(*Acta Anaesth. Belg., 2020, 71, 51-65*)

Anesthetic neurotoxicity in the pediatric population : a systematic review of the clinical evidence

R. Foubert (*), S. Devroe (**), L. Foubert (***), M. Van de Velde (****), S. Rex (****)

Abstract : *Background* : Exposure to general anesthesia (GA) in early life is known to be neurotoxic to animals.
Objectives : To evaluate the risk of GA inducing long-term neurodevelopmental deficits in human children.
Design : Systematic review.
Methods : We included observational and randomized studies that compared the long-term neurodevelopment of postnatal children exposed to GA to the long-term neurodevelopment of children not exposed to GA. We searched MEDLINE, Embase and Web of Science for relevant studies published in the year 2000 or later. We screened all the identified studies on predetermined inclusion and exclusion criteria. A risk of bias assessment was made for each included study. We identified 9 neurodevelopmental domains for which a sub-analysis was made: intelligence; memory; learning; language/speech; motor function; visuospatial skills; development/emotions/behavior; ADHD/attention; autistic disorder.
Results : We included 26 studies involving 605.391 participants. Based on AHRQ-standards 11 studies were of poor quality, 7 studies were of fair quality and 8 studies were of good quality. The major causes of potential bias were selection and comparability bias. On 2 neurodevelopmental domains (visuospatial skills and autistic disorder), the available evidence showed no association with exposure to GA. On 7 other neurodevelopmental domains, the available evidence showed mixed results. The 4 studies that used a randomized or sibling-controlled design showed no association between GA and neurodevelopmental deficits in their primary endpoints.
Limitations : The absence of a meta-analysis and funnel plot.
Conclusions : Based on observational studies, we found an association between GA in childhood and neurodevelopmental deficits in later life. Randomized and sibling-matched observational studies failed to show the same association and therefore no evidence of a causal relationship exists at present. Since GA seems to be a marker, but not a cause of worse neurodevelopment, we argue against delaying or avoiding interventional or diagnostic procedures requiring GA in childhood based on the argument of GA-induced neurotoxicity.

Keywords : Anesthesia ; toxicity ; pediatrics ; growth and development.

Introduction

Anesthesia in children is known to be safe within short-term observation and has enabled children to undergo a wide range of surgical and diagnostic procedures. However, over the past 15 years, much concern has been raised over potential long-term neurotoxic effects of anesthetics to the developing brain and associated neurodevelopmental deficits in children. In earlier systematic reviews (dating from 2012 and 2014), DiMaggio et al. (1) and Wang et al. (2) both showed a modestly increased risk of adverse neurodevelopmental outcomes in children that were exposed to general anesthesia. Since then more and higher quality clinical evidence has become available. This current systematic review aims to re-evaluate the impact of anesthesia exposure in childhood on the occurrence of later deficits in a range of neurodevelopmental outcomes.

In 2003, Jevtovic-Todorovic and colleagues showed that the combined administration of midazolam, nitrous oxide and isoflurane to 7-day old infant rats caused widespread apoptotic neurodegeneration in the developing brain and persistent deficits in memory and learning (3). These effects

R. Foubert, M.D.; S. Devroe, M.D.; L. Foubert, MD, DSc, PhD; M. Van de Velde MD, PhD; S. Rex, MD, PhD
(*) Department of Anesthesiology, University Hospitals Leuven, Leuven, Belgium. Faculty of Medicine, Biomedical Sciences Group, KU Leuven, Leuven, Belgium.
(**) Department of Anesthesiology, University Hospitals Leuven, Leuven, Belgium.
(***) Department of Anesthesiology and Critical Care and Emergency Care Medicine, Onze-Lieve-Vrouw Ziekenhuis, Aalst, Belgium.
(****) Department of Anesthesiology and Cardiovascular Sciences, University Hospitals Leuven, KU Leuven, Leuven, Belgium.
Corresponding author : R. Foubert, MD, Department of Anesthesiology, University Hospitals Leuven, 3000 Leuven, Belgium.
E-mail : Ruben.Foubert@student.kuleuven.be

Paper submitted on Apr 28, 2020 and accepted on May 08, 2020
Conflict of interest : None

were found to be mediated through stimulation of $GABA_A$-receptors and inhibition of NMDA-receptors by the anesthetics. Later evidence showed that both midazolam (a $GABA_A$-receptor agonist) alone and ketamine (a NMDA-receptor antagonist) alone can cause apoptosis in the developing rodent brain (4, 5). Unfortunately, most known anesthetics act on at least one of both receptors (6) and are therefore potentially neurotoxic (7), α_2-agonists such as dexmedetomidine being the only exception. Other studies discovered additional pathophysiological mechanisms (see table 1) that might cause anesthetics to be harmful to the developing brain (8-13). These concerning laboratory findings formed the basis for a multitude of clinical trials, but have recently also been met with much criticism. The circumstances in which animals were investigated did not always represent clinical practice in humans. Some reports showed respiratory depression during anesthesia which might also cause adverse neurodevelopment (7, 14, 15). Other studies did not apply noxious stimuli. New evidence shows that noxious stimuli such as a surgical incision might ameliorate anesthetic neurotoxicity (7, 16). Rodents were often under anesthesia for hours which would correspond to multiple days of anesthesia in children, considering their mutual lifespans. These conditions almost never occur in clinical practice (7). Lastly, animal studies were often performed on rodents during the first 2 weeks of life to mimic the stage of brain development of young children. Recent evidence however, shows that it corresponds better to the brain development of human fetuses of 17-22 weeks postmenstrual age (17-19).

The first clinical trials on the subject of anesthetic neurotoxicity in childhood were observational and retrospective by design and thus inherently prone to confounding factors such as comorbidity and genetic predisposition. They also failed to distinguish between anesthesia and surgery as the attributable cause for potentially observed deficits. The only prospective evidence to date was published in 2016 and 2019 by Davidson and McCann in the GAS-trial (20, 21). The GAS-trial is a randomized controlled trial (RCT) that randomized children scheduled for inguinal hernia repair to either general anesthesia (GA) (plus an optional locoregional block) or to locoregional anesthesia alone. It was found that children exposed to less than 1 hour of sevoflurane before 60 weeks of postmenstrual age did not differ from unexposed children in intelligence, memory, learning, language, motor skills, attention or behavior at 5 years of age.

Despite a growing body of evidence, no definitive conclusions can be drawn towards the long-term safety of anesthesia in children. Much of the published evidence is contradictory, with some studies finding associations between exposure to anesthetics in childhood and a diversity of neuro-cognitive and behavioral problems in later life and other studies finding no resulting deficits at all.

Despite the concerns that the pre-clinical evidence might not accurately represent the clinical practice of anesthesia in children and the fact that the clinical evidence is unequivocal, the FDA did issue a warning regarding the potential neurotoxic effects of anesthesia in young children. They stated that "health care professionals should balance the benefits of appropriate anesthesia in young children and pregnant women against the potential risks" (22, 23). Clearly, with millions of children undergoing anesthesia each year, the subject of anesthetic neurotoxicity in the pediatric population is of great importance, in particular as a relevant proportion of children undergo repeated and prolonged (i.e., a duration > 3h) anesthetic procedures before age 3, which is exactly the population at risk as defined by the FDA (24).

Anesthetic neurotoxicity is not only relevant for the individual child, but also for our public health and our health policy. For example, should elective surgery in childhood be delayed or avoided? To examine whether anesthesia induces neurotoxicity in the pediatric population and if we should change our health policy concerning elective surgery, we aim to systematically review the available clinical studies.

Table 1

Neurological changes noted in infant animals following anaesthetic exposure

Type of injury	Animal species
Neuroapoptosis (8)	Rodents, non-human primates, nematodes
Alterations in dendritic spines (9)	Rodents
Altered neurogenesis (10)	Rodents
Impaired astroglial development (11)	Rodents (astroglial cultures)
Mitochondrial degeneration (12)	Rodents
Decreases in trophic factors (13)	Rodents (primary neurons)

Methods

Ethics approval

This systematic review received ethical approval by OBC, the ethics committee of the group biomedical sciences, KU Leuven, chairman Pascal

Borry, Herestraat 49, 3000 Leuven on March 13, 2018. The internal reference number of our ethics approval is MP005536.

Eligibility criteria

Studies : Primary retrospective, ambidirectional or prospective clinical studies examining the effect of general anesthesia on the neurodevelopment of children exposed before 18 years of age. Only studies with a free full text in English were included. In accordance with earlier systematic reviews on anesthetic neurotoxicity in children by DiMaggio (1) and Wang (25), only studies published in the year 2000 or later were eligible to better reflect current clinical practice and to avoid a wide array of different anesthetic techniques and agents used in the past.

