	75
With standard error as covariate										**R^2 (%)****
Affective*	13	38219†	0.043	0.041	1.1	0.2935	11	57	<0.0001	19
Cognitive	13	29269	0.132	0.031	4.3	<0.0001	11	12	0.3779	0
Summary effect size	26	63861†	0.080	0.029	2.7	0.0062	24	93	<0.0001	7

* Overall domain for all the affective dimensions (Externalizing, Internalizing and Socio-emotional)
** Heterogeneity accounted for by standard error
† Not the sum across relevant dimensions as some studies are included in several dimensions
†† includes original and imputed studies

Heterogeneity

As can be seen from the Q and I^2 statistics in Table 4, there is a large and significant degree of heterogeneity present for the overall summary effect size, indicating a systematic difference in effect sizes between the studies, but only pertaining to the studies included in the affective domain. As possible sources of this heterogeneity, we investigated publication bias and performed a moderator analysis.

Publication bias

The Egger's regression test for funnel plot asymmetry was not significant (p=0.161), but this might be mainly due to low power. To get a visualisation of possible publication bias, a funnel plot is depicted in Figure 3[4].

Figure 3 approximately here

If no publication bias was present, approximately 95 % of the points for the original effect sizes should be located within the white funnel area[25] and should be roughly distributed evenly to the left and to the right of the vertical line illustrating the overall summary effect size. Because this is not the case, a trim and fill analysis was performed, and the results are displayed in Figure 4[4].

Figure 4 approximately here

The imputed effect sizes (open circles) are all smaller than the summary effect size, and the trim and fill analyses suggest that the adjusted overall summary effect size would be 0.075 (c.f., Table 4). It is still significant ($p<0.0001$), but smaller than the originally calculated summary effect size ($g=0.112$). Additionally, even with the imputed effect sizes, there are still significant levels of heterogeneity present, which indicate that other factors than publication bias are contributing to the observed heterogeneity. Table 4 also shows that only the studies in the affective domain appear to be afflicted by publication bias, as the summary effect size for the cognitive domain remains unchanged with the trim and fill analysis. The results from regression-based adjustment for publication bias are consistent with the trim and fill analysis in that they show a similar overall effect size, a clear association in the cognitive domain, and a noticeably weaker association in the affective domain.

Moderator analysis

Considering that performing a moderator analysis is generally not advisable with less than ten studies,[53] we performed a moderator analysis for the whole sample of studies, rather than separately for the affective domain. If significant moderators were found, they would, at least in part, explain the heterogeneity present in the affective domain, considering that there were no significant levels of heterogeneity present between the studies in the cognitive domain. The following moderators were included in single-predictor models: publication year and diet category. Instrument category was initially considered, but only questionnaires were utilized in the studies within the affective domain and this was therefore deemed unnecessary. Only diet category was found to have a significant moderating effect, explaining approximately 35% of the heterogeneity ($p=0.0051$) where the effect sizes seemed to be systematically larger for the studies where a proxy for maternal diet quality (e.g. fish consumption) was used. However, there was still a significant degree of heterogeneity present, indicating that other moderators not considered in the model were influencing the outcome effect sizes.

DISCUSSION

The aim of this meta-analysis was to systematically review and summarize the currently existing literature about the association between maternal diet quality and different child neurodevelopmental outcomes. When dietary exposures believed to be appropriate proxies for maternal diet quality were included, a total of 18 studies comprising 63861 participants were found relevant for inclusion in this meta-analysis.

The meta-analysis showed that a better maternal diet quality had a small, statistically significant association with child neurodevelopment. The summary effect size for the cognitive domain was larger than the overall summary effect size, with no significant presence of heterogeneity. This positive association with cognitive outcomes is in line with findings from a recent narrative review investigating the association between maternal fish intake and child cognitive outcomes.[56] The important contribution of our quantitative meta-analysis is the calculation of average effect sizes, which shows, also after correcting for publication bias, a small but robust association. The summary effect size for the affective domain was smaller than the overall summary effect size, with a large and significant degree of heterogeneity present. Considering that an overall summary effect size is most appropriate to use for studies with little heterogeneity,[22] the summary effect size should be interpreted with caution. If we look at the effect sizes for all four outcome dimensions we find that maternal diet quality is associated with all neurodevelopmental dimensions except for the internalizing dimension, with the strongest associations seen for socio-emotional and general cognitive functioning. However, these effect sizes are still considered small according to Cohens interpretative guidelines.[57]

In the moderator analysis, only type of dietary classification (diet category - dietary pattern or its proxies (fish intake, fruit intake, saturated fat intake or Ω-6/Ω-3 fatty acid ratio)) contributed significantly to the heterogeneity present in the total sample of studies, explaining 35% of the heterogeneity. However, a large degree of heterogeneity remained. As only the fully adjusted effect sizes from each study were included in the meta-analysis, unmeasured or unreported variables may have contributed to the remaining heterogeneity. Additionally, the moderator analysis may have underestimated the amount of heterogeneity explained by the diet category, as studies varied considerably in how they classified the dietary intake, particularly with regards to fish intake.

The majority of the studies included in this meta-analysis using proxies of a maternal dietary pattern during pregnancy had fish or seafood intake as their exposure measure. Although fish intake most likely is a good marker for diet quality, there are several limitations involved. Firstly, the studies investigating fish intake varied greatly with regards to intake group definitions; with division into two, three or four groups, where most groups were compared to a reference group (generally those who never or rarely consumed fish), or included as a continuous variable in a regression model. Some studies also compared extreme groups (lowest vs highest quintile), which were the studies reporting the largest effect sizes. Due to this varying dietary exposure definition it is likely that the amount of heterogeneity the diet category accounts for is underestimated in the moderator analysis. Ideally, we could have used a more elaborate classification of categories, to reflect the actual diversity of the exposure measures, but this was not appropriate considering the small number of studies included in this meta-analysis.[58]

As seen from the funnel plot in Figure 3, there is a clear negative correlation between effect sizes and standard error, indicating that the larger the sample size, the smaller the association between maternal diet quality and the outcome measure. This is not surprising, considering that the effect size of a study with a small sample needs to be large to reach significance in comparison to studies with a large sample size where only very small effect sizes are required to reach statistical significance. However, even if the observed pattern has a statistical explanation, a clear visual indication of publication bias remains. Accordingly, analyses that corrected for publication bias through a trim and fill procedure and meta-regression resulted in overall effect size estimates that were around 30% lower compared to uncorrected overall effect sizes.

Maternal diet – direct effect or marker for child diet?

An important issue that cannot be resolved by this meta-analysis is whether the observed association is based on direct effects of maternal diet quality or whether it is a marker for the child's diet, which is a competing exposure that also influences child development. Not surprisingly, maternal diet quality, as well as maternal post-natal diet, and child diet during infancy and early toddlerhood have been found to be highly correlated.[59-61] Therefore, it remains possible that the observed associations between maternal diet quality and child development are due to the child's diet after pregnancy. Ultimately, the interpretation of the reported effect sizes depends on the assumed causal model. If it is assumed that child diet is a mediator between maternal diet quality and child development, then one has to control for child diet if interested in the direct effect of maternal diet quality, and one must not control for child diet if interested in the total effect of maternal diet quality.[62, 63] However, we suggest that maternal diet quality and child diet have a common cause—e.g. parental education—and an unbiased estimate of the direct effects of maternal diet quality on child development requires controlling for child diet. One mitigating fact is that child diet varies with sociodemographic variables[61] so that controlling for maternal postnatal diet and sociodemographic factors is likely to, at least in part, control for child diet. Still, child diet should ideally be assessed as a distinct factor. Indeed, the only study[35] among the three

studies that reported controlling for the child's diet,[35, 36, 39] which also provided unadjusted and adjusted effect sizes, found that the association between maternal diet quality and both externalizing and internalizing problems in the child was mediated by child diet, reducing the effect of maternal diet quality in three out of four analyses.

An important aspect of child diet during the earliest stages of life is breastfeeding. Previous studies exploring the association between breastfeeding and different cognitive development measures have found associations between longer breastfeeding duration and better general cognitive development,[64-66] higher IQ,[67] better educational attainment[68] and language development,[69] as well as a lower risk of having ADHD.[70] Nine of the studies included in this meta-analysis adjusted for breastfeeding duration,[31, 33, 34, 37, 40, 41, 43, 46, 47] but none provided information on both unadjusted and adjusted effect sizes related to breastfeeding specifically. However, some indicated that breastfeeding did not display a significant confounding effect.

In studies that stratified their sample by breastfeeding practices,[36, 43] a significant association with maternal diet quality was only seen when the child had been breastfed for less than 6 months. This suggests that a better maternal diet quality might serve as a protective or beneficial factor to a larger degree for children who are not breastfed or breastfed less than the recommended period.

Maternal obesity is another factor related to both breastfeeding and the outcomes. Studies consistently show reduced breastfeeding rates in obese mothers[71, 72] and obesity is also linked to impaired cognition and increased behavioural problems in children.[73-75] This highlights the importance of also accounting for maternal BMI in the analysis, but only four studies did this in their analyses.[37, 38, 40, 41]

The effect of maternal diet quality on child development is an exemplary research topic where causal knowledge has to be extracted from observational studies because experimental studies are either unethical or impractical.[76] However, if causal information is to be gleaned from observational data, care must be taken to control for biases due to e.g. self-selection into studies or selective drop-out, for example by using inverse probability weights so that the effective sample better resembles the target population. Importantly, controlling by adding covariates is typically not sufficient to control for e.g. selection bias.[77] The appraisal of the summarized studies with the NOS suggests that while all summarized studies controlled, to varying degrees, for potential confounders by adding covariates to their analysis, systematic control for selection bias is not yet part of routine analysis.

Limitations

Only the fully adjusted effect sizes from each study were selected for inclusion into the meta-analysis, however, the confounders considered by each study varied greatly. This adds to the uncertainty of the results.

Only four of the studies incaluded in this meta-analysis used maternal dietary patterns as their exposure measure for diet quality. Even though investigation of dietary patterns has received increasing interest over the past decade, there is still a lack of research in this area, particularly in relation to child cognitive outcomes. Additionally, there should be more focus on dietary patterns as opposed to global indices of healthy diets, as these indices inherently assume substitutability of different aspects of healthy diets, which might not be valid assumptions.

In the majority of the studies included with fish or seafood intake as the exposure measure, fish intake was not a part of the main exposure or primary investigation – it was often included as a covariate, with the main study focus being investigation of mercury exposure. This, together with the varying definition of fish intake between the studies, probably contributes substantially to the observed

heterogeneity and illustrates the limitations involved in conducting a meta-analysis on studies with very heterogeneous exposure measures.

None of the studies reporting more than one effect size included information on the corresponding correlation between effect sizes. It is likely to assume that at least some of the effect sizes within each study were interdependent as they are measuring different aspects of the same overall cognitive or affective domain. It is widely recommended to aggregate dependent effect sizes to avoid biased estimates,[22] but by grouping the effect sizes into different outcome domains, rather than computing one summary effect size, this was somewhat accounted for.

Lastly, there are many challenges relating to the FFQ as a measurement tool, which is well-known within the nutritional research field.[78, 79] This is mainly due to different types of bias that can arise from using self-report measures of dietary intake, which creates further difficulties in relation to analysis and interpretation of dietary data.

Taken together, these limitations suggest that while the effect sizes reported here provide some information about the association between maternal diet quality and child cognitive and affective outcomes, more research is needed to obtain reliable estimates of such associations.

CONCLUSION

Comparing studies looking at an overall maternal diet quality rather than specific nutrients brings with it many challenges, mainly due to heterogeneous methods for measuring intake, failure to account for child diet during early childhood, as well as the vast number of confounders needed to be considered, both genetic and environmental.

Additionally, the number of studies available for inclusion in this meta-analysis is limited and they are heterogeneous, both with regards to exposure and outcome measures, indicating that results should be interpreted with caution. However, the results point in the direction that a better maternal diet quality is weakly, but robustly associated with a more favourable cognitive development and fewer affective problems in the child.

Lastly, it is important to emphasize that the studies included in this meta-analysis are all observational and do often take only limited steps towards causal identification. Therefore, causal interpretations of the results have to be avoided.

Suggestions for future research

The results of this meta-analysis highlight the need for more research on the effects of maternal diet quality on child cognitive and affective outcomes. The heterogeneity present in this sample of studies, particularly with regards to the definition of the exposure measures, makes comparison of results across studies particularly challenging. To better enable for between-study comparisons in the future, careful consideration should be taken to develop standardized instruments for the measurement of diet quality, which can be applicable, with minor modifications, across different populations. Additionally, studies should aim for the use of validated and recognized instruments for the measurement of cognitive and behavioural outcomes, rather than self-developed or obsolete instruments. With regards to the outcome measures, available reliability data for the instruments used by the included studies does not provide any clear distinction in test-retest reliability between questionnaires and clinical tests,[80-94] but a comparison of effect sizes for outcomes assessed through questionnaires and clinical neuropsychological tests, within the same study, will help to settle this important issue.

Furthermore, it is crucial that future studies investigating the effect of maternal diet quality on child neurodevelopmental outcomes also consider the child's diet, as failure to recognize child diet as an

important contributing factor limits the interpretability of such studies. More generally, more attention to and controlling of confounders, potential competing exposures, and potential bias due to self-selection into studies or selective drop out will be important to better justify a causal interpretation of observational studies.

Finally, greater emphasis should be put on research transparency by means of describing the methodology used more exhaustively and by reporting complete results for both significant and non-significant results.

Contributors: TCB, GB, and HA designed the study. TCB and ALB prepared the data in conjunction with GB. TCB and GB conducted the statistical analysis. TCB drafted the manuscript and had the primary responsibility for the final content. All authors critically reviewed, read and approved the final version of the manuscript. TCB is the guarantor.

Competing interests: None

Funding: The present study was supported by a grant from NevSom; National Centre for Neurodevelopmental Disorders and Hypersomnia - ADHD, Autism, Tourette's syndrome and Hypersomnia (grant no. 51379 – 001/15-473).

Data sharing statement: Table with raw effect measures from included studies and R script of all analysis performed for this meta-analysis is available upon request from the corresponding author Tiril Cecilie Borge; tibo@fhi.no

REFERENCE LIST

1. Barker DJP. Mothers, babies and health in later life. 2nd ed. Edinburgh: Churchill Livingstone 1998.
2. Harding JE. The nutritional basis of the fetal origins of adult disease. *Int J Epidemiol* 2001;30: 15-23 [published Online First: 2001/02/15].
3. Heindel JJ, Vandenberg LN. Developmental origins of health and disease: a paradigm for understanding disease cause and prevention. *Curr Opin Pediatr* 2015;27: 248-53 doi: 10.1097/mop.0000000000000191 [published Online First: 2015/01/31].
4. Thapar A, Cooper M, Jefferies R, et al. What causes attention deficit hyperactivity disorder? *Arch Dis Child* 2012;97: 260-65
5. Roseboom TJ, Painter RC, van Abeelen AF, et al. Hungry in the womb: what are the consequences? Lessons from the Dutch famine. *Maturitas* 2011;70: 141-5 doi: 10.1016/j.maturitas.2011.06.017 [published Online First: 2011/08/02].
6. Prado EL, Dewey KG. Nutrition and brain development in early life. *Nutr Rev* 2014;72: 267-84 doi: 10.1111/nure.12102 [published Online First: 2014/04/02].
7. Zeisel SH. Is maternal diet supplementation beneficial? Optimal development of infant depends on mother's diet. *Am J Clin Nutr* 2009;89: 685s-7s doi: 10.3945/ajcn.2008.26811F [published Online First: 2009/01/01].
8. Englund-Ogge L, Brantsaeter AL, Sengpiel V, et al. Maternal dietary patterns and preterm delivery: results from large prospective cohort study. *Br Med J* 2014;348: 1446-64 doi: 10.1136/bmj.g1446 [published Online First: 2014/03/13].
9. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol* 2002;13: 3-9
10. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery* 2010;8: 336-41 doi: http://dx.doi.org/10.1016/j.ijsu.2010.02.007
11. Norwegian National Nutrition Council. Diet recommendations to promote public health and prevent chronic diseases - Methodology and evidence foundation. Oslo: Norwegian Health Directorate, 2011.

