openheart

Neurocognition after paediatric heart surgery: a systematic review and meta-analysis

Caroline Sterken,[1] Jurgen Lemiere,[2,3] Ilse Vanhorebeek,[1] Greet Van den Berghe,[1] Dieter Mesotten[1]

To cite: Sterken C, Lemiere J, Vanhorebeek I, *et al*. Neurocognition after paediatric heart surgery: a systematic review and meta-analysis. *Open Heart* 2015;**2**:e000255. doi:10.1136/openhrt-2015-000255

► Additional material is available. To view please visit the journal (http://dx.doi:10.1136/openhrt-2015-000255).

Received 18 February 2015
Revised 18 September 2015
Accepted 29 September 2015

[1]Division of Cellular and Molecular Medicine, Clinical Department and Laboratory of Intensive Care Medicine, KU Leuven, Leuven, Belgium
[2]Child and Adolescent Psychiatry, University Hospitals Leuven, Leuven, Belgium
[3]Paediatric Haemato-Oncology, University Hospitals Leuven, Leuven, Belgium

Correspondence to
Dr Dieter Mesotten; dieter.mesotten@med.kuleuven.be

ABSTRACT

Children with congenital heart disease (CHD) often experience difficulties in academic and daily functioning, which have been associated with intelligence and neurocognitive skills, including executive functions (EFs), attention and memory. We report the neurocognitive data of children with CHD who were included in the Leuven glucose control trial (LGC trial). Through a systematic review and meta-analysis, we aimed to find which neurocognitive functions are most consistently and prominently affected. 365 children with CHD and 216 healthy control children underwent extensive neurocognitive testing in the LGC trial. A comprehensive search of electronic databases PubMed, EMBASE and Cochrane was conducted for studies measuring intelligence, EFs, attention and memory in children who underwent heart surgery for CHD. Standardised mean differences (SMDs) between the CHD group and a healthy control group were calculated for these neurocognitive functions. LGC trial data were included in the meta-analysis. Twelve studies with a healthy control group were included in the meta-analysis, involving 647 patients with CHD and 633 controls. The CHD group (median age 7.35 years at testing) had worse scores than healthy control children, for all investigated neurocognitive functions. A medium SMD was found for intelligence (SMD=−0.53 (95% CI −0.68 to −0.38), p<0.00001). Alertness, an attentional function, was also consistently poorer in the CHD group. Memory was less affected, while EF had a medium SMD with large heterogeneity. Children with CHD risk displayed lower performance on intelligence and alertness assessment, which may contribute to difficulties in daily life and school. Heterogeneity in neurocognitive assessment and small sizes in most studies limit the interpretation.

Trial registration number: clinicaltrials.gov Identifier NCT00214916.

INTRODUCTION

As the survival rate of children with congenital heart disease (CHD) continues to improve thanks to medical advances,[1] public interest and research have been focusing more on how children with CHD survive.[2] Children with CHD may experience difficulties in daily life[3] and academic functioning,[4] which may persist into adulthood.[5] In this systematic review and meta-analysis, we aim to examine intelligence and specific neurocognitive skills, especially executive functions (EFs), as important predictors for academic and daily functioning in children with CHD after heart surgery.

Intelligence scores (IQs) are good predictors of academic performance.[6] However, children with CHD often have intelligence scores within average range compared with population norms.[7] Research in other paediatric conditions, such as traumatic brain injury, has demonstrated that intelligence scores are rough measures and fall short in detecting more specific neurocognitive skills, such as EFs.[8] EFs cover a variety of cognitive functions, such as planning, organisation, flexibility, cognitive control and working memory. These are essential in many domains of daily life[9] and contribute to academic performance.[10 11] Other neurocognitive skills, such as memory and attention, which are inter-related with EFs, also contribute to academic performance.[9]

There is growing evidence that brain development in children with CHD can differ from normal brain development. This misdevelopment may even start prenatally due to impaired cerebral blood flow.[12] Postnatally, infants and older children with CHD may have preoperative[13] and postoperative white matter abnormalities,[14] which may relate to worse neurocognitive outcome in children with CHD.[15] Surgery seems to impact neurocognitive outcome in CHD as well.[15] Considering the vulnerability of their brain, neurocognitive functions may be worse in children with CHD than those in healthy control children without CHD. Some relevant reviews have discussed the importance of these neurocognitive functions.[2 16 17]

However, to our knowledge, the outcome data of specific neurocognitive functions have not been analysed systematically in children with CHD. One systematic review[18] and one meta-analysis[19] examined intellectual outcome but not more specific neurocognitive skills in children with CHD after heart surgery.