Participants : Postnatal children younger than 18 years.

Intervention : General anesthesia without limitations in drugs, depth or duration of anesthesia. The indication for general anesthesia could be any surgical or diagnostic procedure. The intervention group should be compared to a control group that was solely exposed to locoregional anesthesia for the same procedure or to a control group that was not exposed to any form of anesthesia at all.

Outcome measures : Neurodevelopmental outcomes measured at least 6 months after exposure to GA, including intelligence; memory; learning; language/speech; visuospatial skills; motor function; development/behavior/emotions; ADHD/attention; and autistic disorder. Some studies compared a neurocognitive parameter such as IQ before and after the administration of GA in the same children. We excluded self-controlled studies not comparing the observed evolution with a control group not exposed to GA.

Information sources

Studies were identified by searching electronic databases and by scanning reference lists of relevant articles. We searched MEDLINE using PubMed, Embase and Web of Science from the 1st of January 2000 to the 1st of August 2018.

We used the following term to search MEDLINE through PubMed: "(pediatrics[Mesh] OR "pediatric*"[tiab] OR "pediatric*"[tiab] OR child[Mesh] OR child*[tiab] OR kid[tiab] OR kids[tiab]) AND (anesthesia[Mesh] OR anesthesia[tiab] OR anesthetics[Mesh] OR anesthetic*[tiab] OR "anesthetic drug*"[tiab] OR "anesthetic agent*"[tiab]) AND (cognition[Mesh] OR cogniti*[tiab] OR neurocogniti*[tiab] OR learning[Mesh] OR learn*[tiab] OR development*[tiab] OR neurodevelopment*[tiab] OR language[Mesh] OR language[tiab] OR behavior[tiab] OR behavior[tiab])".

Filters for publication date, language and availability of full texts were applied as described above. The "humans" filter was also applied.

The search terms for Embase and Web of Science are listed in the appendix.

Study selection

Two authors (RF and SR) independently selected studies, assessed the risk of bias and extracted data. We resolved all disagreements by consensus. If we could not reach a consensus, then a third author (LF) was available to give a final decision.

We screened all articles on their titles and removed all duplicates. We screened the remaining articles using the inclusion and exclusion criteria on their abstract and then again on their full text.

Data collection process

We developed a data sheet to extract relevant data from each included study.

Data items

The following information was extracted from each included study: (a) author, (b) year of publication, (c) design type (observational/randomized/retrospective/ambidirectional/prospective), (d) location of the researched population (city, state, country), (e) time of exposure to GA, (f) study quality as assessed by the 9-star Newcastle-Ottawa scale for non-randomized studies or by the Cochrane Collaboration's tool for randomized studies, (g) type of surgery, (h) anesthetics used, (i) age at exposure, (j) number of times of exposure, (k) control group, (l) number of exposed children versus number of controls, (m) main outcomes, (n) time of evaluation of neurocognitive outcomes, (o) main results and (p) independent risk factors for worse neurocognitive outcomes.

Risk of bias in individual studies

We assessed the risk of bias in the non-randomized studies with the 9-star Newcastle-Ottawa Scale, as recommended by The Cochrane

Handbook for Systematic Reviews of Interventions (26). An individual study can have 0-9 stars with a higher score corresponding to a lower risk of bias. In the randomized studies, we assessed the risk of bias with the Cochrane Collaboration's tool. It uses 6 entries with high, low or unclear risk of bias for each.

An example of a completed form is available in the appendix for both tools.

The Newcastle-Ottawa scale allows conversion to AHRQ (Agency for Healthcare Research and Quality) standards (good, fair and poor) as described in the appendix.

Summary measures

The primary measure of intervention effect was the frequency of studies that found a significant effect for 9 outcomes : (a) intelligence ; (b) memory ; (c) learning ; (d) language/speech ; (e) motor function ; (f) visuospatial skills ; (g) development/ emotions/behavior ; (h) ADHD/attention ; and (i) autistic disorder.

Planned methods of analysis

We provide a narrative summary for each of the 9 summary measures.

Results

Study selection

A flow diagram of the study selection process is presented in Figure 1.

The two additional articles that were obtained by screening reference lists were those by O'Leary et al. and Warner et al. (27, 28). The study by Davidson et al. (20) was removed because it reported on the secondary outcome of the GAS-trial and the study that reported on the primary outcome of the GAS-trial (29) was already included.

Study characteristics

Methods: Of the 26 included studies, 25 were observational and 1 was prospectively randomized. Of the former 15 were retrospective and 10 were ambidirectional. The duration of follow-up ranged from 6 months to 20 years after exposure. All outcomes were measured during childhood or adolescence with a maximum age of 20 years.

Participants: This systematic review reports on a total of 605.391 participants: 119.593 exposed children and 485.798 controls. The age of exposure ranged from the first days of life up until 16 years of age, with the majority being exposed before 4 years of age. Most studies were monocentric. The most studied locations were North America (13/26) and Europe (5/26).

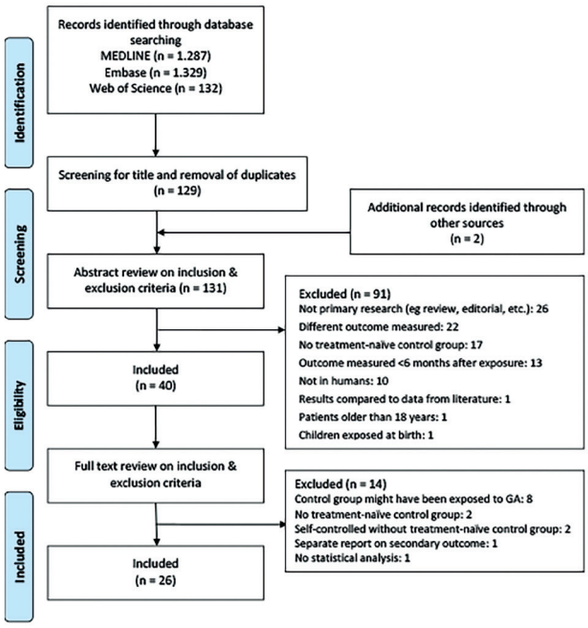

Fig. 1. — The literature search process.

Intervention: All studies had GA as intervention, usually compared to a matched, treatment-naïve control group. The degree of matching varied considerably, ranging from using only gender and age to using 50 variables. A minority of studies compared the exposed children to siblings (30-32) or to children undergoing surgery with locoregional anesthesia (29, 33).

Most studies included single exposures, but 14 studies also explicitly examined multiple exposures.

Outcomes: Of 26 trials, 14 reported on intelligence, 6 reported on memory, 7 on learning, 13 on language/speech, 7 on motor skills, 3 on visuospatial skills, 16 on development/behavior/ emotions, 8 on ADHD/attention and 3 on autistic disorder.

Risk of bias within studies

The quality assessment of the non-randomized studies is shown in figures 2 and 3. The randomized study (29) had 5 items with a low risk of bias and 1 item with a high risk of bias. We converted the 9-star Newcastle-Ottawa scale to AHRQ-standards. According to the AHRQ-standards, 11 studies were of good quality, 7 were

Study ID	A	B	C	D	E	F	G	H	I
Bartels, 2009	*	*		*		*	*	*	*
Wilder, 2009	*	*	*		*	*	*	*	*
DiMaggio, 2011		*	*	*	*	*	*	*	*
Flick, 2011	*	*			*	*	*	*	*
Ing, 2012	*	*			*	*	*	*	*
Bong, 2013	*		*					*	*
Ko, 2014	*	*	*	*	*	*	*	*	*
Stratmann, 2014	*		*		*		*	*	*
Ing, 2014	*	*			*	*	*	*	*
Chernsly, 2014	*	*	*					*	*
Bakri, 2015	*		*					*	*
Ko, 2015	*	*	*	*	*	*	*	*	
Backeljauw, 2015		*	*		*	*	*	*	*
Doberschuetz, 2016			*		*	*	*	*	*
Sun, 2016	*	*	*		*	*	*	*	*
Kermany, 2016	*	*	*				*	*	*
O'Leary, 2016	*	*	*		*	*	*	*	*
Graham, 2016	*	*	*		*	*	*	*	*
Glatz, 2017	*	*	*		*	*	*	*	*
de Haar, 2017	*	*					*	*	
Nestor, 2017	*		*		*	*	*	*	
Hu, 2017	*	*			*	*	*	*	*
Ing, 2017	*	*	*	*	*	*	*	*	*
Warner, 2018	*	*	*		*	*	*	*	*
Khochfe, 2018	*	*	*					*	*

Fig. 2. — Risk of bias evaluation according to the Newcastle-Ottawa Scale for assessing the quality of non-randomized studies.
A-I represents rating categories according to the Newcastle-Ottawa Scale for assessing the quality of non-randomized studies. A, Representativeness of the exposed cohort ; B, Selection of the non-exposed cohort ; C, Ascertainment of exposure ; D, Outcome of interest was not present at start of study ; E, Comparability of cohorts - major factor controlled for ; F, Comparability of cohorts - any additional factor controlled for ; G, Assessment of outcome ; H, Follow-up long enough for outcome to occur? I, Adequacy of follow-up of cohorts.