12. Simopoulos AP. The importance of the omega-6/omega-3 fatty acid ratio in cardiovascular disease and other chronic diseases. *Exp Biol Med (Maywood)* 2008;233: 674-88 doi: 10.3181/0711-mr-311 [published Online First: 2008/04/15].
13. Hooper L, Martin N, Abdelhamid A, et al. Reduction in saturated fat intake for cardiovascular disease. *Cochrane Database Syst Rev* 2015: Cd011737 doi: 10.1002/14651858.cd011737 [published Online First: 2015/06/13].
14. Anderson JW, Baird P, Davis RH, et al. Health benefits of dietary fiber. *Nutr Rev* 2009;67: 188-205 doi: 10.1111/j.1753-4887.2009.00189.x [published Online First: 2009/04/02].
15. Balder HF, Virtanen M, Brants HA, et al. Common and country-specific dietary patterns in four European cohort studies. *J Nutr* 2003;133: 4246-51 [published Online First: 2003/12/04].
16. Knudsen VK, Orozova-Bekkevold IM, Mikkelsen TB, et al. Major dietary patterns in pregnancy and fetal growth. *Eur J Clin Nutr* 2008;62: 463-70 doi: 10.1038/sj.ejcn.1602745 [published Online First: 2007/03/30].
17. Willett WC, Sacks F, Trichopoulou A, et al. Mediterranean diet pyramid: a cultural model for healthy eating. *Am J Clin Nutr* 1995;61: 1402S-06S [published Online First: 2013/02/13].
18. Guenther PM, Casavale KO, Reedy J, et al. Update of the Healthy Eating Index: HEI-2010. *J Acad Nutr Diet* 2013;113: 569-80 doi: 10.1016/j.jand.2012.12.016
19. Wells G, Shea B, O'Connell D, et al. The Newcastle-Ottawa scale (NOS) for assessing the quailty of nonrandomised studies in meta-analyses. 2009. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
20. The Cochrane Collaboration. 13.5.2.3 Tools for assessing methodological quality or risk of bias in non-randomized studies. In: Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011] ed, 2011.
21. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol* 2010;25: 603-5 doi: 10.1007/s10654-010-9491-z [published Online First: 2010/07/24].
22. Del Re A. A Practical Tutorial on Conducting Meta-Analysis in R. *Quant Meth Psych* 2015;11: 37-50
23. Viechtbauer W. Meta-Analysis Package for R; 2015. https://cran.r-project.org/web/packages/metafor/index.html
24. Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. *J Stat Softw* 2010;36: 1-48 doi: 10.18637/jss.v036.i03 [published Online First: 2010/08/05].
25. Borenstein M, Hedges LV, Higgins JPT, et al. Introduction to Meta-Analysis. Chicester, United Kingdom: Wiley 2009.
26. Siddiqi N. Publication bias in epidemiological studies. *Cent Eur J Public Health* 2011;19: 118-20 [published Online First: 2011/07/12].
27. Sterne JAC, Harbord RM. Funnel plots in meta-analysis. *SJ* 2004;4: 127-41
28. Egger M, Smith GD, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315: 629
29. Moreno SG, Sutton AJ, Ades AE, et al. Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. *BMC Med Res Methodol* 2009;9: 2 doi: 10.1186/1471-2288-9-2 [published Online First: 2009/01/14].
30. Barker ED, Kirkham N, Ng J, et al. Prenatal maternal depression symptoms and nutrition, and child cognitive function. *British Journal of Psychiatry* 2013;203: 417-21
31. Daniels JL, Longnecker MP, Rowland AS, et al. Fish intake during pregnancy and early cognitive development of offspring. *Epidemiology* 2004;15: 394-402 doi: 10.1097/01.ede.0000129514.46451.ce
32. Davidson PW, Strain JJ, Myers GJ, et al. Neurodevelopmental effects of maternal nutritional status and exposure to methylmercury from eating fish during pregnancy. *Neurotoxicology* 2008;29: 767-75 doi: 10.1016/j.neuro.2008.06.001 [published Online First: 2008/06/11].
33. Gale CR, Robinson SM, Godfrey KM, et al. Oily fish intake during pregnancy - Association with lower hyperactivity but not with higher full-scale IQ in offspring. *J Child Psychol Psychiatry* 2008;49: 1061-68 doi: 10.1111/j.1469-7610.2008.01908.x [published Online First: 2008/04/15].
34. Hibbeln JR, Davis JM, Steer C, et al. Maternal seafood consumption in pregnancy and neurodevelopmental outcomes in childhood (ALSPAC study): an observational cohort study. *The Lancet* 2007;369: 578-85 doi: 10.1016/S0140-6736(07)60277-3 [published Online First: 2007/02/15].
35. Jacka FN, Ystrom E, Brantsaeter AL, et al. Maternal and early postnatal nutrition and mental health of offspring by age 5 years: a prospective cohort study. *J Am Acad Child Adolesc Psychiatry* 2013;52: 1038-47 doi: 10.1016/j.jaac.2013.07.002 [published Online First: 2013/10/01].
36. Mendez MA, Torrent M, Julvez J, et al. Maternal fish and other seafood intakes during pregnancy and child neurodevelopment at age 4 years. *Public Health Nutr* 2009;12: 1702-10 doi: 10.1017/S1368980008003947 [published Online First: 2008/11/25].

37. Oken E, Radesky JS, Wright RO, et al. Maternal fish intake during pregnancy, blood mercury levels, and child cognition at age 3 years in a US cohort. *Am J Epidemiol* 2008;167: 1171-81 doi: 10.1093/aje/kwn034 [published Online First: 2008/03/18].
38. Oken E, Østerdal ML, Gillman MW, et al. Associations of maternal fish intake during pregnancy and breastfeeding duration with attainment of developmental milestones in early childhood: a study from the Danish National Birth Cohort. *Am J Clin Nutr* 2008;88: 789-96 [published Online First: 2010/05/24].
39. Pina-Camacho L, Jensen SK, Gaysina D, et al. Maternal depression symptoms, unhealthy diet and child emotional-behavioural dysregulation. *Psychol Med* 2015;45: 1851-60 doi: 10.1017/s0033291714002955 [published Online First: 2014/12/19].
40. Steenweg-de Graaff J, Tiemeier H, Steegers-Theunissen RP, et al. Maternal dietary patterns during pregnancy and child internalising and externalising problems. The Generation R Study. *Clin Nutr* 2013;33: 115-21 doi: 10.1016/j.clnu.2013.03.002 [published Online First: 2013/03/14].
41. Valent F, Mariuz M, Bin M, et al. Associations of Prenatal Mercury Exposure From Maternal Fish Consumption and Polyunsaturated Fatty Acids With Child Neurodevelopment: A Prospective Cohort Study in Italy. *J Epidemiol* 2013;23: 360-70 doi: 10.2188/jea.JE20120168 [published Online First: 2013/08/10].
42. Sagiv SK, Thurston SW, Bellinger DC, et al. Prenatal exposure to mercury and fish consumption during pregnancy and attention-deficit/hyperactivity disorder-related behavior in children. *Arch Pediatr Adolesc Med* 2012;166: 1123-31 doi: 10.1001/archpediatrics.2012.1286 [published Online First: 2012/12].
43. Bernard JY, De Agostini M, Forhan A, et al. The Dietary n6:n3 Fatty Acid Ratio during Pregnancy Is Inversely Associated with Child Neurodevelopment in the EDEN Mother-Child Cohort. *J Nutr* 2013;143: 1481-88 doi: 10.3945/jn.113.178640 [published Online First: 2013/07/31].
44. Bolduc FV, Lau A, Rosenfelt CS, et al. Cognitive Enhancement in Infants Associated with Increased Maternal Fruit Intake During Pregnancy: Results from a Birth Cohort Study with Validation in an Animal Model. *Ebiomedicine* 2016;8: 331-40 doi: 10.1016/j.ebiom.2016.04.025
45. Gustafsson HC, Kuzava SE, Werner EA, et al. Maternal dietary fat intake during pregnancy is associated with infant temperament. *Dev Psychobiol* 2016;58: 528-35
46. Julvez J, Mendez M, Fernandez-Barres S, et al. Maternal Consumption of Seafood in Pregnancy and Child Neuropsychological Development: A Longitudinal Study Based on a Population With High Consumption Levels. *Am J Epidemiol* 2016;183: 169-82 doi: 10.1093/aje/kwv195
47. Oken E, Rifas-Shiman SL, Amarasiriwardena C, et al. Maternal prenatal fish consumption and cognition in mid childhood: Mercury, fatty acids, and selenium. *Neurotoxicol Teratol* 2016;57: 71-78
48. Del Re A. Compute Effect Sizes (Package "compute.es"); 2014. https://cran.r-project.org/web/packages/compute.es/
49. Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. *Stat Med* 2000;19: 3127-31 [published Online First: 2000/12/13].
50. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front Psychol* 2013;4: 863 doi: 10.3389/fpsyg.2013.00863 [published Online First: 2013/12/11].
51. Altman DG, Bland JM. How to obtain the confidence interval from a P value. *Br Med J* 2011;343 doi: 10.1136/bmj.d2090 [published Online First: 2011/08/08].
52. Altman DG, Bland JM. How to obtain the P value from a confidence interval. *Br Med J* 2011;343 doi: 10.1136/bmj.d2304 [published Online First: 2011/08/08].
53. The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions, Higgins J, Green S, editors; 2011. www.cochrane-handbook.org
54. Sjowall D, Backman A, Thorell LB. Neuropsychological heterogeneity in preschool ADHD: investigating the interplay between cognitive, affective and motivation-based forms of regulation. *J Abnorm Child Psychol* 2015;43: 669-80 doi: 10.1007/s10802-014-9942-1 [published Online First: 2014/09/23].
55. Gutierrez-Cobo MJ, Cabello R, Fernandez-Berrocal P. The Relationship between Emotional Intelligence and Cool and Hot Cognitive Processes: A Systematic Review. *Front Behav Neurosci* 2016;10: 101 doi: 10.3389/fnbeh.2016.00101 [published Online First: 2016/06/16].
56. Starling P, Charlton K, McMahon AT, et al. Fish intake during pregnancy and foetal neurodevelopment-a systematic review of the evidence. *Nutrients* 2015;7: 2001-14 doi: 10.3390/nu7032001 [published Online First: 2015/03/21].
57. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, N. J: Laurence Erlbaum 1988.
58. Hedges LV, Pigott TD. The power of statistical tests for moderators in meta-analysis. *Psychol Methods* 2004;9: 426-45 doi: 10.1037/1082-989x.9.4.426 [published Online First: 2004/12/16].

59. Ashman AM, Collins CE, Hure AJ, et al. Maternal diet during early childhood, but not pregnancy, predicts diet quality and fruit and vegetable acceptance in offspring. *Matern Child Nutr* 2014 doi: 10.1111/mcn.12151 [published Online First: 2014/10/08].
60. Fisk CM, Crozier SR, Inskip HM, et al. Influences on the quality of young children's diets: the importance of maternal food choices. *Br J Nutr* 2011;105: 287-96 doi: 10.1017/s0007114510003302 [published Online First: 2010/09/03].
61. Emmett PM, Jones LR, Northstone K. Dietary patterns in the Avon Longitudinal Study of Parents and Children. *Nutr Rev* 2015;73 Suppl 3: 207-30 doi: 10.1093/nutrit/nuv055 [published Online First: 2015/09/24].
62. Hernán MA, Robins JM. Causal Inference. Boca Raton: Chapman & Hall/CRC, forthcoming; 2016. https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/
63. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999;10: 37-48 [published Online First: 1999/01/15].
64. Julvez J, Guxens M, Carsin A-E, et al. A cohort study on full breastfeeding and child neuropsychological development: The role of maternal social, psychological, and nutritional factors. *Dev Med Child Neurol* 2014;56: 148-56 doi: 10.1111/dmcn.12282 [published Online First: 2013/10/01].
65. Quigley M, Hockley C, Carson C, et al. Breastfeeding is Associated with Improved Child Cognitive Development: A Population-Based Cohort Study. *J Pediatr* 2012;160: 25-32 doi: 10.1016/j.jpeds.2011.06.035 [published Online First: 2011/08/11].
66. Kramer MS, Aboud F, Mironova E, et al. Breastfeeding and Child Cognitive Development: New Evidence From a Large Randomized Trial. *Arch Gen Psychiatry* 2008;65: 578-84 doi: 10.1001/archpsyc.65.5.578 [published Online First: 2008/05].
67. Brion M-JA, Lawlor DA, Matijasevich A, et al. What are the causal effects of breastfeeding on IQ, obesity and blood pressure? Evidence from comparing high-income with middle-income cohorts. *Int J Epidemiol* 2011;40: 670-80 doi: 10.1093/ije/dyr020 [published Online First: 2011/02/24].
68. Groen-Blokhuis MM, Franic S, van Beijsterveldt CE, et al. A prospective study of the effects of breastfeeding and FADS2 polymorphisms on cognition and hyperactivity/attention problems. *Am J Med Genet B Neuropsychiatr Genet* 2013;162: 457-65 doi: 10.1002/ajmg.b.32175 [published Online First: 2013/06/04].
69. Smith JM. Breastfeeding and language outcomes: A review of the literature. *J Commun Disord* 2015;57: 29-40 doi: 10.1016/j.jcomdis.2015.04.002 [published Online First: 2015/05].
70. Stadler DD, Musser ED, Holton KF, et al. Recalled initiation and duration of maternal breastfeeding among children with and without ADHD in a well characterized case-control sample. *J Abnorm Child Psychol* 2016;44: 347-55 doi: 10.1007/s10802-015-9987-9 [published Online First: 2016/02].
71. Newby RM, Davies PS. Antenatal breastfeeding intention, confidence and comfort in obese and non-obese primiparous Australian women: associations with breastfeeding duration. *Eur J Clin Nutr* 2016;70: 935-40 doi: 10.1038/ejcn.2016.29 [published Online First: 2016/03/24].
72. Bever Babendure J, Reifsnider E, Mendias E, et al. Reduced breastfeeding rates among obese mothers: a review of contributing factors, clinical considerations and future directions. *Int Breastfeed J* 2015;10: 21 doi: 10.1186/s13006-015-0046-5 [published Online First: 2015/07/04].
73. Pugh SJ, Hutcheon JA, Richardson GA, et al. Gestational weight gain, prepregnancy body mass index and offspring attention-deficit hyperactivity disorder symptoms and behaviour at age 10. *BJOG* 2016 doi: 10.1111/1471-0528.13909 [published Online First: 2016/03/22].
74. Pugh SJ, Richardson GA, Hutcheon JA, et al. Maternal Obesity and Excessive Gestational Weight Gain Are Associated with Components of Child Cognition. *J Nutr* 2015;145: 2562-9 doi: 10.3945/jn.115.215525 [published Online First: 2015/10/02].
75. Veena SR, Gale CR, Krishnaveni GV, et al. Association between maternal nutritional status in pregnancy and offspring cognitive function during childhood and adolescence; a systematic review. *BMC Pregnancy Childbirth* 2016;16: 220 doi: 10.1186/s12884-016-1011-z [published Online First: 2016/08/16].
76. Rothman KJ. Epidemiology : an introduction. Oxford: Oxford University Press 2002.
77. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;15: 615-25 [published Online First: 2004/08/17].
78. Freedman LS, Schatzkin A, Midthune D, et al. Dealing With Dietary Measurement Error in Nutritional Cohort Studies. *JNCI Journal of the National Cancer Institute* 2011;103: 1086-92 doi: 10.1093/jnci/djr189
79. Kipnis V, Midthune D, Freedman L, et al. Bias in dietary-report instruments and its implications for nutritional epidemiology. *Public Health Nutr* 2002;5: 915-23 doi: 10.1079/phn2002383 [published Online First: 2003/03/14].
80. Adams W, Sheslow D. Wide Range Assessment of Visual Motor Abilities (WRAVMA). Novato, CA: Academic Therapy Publications 1995.

81. Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986;51: 1173-82 doi: 10.1037/0022-3514.51.6.1173
82. Bayley N. Bayley Scales of Infant and Toddler Development, Technical Manual. 3rd ed. San Antonio TX: Harcourt Assessment 2006.
83. Bilenberg N. The Child Behavior Checklist (CBCL) and related material: standardization and validation in Danish population based and clinically based samples. *Acta Psychiatr Scand Suppl* 1999;398: 2-52
84. Canning EH, Kelleher K. Performance of screening tools for mental health problems in chronically ill children. *Arch Pediatr Adolesc Med* 1994;148: 272-8
85. Fenson L, Dale PS, Reznick JS, et al. The MacArthur Communicative Development Inventories: User's guide and technical manual. San Diego: Singular Publishing Group 1993.
86. Frankenburg WK, Dodds J, Archer P, et al. The Denver II: A Major Revision and Restandardization of the Denver Developmental Screening Test. *Pediatrics* 1992;89: 91-97
87. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. *J Am Acad Child Adolesc Psychiatry* 2001;40: 1337-45 doi: 10.1097/00004583-200111000-00015 [published Online First: 2001/11/09].
88. McCarthy D. Manual for the McCarthy scales of children's abilities. New York: The Psychological Corporation 1972.
89. McDevitt SC, Carey WB. The measurement of temperament in 3--7 year old children. *J Child Psychol Psychiatry* 1978;19: 245-53
90. Novik TS. Validity of the Child Behaviour Checklist in a Norwegian sample. *Eur Child Adolesc Psychiatry* 1999;8: 247-54
91. Squires J, Twombly E, Bricker D, et al. Ages and Stages Questionnaire - Technical report. Baltimore, MD: Brookes Publishing Co. Inc 2009.
92. Stone LL, Otten R, Engels RC, et al. Psychometric properties of the parent and teacher versions of the strengths and difficulties questionnaire for 4- to 12-year-olds: a review. *Clin Child Fam Psychol Rev* 2010;13: 254-74 doi: 10.1007/s10567-010-0071-2 [published Online First: 2010/07/01].
93. Williams KT, Wang J. Technical references to the Peabody Picture Vocabulary Test- Third Edition (PPVT-III). Circle Pines, MN: American Guidance Service 1997.
94. Williams PE, Weiss LG, Rolfhus EL. WISC–IV Technical Report #2 Psychometric Properties: The Psychological Corporation, 2003.

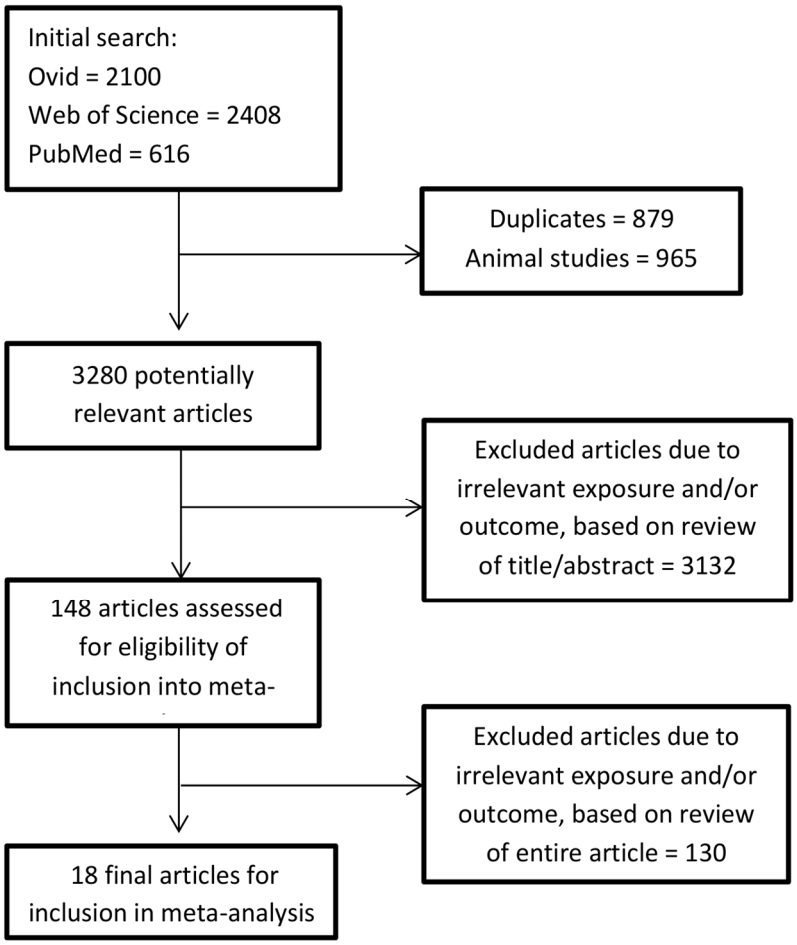

Figure 1: Flow diagram of the study selection process

139x169mm (300 x 300 DPI)

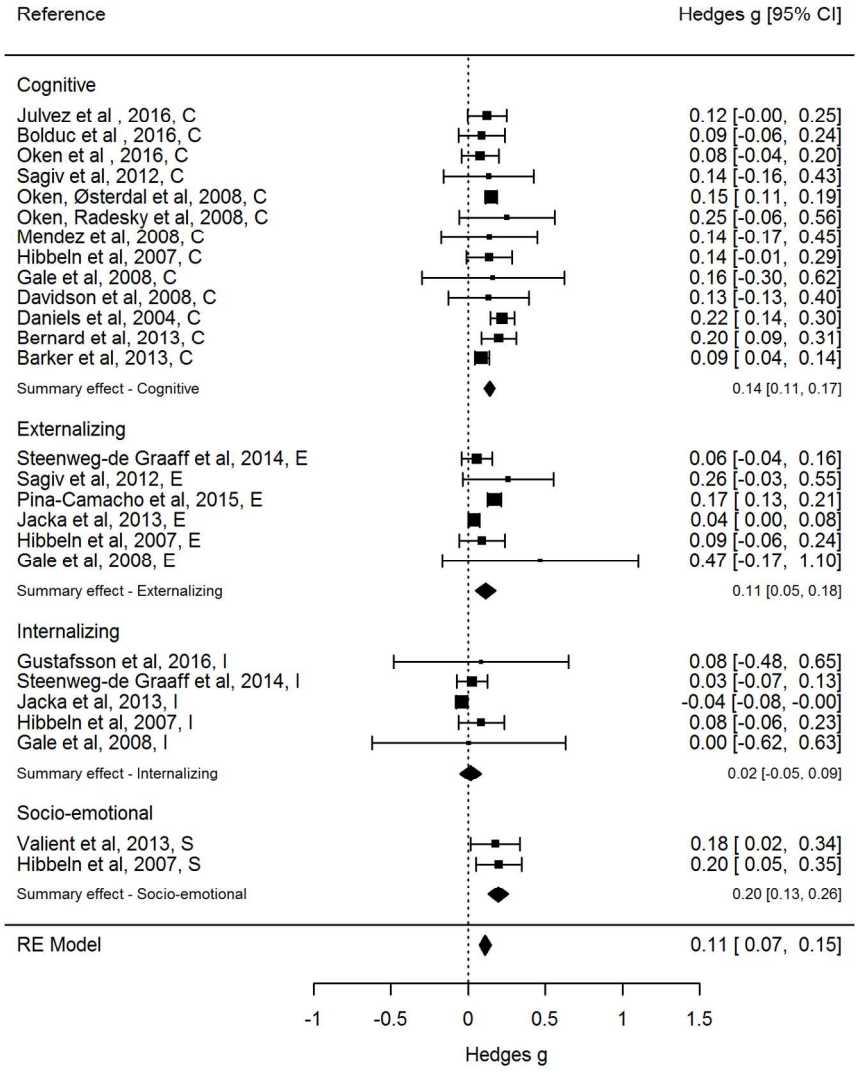

Figure 2: Forest plot of REM of included studies and their respective effect size, with summary effect size for the cognitive and affective domains, as well as an overall summary effect size

127x169mm (300 x 300 DPI)

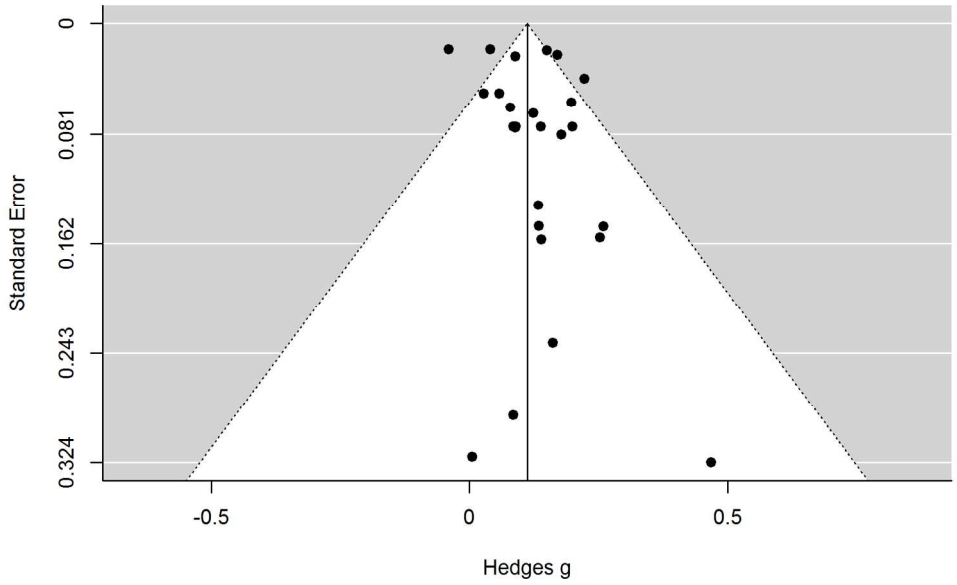

Figure 3: Funnel plot of the REM of included studies and their respective effect size and standard errors

169x118mm (300 x 300 DPI)

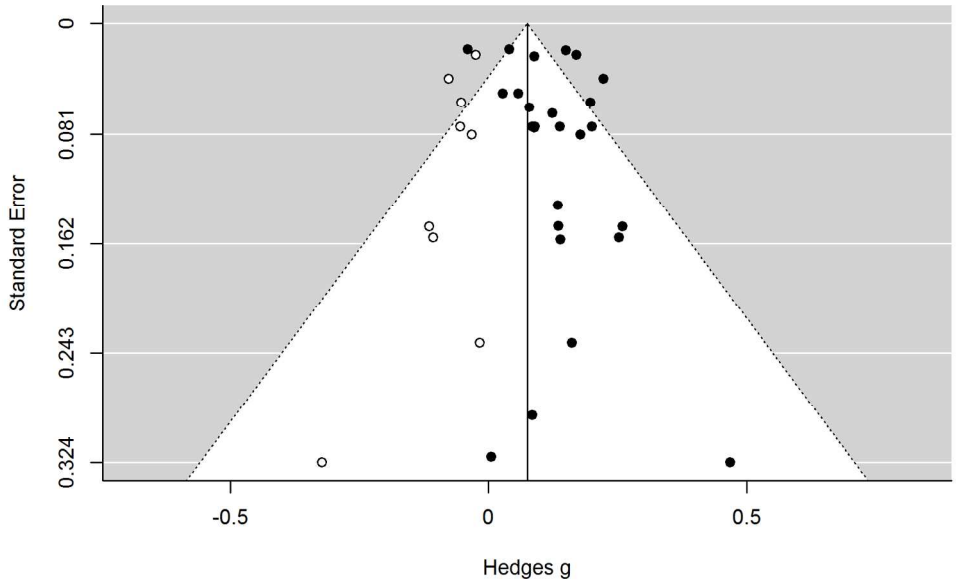

Figure 4: Funnel plot based on trim and fill analysis, showing original studies (closed circle) and imputed studies (open circle)

169x118mm (300 x 300 DPI)

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	3
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	3
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	3-4
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	4
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	4
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	4
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	4
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	4
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	4
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	4
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	4-5

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	5
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	5
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	5-6
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	6-8
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	8-9
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	9-10
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	11
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	11-12
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	11-12
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	12-14
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	14-15
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	15-16
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	16

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

Supplementary table 1: Example of full search string (Ovid)

	Searches	Results
1	(Maternal* or prenatal* or perinatal* or gestational* or pregnan*)	1768624
2	(diet*3 or nutrition or fiber or fibre or protein or fat or fatty or carbohydrate or fruit or vegetable or fish or seafood)	7878460
3	child* or toddler* or offspring	3862890
4	(behavior?r or behavior?r disorder* or externali?ing or internali?ing or mental health or mental development or learning disorder* or cogniti*3 or neurocogniti*3 or memory or IQ or executive or ADHD or attention deficit hyperactivity disorder or Attention deficit disorder or oppositional defiant disorder or conduct disorder or development* disability or neurodevelopment* or autism spectrum disorder or hyperkinetic disorder or hyperactivity disorder or language or communication or affective or developmental milestone*)	6446778
5	Combine 1-4	11512
6	Limit 5 to appropriate age group (infant – 12 years of age)	8003
7	Limit 6 to pregnancy	4632
8	Limit 7 to humans	3316
9	Limit to original articles	2974
10	Remove duplicates from 9	2100

Supplementary table 1: Overview of control variables for each study

Confounders	Barker et al (2013)	Bernard et al (2013)	Bolduc et al (2016)	Daniels et al (2004)	Davidson et al (2008)	Gale et al (2008)	Gustafsson et al (2016)**	Hibbeln et al (2007)	Jacka et al (2013)	Julvez et al (2016)	Mendez et al (2008)	Oken, Radensky et al (2008)	Oken et al (2016)	Oken, Østerdal et al (2008)	Pina-Camacho et al (2015)	Sagiv et al (2012)	Steenweg-de Graaff et al (2014)	Valent et al (2013)
Alcohol		x		x		x		x				x		x		x		x
Biomarkers																x		
Birth complications	x														x			
Breastfeeding				x		x		x		x	x	x	x	x				x
Child gender		x		x	x		x	x	x	x	x	x	x	x		x	x	x
Child ADHD medication																x		
Child age at assessment		x		x						x	x	x	x	x		x	x	
Child birth weight					x	x	x			x			x	x				x
Child dietary pattern									x									
Child fish intake											x							x
Child sugary snacks/drinks intake																	x	
Daycare attendance		x				x												x
Ethnicity								x		x		x	x			x	x	x
Fetal growth												x						
Home environment	x			x	x			x					x		x	x		x
Length of gestation		x						x			x	x	x	x				
Marital status	x				x			x	x			x		x	x	x	x	x
Maternal age		x	x		x	x	x	x	x	x		x		x		x	x	x
Maternal energy intake		x					x			x							x	
Maternal diet*			x					x								x		x
Maternal gestational diabetes			x															
Maternal IQ					x	x							x			x		x
Maternal mental health							x		x					x		x	x	
Maternal pre pregnancy BMI												x		x			x	x
Maternal pregnancy weight/weight gain		x					x			x								x
Maternal supplement use			x														x	
Parental learning difficulties														x				
Parity	x	x		x				x		x	x		x	x	x		x	
Paternal age									x									
SES	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Smoking		x		x		x		x	x			x	x	x		x	x	x

*For some studies where fish intake was the exposure, other maternal dietary components were included in the analysis as confounders.