A large randomised controlled trial (n=700), in which neurocognitive development of children was assessed 4 years after critical illness and treatment with tight glucose control, has recently been completed.[20] The results demonstrated that tight glucose control in critically ill children improved motor coordination and cognitive flexibility in comparison with children in whom blood glucose levels up to 215 mg/dL were tolerated.[20] Seventy-five per cent of the study population in this Leuven glucose control trial (LGC trial) underwent heart surgery for congenital heart defects. Neurocognitive data of the heart surgery subgroup have not yet been analysed.

Thus, the first aim of this paper was to report the neurocognitive data of this large cohort of children with heart surgery for CHD, included in the LGC trial, and healthy controls. The second aim was to carry out a systematic review and meta-analysis for intelligence, EFs, attention and memory in children with CHD after heart surgery. We hypothesised that specific neurocognitive skills such as EFs are more impaired than is intelligence, in children with CHD.

METHODS

Analysis of the LGC trial data

Data of all children who underwent neurocognitive testing and for whom a full-scale IQ was available, were analysed (children with CHD n=361, healthy controls n=215). For four children of the CHD group and one child of the control group, a full-scale IQ could not be calculated. Baseline neurocognitive data, at the time patients were included into the LGC trial, were not available. Demographic, clinical and neurocognitive data are reported as numbers and percentages, or as median and IQR. Because of imbalance between the CHD and control group for gender, presence of a syndrome, socioeconomic status and age at follow-up, propensity score matching was performed using IBM SPSS Statistics V.22.0.0.1 and R statistical software V.2.15.3. For more details on propensity score matching, we refer to eMethods 1. Demographic, clinical and neurocognitive data of the tested and matched population were further analysed using χ^2 test for dichotomous variables and unpaired non-parametric Wilcoxon rank-sum tests for continuous variables (JMP V.11.2.0 (SAS Institute)). The details of the study protocol and neurocognitive test battery have previously been reported.[20]

Systematic review and meta-analysis

Data sources and searches

A comprehensive search of electronic databases PubMed, EMBASE and Cochrane was conducted for studies published between the beginning of each database and December 2014. Each search strategy consisted of four major parts: cognition, heart, child and tests. Both index language terms (MeSH, Emtree) and keywords were used in every part of the search strategy (see eMethods 2). We also manually screened reference lists of studies identified through database search.

Study selection

Two selection criteria were premised for title and abstract screening. First, the study population consisted of infants, children, adolescents and/or young adults (<24 years old) with CHD. Second, they needed to have, at least, intelligence testing with an indication of a specific measure of EF, memory or attention, or a broader neurocognitive assessment. For the full-text screening, we introduced two more selection criteria. First, the participants had neurocognitive testing after heart surgery or an interventional cardiac procedure. Studies that tested participants in the first six postoperative months were excluded to avoid interference with the acute medical phase. Second, at least one test that measured EF, memory and/or attention was needed.

Data extraction

First, we gathered data on the sample size, age at surgery and age at testing of the tested groups. We also investigated whether the same study population was tested in already included studies of the same research group. Furthermore, for comparison of the CHD group versus a healthy control group, we collected data necessary to quantify differences between the CHD and healthy control group. The collected data are summarised in box 1. In case of missing data for the quantitative analysis, authors were contacted.

Data analysis

As appropriate, mean (SD) or median (IQR) sample size, age at surgery and age at testing were calculated for all studies. Risk of bias of all included studies[20–31] was assessed by two raters (CS and JL) independently by means of a modified version of the Downs and Black

Box 1 Data collected for quantitative analysis

Congenital heart disease (CHD) group: type of CHD, sample size
Control group: type of control group, sample size
Age surgery, age testing
Overlap study population with another included study of the same research group: yes/no and if yes, decision (inclusion/exclusion)
Intelligence: name of test(s), intelligence scores (mean, SD) of CHD and healthy control group
Memory: name of test(s), memory scores (mean, SD) of CHD and healthy control group
Attention: name of test(s), attention scores (mean, SD) of CHD and healthy control group
Executive functions: name of test(s), executive function scores (mean, SD) of CHD and healthy control group

checklist.[32] Six items that focus on RCTs (items 4, 8, 14, 19, 23, 24) were omitted, leading to a maximum total score of 22. In case of different scores by the two raters, a consensus was reached through discussion. Inter-rater reliability was good (Spearman's r 0.722; p=0.008). Depending on the number of included studies, mean (SD) or median (IQR) total risk of bias and subscores were calculated for all studies.