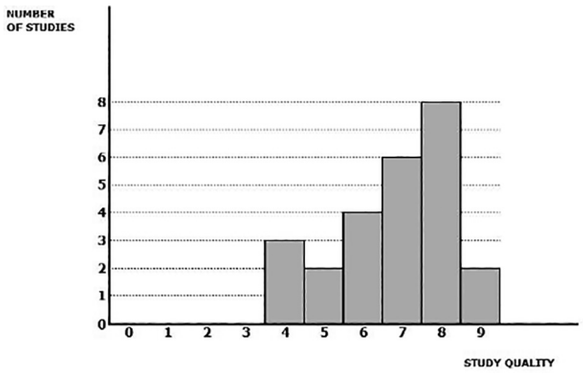

Fig. 3. — The study quality of the observational trials as assessed by the 9-star Newcastle-Ottawa scale.

of fair quality and 7 were of poor quality. The randomized trial generally showed a low risk of bias. To summarize, the evidence presented in this trial is of mixed quality.

Results of individual studies

Results and additional extracted data are presented in table 2.

Synthesis of results

Intelligence *(14 studies)* : Five studies found that intelligence was significantly reduced in exposed children (34-38), including the very large observational cohort of Glatz (36). A smaller study (38) found decreased intelligence only in multiply exposed children. Among the negative studies (28, 29, 31, 32, 39-43) were the twin study by Bartels (31), the sibling-controlled PANDA-study (32) and the GAS-trial (29).

Memory *(6 studies)* : Multiple, but not single exposures affected memory in the MASK-trial (28) (only 1 of 5 tests) and in the trial by Flick (39). Stratmann (41) found recollection, but not familiarity to be impaired in exposed children, whereas Sun (32), Kermany (44) and the GAS-trial (29) found no effect.

Learning *(7 studies)* : Three studies (38, 39, 45) showed that multiple, but not single exposures were associated with learning disability. The MASK-trial (28) confirmed these results in the reading-domain, but not in the mathematics-domain. Bong (40) confirmed an association without differentiating between single and multiple exposures. The GAS-trial (29) and the PANDA-trial (32) did not show an increased risk. Exposure in all 7 studies occurred before 4 years of age. Missed school days did not play a role in these results.

Language/speech *(13 studies)* : Three small studies (34, 43, 46) showed worse outcomes and 5 studies (27, 28, 35, 39, 47) showed mixed results. Children exposed before 2 years of age were at risk in one study, but safe when exposed between 2-4

Table 2

Results of individual studies

Author	Year	Design type	Location	Time of exposure	Quality	Surgery	Anesthetics	Age at exposure	Times of exposure	Control group	#exposed/ #controls	Main outcomes	Time of evaluation	Main results	Independent risk factors
Bartels	2009	Observational, retrospective; monozygotic twin design	Netherlands	1986-1998	7	NS	NS	< 3 years & < 12 years	NS	Monozygotic twin pairs	400/1,756 & 914/1,372	Intelligence (EA, CP)	12 years of age	Twin pairs discordant for exposure: lower EA and more CP than concordant non-exposed twins. Within discordant twin pairs: no difference in EA or CP.	
Wilder	2009	Observational, retrospective cohort	Olmsted County, Minnesota, USA	1976-1986	8	NL	N_2O 91%, Halo 89%, Ket 9%	< 4 years	Single and multiple	Cohort	593/4,764	Learning	5 - 19 years of age	Multiple (HR = 1.59 & 2.60 for 2 & 3 or more exposures, respectively), but not single exposures were associated with more LD at 5 years of age.	Longer cumulative duration of anesthesia
DiMaggio	2011	Observational, retrospective cohort; sibling-matched sub-analysis	New York State, USA	1999-2005	8	NL	NS	< 3 years	Single and multiple	Cohort + sub-analysis in siblings	304/10,146	Developmental or behavioral disorders	At diagnosis, loss to follow-up or end of 2005	Multiple (HR = 2.9 & 4.0 for 2 & 3 or more exposures, respectively), but not single exposures were associated with more devolopmental and behavioral disorders. A sub-analysis within discordant twin pairs showed no difference in outcomes for any number of exposures compared to no exposure.	Male gender, birth complications
Flick	2011	Observational, retrospective cohort	Rochester, Minnesota, USA	1976-1984	7	NL	N_2O 88.1%, Halo 87.5%	< 2 years	Single and multiple	1:2 matched on gender, mother's education, birth weight, gestational age, birth date	350/700	Cognitive functioning & achievement, learning, language & speech, behavior & emotions	At emigration, death, loss follow-up or 19 years old	Cognitive functioning & achievement: no association except on the subscale memory where multiple, but not single exposures were associated with worse outcomes. Learning: multiple (HR = 2.12), but not single exposures were associated with more LD. Language & speech: multiple (HR = 3.49), but not single exposures were associated with worse speech/language. Behavior & emotions: no association.	
Ing	2012	Observational, retrospective cohort	Perth, Australia	1989-1995	7	NL	NS	< 2 years	Single and multiple	Cohort	321/2,287	Intelligence (cognition), language, motor function, behavior	10 years of age	Intelligence: single (aRR = 1.73), but not multiple (aRR = 1.92) exposures were associated with increased risk of disability. Language: both single (aRR = 2.36) and multiple (aRR = 2.68) exposures were associated with an increased risk of disability. Motor function: no association. Behavior: no association.	
Bong	2013	Observational, retrospective case control	Singapore	1998-2000	4	Minor surgery	Sevo 100%, Thio 11.7%, Prop 0.5%	< 1 year	NS	Cohort	100/106	Intelligence (EA), learning	12 years of age	EA: no association. Learning: more LD (OR = 4.5) in exposed children.	Male gender, formal LD, lower maternal education and lower housing category were associated with worse EA
Ko	2014	Observational, retrospective cohort	Taiwan	2001-2008	9	NS	NS	< 3 years	Single and multiple	1:4 matched	3,293/13,172	ADHD	31 december 2010	GA was not associated with a higer risk of developing ADHD for any number of exposures or any age at exposure.	Congenital anomalies. Class II, III, IV and V class of residence were related to a lower risk of ADHD in comparison to group I class of residence.
Stratmann	2014	Observational, ambidirectional case-control	California, USA	Billing in 2004	6	Excluded: neuro and CHD	Sevo 92.9%, Iso 46.4%, Halo 25%, Prop 10.7%	< 1 year	NS	Matched on age and gender from a registry of parents interested in pediatric research	28/28	Intelligence (IQ), memory, behavior	6-11 years of age	IQ: no associaton. Memory: recollection, but not familiarity was lower in exposed children. Behavior: no associaton.	
Ing	2014	Observational, retrospective cohort	Perth, Australia	1992-2002	7	NL	NS	3-5 years & 5-10 years	NS	No GA < 5 years & no GA < 10 years	220/2,107 & 155/1,952	Intelligence (cognition), language, motor function, behavior	10 years of age	Intelligence: no associaton. Language: no associaton. Motor function: both exposure between 3-5 years (aRR = 2.32) and exposure between 5-10 years (aRR = 2.33) was associated with motor deficit. A subsequent suba nalysis showed this was true for both fine and gross motor skill. Behavior: no associaton.	
Chemaly	2014	Observational, ambidirectional, case control	Beirut, Lebanon	2004-2005	5	Surgery or diagnostic procedure	NS	< 4 years	Single	Children from the same institution	292/300	Behavior	10-11 years of age	Parent-reported behavioral problems were more frequent in the exposed than in the control group (28.4% vs 5.7%).	Exposure to GA for surgery compared to exposure for diagnostic procedures, longer duration of GA, younger age at exposure (especially < 6 months), multiple anesthetic agents.