**All covariates did not significantly correlate with outcome, so these were not included in the final analysis

Supplementary table 3: Evaluation of individual study quality with The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analyses

		Barker et al (2013)	Bernard et al (2013)	Bolduc et al (2016)	Daniels et al (2004)	Davidson et al (2008)	Gale et al (2008)	Gustafsson et al (2016)	Hibbeln et al (2007)	Jacka et al (2013)	Julvez et al (2016)	Mendez et al (2008)	Oken, Radensky et al (2008)	Oken, Østerdal et al (2008)	Oken et al (2016)	Pina-Camacho et al (2015)	Sagiv et al (2012)	Steenweg-de Graaff et al (2014)	Valent et al (2013)
Selection	**1) Representativeness of the exposed cohort:** a) truly representative of the average pregnant woman in the community*; b) somewhat representative of the average pregnant woman in the community*; c) selected group of users e.g. nurses, volunteers; d) no description of the derivation of the cohort	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c
	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	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	c	a	a
	3) Ascertainment of exposure: a) secure record (e.g. surgical records)*; b) structured interview*; c) written self-report; d) no description	c	c	c	c	a	c	c	c	c	c	c	c	c	c	c	a	c	c
	4) Demonstration that outcome of interest was not present at start of study: a) yes*; b) no	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	c	a	a
Comparability	**1) Comparability of cohorts on the basis of the design or analysis:**																		
	a) study controls for SES (maternal education and/or income)*	x	x	x	x	x	x	x	x	x	x	x	x	x	a	x	x	x	x
	b) study controls for child dietary factors other than breastfeeding (e.g. dietary patterns, fish intake)*									x		x				x	x		
Outcome	**1) Assessment of outcome:** a) independent blind assessment*; b) record linkage*; c) self-report; d) no description	a	c	a	c	a	c	c	a/c	c	c	a	a	c	a	c	a	c	a
	2) Was follow-up long enough for outcomes to occur: a) yes*; b) no	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
	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 - less than 20 % lost or description of those lost suggested no difference from those followed*; c) follow up rate < 80% and no description of those lost; d) no statement	c	c	c	c	c	c	c	c	c	c	d	c	c	d	c	c	c	c
Total number of stars		**5**	**4**	**5**	**4**	**5**	**4**	**4**	**5**	**5**	**4**	**6**	**5**	**4**	**5**	**5**	**6**	**4**	**5**
Quality rating according to guideline**		**fair**	**poor**	**fair**	**poor**	**fair**	**poor**	**poor**	**fair**	**poor**	**poor**	**fair**	**fair**	**poor**	**fair**	**poor**	**fair**	**poor**	**fair**

*=one star (marked in yellow when each respective study were given a star)

**Thresholds for converting the NOS rating to Agency for Healthcare Research and Quality - AHRQ - standards (good, fair, and poor):

Good quality: 3 or 4 stars in Selection domain AND 1 or 2 stars in Comparability domain AND 2 or 3 stars in Outcome domain

Fair quality: 2 stars in Selection domain AND 1 or 2 stars in Comparability domain AND 2 or 3 stars in Outcome domain

Poor quality: 0 or 1 star in Selection domain OR 0 stars in Comparability domain OR 0 or 1 stars in Outcome domain

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

BMJ Open

The Importance of Maternal Diet Quality During Pregnancy on Cognitive and Behavioural Outcomes in Children – A Systematic Review and Meta-Analysis

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-016777.R1
Article Type:	Research
Date Submitted by the Author:	24-May-2017
Complete List of Authors:	Borge, Tiril; Nasjonalt folkehelseinstitutt, Department of Child Health Aase, Heidi; Nasjonalt folkehelseinstitutt, Department of Child Development Brantsaeter, Anne-Lise; Nasjonalt folkehelseinstitutt, Department of Environmental exposure and Epidemiology Biele, Guido; Nasjonalt folkehelseinstitutt, Department of Child Health
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Mental health, Nutrition and metabolism, Public health
Keywords:	EPIDEMIOLOGY, MENTAL HEALTH, NUTRITION & DIETETICS, Child & adolescent psychiatry < PSYCHIATRY, PUBLIC HEALTH

TITLE PAGE

Title: The Importance of Maternal Diet Quality During Pregnancy on Cognitive and Behavioural Outcomes in Children – A Systematic Review and Meta-Analysis

Corresponding author:
Tiril Cecilie Borge
Norwegian Institute of Public Health, Department of Child Health
P.O. Box 4404 Nydalen
0403 OSLO
Norway

Email address of the corresponding author: tibo@fhi.no
Tel: +4797613762

Co-authors (in correct author order):
Heidi Aase, Norwegian Institute of Public Health, Department of Child Development, Oslo, Norway
Anne Lise Brantsæter, Norwegian Institute of Public Health, Department of Environmental Exposure and Epidemiology, Oslo, Norway
Guido Biele, Norwegian Institute of Public Health, Department of Child Health, Oslo, Norway

Keywords: maternal diet quality, pregnancy, cognitive development, affective functioning, systematic, review, meta-analysis

Word count: 6443
Number of figures: 4
Number of tables: 3
Number of references: 95
Number of supplementary files for online only publication: 1

ABSTRACT

Objectives: This systematic review and meta-analysis provides a quantitative summary of the literature exploring the relationship between maternal diet quality during pregnancy and child cognitive and affective outcomes. We investigate whether there are indications for robust associations and aim to identify methodological strengths and challenges of the current research to provide suggestions of improvement for future research.

Design and participants: Relevant studies were identified through a systematic literature search in relevant databases. All studies investigating maternal diet quality during pregnancy in relation to child cognitive or affective functioning in children of elementary school age or younger were assessed for inclusion.

Results: 18 relevant studies, comprising 63861 participants were identified. The results indicated a small positive association between better maternal diet quality during pregnancy and child functioning. We observed publication bias and significant heterogeneity between studies, where type of diet classification, publication year and outcome domain together accounted for about 30% of this heterogeneity. Trim and fill analysis substantiated the presence of publication bias for studies in the affective domain and showed an adjusted effect size of Hedge's g=0.088 (p=0.0018) [unadjusted g=0.093 (p=0.03)]. We observed no publication bias in the cognitive domain, where results indicated a slightly larger effect size (g=0.14 (p<0.0001)) compared to that of the affective domain. The overall summary effect size was g=0.075 (p<0.0001) adjusted for publication bias [unadjusted g=0.112 (p=0.0001)]. Child diet was not systematically controlled for in the majority of the studies.

Conclusion: The results indicated that a better maternal diet quality during pregnancy has a small positive association with child neurodevelopment, with more reliable results seen for cognitive development. These results warrant further research on the association between maternal diet quality during pregnancy and cognitive and affective aspects of child neurodevelopment, whereby it is crucial that future studies account for child diet in the analysis.

Strengths and limitations of this study

- This is the first article to summarize research into the association between maternal diet quality during pregnancy and child neurodevelopment
- Major strengths of this research are the use of meta-analytic methods for calculation of average effect sizes and investigation of publication bias
- This study highlights strengths and challenges of an emerging research field, thus building the foundation for improved future research
- A limitation is the relatively small number of relevant studies identified for inclusion in the meta-analysis
- Since this meta-analysis is based on observational studies, no strong causal interpretations about the association of maternal diet quality during pregnancy and child neurodevelopment can be made

INTRODUCTION

The importance of adequate nutrition during foetal life for long-term physical health is well documented.[1, 2] However, the relationship between maternal nutrition during pregnancy and child mental health is less established.[3] The prenatal environment is crucial in relation to cognitive development of the child, particularly during critical periods of brain development, which highlights the foetus' need for optimal nutrition.[4] There are documented detrimental effects of severe maternal malnutrition during pregnancy,[5] and severe deficiencies of certain micronutrients, like iron and iodine[6] on child neurodevelopment and general cognitive functions, as well as severe deficiencies of folate and choline on child neural tube defects,[7] but the impact of more subtle variations in maternal diet quality[1] on child neurodevelopment has received little attention until recently.

It has become increasingly recognized that investigating the impact of diet on most disease outcomes cannot be done solely by investigating single nutritional components separately. Considering that the human diet consists of at least 25 000 biologically active components, with only a fraction having been defined as nutrients, it is likely that the majority of the dietary constituents effect human health in an interdependent manner. Looking at overall diet quality, e.g. through dietary patterns, is believed to represent a valid and meaningful measure of overall nutrient intake[8] and is a promising approach when aiming to study diet related associations.[9]

To date, no meta-analysis has summarized research on maternal diet quality and child neurodevelopment. As the research interest for this topic is rapidly increasing, it is valuable to summarize the research to date on this topic using statistical procedures. The aim of this meta-analysis is to provide a quantitative summary of the existing literature exploring the relationship between maternal diet quality and child cognitive and affective outcomes. The goals are to investigate whether there are indications for robust associations, despite the limited amount of studies available, and to identify methodological strengths and challenges of the current research to provide suggestions of improvement for future research.

METHODS

As a scientific guideline for this manuscript, we followed the PRISMA statement.[10]

Defining exposure and outcome measures

Despite the increased interest in studying dietary patterns as a measure for diet quality, the current literature regarding the associations between maternal dietary patterns during pregnancy and child neurodevelopmental outcomes is sparse. A preliminary literature search resulted in only four articles with defined maternal dietary patterns as exposure relevant for inclusion into the meta-analysis. Consequently, in order to increase the basis for analysis, articles with dietary exposures believed to be good proxies for maternal diet quality were included. Based on the existing literature, fish intake,[11] Ω-6/Ω-3 fatty acid ratio,[12] saturated fat intake[13] and dietary fibre (reflecting intake of whole-grain foods, vegetables, fruits, legumes, and nuts)[14] were considered good proxies for maternal diet quality. In previous research, based on dietary data from large cohorts where dietary patterns have been identified with data driven methods, consumption of fish and fibre rich foods, as well as limited intake of saturated fats, have consistently been associated with a healthier dietary pattern, both in the general population,[15] and in pregnant women.[8, 16] Additionally, fibre rich foods, fat quality and fish are incorporated into established healthy food indices, like the Mediterranean diet index[17] and the Healthy Eating Index (HEI-2010).[18] Considering the already limited amount of available relevant literature, few limitations were put on the possible outcome as long as it covered a child

[1] When using the term "maternal diet quality" in this paper we are always referring to the maternal diet quality during pregnancy, unless otherwise stated

neurodevelopmental domain, like cognition (IQ and language) or affect (externalizing and internalizing difficulties).

Search criteria and strategies

An extensive search string was developed as to not exclude any relevant literature, and adapted to each database, including key words relating to maternal diet quality, child mental health, cognitive function (language, communication skills, IQ) neurodevelopmental disorders (ADHD, ODD, CD, ASD) and affective functioning. The following exclusion criteria were applied: children with very low birth weight; children older than elementary school age; and studies focusing on single micronutrients and/or supplements. The search string was developed by TCB in collaboration with a specialist librarian. For full search string, see supplementary table 1.

Data collection and extraction process

After identification of original articles through the initial search, excluding duplicates, TCB and ALB independently screened title and/or abstract of each study. TCB's and ALB's final list of eligible and possibly eligible studies were then crosschecked and read in full text by both. If both reviewers were unsure whether an article was eligible for inclusion, GB was consulted to assure coherence regarding the final selection of articles for inclusion. After identification of the eligible articles the relevant information from each study was extracted by TCB (e.g. year of publication, total number of participants, dietary exposure and outcome measures assessed, confounders controlled for, and reported effect sizes) in collaboration with GB and ALB. TCB and GB then assessed individual study quality with the Newcastle-Ottawa Scale (NOS) for assessing the quality of cohort studies in meta-analysis.[19]

Individual study quality assessment

For each eligible study included in the meta-analysis we performed an individual study quality assessment using the NOS.[19] The NOS provides an easy to use study quality checklist and is recognized by Cochrane.[20] The scoring system is based on the assessment of three aspects of a study; Selection (representativeness of cohort and exposure assessment); Comparability (ascertainment of confounding); and Outcome (assessment of outcome and follow-up). The scoring system categorizes studies as being of good, fair or poor methodological quality, whereby insufficiency in one of the domains results in a "poor" rating. While the NOS has been criticized for an overly general definition of quality criteria,[21] this generality allows for a wide application of the scale. Moreover, the intent of the scale is clear: A good rating of the Selection dimensions requires a representative sample and high quality measurement; a good rating of the Comparability dimension requires control of appropriate confounders; and a good rating of the Outcome dimensions requires a high quality measurement of outcomes and/or high follow up rates or correction for non-random drop out.

Analysis of reported effect sizes

To be able to compare the results of the studies, association measures reported in each individual study had to be transformed into a standardized effect size[2]. The effect size measure utilized for this meta-analysis was Hedges' g, which is a more conservative effect size measure compared to Cohen's d.[22] Effect sizes were calculated to reflect the association between better maternal diet quality and the different cognitive and affective outcomes, where a positive value indicate a better outcome, hence better language development or general cognitive functioning, or less affective problems.

The meta-analysis was conducted with the R statistical software (version 3.2.2.), using the Metafor package, version 1.9-8.[23, 24] Because the included studies were heterogeneous with regards

[2] When referring to "effect size" in this paper we are not indicating causality – it is merely the statistical term of the reported outcome measures.

to neurodevelopmental outcomes, choice of statistical procedures, and effect sizes, we used a random-effects model (REM) for analyses.[22] This approach models variance between the included studies, and assumes that observed differences in effect sizes are due to both sampling error and true effect size differences in the studies' background populations.[22]

A restricted maximum-likelihood method for estimation of heterogeneity was used to compute relevant Q-statistics, with corresponding I^2-statistics. The Q-statistics indicate whether there is statistically significant heterogeneity across the studies' effect sizes, whereas the I^2-statistics indicate the extent of heterogeneity. A significant Q-statistic indicates systematic (as opposed to random) variation of effect sizes between studies.

Possible moderators

If the REM analyses indicate presence of heterogeneity, moderator analyses can be used to investigate potential causes of this heterogeneity. We used meta-regressions where we added potential moderator variables individually to separate regression models in order to assess their effect on the association between exposure and outcome. The following factors were available for consideration as possible moderators: Publication year, child age at assessment, outcome domain (cognitive, affective), diet category (type of diet classification - whether the exposure is defined as maternal dietary pattern or a proxy for maternal dietary pattern) and instrument category (measurement of outcome - questionnaire or neuropsychological test), as they are all factors which might moderate the association between exposure and outcome. The categorical factors were dichotomous.

Publication bias

Publication bias describes a situation in which the decision to publish research results depends on obtaining statistically significant results.[25] Indeed, studies reporting statistically significant results are more likely to be published than studies reporting results that are not statistically significant.[26] One visual meta-analytic tool traditionally used to investigate publication bias is the funnel plot.[27] To complement the potential subjectivity of visual inspection of funnel plots the Egger's regression test for funnel plot asymmetry[28] can be performed.

In the presence of publication bias, a "trim and fill" approach can be used to correct for it. The trim and fill method uses effect sizes and their standard error to generate a "complete" distribution of effect sizes that likely would have been reported without publication bias by adding imputed studies to the reported studies. If the average effect sizes calculated from published and "complete" effect sizes do not differ noticeably, one can have more confidence in the average effect size from a group of studies that appear afflicted by publication bias.[25] As there is some discussion in the literature regarding the optimal methods for adjustment for publication bias,[29] a meta-regression to adjust for publication bias using the standard error of effect sizes as a moderator should also be performed.

RESULTS

Study sample and selection

Ovid (Embase, Psychinfo, Medline), PubMed and ISI: WEB of science were searched on November 16th 2016 using the full search string. Additionally, we explicitly searched for relevant studies that employed a priori dietary quality indices (like the HEI-2010 and Mediterranean diet index).

A total of 18 studies fit the inclusion criteria,[30-47] the majority reporting several outcome measures. The study selection process is visualized in Figure 1.

Figure 1 approximately here

All included studies were observational in nature and based on a prospective cohort design or case-control design, with baseline measures of maternal dietary intake during pregnancy and subsequent measurement of child cognitive or affective functioning, at one or more time points.

All studies collected information on maternal dietary intake with the use of a food frequency questionnaire (FFQ), either self-administered or by a trained interviewer, with some using validated FFQs; and/or a food-diary. Data obtained with these instruments was used the basis for the definition of dietary patterns, estimation of fish/seafood intake, fruit intake, saturated fat intake and estimation of Ω-6/Ω-3 fatty acid ratio based on intake.

Four studies used maternal dietary patterns as exposure variables,[30, 35, 39, 40] either defined by the use of principal component analysis (PCA) or confirmatory factor analysis (CFA). In three studies,[30, 35, 40] two distinct dietary patterns were identified; one "healthy" and one "unhealthy", while only an unhealthy dietary pattern was defined in the fourth.[39] The healthy dietary patterns were generally characterized by higher intakes of vegetables, fish, legumes, wholegrains and vegetable oils, while the unhealthy dietary patterns consisted of higher intakes of processed foods (fried foods, French fries, meats) confectionary foods (cakes, candy, sugary drinks), refined cereals, and salty snacks.

Eleven studies used maternal fish intake as exposure,[31-34, 36-38, 41, 42, 46, 47] where the studies categorized fish intake into groups based on meals/portions or grams eaten per day or week. The remaining three studies used Ω-6/Ω-3 fatty acid ratio,[43] saturated fat intake[45] and fruit intake[44] as their exposure variable. All studies were published in the period from 2004-2016, with study populations ranging from 48 to 23020 mother-child pairs (Table 1).