Analyses were performed on neurocognitive data of studies that used a healthy control group. To combine these continuous data measured by different instruments, effect sizes, that is, standardised mean differences (SMDs), were calculated in Review Manager V.5.2 by means of Inverse Variance, random effects analysis (http://ims.cochrane.org/revman). SMDs were calculated for the following neurocognitive functions: intelligence, alertness (attentional function), memory and inhibition (EF). Neurocognitive data of the LGC trial were included in the SMD calculation. A two-sided p value <0.05 was considered statistically significant. SMDs were classified according to Cohen's guidelines: d=0.2 defined as small, d=0.5 as medium and d=0.8 as large.[33] For clinical interpretation, the overall SMD for intelligence scores was multiplied by the typical SD of the normal IQ distribution (mean 100±SD 15). The I^2 statistic was used to evaluate heterogeneity. I^2 ≥50% was defined as substantial heterogeneity. Funnel plots were also made in Review Manager when at least 10 studies were available, to explore small-study effects.

For more details on study selection and data analysis, we refer to eMethods 1.

RESULTS

Analysis of the LGC trial data

Demographic and clinical data of tested and matched postheart surgery population and healthy controls are presented in table 1. Neurocognitive data are presented in table 2. Children of the tested postheart surgery population have worse scores for intelligence, visual-motor integration and all measures of alertness, motor coordination, inhibition, flexibility (except for ΔN of errors), memory and behaviour, compared with healthy controls. After propensity score matching, the CHD group has worse scores for intelligence, visual-motor integration, motor coordination (alternating taps), inhibition (ΔN of errors), memory (verbal working memory and immediate memory) and behaviour (internalising and total problems).

Systematic review and meta-analysis

Flow diagram

Figure 1 shows the article screening phases and the reasons for exclusion, according to the PRISMA guidelines (Preferred Reporting Items for Systematic reviews and Meta-Analysis).[34]

An overview of the 12 included studies is presented in etable 1.

Sample size, and age at surgery and at testing

The median sample size of the CHD groups in all 12 studies included in the meta-analysis was 31 (IQR 18–43) The median sample size of the healthy control groups was 33 (IQR 20–42). The median age at surgery was 1.25 years (IQR 0.0.7–3.85) and the median age at testing was 7.35 years (IQR 5.70–8.37).

Quality assessment

The median (IQR) risk of bias of all 12 included studies was 16 (IQR 15–17). Median score for reporting was 8 (IQR 7–8) with a maximum score of 9. Median score for external validity was 2 (IQR 1–2) with a maximum score of 3 and median score for internal validity was 7 (IQR 5–7) with a maximum score of 9. Only 1 out of 12 studies mentioned that the study had sufficient power to detect a clinically relevant difference.

Meta-analyses: CHD group versus healthy controls

Total SMDs for all neurocognitive domains indicate lower scores for the CHD group. A medium SMD with low heterogeneity was found for intelligence (SMD=−0.53 (95% CI −0.68 to −0.38), p<0.00001, I^2=32%; figure 2), indicating a drop of 0.53 times the SD of the normal IQ distribution: −8 IQ points. A medium SMD was also found for alertness non-reaction time (SMD=−0.47 (95% CI −0.67 to −0.27), p<0.00001, I^2=19%; figure 3) and a smaller SMD for alertness reaction time (SMD=0.25 (95% CI 0.08 to 0.42), p=0.004, I^2=0%; efigure 1). Verbal memory showed a smaller SMD (SMD=−0.35 (95% CI −0.54 to −0.15), p=0.0004, I^2=0%; figure 4), while non-verbal memory did not differ between children with CHD and healthy controls (efigure 2). EF reaction time, examining the inhibition function, had the largest SMD, but a high level of heterogeneity (SMD=0.57 (95% CI 0.10 to 1.04), p=0.02, I^2=80%; figure 5). EF non-reaction time for inhibition also showed a medium SMD but with low heterogeneity (SMD=−0.51 (95% CI −0.74 to −0.29), p<0.00001, I^2=0%; efigure 3). The funnel plot for intelligence was slightly asymmetrical (efigure 4). Owing to the limited availability of studies on the other neurocognitive functions, funnel plots could not be built for alertness, memory and inhibition.

DISCUSSION

This meta-analysis and the results of the LGC trial provide sound evidence that not only intelligence but also more specific neurocognitive functions are impaired in children with CHD who underwent heart surgery, compared with healthy controls. Contrary to our hypothesis, intelligence, EFs and alertness seem to be equally affected.