Author	Year	Design type	Location	Time of exposure	Quality	Surgery	Anesthetics	Age at exposure	Times of exposure	Control group	#exposed/ #controls	Main outcomes	Time of evaluation	Main results	Independent risk factors
Bakri	2015	Observational, ambidirectional, case control	Egypt	NS	4	Day case	Sevo 100%	< 5 years	2	Age-matched from the vaccination clinic	35/35	Behavior	1,5-5 years of age	Increased risk of being anxious or depressed (RR = 11), of having sleep problems (RR = 4.5), attention problems (RR = 8), anxiety problems (RR = 3.7) and attention deficit/hyperactivity problems (RR = 3). No difference for 9 other items.	
Ko	2015	Observational, retrospective cohort	Taiwan	2001-2009	8	NL	NS	< 2 years	Single and multiple	1:4 matched	5,197/20,788	Autistic disorder	3-5 years of age	No association with autistic disorder. Neither age at first exposure or total number of exposures influenced the risk of autistic disorder. The exposed group had more perinatal conditions, congenital anomalies, neurological diseases and endocrine diseases.	Home location III - IV were associated with a lower risk of autistic behavior. Congenital anomalies and neurological diseases were associated with a greater risk of autistic behavior.
Backeljauw	2015	Observational, ambidirectional cohort	Cincinnati, Ohio, USA	1986-2003	7	NL, but 65% ENT	N_2O 88%	< 4 years	1 - 3	1:1-matched on age, gender, handedness, socioeconomic status	53/53	Intelligence, language	5-18 years of age	Intelligence: exposed children scored lower on performance IQ, but normal on verbal and full scale IQ. Language: exposed children had worse listening comprehension but scored normal on two other language tests.	
Doberschuetz	2016	Observational, ambidirectional cohort	Frankfurt, Germany	2008-2011	6	GI mal-formations	NS	< 28 days	Single	Healthy twin if available. Ohterwise from hospitals and pediatric	40/40	Intelligence (cognition), language, motor function	2 years of age	Intelligence: no association. Language: lower index in exposed children. Motor function: no associaton.	
Sun	2016	Observational, ambidirectional sibling-matched cohort	USA	2000-2010	7-8	Inguinal hernia repair	Sevo 95.3%, Iso 59.1%, IV anesthetics 26.7%	< 3 years	Single	1:1 sibling (closest in age, < 3 years difference) matched	105/105	Intelligence (IQ), cognition (memory, learning, motor, visuospatial, language, attention, executive function) and behavior	8-15 years of age	Intelligence: no associaton for exposure in general, for any specific age at exposure or any duration of exposure. Cognition: no associaton. Behavior: exposure was associated with a higher chance of having an abnormally bad internalizing score.	Race, study site and socioeconomic status was associated with IQ. Gender was associated with most domain-specific cognitive functions and behavior.
Kermany	2016	Observational, ambidirecitonal cohort	Teheran, Iran	2014	6	Glaucoma	Sevo 100%	5-16 years	Single and multiple	From same glaucoma clinic	68/47	Memory, language	> 6 months after exposure	Memory, language: no associaton overall, for any specific age groups or for any number of exposures.	Gender (females performing better on the FDS)
O'Leary	2016	Observational, retrospective cohort	Ontario, Canada	1998-2012	8	NL	NS	< 7 years	Single and multiple	Cohort	28,366/ 55,910	Any domain of EDI (5 developmental domains) in lowest 10th percentile.	5-6 years of age	Children < 2 years at first surgery did not have increased odds of early developmental vulnerability, but children > 2 years at first exposure did (OR= 1.05). Increasing frequency of exposure was not associated with a greater risk of early developmental vulnerability.	
Graham	2016	Observational, retrospective cohort	Manitoba, Canada	2000-2009	8	NL	NS	0-2 years & 2-4 years	Single and multiple	1:3-matched on birth year, gender, mother's age at birth of first child, income quintile, urban vs rural residence	4,470/13,586	EDI-domains (communication/general knowledge, emotional, language/cognition, physical and social)	Kindergarten	Children < 2 years at first exposure: no associaton for single or multiple exposures. Single exposure between 2-4 years: incresed risk of deficits in communication/general knowledge, language/cognition, physical and social domains. Multiple exposures did not further increase this risk. Children exposed to GA < 4 years were more likely to be born prematurely or large for gestational age, needed more health care resources, were more likely to require income assistance or to be taken into care by the child welfare system prior to exposure.	Income assistance, child and Family Services involvment, mother's age birth of first child, gestational age, more utilization of health care resources prior to exposure.
Glatz	2017	Observational, retrospective cohort	Sweden	1973-1997	8	Excluded: neuro, cardiac, thoracic, cancer surgery	NS	< 4 years	Single and multiple	1:5-matched children on gender, month of birth, maternal parity, county of residence at delivery	33,514/ 159,619	Intelligence (EA, IQ)	16 & 18 years of age	Single exposure < 4 years: 0.41% lower mean school grades at 16 years of age and 0.97% lower mean IQ test score at 18 years of age. These differences did not increase with multiple exposures. The overall difference was 5 to 10 times smaller than the differences associated with gender, maternal education or month of birth. Exposure did no increase the risk of having school grades below the 10th percentile. Exposure was associated with an increased risk of not having recorded school grades at age 16 (OR = 1.29).	Gender, maternal educational level and month of birth (all had bigger influence than GA exposure).

Author	Year	Design type	Location	Time of exposure	Quality	Surgery	Anesthetics	Age at exposure	Times of exposure	Control group	#exposed/ #controls	Main outcomes	Time of evaluation	Main results	Independent risk factors
de Heer	2017	Observational, retrospective cohort	Rotterdam, Netherlands	2002-2011	4	NS	NS	< 5 years	NS	Cohort	415/3,026	Intelligence (IQ)	6 years of age	The average IQ was 2.1 points lower in children exposed to anesthesia. Children who required anesthesia before 5 years of age were more likely to be male, to have mothers with lower educational level or who smoked during pregnancy.	Born < 32 weeks (-9,8 IQ points), lower maternal education level, lower maternal IQ and smoking while pregnant.
Nestor	2017	Observational, ambidirectional case control	Florida, USA	2010-2011	6	Surgery, diagnostic	NS	< 1 year	Single and multiple	Matched on gender, race, gestational age, birth weight, comorbidities	298/129	Language & speech, motor function, development, behavior, ADHD, autistic disorder	Median follow-up of 5,1 years	Language & speech: more need for speech therapy. Motor: no association. Development & behavior: more developmental delay. ADHD: no associaton. Autistic disorder: no association. Patients who were exposed to GA for MRI had the poorest outcomes.	
Hu	2017	Observational, retrospective, birth cohort	Olmsted County, Minnesota, USA	1996-2003	7	NL	N_2O 85%, Sevo 72%, Iso 36%, Halo 16%, Thio 6%, Prop 4%	< 3 years	Single and multiple	Matched on 50 variables	573/463	Intelligence (EA), learning, language & speech, emotions & behavior, ADHD	At emigration, death, last follow-up or december 31, 2014.	EA: multiple, but not single exposures were associated with decreased test scores. Learning: multiple (HR = 2.17) but not single exposures were associated with an increased risk of LD overall. A subanalysis showed that for reading also single exposures were associated with worse test scores. Language & speech: no association. Emotions & behavior: no associaton. ADHD: multiple, but not single exposures were associated with an increased risk.	
Ing	2017	Observational, retrospective, cohort	Texas & New York, USA	1999-2010	9	Minor surgery	NS	< 6 years	Single	Cohort, 1:5-matched on 50 variables	38,493/ 192,465	Mental disorders, sub-analysis of developmental delay and of ADHD	12 years of age	Increased risk of any mental disorder (HR = 1.26), of developmental delay (HR = 1.26) and of ADHD (HR = 1.31). A sub-analysis of age at exposure showed no significant differences for any age groupes.	
Warner	2018	Observational, ambidirectional cohort	Olmsted County, Minnesota, US	1994-2010	8	NL	N_2O 90%, Sevo 79%	< 3 years	Single and multiple	Matched on 40 variables	586/411	Intelligence (IQ), cognition (attention, memory, executive function, language, visual, motor), parent-reported executive function	8-12 & 15-19 years of age	IQ: no associaton for single or multiple exposures. Cognition: single exposure was not associated with any deficits. Multiple exposed children scored worse on 1 reading and 1 fine motor scale, but not on any other scale. Parent-reported executive function: more problems reported for global executive function and learning in singly and multiply exposed children. More problems reported for behavior in multiply exposed children.	
Khochfe	2018	Observational, ambidirectional cohort	Beirut, Lebanon	2002-2006	5	NS	NS	< 2 years	Single	Locoregional block	168/226	Behavior	8-15 years of age	GA was associated with a higher chance of parent-reported behavioral problems.	Younger age at time of exposure, longer surgery duration
McCann	2019	RCT	Australia, Canada, Italy, Netherlands, New Zeland, UK, USA	2007-2013	Good	Inguinal hernia repair	Sevo (+/- locoregional block)	< 60 weeks post-menstrual age	Single	Spinal, caudal or combinded spinal caudal block	242/205	Intelligence (IQ), memory, learning, language, motor skills, visuospatial skills, behavior, attention, autistic disorder	2 years of age	No associaton wtih any of the measured outcomes. In a secondary endpoint measured at 2 years of age, no association was found either.	