Table 1: Description of the studies and assessment of the exposure

Reference	Dietary data collection period	Country	Cohort name	Total n *	Dietary assessment	FFQ items	Exposure variable	Exposure variable scale
Barker et al, 2013[30]	1990-1992	UK	ALSPAC	6979	FFQ	103	"Healthy" and "unhealthy" patterns, grouped using CFA	Dietary patterns as continuous variable
Bernard et al, 2013[43]	2003-2005	France	Eden cohort	1335	FFQ	137	Estimated maternal intake of total n6 and total n3 (in g/d), then calculation of n6:n3 fatty acid ratio.	Fatty acid ratio as continuous variable
Bolduc et al, 2016[44]	2008-2012	Canada	CHILD Edmonton sub-cohort	688	FFQ	175	Total fruit intake estimated from FFQ	Maternal fruit intake as continuous predictor of neurodevelopment
Daniels et al, 2004[31]	1990-1992	UK	ALSPAC	7421	FFQ	103	Maternal fish intake as servings/week	Categorical variable; Four intake groups (of which one referent group)
Davidson et al, 2008[32]	2001-2002	Seychelles	Seychelles Child Development Study	229	FFQ and 4-day diet diary	NS	Maternal fish intake as g/day	Maternal fish intake as continuous variable
Gale et al, 2008[33]	1991-1992	UK	NS	217	FFQ	100	Maternal fish intake as times eaten/week	Categorical variable; Three intake groups (of which one referent group)
Gustafsson et al, 2016[45]	NS	USA	Ongoing longitudinal study	48	ASA24**	NS	Total and saturated fat intake	Fat intake as continuous variable
Hibbeln et al, 2007[34]	1990-1992	UK	ALSPAC	11875	FFQ	103	Maternal seafood intake as g/week	Categorical variable; Three intake groups (of which one referent group)
Jacka et al, 2013[35]	2002-2008	Norway	MoBa	23020	validated FFQ	255	"Healthy" and "unhealthy" data driven dietary pattern scores	Dietary patterns scores as continuous variables
Julvez et al, 2016[46]	2004-2008	Spain	INMA	1892	FFQ	101	Maternal seafood intake	Maternal seafood intake as Quintiles
Mendez et al, 2008[36]	1997-1998	Menorca	Prospective birth cohort	392	FFQ	42	Maternal fish intake	Categorical variable; Four intake groups (of which one referent group)
Oken, Radesky et al, 2008a[37]	1999-2002	USA	Project Viva	341	validated FFQ	>140	Maternal fish intake	Categorical variable; Three intake groups (of which one referent group)
Oken, Østerdal et al, 2008[38]	1997-2003	Denmark	Danish birth cohort	25446	Validated FFQ	>360	Maternal fish intake as servings(g)/week	Categorical variable; Three quintiles of intake (lowest, middle, highest)
Oken et al, 2016[47]	1999-2002	USA	Project Viva	1068	validated FFQ	>140	Maternal fish intake	Categorical variable; Three intake groups (of which one (no fish) referent group)
Pina-Camacho et al, 2015[39]	1990-1992	UK	ALSPAC	7814	FFQ	103	General unhealthy diet (second-order latent factor) generated with CFA	Unhealthy dietary pattern as continuous variable
Sagiv et al, 2012[42]	1993-1998	USA	New Bedford cohort	362	FFQ	NS	Maternal fish intake, expressed as total servings/week	Maternal fish intake as continuous variable
Steenweg-de Graaff et al, 2014[40]	2001-2006	Netherlands	Generation R	3104	Validated FFQ	293	"Healthy" and "unhealthy" data driven dietary pattern scores	Dietary patterns scores as continuous variables
Valent et al, 2013[41]	2007-2009	Italy	Set within project (PHIME).	606	FFQ (Adapted from a validated FFQ)	138	Maternal fish intake as servings/week	Maternal fish intake as continuous variable

* differs from total n used in analysis (due to missing data, excluded participants, twin births etc.) **Automated Self-Administered 24-Hour (ASA24®) Dietary Assessment Tool

Outcome measures

Table 2 summarizes the wide range of different neuropsychological instruments that were used across the studies to assess cognitive and behavioural functions. A total of 18 original instruments were used in addition to one self-developed instrument, comprising both questionnaires and neuropsychological tests.

Table 2: Overview of outcomes assessment methods

Reference	Cognitive outcome assessment*†	Affective outcome assessment*†	Child age at assessment
Barker et al, 2013[30]	WISC-III[n]		8 years
Bernard et al, 2013[43]	MCDI[q], ASQ[q]		2 & 3 years
Bolduc et al, 2016[44]	BSID-III (cognitive subscale)[n]		1 year
Daniels et al, 2004[31]	MCDI[q],		1.25 & 1.5 years
Davidson et al, 2008[32]	BSID-II (Psychomotor Developmental Index)[n]		2.5 years
Gale et al, 2008[33]	WASI[n]	SDQ[q]	9 years
Gustafsson et al, 2016[45]		IBQ-R[q]	4 months
Hibbeln et al, 2007[34]	WISC-III[n],	SDQ[q]	7 years
Jacka et al, 2013[35]		CBCL[q]	1.5 years
Julvez et al, 2016[46]	BSID (mental and psychomotor developmental index)[n], MCSA[n]		14 months & 5 years
Mendez et al, 2008[36]	MCSA[n]		4 years
Oken, Radesky et al, 2008[37]	PPVT[n], WRAVMA[n]		3 years
Oken, Østerdal et al, 2008[38]	Self-developed instrument (9 q's regarding developmental milestones)[q]		1.5 years
Oken et al, 2016[47]	WRAML[n], KBIT-II [n]		7.7 years**
Pina-Camacho et al, 2015[39]		CITS[q]	2 years
Sagiv et al, 2012[42]	WISC-III[n]	CRS-T[q]	8 years
Steenweg-de Graaff et al, 2014[40]		CBCL[n]	3.5 years
Valent et al, 2013[41]		BSID-III (socio emotional subscale)[n]	1.5 years

* Outcomes are administered either as a questionnaire (q) or neuropsychological test (n)

**median age in years of children at outcome assessment

†Full name of instruments in alphabetical order: ASQ: Ages and stages questionnaire, BSID: Bayley Scales of Infant Development, CAST: Childhood Asperger Syndrome Test, CBCL: Child behaviour checklist, CITS: Carey infant temperament scale, CRS-T: Conners rating scale – teacher, DDST: Denver Developmental Screening Test, IBQ-R: Revised infant behaviour questionnaire, KBIT-II: Kaufman Brief Intelligence Test 2nd edition, MCDI: MacArthur Communicative Development Inventory, MCSA: McCarthy Scales of Children's Abilities, PDI: Psychomotor Developmental Index, PPVT: Peabody Picture Vocabulary Test, SDQ: Strength and Difficulties Questionnaire, WASI: Wechsler Abbreviated Scale of Intelligence, WISC: Wechsler Intelligence Scale for Children, WRAML: Wide Range Assessment of Memory and Learning, WRAVMA: Wide Range Assessment of Visual Motor Abilities

Confounders

Overall, the studies controlled for a number of different factors, depicted in supplementary table 2. Studies varied greatly in which confounders they included in the analysis, with SES being the only confounder considered by all studies.

Individual study quality assessment

Each study was evaluated with the NOS checklist, please see supplementary table 3 for individual study scoring information. Of the 18 included studies, 9 were rated as of "fair" quality and 9 as of "poor" quality. No study received the rating "good" because none of the studies that used high quality measurements also adequately dealt with self-selection into studies and selective dropout of participants. All "poor" ratings were due to insufficiencies in the "Outcome" dimension of the NOS in studies that measured outcomes through self-reports (independent blind assessments or record linkage is preferred by the NOS) and that additionally did not account for selective dropout between exposure and outcome assessments.

Computation of Effect Sizes

None of the included studies reported Hedges' g as their effect size. The compute.es package[48] was used to calculate Hedges' g for the studies reporting the following: 1) For odds ratio (OR) Hedges' g was calculated using the "lores" function (based on log of OR and its corresponding variance).; 2) for p-values (with information of group sample sizes), Hedges' g was calculated using the "pes" function; 3) for correlation coefficient (r), Hedges' g was calculated using the "res" function.

Some studies did not report all the required information to calculate Hedges' g in the compute.es package. For the studies reporting OR, where group sample sizes were not reported, the method for converting OR to Cohens' d proposed by Chinn[49] was used. For mean difference in standardized test score, Hedges' g was calculated using standard deviation (SD) and mean difference. For regression coefficient (β), Hedges' g was calculated using p-value, standard error (SE) and z-statistics. For Cohens' d, Hedges' g was calculated using the formula proposed by Lakens.[50] For studies lacking p-value, confidence intervals (CIs), standard deviations or standard errors, the required inferential statistics were calculated in advance of the final effect size estimation by using appropriate formulas.[51-53]

Summarizing effect sizes

After Hedges' g had been calculated, the effect sizes were further summarized as many studies reported several effect sizes for the same exposure-outcome combination. Preferably, only one effect size per study should be retained,[22] however this was considered inapplicable, due to the large variations in neurodevelopmental outcomes assessed in the different studies. After careful consideration, four outcome dimensions (externalizing, internalizing, socio-emotional, cognitive) covering the affective and general cognitive domains were chosen. Selection of the outcome measures into each respective domain was based on 1) a thorough review of the properties of each instrument with regards to what area of development the instrument is aimed at measuring based on the manual for each instrument, and 2) research indicating that language, cognition, and executive functions are more strongly correlated with each other than with affective functioning.[54, 55]

Additionally, we applied the following rules to reduce the number of effect sizes per study, aiming to obtain only one effect size per outcome dimension per study:

1. For fish/seafood intake, if more than two intake groups had been defined, we only included the group which best corresponded to what is considered a healthy diet by the national health authorities, which is 2-3 servings per week for total fish intake, where about half should be fatty fish[11]
2. If studies reported statistics for a total score as well as sub scales of an outcome measure, only the effect size for the total score was included in the analyses

3. If the effect sizes were based on sample stratification (e.g. by breastfeeding duration), measures were collapsed. Collapsed effect sizes were calculated as weighted means of Hedges' g over comparisons, whereby weights depended on the n of each comparison
4. If OR was calculated for both high and low test score in the original article, only the effect sizes corresponding to the high test score was retained
5. If effect sizes were reported for all types of fish as well as oily fish - only all types of fish were included in the final analysis, as to make the exposure definition as homogenous as possible
6. If studies reported associations with both an unhealthy and a healthy dietary pattern, the effect size for the unhealthy dietary pattern was reversed and then averaged with the healthy dietary pattern
7. For studies reporting several eligible effect sizes for an outcome dimension, the effect size based on the most valid measurement instrument were used (e.g. effect size from a neuropsychological test vs questionnaire or a validated questionnaire vs a not validated questionnaire).
8. For studies that report more than one effect size relevant for inclusion in one domain based on the same type of instrument, the average weighted effect size across those reported were included after Hedges g had been calculated
9. For reported effect sizes where corresponding outcome dimensions were unclear (e.g. based on inadequate reporting of instrument properties) and not resolved with discussion among the reviewers, these effect sizes were excluded. This occurred for three separate instruments, each from different studies[43, 44, 46]

Application of these rules resulted in a total of 26 separate effect sizes. Only the fully adjusted effect sizes from each study were chosen. We initially considered including the corresponding unadjusted effect sizes for each study, however only four studies provided this information,[33, 35, 41, 42] and the studies that reported minimally adjusted results adjusted for different variables.

The 18 studies included in the final meta-analysis comprised a total of 63861[3] participants and 26 separate effect sizes divided into four different cognitive or affective dimensions. These effect sizes with corresponding confidence intervals (CI) are depicted in the forest plot in Figure 2[4]. The size of each square reflects the precision of the effect size estimate by means of the weight that is assigned to each respective study when the summary effect size is computed. A larger square equals larger weight assigned to that study.

Figure 2 approximately here

Effect size dependencies

There are several sources of effect size dependencies within our sample of studies. Firstly, some of the included studies use sub populations of the same cohort sample, while using different outcome measures at different time points: Four of the studies[30, 31, 34, 39] were based on subsamples of the ALSPAC cohort, and two studies[37, 47] were based on subsamples from the Project Viva cohort. Secondly, some of the studies included in this meta-analysis[33-35, 40] report multiple effect sizes.

We used a weighting scheme and calculated robust standard errors to account for these sources of dependencies.[56] Weights were adjusted for studies that contribute multiple effect sizes by recalculating them such that the sum of the weights of all effect sizes from a study reflect the sample size of that study.

[3] As more than one effect size per study was incorporated into the meta-analysis, with sometimes differing n between neurocognitive domains, the largest n for each study was used as a basis for total n.

[4] The data in Figure 2, 3 and 4 is based on adjusted standard errors to account for studies contributing with multiple outcomes as described in the "Effect size dependencies" section

When using the Metafor package which calculates weights from effect size variances or standard errors, this can be achieved by calculating effect size variances with adjusted N. In particular, we adjusted N for study i such that: $aN_i = N_i / \sum_j^k N_j$. Here k is the number of effect sizes from a study sample (e.g. ALSPAC) and N_j are the sample sizes for the different effect sizes. When estimating average effect sizes for specific domains, this approach corrects for multiple effect sizes for one domain coming from one study sample. When estimating the overall effect size this approach corrects for multiple contributions of effect sizes from one or more domains from one study sample. Hence, one study could be allocated different weights, depending on the meta-analytic model. Optimally the calculation of overall effect sizes would also account for the covariance between effects in different domains, however the reviewed articles did not provide this information. The employed weighting scheme implies an assumed correlation of rho = 0.5.

For the two REMs investigating for publication bias (trim and fill and with standard error as moderator) the adjusted standard errors based on the above formulae was used. For the original REM we used the reported effect sizes and corresponding adjusted variance as a basis for the calculations and obtained robust standard errors[56] using Metafor's "robust" function. We chose this robust estimator function as it is appropriate to use for models with unspecified heteroscedasticity,[23] which is the case with all studies reporting multiple effect sizes that are included in this meta-analysis.

REM

Three separate REM's were fit: One across all studies to yield an overall summary effect size, one for the cognitive domain and one for the affective domain. Table 3 provides a summary of the three REM's, including test results for heterogeneity (Q- and I^2-statistics). For the original REM the effect sizes are typically larger for the cognitive domain (g=0.14), compared to the affective domain (g=0.093), while the summary effect size across both domains is g=0.112.

Table 3: REM statistics for separate meta-analyses for overall summary effect size, cognitive domain and affective domain, including test for heterogeneity

#	Outcome	Model type	N^s	N^p	Hedges g	SE	z	p-value	df	Q	Qp	I^2 (%)
1	Summary Effect Size	Original REM	26	63861†	0.112	0.023	4.9	0.0001	25	102	<.0001	69
		REM with SE as moderator			0.079	0.029	4	0.0065	24	93	<.0001	68
		REM with Trim & fill	35††		0.075	0.019	2.7	<.0001	34	150	<.0001	75
2	Cognitive domain	Original REM	13	29269	0.14	0.016	8.5	<.0001	12	12	0.451	27
		REM with SE as moderator			0.132	0.030	4.3	<.0001	11	12	0.3836	29
		REM with Trim & fill			0.140	0.017	8.2	<.0001	12	12	.4510	27
3	Affective domain	Original REM	13	38219	0.093	0.034	2.7	0.03	12	67	<.0001	77
		REM with SE as moderator			0.043	0.041	1.1	0.2935	11	57	<.0001	74
		REM with Trim & fill	14††		0.088	0.028	3.1	0.0018	13	68	<.0001	76

N^s = number of included studies, N^p = number of total included participants

† Not the sum across the two domains as some studies are included in both domains

†† includes original and imputed studies

Heterogeneity

As can be seen from the Q and I^2 statistics in Table 3, there is a significant degree of heterogeneity present for the overall summary effect size, indicating a systematic difference in effect sizes between the studies. As possible sources of this heterogeneity, we investigated publication bias and performed a moderator analysis on all included studies.

Publication bias

The Egger's regression test for funnel plot asymmetry was not significant (p=0.1581), which might be mainly due to low power. To get a visualisation of possible publication bias, a funnel plot is depicted in Figure 3[4].

Figure 3 approximately here

If no publication bias was present, approximately 95 % of the points for the original effect sizes should be located within the white funnel area[25] and should be roughly distributed evenly to the left and to the right of the vertical line illustrating the overall summary effect size. Because this is not the case, a trim and fill analysis was performed, and the results are displayed in Figure 4[4].

Figure 4 approximately here

The imputed effect sizes (open circles) are all smaller than the summary effect size, and the trim and fill analyses suggest that the adjusted overall summary effect size would be 0.075 (c.f., Table 3). It is still significant (p<0.0001), but smaller than the originally calculated summary effect size (g=0.112). Additionally, even with the imputed effect sizes, there are still significant levels of heterogeneity present, which indicate that other factors than publication bias are contributing to the observed heterogeneity. Table 3 also shows that only the studies in the affective domain appear to be afflicted by publication bias, as the summary effect size for the cognitive domain remains unchanged with the trim and fill analysis. The results from regression-based adjustment for publication bias are consistent with the trim and fill analysis in that they show a similar overall effect size, a clear association in the cognitive domain, and a noticeably weaker association in the affective domain.