RCT = randomized controlled trial. NS = not specified, NL = no limitations. CHD = congenital heart disease, ENT = ear-nose-throat, GI = gastrointestinal. N2O = nitrous oxide, Halo = halothane, Ket = ketamine, Sevo = sevoflurane, Thio = thiopental, Prop = propofol, Iso = isoflurane. GA = general anesthesia. EA = educational achievement, CP = cognitive performance, LD = learning disability, ADHD = attention deficit hyperactivity disorder, IQ = intelligence quotient, EDI = Early Development Instrument. HR = hazard ratio, aRR= adjusted risk ratio, OR = odds ratio, RR = risk ratio, MRI = magnetic resonance imaging, RA = regional anesthesia. FDS = forward digit span test.

years of age (27). Exactly the opposite was shown by another study (47). Flick (39) only showed worse outcomes for multiply exposed children. The MASK-trial (28) showed that multiple, but not single exposures were associated with worse scores on a reading test, but not on another language test. Both multiply and singly exposed children scored normally on another language test. The other 5 studies (20, 32, 34, 38, 44), including GAS (29) and PANDA (32), showed no association between exposure to GA and language and speech.

Motor skills *(7 studies)* : The MASK-trial (28) showed that multiply exposed children scored worse on fine motor skills, whereas singly exposed children did not. Another trial by Ing (42) also showed worse motor skills in exposed children. The GAS (29) and PANDA-trials (32) and three smaller trials (34, 43, 46) were all negative.

Visuospatial skills *(3 studies)* : The GAS (29), PANDA (32) and MASK-trials (28) were the only studies to examine the visuospatial abilities of children exposed to GA. They all showed no association.

Development/behavior/emotions *(16 studies)* : In 12 studies Child Behavior Checklists were administered to parents or teachers. Three studies (33, 48, 49) showed worse results, 3 studies (27, 32, 47) showed mixed results, 1 study (28) showed worse results in multiply exposed children and 5 studies (29, 34, 39, 41, 42) showed non-inferior scores in exposed children. Another 3 studies (30, 46, 50) used database-codes to detect diagnoses of developmental problems. They showed that those codes were more prevalent in exposed children. In 1 of those, a sub-analysis showed no difference within discordant twin pairs (30). Individualized education plans were not more prevalent in exposed children (38).

ADHD/attention *(8 studies)* : Three studies (38, 49, 50) showed an association between exposure and ADHD, but 2 of them only in multiply exposed children (38, 50). These results are contradicted by Ko (51) and the GAS-trial (29) who showed no such association. The GAS (29), PANDA (32) and MASK-trials (28) showed non-inferior results on neuropsychological tests of attention. Bakri (49) did find more attention problems in exposed children.

Autistic disorder *(3 studies)* : The only three studies (29, 46, 52) to examine the risk of autistic disorder in relation to exposure to GA showed no association.

Of the 13 studies that explicitly differentiated between single and multiple exposures, 5 showed multiple exposures to be associated with worse outcomes on one or more neurodevelopmental domains as compared to single exposures.

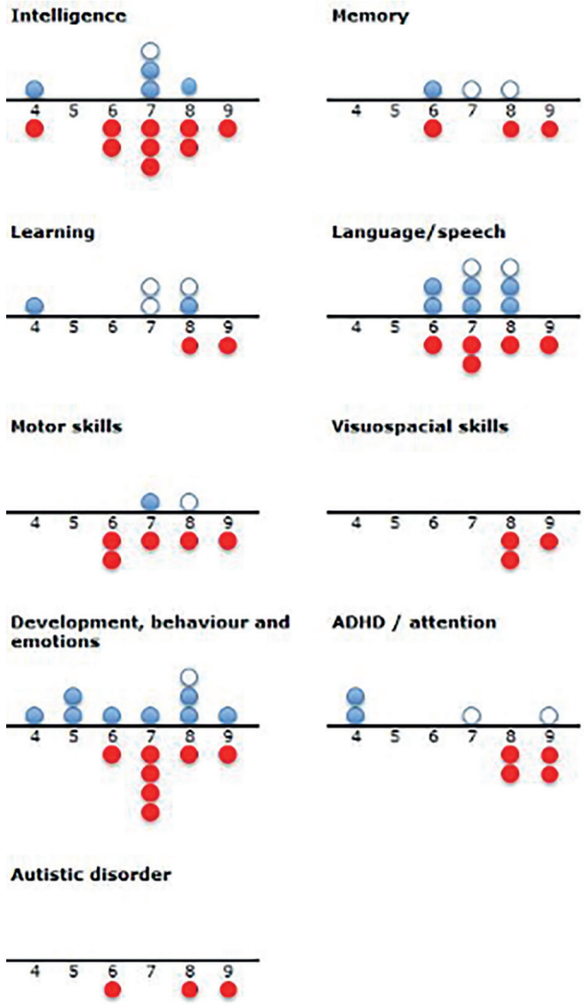

Fig. 4. — Graphical representation of the results for each neurodevelopmental domain. The X-axis represents the risk of bias assessment based on the 9-Star Newcastle-Ottawa Scale for assessing the risk of bias in observational studies. A higher number indicates a lower risk of bias. For the purpose of this graphic, the RCT by McCann (29) is attributed the highest score of 9 because it is a RCT with a low risk of bias according to the Cochrane Tool. A full blue circle represents a study that showed an association between exposure to GA and a neurodevelopmental deficit (both in singly and multiply exposed children or in a study that made no distinction between the two). An open blue circle represents a study that only showed an association in multiply, but not in singly exposed children. A full red circle represents a study that showed no association (either in both singly and multiply exposed children or in a study that made no distinction between the two).

Figure 4 graphically represents the results for each neurodevelopmental domain.

Potentially observed differences in exposed versus non-exposed children might be due to inherent neurotoxicity of anesthetics (through

$GABA_A$-agonism and NMDA-antagonism) and due to confounding modifiers such as (a) the conduct of anesthesia (occurrence of hypotension (53), hypoxemia, emergence delirium (54) and pain (55), (b) the effect of surgical stress (type and extent of surgery (56, 57), number of interventions) and (c) the patient's characteristics (age at exposure, nutritional status, comorbidities, genetics (57, 58) and neurodevelopmental status before exposure to GA).

Unfortunately, none of the included studies analyzed anesthesia records, so no data is available on the impact of (a) the conduct of anesthesia and (b) the effect of surgical stress on neurodevelopmental outcomes. Many observational studies did not specify which surgeries were performed in the exposed children and defined no exclusion criteria based on the type of surgery. Most of the observational studies did control for many (c) patient characteristics (such as age and gender) but were by design incapable of excluding possible confounding modifiers such as genetics and neurodevelopmental status before exposure to GA.

In 15/24 of studies, the specific anesthetics to which participants were exposed were not mentioned.

Risk of bias across studies

Selection bias : A major concern of bias is that none of the included studies showed that the outcome of interest was not present at the start of the study. At best, the study showed that the diagnosis of interest (e.g. an ICD-9 code for ADHD) had not been made before exposure to GA. In a retrospective study, the absence of such diagnosis prior to exposure and the presence of said diagnosis after the exposure does not prove the disease would not have occurred if there had never been exposure to GA. Many studies, however, did not address this methodological problem. The majority of studies used an exposed cohort that was representative of the surgical pediatric patient in general. The non-exposed cohort was usually drawn from the same community. The ascertainment of exposure was usually done by screening secure hospital or insurance records.

Comparability bias : Some studies used up to 50 elements to match exposed children to controls, whereas others barely performed any matching at all. Even the best-matched studies still fail to control for what is arguably the most important cause of bias, namely that sick children are compared to healthy children. Only 8 out of 26 studies (27, 29, 30, 32, 36, 40, 50, 52) excluded patients who required GA for major surgeries such as cardiac, oncological or intracranial operations, of which the underlying disease can reasonably be assumed to also interfere with neurodevelopment.

Children exposed to surgery necessitating GA are more likely to live in poorer socioeconomic circumstances (47, 51) or to be born prematurely (47, 51), which has also been shown to interfere with neurodevelopment (37, 47). Fortunately, 21 of the included studies (27-40, 42-44, 47, 50-52) controlled for socioeconomic circumstances and 21 for prematurity and/or low birth weight (27-29, 31-34, 36-48, 50).

In a twin study (31), strong evidence was provided that genetic predisposition can predict worse educational achievement and the need for GA. Only 4 of the 26 included studies (29, 31, 32, 43) (partially) controlled for genetic predisposition. Ko (51, 52) showed that congenital anomalies and chromosomal abnormalities were more prevalent in children who required GA.

Detection bias : The outcome of interest was usually ascertained by independent blind assessment or by linkage of secure hospital or insurance records. An important exception constitutes studies that used Child Behavior Checklists. The parents who completed these checklists were aware that their child had undergone GA and that their child took part in a study that examined the neurotoxic effects of GA. Parents who suspected their child's behavioral problems to be caused by exposure to GA might also have been more likely to take part in these studies.

Attrition bias : The follow-up was more than 6 months and the follow-up rate was usually more than 80%. Some studies showed, however, that patients lost to follow-up were different in baseline characteristics to patients who were included in the final analysis (28, 34, 42).

Publication bias : Assessing publication bias was outside the scope of this systematic review.

Discussion

Summary of evidence

We systematically reviewed the available clinical evidence on anesthetic neurotoxicity in the pediatric population and divided neurodevelopment into 9 subdomains to better systemize the available data.

For 2 of the studied subdomains (visuospatial skills and autistic disorder), the evidence is equivocally negative.

For 7 other subdomains (intelligence; memory ; learning; language/speech ; motor skills, development/behavior/emotions ; ADHD/attention), the evidence is mixed. Methodological differences in studies may contribute to the disagreement between human studies. Multiple factors, such as sample size, selection of exposed cohort, duration of exposure, age at exposure, type of surgery, indication for surgery, follow-up as well as the use of different effect size/outcome measures could likely influence results (59, 60). The general picture on these subdomains is that one hand the RCT by McCann et al., the GAS-trials (29) and the 3 observational trials that used a twin- or sibling-matched design (30-32), were negative on all primary endpoints. On the other hand, many observational studies that did not use a twin- or sibling-matched design to reduce confounding, but generally had large cohort sizes, often show an association between exposure to GA and a worse neurodevelopmental outcome. A few concerns regarding these observational studies should be raised.

First, many of the clinical trials used a neuropsychological battery rather than examining a specific domain. It was not uncommon for a trial to perform multiple or even tens of tests on exposed children. When a few of these tests turn out to be significantly worse in exposed children, does this represent a true difference? It should at least be considered that some part of the observed deficits is to be attributed to multiple testing rather than to a true difference.

Second, the size of the effects in these observational trials is usually quantitatively small.

Third, this association in observational studies does not prove that the observed deficits are caused by anesthetic drugs (61). Worse neurodevelopment might also be caused by (a) the conduct of anesthesia [occurrence of hypotension (53), hypoxemia, emergence delirium (54) and pain(55)], (b) the effect of surgical stress [type and extent of surgery (56, 57), number of interventions] and (c) the patient's characteristics [age at exposure, nutritional status, comorbidities, genetics (57, 58) and neurodevelopmental status before exposure to GA]. It is insufficiently clear to which extent these confounding factors can modify synaptic connections and plasticity required for neurodevelopment. If hypothetically all confounding factors could be eliminated from these observational trials, then the remaining neurodevelopmental effect could solely be attributed to GA. It is unclear whether the remaining result would still be significant.

These concerns should be taken into account when trying to make definitive conclusions from these two different groups of studies.

In 15/24 of studies, the specific anesthetics to which participants were exposed were not mentioned. We are therefore unable to compare the effects of different anesthetic agents.

In 2014, Wang et al. (25) published an excellent systematic review on the potential neurotoxic effects of GA based on 7 trials. A key finding at the time was that the risk of neurodevelopmental deficits seemed to decrease with increasing age at the time of exposure. From our review of 26 trials, we cannot provide evidence to support this finding. Most of the included trials in this review made no distinction between different ages of exposure. The ones that made a distinction present conflicting evidence, for example on the domain of language & speech (see above)(27, 47). From our data, we cannot advocate delaying procedures requiring GA in childhood, as has been proposed in the past (62).

In the same review, it was shown that an increasing number of exposures was associated with an increased hazard ratio for having a poorer neurodevelopmental outcome. Again, we cannot present evidence to support this finding from our systematic review.

Even if we postulate that the neurotoxic effects of anesthetics in childhood are small or non-existent, we have to remember this statement is based on comparison of average between groups. This does not exclude the possibility that a small minority of exposed children is particularly vulnerable and suffers from severe consequences of exposure to GA. Perhaps their number is simply too small to influence the average. In the section on attrition bias, we already mentioned a difference in baseline characteristics between patients lost to follow-up and patients who remained in follow-up. Interestingly, Glatz et al. (36) showed that exposure to GA before 4 years of age was not associated with an increased risk of having school grades below the 10th percentile. Exposure to GA before 4 years of age was however, associated with an increased risk of having no recorded school grades at age 16 (OR = 1.26). Two studies by Hansen similarly showed an increase in test score nonattainment in children exposed to GA (63, 64). The explanation for these findings remains unclear.

Limitations

A major limitation of this systematic review is that we did not perform a meta-analysis. The primary

obstacle was the heterogeneity in outcome measures used to assess "neurodevelopment". In our review, we identified 9 different domains. Even within these domains, the outcome was often measured in different ways. For example, intelligence was measured as IQ by some and as school results by others. These qualitative differences in outcome measures proved too difficult to overcome in trying to compare them mathematically in a meta-analysis.

It could also be argued that we should have been more restrictive in determining which studies we included in our systematic review. One could think of arguments to only include the randomized evidence and the evidence that used twin- or sibling-matched designs. The resulting evidence would have been less prone to bias. On the other hand, the validity of our systematic review would have dramatically declined as it would merely have reported on a few thousand subjects instead of the more than 600.000 subjects we can report on now.

Although a well-designed RCT could provide a clear estimation of the effect on GA on neurodevelopment, only one such study could be found (29). Therefore, we rely mostly on observational studies which have the advantage of dealing with large sample sizes but the disadvantage of numeral confounding factors.

Another limitation is that we are not able to account for possible publication bias. Constructing a funnel plot was outside the scope of this systematic review. Therefore, the evidence that is presented here might overestimate the real effect.

We also introduced what is sometimes called ''English language bias'' by excluding all studies published in different languages. We are however confident that important findings would have been published in English.

Many of the included observational trials examined cohorts that were in follow up for years or even decades and practice has evolved over time. For example, halothane was the most commonly used anesthetic agent (together with nitrous oxide) in some of the included studies (39, 45), but is rarely used now in the majority of centers. Extrapolation of the results of these studies to current practice might therefore be limited. Furthermore, 15/24 studies did not mention which anesthetic agents were used.

Conclusions

From this review, we can conclude that exposure to general anesthesia in childhood, especially when multiple, is associated with worse neurodevelopment in later life. We argue against a causative role of anesthesia in the observed deficits in exposed children, but we recognize there is at this time insufficient evidence to definitively prove which are the culprit factors. We can, however, provide evidence that even if anesthetics play a direct role in the observed deficits, the size effect that can be attributed to them is small. The trials that eliminated much of the methodological bias by using a sibling-matched or randomized design seem to support our view that general anesthesia in children does not directly harm neurodevelopment. Bartels et al. (31) came to the same conclusion in their landmark twin study in 2009 and stated: "Thus, there is no evidence for a causal relationship between anesthesia administration and later learning-related outcomes in this sample. Rather, there is evidence for early anesthesia being a marker of an individual's vulnerability for later learning problems, regardless of their exposure to anesthesia". Studies performed after this landmark study that are included in our review do not contradict this statement.

Ultimately, we have to decide whether we would withhold general anesthesia for surgeries and diagnostic procedures in children because of a doubtful direct effect on neurodevelopment. Indications for anesthesia must be the result of a well-considered risk-benefit balance. While many possible arguments could be raised in favor of or against anesthesia, using the argument of anesthetic neurotoxicity seems unwise, since it is merely based on circumstantial evidence. All the studies that could have pointed towards a causative role of anesthetics failed to do so.

In the future, clinicians and parents will await clear answers regarding the potential neurotoxicity of GA on children. Therefore, we will have to focus on populations at risk, dose and duration of exposure to anesthetics and potential reliable biomarkers to identify neurodevelopmental changes.

Studies that use siblings as a control or that randomize children to general or loco-regional anesthesia provide us with a unique opportunity to study the direct effects of sur-gery and anesthesia on neurodevelopment.

The best possible study would construct a cohort prospectively from birth and follow these participants through childhood. Crucially, such a study should test neurodevelopmental outcomes before exposure to GA. The evolution of neurodevelopmental outcomes in relation to exposure to GA could then be examined and compared to the evolution of the children in the non-exposed part of the cohort. We hypothesize that children who will require a GA will have worse

neurodevelopmental outcomes already before exposure to GA (for example, for genetic or socio-economic reasons). GA in itself may not influence the evolution of the neurodevelopmental outcomes in exposed children.