Moderator analyses

Considering that performing a moderator analysis is generally not advisable with less than ten studies,[53] we performed a moderator analysis for the whole sample of studies, rather than separately for the affective domain. The following moderators were initially included in single-predictor models: publication year, child age at assessment, outcome domain and diet category. Instrument category was originally considered, but only questionnaires were utilized in the studies within the affective domain and this was therefore deemed unnecessary. Child age at assessment explained none of the heterogeneity and were excluded from further analysis. Separately, outcome domain, publication year and diet category accounted for some of the heterogeneity present, and when included together in a moderator analysis they explained approximately 30 % of the heterogeneity (p= 0.0471). However, there was still a significant degree of heterogeneity present (p<0.0001), indicating that other moderators not considered in the model were influencing the outcome effect sizes.

DISCUSSION

The aim of this meta-analysis was to systematically review and summarize the currently existing literature about the association between maternal diet quality and different child neurodevelopmental outcomes. When dietary exposures believed to be appropriate proxies for maternal diet quality were included, a total of 18 studies comprising 63861 participants were found relevant for inclusion in this meta-analysis.

The meta-analysis showed that a better maternal diet quality had a small, statistically significant association with child neurodevelopment. The summary effect size for the cognitive domain was larger than the overall summary effect size, with no significant presence of heterogeneity. This positive association with cognitive outcomes is in line with findings from a recent narrative review investigating the association between maternal fish intake and child cognitive outcomes.[57] The important contribution of our quantitative meta-analysis is the calculation of average effect sizes, which shows, also after correcting for publication bias, a small but robust association. The summary effect size for the affective domain was smaller than the overall summary effect size, with a large and significant degree of heterogeneity present. Considering that an overall summary effect size is most appropriate to use for studies with little heterogeneity,[22] the summary effect size should be interpreted with caution. If we look at the effect sizes for all four outcome dimensions (c.f. Figure 2) we find that maternal diet quality is associated with all neurodevelopmental dimensions except for the internalizing dimension, with the strongest associations seen for socio-emotional and general cognitive functioning. However, these effect sizes are still considered small according to Cohens interpretative guidelines.[58]

In the moderator analysis, outcome domain, publication year and diet category (type of dietary classification - dietary pattern or its proxies (fish intake, fruit intake, saturated fat intake or Ω-6/Ω-3 fatty acid ratio)) contributed significantly to the heterogeneity present in the total sample of studies, explaining 30% of the heterogeneity. However, a large degree of heterogeneity remained. As only the fully adjusted effect sizes from each study were included in the meta-analysis, unmeasured or unreported variables may have contributed to the remaining heterogeneity. Furthermore, these results might indicate that maternal diet might be of more importance for certain neurodevelopmental outcomes. However, we emphasize that this moderator analysis is only exploratory and the results should be seen as preliminary given the small number of eligible studies included in the meta-analysis and the possible number of potential moderators.

The majority of the studies included in this meta-analysis using proxies of a maternal dietary pattern during pregnancy had fish or seafood intake as their exposure measure. Although fish intake most likely is a good marker for diet quality, there are limitations involved. One major limitation is that the studies investigating fish intake varied greatly with regards to intake group definitions; with division into two, three or four groups, where most groups were compared to a reference group (generally those who never or rarely consumed fish), or included as a continuous variable in a linear regression model. Some studies also compared extreme groups (lowest vs highest quintile), which were the studies reporting the largest effect sizes. Due to this varying dietary exposure definition it is likely that the amount of heterogeneity the diet category accounts for is underestimated in the moderator analysis. Ideally, we could have used a more elaborate classification of categories, to reflect the actual diversity of the exposure measures, but this was not appropriate considering the small number of studies included in this meta-analysis.[59]

As seen from the funnel plot in Figure 3, there is a clear negative correlation between effect sizes and standard error, indicating that the larger the sample size, the smaller the association between maternal diet quality and the outcome measure. This is not surprising, considering that the effect size of a study with a small sample needs to be large to reach significance in comparison to studies with a large sample

size where only very small effect sizes are required to reach statistical significance. However, even if the observed pattern has a statistical explanation, a clear visual indication of publication bias remains. Accordingly, analyses that corrected for publication bias through a trim and fill procedure and meta-regression resulted in overall effect size estimates that were around 30% lower compared to the effect size estimates from the original REM.

Maternal diet – direct effect or marker for child diet?

An important issue that cannot be resolved by this meta-analysis is whether the observed association is based on direct effects of maternal diet quality or whether it is a marker for the child's diet, which is a competing exposure that also influences child development. Not surprisingly, maternal diet quality, as well as maternal post-natal diet, and child diet during infancy and early toddlerhood have been found to be highly correlated.[60-62] Therefore, it remains possible that the observed associations between maternal diet quality and child development are due to the child's diet after pregnancy. Ultimately, the interpretation of the reported effect sizes depends on the assumed causal model. If it is assumed that child diet is a mediator between maternal diet quality and child development, then one has to control for child diet if interested in the direct effect of maternal diet quality, and one must not control for child diet if interested in the total effect of maternal diet quality.[63, 64] However, we suggest that maternal diet quality and child diet have a common cause—e.g. parental education—and an unbiased estimate of the direct effects of maternal diet quality on child development requires controlling for child diet. One mitigating fact is that child diet varies with sociodemographic variables[62] so that controlling for maternal postnatal diet and sociodemographic factors is likely to, at least in part, control for child diet. Still, child diet should ideally be assessed as a distinct factor. Indeed, the only study[35] among the three studies that reported controlling for the child's diet,[35, 36, 39] which also provided unadjusted and adjusted effect sizes, found that the association between maternal diet quality and both externalizing and internalizing problems in the child was mediated by child diet, reducing the effect of maternal diet quality in three out of four analyses.

An important aspect of child diet during the earliest stages of life is breastfeeding. Previous studies exploring the association between breastfeeding and different cognitive development measures have found associations between longer breastfeeding duration and better general cognitive development,[65-67] higher IQ,[68] better educational attainment[69] and language development,[70] as well as a lower risk of having ADHD.[71] Nine of the studies included in this meta-analysis adjusted for breastfeeding duration,[31, 33, 34, 37, 40, 41, 43, 46, 47] but none provided information on both unadjusted and adjusted effect sizes related to breastfeeding specifically. However, some indicated that breastfeeding did not display a significant confounding effect.

In studies that stratified their sample by breastfeeding practices,[36, 43] a significant association with maternal diet quality was only seen when the child had been breastfed for less than 6 months. This suggests that a better maternal diet quality might serve as a protective or beneficial factor to a larger degree for children who are not breastfed or breastfed less than the recommended period.

Maternal obesity is another factor related to both breastfeeding and the outcomes. Studies consistently show reduced breastfeeding rates in obese mothers[72, 73] and obesity is also linked to impaired cognition and increased behavioural problems in children.[74-76] This highlights the importance of also accounting for maternal BMI in the analysis, but only four studies did this in their analyses.[37, 38, 40, 41]

The effect of maternal diet quality on child development is an exemplary research topic where causal knowledge has to be extracted from observational studies because experimental studies are either

unethical or impractical.[77] However, if causal information is to be gleaned from observational data, care must be taken to control for biases due to e.g. self-selection into studies or selective drop-out, for example by using inverse probability weights so that the effective sample better resembles the target population. Importantly, controlling by adding covariates is typically not sufficient to control for e.g. selection bias.[78] The appraisal of the summarized studies with the NOS suggests that while all summarized studies controlled, to varying degrees, for potential confounders by adding covariates to their analysis, systematic control for selection bias is not yet part of routine analysis.

Limitations

Only the fully adjusted effect sizes from each study were selected for inclusion into the meta-analysis, however, the confounders considered by each study varied greatly. This adds to the uncertainty of the results.

Only four of the studies included in this meta-analysis used maternal dietary patterns as their exposure measure for diet quality. Even though investigation of dietary patterns has received increasing interest over the past decade, there is still a lack of research in this area, particularly in relation to child cognitive outcomes. Additionally, there should be more focus on dietary patterns as opposed to global indices of healthy diets, as these indices inherently assume substitutability of different aspects of healthy diets, which might not be valid assumptions.

In the majority of the studies included with fish or seafood intake as the exposure measure, fish intake was not a part of the main exposure or primary investigation – it was often included as a covariate, with the main study focus being investigation of mercury exposure. This, together with the varying definition of fish intake between the studies, probably contributes substantially to the observed heterogeneity and illustrates the limitations involved in conducting a meta-analysis on studies with very heterogeneous exposure measures.

None of the studies reporting more than one effect size included information on the corresponding correlation between effect sizes. It is likely to assume that at least some of the effect sizes within each study were interdependent as they are measuring different aspects of the same overall cognitive or affective domain. It is widely recommended to aggregate dependent effect sizes to avoid biased estimates while at the same time account for possible dependencies within the sample of studies.[22] This meta-analysis used non-independent effect sizes due to both multiple effect sizes reported from the same study as well as several studies using the same cohort as a basis for their study sample. However, we took care in accounting for these dependencies by calculating adjusted weights and robust standard errors, and performing an overall meta-analysis as well as individual meta-analyses for both domains. Moreover, three different REMs were fit (c.f. Table 3) to test for plausible moderators. Hence, despite the challenges posed by dependent effect sizes our analyses likely provides unbiased summary effect sizes for the association between maternal diet quality and child development.

Lastly, there are many challenges relating to the FFQ as a measurement tool, which is well-known within the nutritional research field.[79, 80] This is mainly due to different types of bias that can arise from using self-report measures of dietary intake, which creates further difficulties in relation to analysis and interpretation of dietary data.

Taken together, these limitations suggest that while the effect sizes reported here provide some information about the association between maternal diet quality and child cognitive and affective outcomes, more research is needed to obtain reliable estimates of such associations.

CONCLUSION

Comparing studies looking at an overall maternal diet quality rather than specific nutrients brings with it many challenges, mainly due to heterogeneous methods for measuring intake, failure to account for child diet during early childhood, as well as the vast number of confounders needed to be considered, both genetic and environmental.

Additionally, the number of studies available for inclusion in this meta-analysis is limited and they are heterogeneous, both with regards to exposure and outcome measures, indicating that results should be interpreted with caution. However, the results point in the direction that a better maternal diet quality is weakly, but robustly associated with a more favourable cognitive development and fewer affective problems in the child.

Lastly, it is important to emphasize that the studies included in this meta-analysis are all observational and do often take only limited steps towards causal identification. Therefore, causal interpretations of the results have to be avoided.

Suggestions for future research

The results of this meta-analysis highlight the need for more research on the effects of maternal diet quality on child cognitive and affective outcomes. The heterogeneity present in this sample of studies, particularly with regards to the definition of the exposure measures, makes comparison of results across studies particularly challenging. To better enable for between-study comparisons in the future, careful consideration should be taken to develop standardized instruments for the measurement of diet quality, which can be applicable, with minor modifications, across different populations. Additionally, studies should aim for the use of validated and recognized instruments for the measurement of cognitive and behavioural outcomes, rather than self-developed or obsolete instruments. With regards to the outcome measures, available reliability data for the instruments used by the included studies does not provide any clear distinction in test-retest reliability between questionnaires and clinical tests,[81-95] but a comparison of effect sizes for outcomes assessed through questionnaires and clinical neuropsychological tests, within the same study, will help to settle this important issue.

Furthermore, it is crucial that future studies investigating the effect of maternal diet quality on child neurodevelopmental outcomes also consider the child's diet, as failure to recognize child diet as an important contributing factor limits the interpretability of such studies. More generally, more attention to and controlling of confounders, potential competing exposures, and potential bias due to self-selection into studies or selective drop out will be important to better justify a causal interpretation of observational studies.

Finally, greater emphasis should be put on research transparency by means of describing the methodology used more exhaustively and by reporting complete results for both significant and non-significant results.

Contributors: TCB, GB, and HA designed the study. TCB and ALB prepared the data in conjunction with GB. TCB and GB conducted the statistical analysis. TCB drafted the manuscript and had the primary responsibility for the final content. All authors critically reviewed, read and approved the final version of the manuscript. TCB is the guarantor.

Competing interests: None

Funding: The present study was supported by a grant from NevSom; National Centre for Neurodevelopmental Disorders and Hypersomnia - ADHD, Autism, Tourette's syndrome and Hypersomnia (grant no. 51379 – 001/15-473).

Data sharing statement: Table with raw effect measures from included studies and R script of all analysis performed for this meta-analysis is available upon request from the corresponding author Tiril Cecilie Borge; tibo@fhi.no

REFERENCE LIST

1. Barker DJP. Mothers, babies and health in later life. 2nd ed. Edinburgh: Churchill Livingstone 1998.
2. Harding JE. The nutritional basis of the fetal origins of adult disease. *Int J Epidemiol* 2001;30: 15-23 [published Online First: 2001/02/15].
3. Heindel JJ, Vandenberg LN. Developmental origins of health and disease: a paradigm for understanding disease cause and prevention. *Curr Opin Pediatr* 2015;27: 248-53 doi: 10.1097/mop.0000000000000191 [published Online First: 2015/01/31].
4. Thapar A, Cooper M, Jefferies R, et al. What causes attention deficit hyperactivity disorder? *Arch Dis Child* 2012;97: 260-65
5. Roseboom TJ, Painter RC, van Abeelen AF, et al. Hungry in the womb: what are the consequences? Lessons from the Dutch famine. *Maturitas* 2011;70: 141-5 doi: 10.1016/j.maturitas.2011.06.017 [published Online First: 2011/08/02].
6. Prado EL, Dewey KG. Nutrition and brain development in early life. *Nutr Rev* 2014;72: 267-84 doi: 10.1111/nure.12102 [published Online First: 2014/04/02].
7. Zeisel SH. Is maternal diet supplementation beneficial? Optimal development of infant depends on mother's diet. *Am J Clin Nutr* 2009;89: 685s-7s doi: 10.3945/ajcn.2008.26811F [published Online First: 2009/01/01].
8. Englund-Ogge L, Brantsaeter AL, Sengpiel V, et al. Maternal dietary patterns and preterm delivery: results from large prospective cohort study. *Br Med J* 2014;348: 1446-64 doi: 10.1136/bmj.g1446 [published Online First: 2014/03/13].
9. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol* 2002;13: 3-9
10. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery* 2010;8: 336-41 doi: http://dx.doi.org/10.1016/j.ijsu.2010.02.007
11. Norwegian National Nutrition Council. Diet recommendations to promote public health and prevent chronic diseases - Methodology and evidence foundation. Oslo: Norwegian Health Directorate, 2011.
12. Simopoulos AP. The importance of the omega-6/omega-3 fatty acid ratio in cardiovascular disease and other chronic diseases. *Exp Biol Med (Maywood)* 2008;233: 674-88 doi: 10.3181/0711-mr-311 [published Online First: 2008/04/15].
13. Hooper L, Martin N, Abdelhamid A, et al. Reduction in saturated fat intake for cardiovascular disease. *Cochrane Database Syst Rev* 2015: Cd011737 doi: 10.1002/14651858.cd011737 [published Online First: 2015/06/13].
14. Anderson JW, Baird P, Davis RH, et al. Health benefits of dietary fiber. *Nutr Rev* 2009;67: 188-205 doi: 10.1111/j.1753-4887.2009.00189.x [published Online First: 2009/04/02].
15. Balder HF, Virtanen M, Brants HA, et al. Common and country-specific dietary patterns in four European cohort studies. *J Nutr* 2003;133: 4246-51 [published Online First: 2003/12/04].
16. Knudsen VK, Orozova-Bekkevold IM, Mikkelsen TB, et al. Major dietary patterns in pregnancy and fetal growth. *Eur J Clin Nutr* 2008;62: 463-70 doi: 10.1038/sj.ejcn.1602745 [published Online First: 2007/03/30].
17. Willett WC, Sacks F, Trichopoulou A, et al. Mediterranean diet pyramid: a cultural model for healthy eating. *Am J Clin Nutr* 1995;61: 1402S-06S [published Online First: 2013/02/13].
18. Guenther PM, Casavale KO, Reedy J, et al. Update of the Healthy Eating Index: HEI-2010. *J Acad Nutr Diet* 2013;113: 569-80 doi: 10.1016/j.jand.2012.12.016
19. Wells G, Shea B, O'Connell D, et al. The Newcastle-Ottawa scale (NOS) for assessing the quailty of nonrandomised studies in meta-analyses. 2009. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
20. The Cochrane Collaboration. 13.5.2.3 Tools for assessing methodological quality or risk of bias in non-randomized studies. In: Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011] ed, 2011.

21. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol* 2010;25: 603-5 doi: 10.1007/s10654-010-9491-z [published Online First: 2010/07/24].
22. Del Re A. A Practical Tutorial on Conducting Meta-Analysis in R. *Quant Meth Psych* 2015;11: 37-50
23. Viechtbauer W. Meta-Analysis Package for R; 2016. https://cran.r-project.org/web/packages/metafor/index.html
24. Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. *J Stat Softw* 2010;36: 1-48 doi: 10.18637/jss.v036.i03 [published Online First: 2010/08/05].
25. Borenstein M, Hedges LV, Higgins JPT, et al. Introduction to Meta-Analysis. Chicester, United Kingdom: Wiley 2009.
26. Siddiqi N. Publication bias in epidemiological studies. *Cent Eur J Public Health* 2011;19: 118-20 [published Online First: 2011/07/12].
27. Sterne JAC, Harbord RM. Funnel plots in meta-analysis. *SJ* 2004;4: 127-41
28. Egger M, Smith GD, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315: 629
29. Moreno SG, Sutton AJ, Ades AE, et al. Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. *BMC Med Res Methodol* 2009;9: 2 doi: 10.1186/1471-2288-9-2 [published Online First: 2009/01/14].
30. Barker ED, Kirkham N, Ng J, et al. Prenatal maternal depression symptoms and nutrition, and child cognitive function. *British Journal of Psychiatry* 2013;203: 417-21
31. Daniels JL, Longnecker MP, Rowland AS, et al. Fish intake during pregnancy and early cognitive development of offspring. *Epidemiology* 2004;15: 394-402 doi: 10.1097/01.ede.0000129514.46451.ce
32. Davidson PW, Strain JJ, Myers GJ, et al. Neurodevelopmental effects of maternal nutritional status and exposure to methylmercury from eating fish during pregnancy. *Neurotoxicology* 2008;29: 767-75 doi: 10.1016/j.neuro.2008.06.001 [published Online First: 2008/06/11].
33. Gale CR, Robinson SM, Godfrey KM, et al. Oily fish intake during pregnancy - Association with lower hyperactivity but not with higher full-scale IQ in offspring. *J Child Psychol Psychiatry* 2008;49: 1061-68 doi: 10.1111/j.1469-7610.2008.01908.x [published Online First: 2008/04/15].
34. Hibbeln JR, Davis JM, Steer C, et al. Maternal seafood consumption in pregnancy and neurodevelopmental outcomes in childhood (ALSPAC study): an observational cohort study. *The Lancet* 2007;369: 578-85 doi: 10.1016/S0140-6736(07)60277-3 [published Online First: 2007/02/15].
35. Jacka FN, Ystrom E, Brantsaeter AL, et al. Maternal and early postnatal nutrition and mental health of offspring by age 5 years: a prospective cohort study. *J Am Acad Child Adolesc Psychiatry* 2013;52: 1038-47 doi: 10.1016/j.jaac.2013.07.002 [published Online First: 2013/10/01].
36. Mendez MA, Torrent M, Julvez J, et al. Maternal fish and other seafood intakes during pregnancy and child neurodevelopment at age 4 years. *Public Health Nutr* 2009;12: 1702-10 doi: 10.1017/S1368980008003947 [published Online First: 2008/11/25].
37. Oken E, Radesky JS, Wright RO, et al. Maternal fish intake during pregnancy, blood mercury levels, and child cognition at age 3 years in a US cohort. *Am J Epidemiol* 2008;167: 1171-81 doi: 10.1093/aje/kwn034 [published Online First: 2008/03/18].
38. Oken E, Østerdal ML, Gillman MW, et al. Associations of maternal fish intake during pregnancy and breastfeeding duration with attainment of developmental milestones in early childhood: a study from the Danish National Birth Cohort. *Am J Clin Nutr* 2008;88: 789-96 [published Online First: 2010/05/24].
39. Pina-Camacho L, Jensen SK, Gaysina D, et al. Maternal depression symptoms, unhealthy diet and child emotional-behavioural dysregulation. *Psychol Med* 2015;45: 1851-60 doi: 10.1017/s0033291714002955 [published Online First: 2014/12/19].
40. Steenweg-de Graaff J, Tiemeier H, Steegers-Theunissen RP, et al. Maternal dietary patterns during pregnancy and child internalising and externalising problems. The Generation R Study. *Clin Nutr* 2013;33: 115-21 doi: 10.1016/j.clnu.2013.03.002 [published Online First: 2013/03/14].
41. Valent F, Mariuz M, Bin M, et al. Associations of Prenatal Mercury Exposure From Maternal Fish Consumption and Polyunsaturated Fatty Acids With Child Neurodevelopment: A Prospective Cohort Study in Italy. *J Epidemiol* 2013;23: 360-70 doi: 10.2188/jea.JE20120168 [published Online First: 2013/08/10].
42. Sagiv SK, Thurston SW, Bellinger DC, et al. Prenatal exposure to mercury and fish consumption during pregnancy and attention-deficit/hyperactivity disorder-related behavior in children. *Arch Pediatr Adolesc Med* 2012;166: 1123-31 doi: 10.1001/archpediatrics.2012.1286 [published Online First: 2012/12].

43. Bernard JY, De Agostini M, Forhan A, et al. The Dietary n6:n3 Fatty Acid Ratio during Pregnancy Is Inversely Associated with Child Neurodevelopment in the EDEN Mother-Child Cohort. *J Nutr* 2013;143: 1481-88 doi: 10.3945/jn.113.178640 [published Online First: 2013/07/31].
44. Bolduc FV, Lau A, Rosenfelt CS, et al. Cognitive Enhancement in Infants Associated with Increased Maternal Fruit Intake During Pregnancy: Results from a Birth Cohort Study with Validation in an Animal Model. *Ebiomedicine* 2016;8: 331-40 doi: 10.1016/j.ebiom.2016.04.025
45. Gustafsson HC, Kuzava SE, Werner EA, et al. Maternal dietary fat intake during pregnancy is associated with infant temperament. *Dev Psychobiol* 2016;58: 528-35
46. Julvez J, Mendez M, Fernandez-Barres S, et al. Maternal Consumption of Seafood in Pregnancy and Child Neuropsychological Development: A Longitudinal Study Based on a Population With High Consumption Levels. *Am J Epidemiol* 2016;183: 169-82 doi: 10.1093/aje/kwv195
47. Oken E, Rifas-Shiman SL, Amarasiriwardena C, et al. Maternal prenatal fish consumption and cognition in mid childhood: Mercury, fatty acids, and selenium. *Neurotoxicol Teratol* 2016;57: 71-78
48. Del Re A. Compute Effect Sizes (Package "compute.es"); 2014. https://cran.r-project.org/web/packages/compute.es/
49. Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. *Stat Med* 2000;19: 3127-31 [published Online First: 2000/12/13].
50. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front Psychol* 2013;4: 863 doi: 10.3389/fpsyg.2013.00863 [published Online First: 2013/12/11].
51. Altman DG, Bland JM. How to obtain the confidence interval from a P value. *Br Med J* 2011;343 doi: 10.1136/bmj.d2090 [published Online First: 2011/08/08].
52. Altman DG, Bland JM. How to obtain the P value from a confidence interval. *Br Med J* 2011;343 doi: 10.1136/bmj.d2304 [published Online First: 2011/08/08].
53. The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions, Higgins J, Green S, editors; 2011. www.cochrane-handbook.org
54. Sjowall D, Backman A, Thorell LB. Neuropsychological heterogeneity in preschool ADHD: investigating the interplay between cognitive, affective and motivation-based forms of regulation. *J Abnorm Child Psychol* 2015;43: 669-80 doi: 10.1007/s10802-014-9942-1 [published Online First: 2014/09/23].
55. Gutierrez-Cobo MJ, Cabello R, Fernandez-Berrocal P. The Relationship between Emotional Intelligence and Cool and Hot Cognitive Processes: A Systematic Review. *Front Behav Neurosci* 2016;10: 101 doi: 10.3389/fnbeh.2016.00101 [published Online First: 2016/06/16].
56. Hedges LV, Tipton E, Johnson MC. Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods* 2010;1: 39-65 doi: 10.1002/jrsm.5
57. Starling P, Charlton K, McMahon AT, et al. Fish intake during pregnancy and foetal neurodevelopment-a systematic review of the evidence. *Nutrients* 2015;7: 2001-14 doi: 10.3390/nu7032001 [published Online First: 2015/03/21].
58. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, N. J: Laurence Erlbaum 1988.
59. Hedges LV, Pigott TD. The power of statistical tests for moderators in meta-analysis. *Psychol Methods* 2004;9: 426-45 doi: 10.1037/1082-989x.9.4.426 [published Online First: 2004/12/16].
60. Ashman AM, Collins CE, Hure AJ, et al. Maternal diet during early childhood, but not pregnancy, predicts diet quality and fruit and vegetable acceptance in offspring. *Matern Child Nutr* 2014 doi: 10.1111/mcn.12151 [published Online First: 2014/10/08].
61. Fisk CM, Crozier SR, Inskip HM, et al. Influences on the quality of young children's diets: the importance of maternal food choices. *Br J Nutr* 2011;105: 287-96 doi: 10.1017/s0007114510003302 [published Online First: 2010/09/03].
62. Emmett PM, Jones LR, Northstone K. Dietary patterns in the Avon Longitudinal Study of Parents and Children. *Nutr Rev* 2015;73 Suppl 3: 207-30 doi: 10.1093/nutrit/nuv055 [published Online First: 2015/09/24].
63. Hernán MA, Robins JM. Causal Inference. Boca Raton: Chapman & Hall/CRC, forthcoming; 2016. https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/
64. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999;10: 37-48 [published Online First: 1999/01/15].
65. Julvez J, Guxens M, Carsin A-E, et al. A cohort study on full breastfeeding and child neuropsychological development: The role of maternal social, psychological, and nutritional factors. *Dev Med Child Neurol* 2014;56: 148-56 doi: 10.1111/dmcn.12282 [published Online First: 2013/10/01].

66. Quigley M, Hockley C, Carson C, et al. Breastfeeding is Associated with Improved Child Cognitive Development: A Population-Based Cohort Study. *J Pediatr* 2012;160: 25-32 doi: 10.1016/j.jpeds.2011.06.035 [published Online First: 2011/08/11].
67. Kramer MS, Aboud F, Mironova E, et al. Breastfeeding and Child Cognitive Development: New Evidence From a Large Randomized Trial. *Arch Gen Psychiatry* 2008;65: 578-84 doi: 10.1001/archpsyc.65.5.578 [published Online First: 2008/05].
68. Brion M-JA, Lawlor DA, Matijasevich A, et al. What are the causal effects of breastfeeding on IQ, obesity and blood pressure? Evidence from comparing high-income with middle-income cohorts. *Int J Epidemiol* 2011;40: 670-80 doi: 10.1093/ije/dyr020 [published Online First: 2011/02/24].
69. Groen-Blokhuis MM, Franic S, van Beijsterveldt CE, et al. A prospective study of the effects of breastfeeding and FADS2 polymorphisms on cognition and hyperactivity/attention problems. *Am J Med Genet B Neuropsychiatr Genet* 2013;162: 457-65 doi: 10.1002/ajmg.b.32175 [published Online First: 2013/06/04].
70. Smith JM. Breastfeeding and language outcomes: A review of the literature. *J Commun Disord* 2015;57: 29-40 doi: 10.1016/j.jcomdis.2015.04.002 [published Online First: 2015/05].
71. Stadler DD, Musser ED, Holton KF, et al. Recalled initiation and duration of maternal breastfeeding among children with and without ADHD in a well characterized case-control sample. *J Abnorm Child Psychol* 2016;44: 347-55 doi: 10.1007/s10802-015-9987-9 [published Online First: 2016/02].
72. Newby RM, Davies PS. Antenatal breastfeeding intention, confidence and comfort in obese and non-obese primiparous Australian women: associations with breastfeeding duration. *Eur J Clin Nutr* 2016;70: 935-40 doi: 10.1038/ejcn.2016.29 [published Online First: 2016/03/24].
73. Bever Babendure J, Reifsnider E, Mendias E, et al. Reduced breastfeeding rates among obese mothers: a review of contributing factors, clinical considerations and future directions. *Int Breastfeed J* 2015;10: 21 doi: 10.1186/s13006-015-0046-5 [published Online First: 2015/07/04].
74. Pugh SJ, Hutcheon JA, Richardson GA, et al. Gestational weight gain, prepregnancy body mass index and offspring attention-deficit hyperactivity disorder symptoms and behaviour at age 10. *BJOG* 2016 doi: 10.1111/1471-0528.13909 [published Online First: 2016/03/22].
75. Pugh SJ, Richardson GA, Hutcheon JA, et al. Maternal Obesity and Excessive Gestational Weight Gain Are Associated with Components of Child Cognition. *J Nutr* 2015;145: 2562-9 doi: 10.3945/jn.115.215525 [published Online First: 2015/10/02].
76. Veena SR, Gale CR, Krishnaveni GV, et al. Association between maternal nutritional status in pregnancy and offspring cognitive function during childhood and adolescence; a systematic review. *BMC Pregnancy Childbirth* 2016;16: 220 doi: 10.1186/s12884-016-1011-z [published Online First: 2016/08/16].
77. Rothman KJ. Epidemiology : an introduction. Oxford: Oxford University Press 2002.
78. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;15: 615-25 [published Online First: 2004/08/17].
79. Freedman LS, Schatzkin A, Midthune D, et al. Dealing With Dietary Measurement Error in Nutritional Cohort Studies. *JNCI Journal of the National Cancer Institute* 2011;103: 1086-92 doi: 10.1093/jnci/djr189
80. Kipnis V, Midthune D, Freedman L, et al. Bias in dietary-report instruments and its implications for nutritional epidemiology. *Public Health Nutr* 2002;5: 915-23 doi: 10.1079/phn2002383 [published Online First: 2003/03/14].
81. Adams W, Sheslow D. Wide Range Assessment of Visual Motor Abilities (WRAVMA). Novato, CA: Academic Therapy Publications 1995.
82. Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986;51: 1173-82 doi: 10.1037/0022-3514.51.6.1173
83. Bayley N. Bayley Scales of Infant and Toddler Development, Technical Manual. 3rd ed. San Antonio TX: Harcourt Assessment 2006.
84. Bilenberg N. The Child Behavior Checklist (CBCL) and related material: standardization and validation in Danish population based and clinically based samples. *Acta Psychiatr Scand Suppl* 1999;398: 2-52
85. Canning EH, Kelleher K. Performance of screening tools for mental health problems in chronically ill children. *Arch Pediatr Adolesc Med* 1994;148: 272-8
86. Fenson L, Dale PS, Reznick JS, et al. The MacArthur Communicative Development Inventories: User's guide and technical manual. San Diego: Singular Publishing Group 1993.
87. Frankenburg WK, Dodds J, Archer P, et al. The Denver II: A Major Revision and Restandardization of the Denver Developmental Screening Test. *Pediatrics* 1992;89: 91-97
88. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. *J Am Acad Child Adolesc Psychiatry* 2001;40: 1337-45 doi: 10.1097/00004583-200111000-00015 [published Online First: 2001/11/09].

89. McCarthy D. Manual for the McCarthy scales of children's abilities. New York: The Psychological Corporation 1972.
90. McDevitt SC, Carey WB. The measurement of temperament in 3--7 year old children. *J Child Psychol Psychiatry* 1978;19: 245-53
91. Novik TS. Validity of the Child Behaviour Checklist in a Norwegian sample. *Eur Child Adolesc Psychiatry* 1999;8: 247-54
92. Squires J, Twombly E, Bricker D, et al. Ages and Stages Questionnaire - Technical report. Baltimore, MD: Brookes Publishing Co. Inc 2009.
93. Stone LL, Otten R, Engels RC, et al. Psychometric properties of the parent and teacher versions of the strengths and difficulties questionnaire for 4- to 12-year-olds: a review. *Clin Child Fam Psychol Rev* 2010;13: 254-74 doi: 10.1007/s10567-010-0071-2 [published Online First: 2010/07/01].
94. Williams KT, Wang J. Technical references to the Peabody Picture Vocabulary Test- Third Edition (PPVT-III). Circle Pines, MN: American Guidance Service 1997.
95. Williams PE, Weiss LG, Rolfhus EL. WISC–IV Technical Report #2 Psychometric Properties: The Psychological Corporation, 2003.