Finally, it should be noted that the brain development of mammals during the first week of life has now been shown to be similar to that of human fetuses of 17-22 weeks old (17-19). Therefore, the preclinical studies that were designed to study anesthetic neurotoxicity in childhood actually reflect a possible effect on the human fetal brain rather than on young children's brains. Since the neurotoxic effects of anesthetics in mammals have been proven during the first weeks of life, it would be interesting to examine the effects of $GABA_A$-agonists and NMDA-antagonists on the fetal brain. This holds true not only for anesthetics but even more so for benzodiazepines that are often administered on a regular basis rather than once during pregnancy.

Acknowledgments

The authors wish to thank the biomedical reference librarians of the KU Leuven Libraries – 2Bergen – learning Centre Désiré Collen (Leuven, Belgium) for their help in conducting the systematic literature search.

References

1. Di Maggio, C., Sun, L. S., Ing, C. and Li, G. 2012. Pediatric anesthesia and neurodevelopmental impairments: A bayesian meta-analysis. J Neurosurg Anesth. 24 (4): 376-381.
2. Wang, X., Xu, Z. and Miao, C.-H. 2014. Current Clinical Evidence on the Effect of General Anesthesia on Neurodevelopment in Children: An Updated Systematic Review with Meta-Regression. PLoS One 9 (1): e85760.
3. Jevtovic-Todorovic, V., Hartman, R. E., Izumi, Y., Benshoff, N. D., Dikranian, K., Zorumski, C. F., et al. 2003. Early exposure to common anesthetic agents causes widespread neurodegeneration in the developing rat brain and persistent learning deficits. J. Neurosci. 23 (3): 876-882.
4. Ikonomidou, C., Bosch, F., Miksa, M., Bittigau, P., Vöckler, J., Dikranian, K., et al. 1999. Blockade of NMDA Receptors and Apoptotic Neurodegeneration in the Developing Brain. Science (80-.). 283 (5398): 70-74.
5. Young, C., Jevtovic-Todorovic, V., Qin, Y. Q., Tenkova, T., Wang, H., Labruyere, J. and Olney, J. W. 2005. Potential of ketamine and midazolam, individually or in combination, to induce apoptotic neurodegeneration in the infant mouse brain. Br. J. Pharmacol. 146 (2): 189-197.
6. Karmarkar, S. W., Bottum, K. M. and Tischkau, S. A. 2010. Considerations for the use of anesthetics in neurotoxicity studies. Comp. Med. 60 (4): 256-262.
7. Morimoto, Y. 2017. Anesthesia and Neurotoxicity. Neurotox. Anesth. Dev. brain. Anesth. Neurotoxicity.: p 8-98. Tokyo. Springer.
8. Brambrink, A. M., Evers, A. S., Avidan, M. S., Farber, N. B., Smith, D. J., Zhang, X., et al. 2010. Isoflurane-induced neuroapoptosis in the neonatal rhesus macaque brain. Anesthesiology.
9. Briner, A., Nikonenko, I., De Roo, M., Dayer, A., Muller, D. and Vutskits, L. 2011. Developmental stage-dependent persistent impact of propofol anesthesia on dendritic spines in the rat medial prefrontal cortex. Anesthesiology.
10. Stratmann, G., Sall, J. W., May, L. D. V., Bell, J. S., Magnusson, K. R., Rau, V., et al. 2009. Isoflurane differentially affects neurogenesis and long-term neurocognitive function in 60-day-old and 7-day-old rats. Anesthesiology.
11. N., L., C., H., J.R., L., J., S. and V., J.-T. 2011. Isoflurane impairs immature astroglia development in vitro: The role of actin cytoskeleton. J. Neuropathol. Exp. Neurol.
12. Sanchez, V., Feinstein, S. D., Lunardi, N., Joksovic, P. M., Boscolo, A., Todorovic, S. M. and Jevtovic-Todorovic, V. 2011. General anesthesia causes long-term impairment of mitochondrial morphogenesis and synaptic transmission in developing rat brain. Anesthesiology 115 (5): 992-1002.
13. Head, B. P., Patel, H. H., Niesman, I. R., Drummond, J. C., Roth, D. M. and Patel, P. M. 2009. Iof p75 neurotrophin receptor attenuates isoflurane-mediated neuronal apoptosis in the neonatal central nervous system. Anesthesiology.
14. Hansen, T. G. 2015. Anesthesia-related neurotoxicity and the developing animal brain is not a significant problem in children. Paediatr. Anaesth. 25 (1): 65-72.
15. Wu, B., Yu, Z., You, S., Zheng, Y., Liu, J., Gao, Y., et al. 2014. Physiological disturbance may contribute to neurodegeneration induced by isoflurane or sevoflurane in 14 day old rats. PLoS One 9 (1): 1-7.
16. Liu, J. R., Liu, Q., Li, J., Baek, C., Han, X. H., Athiraman, U. and Soriano, S. G. 2012. Noxious stimulation attenuates ketamine-induced neuroapoptosis in the developing rat brain. Anesthesiology 117 (1): 64-71.
17. Loepke, A. W., McGowan, F. X. and Soriano, S. G. 2008. CON: The toxic effects of anesthetics in the developing brain: The clinical perspective. Anesth. Analg. 106 (6): 1664-1669.
18. Clancy, B., Darlington, R. B. and Finlay, B. L. 2001. Translating developmental time across mammalian species. Neuroscience 105 (1): 7-17.
19. Clancy, B., Finlay, B. L., Darlington, R. B. and Anand, K. J. S. 2007. Extrapolating brain development from experimental species to humans. Neurotoxicology 28: 931-937.
20. Davidson, A. J., Disma, N., de Graaff, J. C., Withington, D. E., Dorris, L., Bell, G., et al. 2016. Neurodevelopmental outcome at 2 years of age after general anaesthesia and awake-regional anaesthesia in infancy (GAS): an international multicentre, randomised controlled trial. Lancet 387 (10015): 239-250.
21. Mccann, M. E., Graaff, J. C. De, Dorris, L., Disma, N., Withington, D., Bell, G., et al. 2019. Neurodevelopmental outcome at 5 years of age after general anaesthesia or awake-regional anaesthesia in infancy (GAS): an international , multicentre , randomised , controlled equivalence trial. 393.
22. FDA. 2016. FDA Drug Safety Communication: FDA review results in new warnings about using general anesthetics and sedation drugs in young children and pregnant women. Available at: https://www.fda.gov/Drugs/DrugSafety/ucm532356.htm (accessed February 18, 2018).
23. FDA. 2017. FDA Drug Safety Communication: FDA approves label changes for use of general anesthetic and sedation drugs in young children. Available at: https://www.fda.gov/Drugs/DrugSafety/ucm554634.htm (accessed February 18, 2018).
24. Shi, Y., Hu, D., Rodgers, E. L., Katusic, S. K., Gleich, S. J. and Andrew, C. 2018. Epidemiology of general anesthesia

prior to age 3 in a population-based birth cohort. Paediatr Anaesth. 28 (6): 513-519.
25. Wang, X., Xu, Z. and Miao, C.-H. 2014. Current Clinical Evidence on the Effect of General Anesthesia on Neurodevelopment in Children: An Updated Systematic Review with Meta-Regression. PLoS One 9 (1): e85760.
26. Higgins, J. P. and Green, S. 2011. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. Available at: http://handbook.cochrane.org/.
27. O'Leary, J. D., Janus, M., Duku, E., Wijeysundera, D. N., To, T., Li, P., et al. 2016. A Population-based Study Evaluating the Association between Surgery in Early Life and Child Development at Primary School Entry. Anesthesiology 125 (2): 272-9.
28. Warner, D. O., Zaccariello, M. J., Katusic, S. K., Schroeder, D. R., Hanson, A. C., Schulte, P. J., et al. 2018. Neuropsychological and Behavioral Outcomes after Exposure of Young Children to Procedures Requiring General Anesthesia: The Mayo Anesthesia Safety in Kids (MASK) Study. Anesthesiology 129 (1): 89-105.
29. Mccann, M. E., Graaff, J. C. De, Dorris, L., Disma, N., Withington, D., Bell, G., et al. 2019. Neurodevelopmental outcome at 5 years of age after general anaesthesia or awake-regional anaesthesia in infancy (GAS): an international , multicentre , randomised , controlled equivalence trial. Lancet 393 (10172): 664-677.
30. DiMaggio, C., Sun, L. S. and Li, G. 2011. Early Childhood Exposure to Anesthesia and Risk of Developmental and Behavioral Disorders in a Sibling Birth Cohort. Anesth. Analg. 113 (5): 1143-1151.
31. Bartels, M., Althoff, R. R. and Boomsma, D. I. 2009. Anesthesia and Cognitive Performance in Children: No Evidence for a Causal Relationship. Twin Res. Hum. Genet. 12 (03): 246-253.
32. Sun, L. S., Li, G., Miller, T. L. K., Salorio, C., Byrne, M. W., Bellinger, D. C., et al. 2016. Association Between a Single General Anesthesia Exposure Before Age 36 Months and Neurocognitive Outcomes in Later Childhood. JAMA 315 (21): 2312-2320.
33. Khochfe, A. R., Rajab, M., Ziade, F., Naja, Z. Z., Naja, A. S. and Naja, Z. M. 2019. The effect of regional anaesthesia versus general anaesthesia on behavioural functions in children. Anaesth. Crit. Care Pain Med. 38 (4): 357-361.
34. Ing, C., DiMaggio, C., Whitehouse, A., Hegarty, M. K., Brady, J., von Ungern-Sternberg, B. S., et al. 2012. Long-term Differences in Language and Cognitive Function After Childhood Exposure to Anesthesia. Pediatrics 130 (3): e476-e485.
35. Backeljauw, B., Holland, S. K., Altaye, M. and Loepke, A. W. 2015. Cognition and Brain Structure Following Early Childhood Surgery With Anesthesia. Pediatrics 136 (1): e1-e12.
36. Glatz, P., Sandin, R. H., Pedersen, N. L., Bonamy, A.-K., Eriksson, L. I. and Granath, F. 2017. Association of Anesthesia and Surgery During Childhood With Long-term Academic Performance. JAMA Pediatr. 171 (1): e163470.
37. De Heer, I. J., Tiemeier, H., Hoeks, S. E. and Weber, F. 2017. Intelligence quotient scores at the age of 6 years in children anaesthetised before the age of 5 years. Anaesthesia 72 (1): 57-62.
38. Hu, D., Flick, R. P., Zaccariello, M. J., Colligan, R. C., Katusic, S. K., Schroeder, D. R., et al. 2017. Association between exposure of young children to procedures requiring general anesthesia and learning and behavioral outcomes in a population-based birth cohort. Anesthesiology 127 (2): 227-240.
39. Flick, R. P., Katusic, S. K., Colligan, R. C., Wilder, R. T., Voigt, R. G., Olson, M. D., et al. 2011. Cognitive and Behavioral Outcomes After Early Exposure to Anesthesia and Surgery. Pediatrics 128 (5): e1053-e1061.
40. Bong, C. L., Allen, J. C. and Kim, J. T. S. 2013. The effects of exposure to general anesthesia in infancy on academic performance at age 12. Anesth. Analg. 117 (6): 1419-1428.
41. Stratmann, G., Lee, J., Sall, J. W., Lee, B. H., Alvi, R. S., Shih, J., et al. 2014. Effect of general anesthesia in infancy on long-term recognition memory in humans and rats. Neuropsychopharmacology 39 (10): 2275-2287.
42. Ing, C. H., DiMaggio, C. J., Whitehouse, A. J. O., Hegarty, M. K., Sun, M., von Ungern-Sternberg, B. S., et al. 2014. Neurodevelopmental Outcomes After Initial Childhood Anesthetic Exposure Between Ages 3 and 10 Years. J. Neurosurg. Anesthesiol. 26 (4): 377-86.
43. Doberschuetz, N., Dewitz, R., Rolle, U., Schlösser, R. and Allendorf, A. 2016. Follow-up of children with gastrointestinal malformations and postnatal surgery and anesthesia: Evaluation at two years of age. Neonatology 110 (1): 8-13.
44. Poor Zamany Nejat Kermany, M., Roodneshin, F., Ahmadi Dizgah, N., Gerami, E. and Riahi, E. 2016. Early childhood exposure to short periods of sevoflurane is not associated with later, lasting cognitive deficits. Paediatr. Anaesth. 26 (10): 1018-1025.
45. Wilder, R. T., Flick, R. P., Sprung, J. and Katusic, S. K. 2009. Early Exposure to Anesthesia and Learning Disabilities in a Population-based Birth Cohort. Anesthesiology 110 (4): 796-804.
46. Nestor, K. A., Zeidan, M., Boncore, E., Richardson, A., Alex, G., Weiss, M. and Islam, S. 2017. Neurodevelopmental outcomes in infants undergoing general anesthesia. J. Pediatr. Surg. 52 (6): 895-900.
47. Graham, M. R., Brownell, M., Chateau, D. G., Dragan, R. D., Burchill, C. and Fransoo, R. R. 2016. Neurodevelopmental Assessment in Kindergarten in Children Exposed to General Anesthesia before the Age of 4 Years. Anesthesiology 125 (4): 667-677.
48. Chemaly, M., El-Rajab, M. A., Ziade, F. M. and Naja, Z. M. 2014. Effect of one anesthetic exposure on long-term behavioral changes in children. J. Clin. Anesth. 26 (7): 551-556.
49. Bakri, M. H., Ismail, E. A., Ali, M. S., Elsedfy, G. O., Sayed, T. A. and Ibrahim, A. 2015. Behavioral and emotional effects of repeated general anesthesia in young children. Saudi J. Anaesth. 9 (2): 161-166.
50. Ing, C., Sun, M., Olfson, M., DiMaggio, C. J., Sun, L. S., Wall, M. M. and Li, G. 2017. Age at Exposure to Surgery and Anesthesia in Children and Association With Mental Disorder Diagnosis. Anesth. Analg. 125 (6): 1988-1998.
51. Ko, W. R., Liaw, Y. P., Huang, J. Y., Zhao, D. H., Chang, H. C., Ko, P. C., et al. 2014. Exposure to general anesthesia in early life and the risk of attention deficit/hyperactivity disorder development: A nationwide, retrospective matched-cohort study. Paediatr. Anaesth. 24 (7): 741-748.
52. Ko, W. R., Huang, J. Y., Chiang, Y. C., Nfor, O. N., Ko, P. C., Jan, S. R., et al. 2015. Risk of autistic disorder after exposure to general anaesthesia and surgery. Eur. J. Anaesthesiol. 32 (5): 303-310.
53. Weber, F., Honing, G. H. M. and Scoones, G. P. Arterial blood pressure in anesthetized neonates and infants : a retrospective analysis of 1091 cases. 26 (2016): 815-822.
54. Mason, K. P. 2017. Paediatric emergence delirium : a comprehensive review and interpretation of the literature. Br. J. Anaesth. 118 (3): 335-343.
55. Grunau, R. E., Holsti, L. and Peters, J. W. B. 2006. Long-term consequences of pain in human neonates. 268-275.
56. Simonsz, H. J. and Kolling, G. H. 2008. Motor development and surgery for infantile esotropia. 115-116.
57. Stolwijk, L. J., Lemmers, M. A. and Harmsen, M. 2019. Neurodevelopmental Outcomes After Neonatal Surgery for Major Noncardiac Anomalies. Pediatrics 137 (2): e20151728.

58. Laing, S., Walker, K., Ungerer, J., Badawi, N. and Spence, K. 2011. Early development of children with major birth defects requiring. J. Paediatr. Child Heal. 47 (3): 140-147.
59. Hansen, T. G., Clausen, N. G. and K, S. 2018. Systematic review of the neurocognitive outcomes used in studies of paediatric anaesthesia neurotoxicity. Br. J. Anaesth. 120 (6): 1255-1273.
60. Davidson, A. and Sun, L. 2018. Clinical Evidence for Any Effect of Anesthesia on the Developing Brain. Anesthesiology 126 (4): 840-753.
61. Mann, C. 2003. Observational research methods . Research design II : Emerg. Med. J. 20 (1): 54-60.
62. Rappaport, B. A., Suresh, S., Hertz, S., Evers, A. S. and Orser, B. A. 2015. Anesthetic Neurotoxicity — Clinical Implications of Animal Models. N. Engl. J. Med. 372 (9): 796-797.
63. Hansen, T. G., Pedersen, J. K., Henneberg, S. W., Morton, N. S. and Christensen, K. 2013. Educational outcome in adolescence following pyloric stenosis repair before 3 months of age: A nationwide cohort study. Paediatr. Anaesth. 23 (10): 883-890.
64. Hansen, T. G., Jacob K Pedersen, J. K., Henneberg, S. W., Pedersen, D. A., Murray, J. C., Morton, N. S. and Christensen, K. 2011. Academic Performance in Adolescence after Inguinal Hernia Repair in Infancy: A Nationwide Cohort Study. Anesthesiology 114 (5): 1076-1085.