Figure legends

Figure 1: Flow diagram of the study selection process

Figure 2: Forest plot of original REM of all included studies, with summary effect size for the cognitive and affective dimensions, as well as the overall summary effect size

Figure 3: Funnel plot of original REM of all included studies with their respective effect size and standard errors

Figure 4: Funnel plot of REM with trim and fill analysis, showing original studies (closed circle) and imputed studies (open circle)

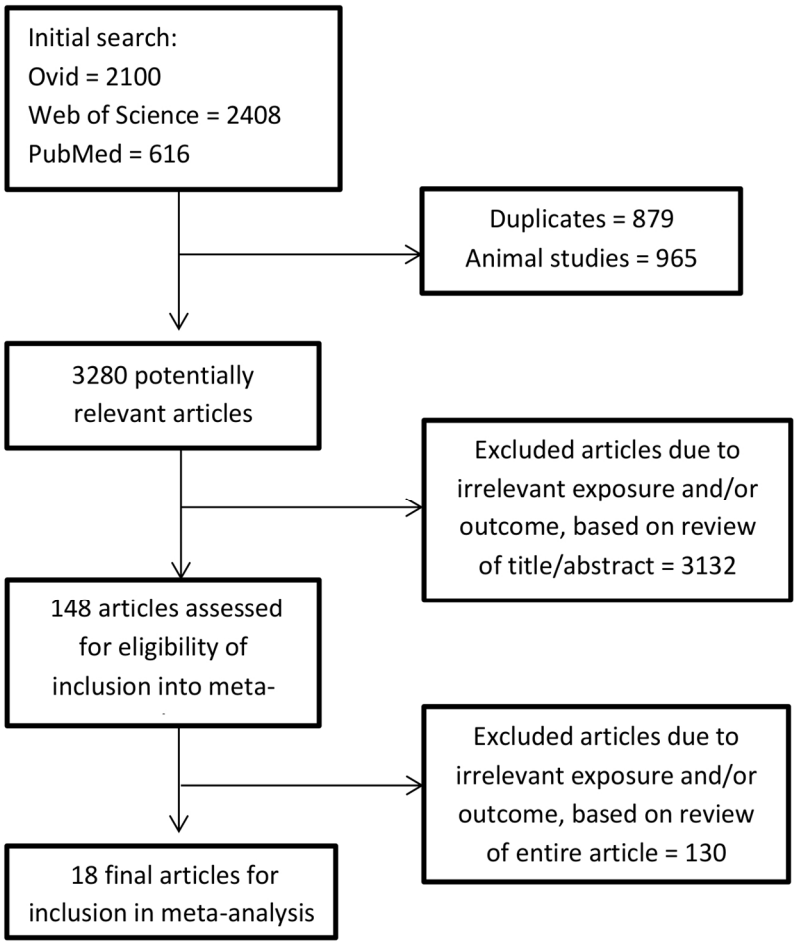

Figure 1: Flow diagram of the study selection process

139x169mm (300 x 300 DPI)

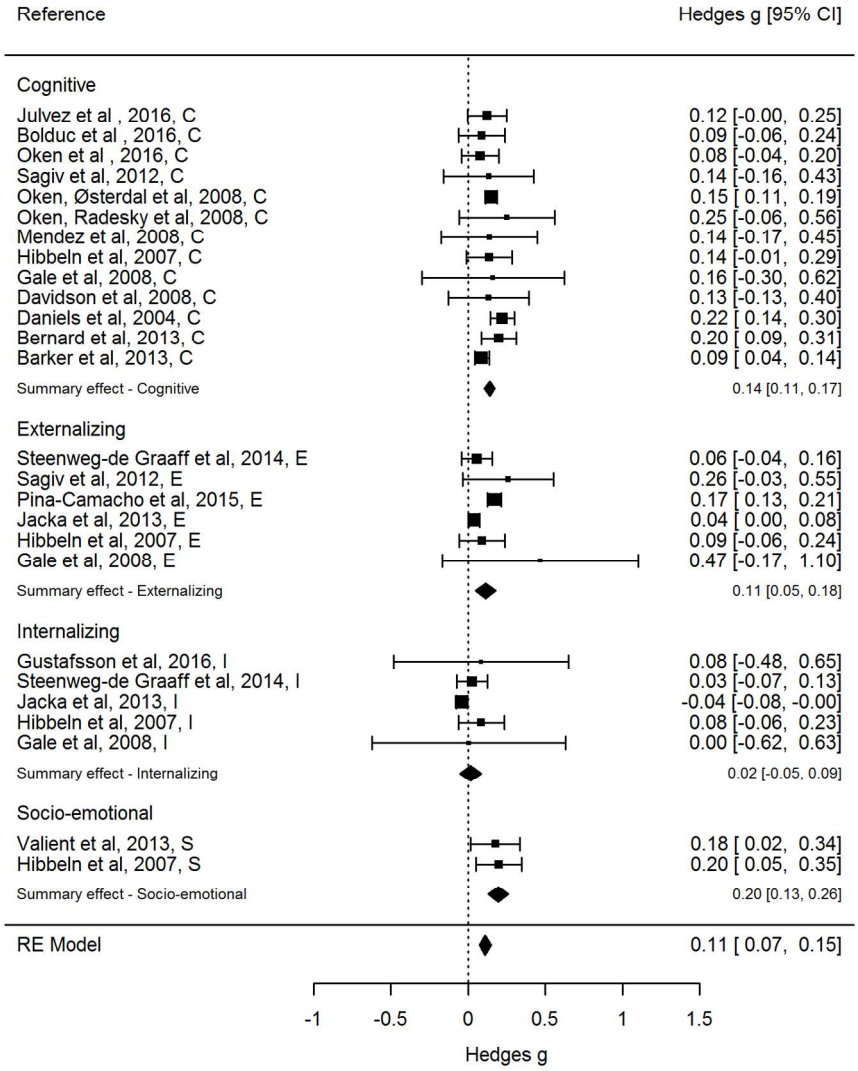

Figure 2: Forest plot of original REM of all included studies, with summary effect size for the cognitive and affective dimensions, as well as the overall summary effect size

127x169mm (300 x 300 DPI)

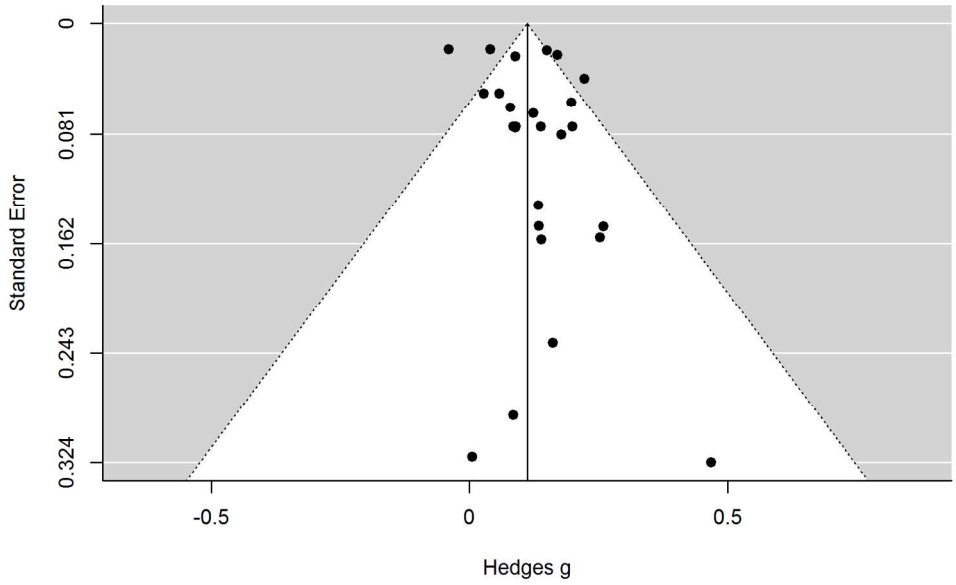

Figure 3: Funnel plot of original REM of all included studies with their respective effect size and standard errors

169x118mm (300 x 300 DPI)

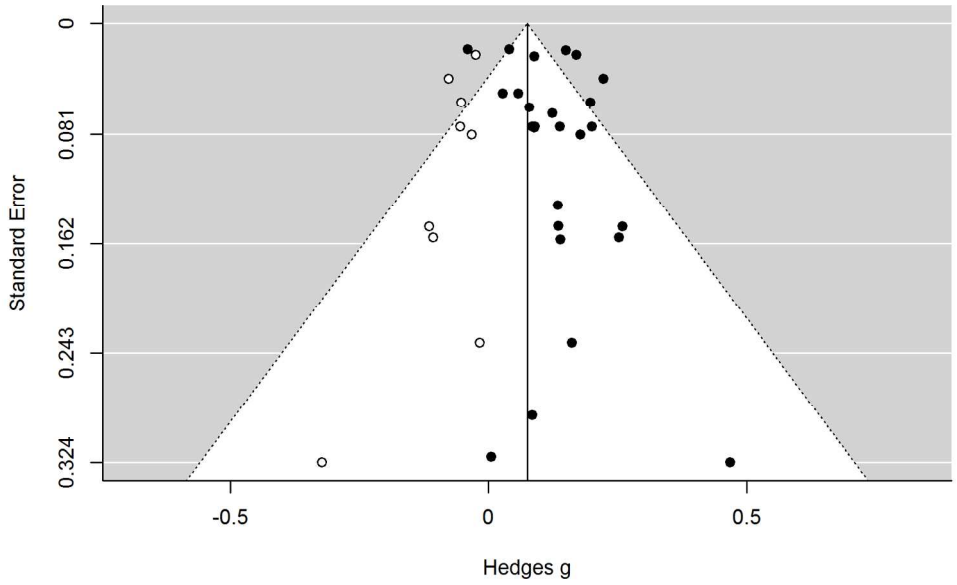

Figure 4: Funnel plot of REM with trim and fill analysis, showing original studies (closed circle) and imputed studies (open circle)

169x118mm (300 x 300 DPI)

Supplementary table 1: Example of full search string (Ovid)

	Searches	Results
1	(Maternal* or prenatal* or perinatal* or gestational* or pregnan*)	1768624
2	(diet*3 or nutrition or fiber or fibre or protein or fat or fatty or carbohydrate or fruit or vegetable or fish or seafood)	7878460
3	child* or toddler* or offspring	3862890
4	(behavior?r or behavior?r disorder* or externali?ing or internali?ing or mental health or mental development or learning disorder* or cogniti*3 or neurocogniti*3 or memory or IQ or executive or ADHD or attention deficit hyperactivity disorder or Attention deficit disorder or oppositional defiant disorder or conduct disorder or development* disability or neurodevelopment* or autism spectrum disorder or hyperkinetic disorder or hyperactivity disorder or language or communication or affective or developmental milestone*)	6446778
5	Combine 1-4	11512
6	Limit 5 to appropriate age group (infant – 12 years of age)	8003
7	Limit 6 to pregnancy	4632
8	Limit 7 to humans	3316
9	Limit to original articles	2974
10	Remove duplicates from 9	2100

Supplementary table 2: Overview of control variables for each study

Confounders	Barker et al (2013)	Bernard et al (2013)	Bolduc et al (2016)	Daniels et al (2004)	Davidson et al (2008)	Gale et al (2008)	Gustafsson et al (2016)**	Hibbeln et al (2007)	Jacka et al (2013)	Julvez et al (2016)	Mendez et al (2008)	Oken, Radensky et al (2008)	Oken et al (2016)	Oken, Østerdal et al (2008)	Pina-Camacho et al (2015)	Sagiv et al (2012)	Steenweg-de Graaff et al (2014)	Valent et al (2013)
Alcohol		x		x		x		x				x		x		x		x
Biomarkers																x		
Birth complications	x														x			
Breastfeeding				x		x		x		x	x	x	x	x				x
Child gender		x		x	x		x	x	x	x	x	x	x	x		x	x	x
Child ADHD medication																x		
Child age at assessment		x		x						x	x	x	x	x		x	x	
Child birth weight					x	x	x			x			x	x				x
Child dietary pattern									x									
Child fish intake											x							x
Child sugary snacks/drinks intake																	x	
Daycare attendance		x				x												x
Ethnicity								x		x		x	x			x	x	x
Fetal growth												x						
Home environment	x			x	x			x					x		x	x		x
Length of gestation		x						x			x	x	x	x				
Marital status	x				x			x	x			x		x	x	x	x	x
Maternal age		x	x		x	x	x	x	x	x		x		x		x	x	x
Maternal energy intake		x					x			x							x	
Maternal diet*			x					x								x		x
Maternal gestational diabetes			x															
Maternal IQ					x	x							x			x		x
Maternal mental health							x		x					x		x	x	
Maternal pre pregnancy BMI												x		x			x	x
Maternal pregnancy weight/weight gain		x					x			x								x
Maternal supplement use			x														x	
Parental learning difficulties														x				
Parity	x	x		x				x		x	x		x	x	x		x	
Paternal age									x									
SES	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Smoking		x		x		x		x	x			x	x	x		x	x	x

*For some studies where fish intake was the exposure, other maternal dietary components were included in the analysis as confounders.

**All covariates did not significantly correlate with outcome, so these were not included in the final analysis

Supplementary table 3: Evaluation of individual study quality with The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analyses

		Barker et al (2013)	Bernard et al (2013)	Bolduc et al (2016)	Daniels et al (2004)	Davidson et al (2008)	Gale et al (2008)	Gustafsson et al (2016)	Hibbeln et al (2007)	Jacka et al (2013)	Julvez et al (2016)	Mendez et al (2008)	Oken, Radensky et al (2008)	Oken, Østerdal et al (2008)	Oken et al (2016)	Pina-Camacho et al (2015)	Sagiv et al (2012)	Steenweg-de Graaff et al (2014)	Valent et al (2013)
Selection	**1) Representativeness of the exposed cohort:** a) truly representative of the average pregnant woman in the community*; b) somewhat representative of the average pregnant woman in the community*; c) selected group of users e.g. nurses, volunteers; d) no description of the derivation of the cohort	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c	c
	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	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	c	a	a
	3) Ascertainment of exposure: a) secure record (e.g. surgical records)*; b) structured interview*; c) written self-report; d) no description	c	c	c	c	a	c	c	c	c	c	c	c	c	c	c	a	c	c
	4) Demonstration that outcome of interest was not present at start of study: a) yes*; b) no	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	c	a	a
Comparability	**1) Comparability of cohorts on the basis of the design or analysis:**																		
	a) study controls for SES (maternal education and/or income)*	x	x	x	x	x	x	x	x	x	x	x	x	x	a	x	x	x	x
	b) study controls for child dietary factors other than breastfeeding (e.g. dietary patterns, fish intake)*									x		x				x	x		
Outcome	**1) Assessment of outcome:** a) independent blind assessment*; b) record linkage*; c) self-report; d) no description	a	c	a	c	a	c	c	a/c	c	c	a	a	c	a	c	a	c	a
	2) Was follow-up long enough for outcomes to occur: a) yes*; b) no	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
	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 - less than 20 % lost or description of those lost suggested no difference from those followed*; c) follow up rate < 80% and no description of those lost; d) no statement	c	c	c	c	c	c	c	c	c	c	d	c	c	d	c	c	c	c
Total number of stars		**5**	**4**	**5**	**4**	**5**	**4**	**4**	**5**	**5**	**4**	**6**	**5**	**4**	**5**	**5**	**6**	**4**	**5**
Quality rating according to guideline**		**fair**	**poor**	**fair**	**poor**	**fair**	**poor**	**poor**	**fair**	**poor**	**poor**	**fair**	**fair**	**poor**	**fair**	**poor**	**fair**	**poor**	**fair**

*=one star (marked in yellow when each respective study were given a star)

**Thresholds for converting the NOS rating to Agency for Healthcare Research and Quality - AHRQ - standards (good, fair, and poor):

Good quality: 3 or 4 stars in Selection domain AND 1 or 2 stars in Comparability domain AND 2 or 3 stars in Outcome domain

Fair quality: 2 stars in Selection domain AND 1 or 2 stars in Comparability domain AND 2 or 3 stars in Outcome domain

Poor quality: 0 or 1 star in Selection domain OR 0 stars in Comparability domain OR 0 or 1 stars in Outcome domain

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	3
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	3
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	3-4
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	4
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	4
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	4
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	4
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	4
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	4
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	4
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	4-5

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	5
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	5
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	5-6
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	6-8
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	9
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	9-10
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	11
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	11-12
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	11-12
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	12-15
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	15
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	15-16
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	16

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.