Intelligence

The SMD of −0.53 (95% CI −0.68 to −0.38) in the meta-analysis and the effect size of 8 IQ points in both

Table 1 Demographic and clinical data of tested and propensity score-matched postheart surgery population and healthy controls in the LGC trial

	Tested population			Propensity score-matched population		
	Tested—post heart surgery population (N=361)	Healthy controls (N=215)	p Value	Tested—post heart surgery population (N=167)	Healthy controls (N=167)	p Value
Caucasian race*	343 (95.01)	211 (98.14)	0.058	162 (97.01)	164 (98.20)	0.474
Exclusive European*	330 (91.41)	201 (93.49)	0.369	160 (95.81)	156 (93.41)	0.332
Exclusive Dutch language*	296 (81.99)	186 (86.51)	0.155	144 (86.23)	142 (85.03)	0.755
Male sex*	205 (56.79)	93 (43.26)	0.001	81 (48.50)	81 (48.50)	1.000
Age at randomisation, years†	0.76 (0.22–4.1)	NA	NA	2.11 (0.34–4.77)	NA	NA
Type of congenital heart disease*						
Obstructive pathology	79 (21.88)	NA	NA	37 (22.15)	NA	NA
Left-right shunt	121 (33.51)	NA	NA	60 (35.92)	NA	NA
Cyanotic and not univentricular	122 (33.79)	NA	NA	59 (35.32)	NA	NA
Cyanotic and univentricular	70 (19.39)	NA	NA	27 (16.16)	NA	NA
Other	11 (3.04)	NA	NA	5 (2.99)	NA	NA
RACHS classification*						
1 or 2	168 (47.72)	NA	NA	79 (48.17)	NA	NA
3 or 4	175 (49.71)	NA	NA	84 (51.21)	NA	NA
5 or 6	9 (2.55)	NA	NA	1 (0.60)	NA	NA
Syndrome, at randomisation*	69 (19.11)	16 (7.44)	<0.001	20 (11.98)	16 (9.58)	0.480
PELOD first 24 h in ICU†	11 (2–12)	NA	NA	11 (2–12)	NA	NA
Socioeconomic status score†	35 (24–48.5)	42.5 (29–54)	<0.001	39.5 (29–50)	39.5 (29–52.5)	0.455
At follow-up†						
Height, cm	107 (103–126)	122 (108–151)	<0.001	116 (104–131)	117 (107–139)	0.069
Weight, kg	18 (15–24)	22 (18–40)	<0.001	20 (16–29)	21 (18–34)	0.072
Head circumference, cm	50.7 (49.2–52.5)	52 (50.8–54)	<0.001	51 (49.5–53)	51.8 (50.5–53.4)	0.001
Age, year	4.67 (4.14–7.93)	6.75 (4.68–11.56)	<0.001	6.02 (4.21–8.78)	5.91 (4.58–9.07)	0.359

*Numbers and percentages.
†Median (IQR).
ICU, intensive care unit; LGC, Leuven glucose control; NA, not applicable; PELOD, paediatric logistic organ dysfunction;[43] RACHS, risk adjustment for congenital heart surgery.[42]

Table 2 Results of neurocognitive test battery in tested and propensity score-matched postheart surgery population and healthy controls in the LGC trial

	Tested population			Propensity score-matched population		
	Tested—post heart surgery population (N=361)	Healthy controls (N=215)	p Value	Tested—post heart surgery population (N=167)	Healthy controls (N=167)	p Value
Clinical neurological evaluation score (range 0–8)*	1 (0–2)	0 (0–1)	**<0.001**	1 (0–2)	0 (0–1)	**<0.001**
Intelligence (range of possible scores, 45–155)†						
Full-scale IQ	90 (75–100)	103 (91–111)	**<0.001**	92 (78–103)	101 (90–110)	**<0.001**
Verbal IQ	90.5 (76.2–101)	102 (92–114)	**<0.001**	94 (79.5–104)	101 (91–112)	**<0.001**
Performance IQ	89 (77–101)	103 (92–112)	**<0.001**	91 (77–102)	101 (89–112)	**<0.001**
Visual-motor integration (range 0.9–20)†	9 (7–10)	10 (8–12)	**<0.001**	9 (7–10)	10 (9–12)	**<0.001**
Attention, motor coordination and executive functions						
Alertness*						
Reaction time dominant hand, ms	691 (447–982)	481 (320–700)	**<0.001**	566 (374–829)	544 (365–749)	0.345
Within-patient SD of repeated tests	404 (143–642)	165 (83–383)	**<0.001**	228 (115–535)	190 (98–433)	0.075
Reaction time non-dominant hand, ms	697 (436–968)	499 (326–721)	**<0.001**	543 (374–791)	542 (375–744)	0.677
Within-patient SD of repeated tests	336 (158–611)	192 (87–379)	**<0.001**	221 (110–482)	216 (105–412)	0.410
Motor coordination (Number of taps in 10 s)†						
Number of unimanual taps dominant hand	28 (22–38)	35 (25–46)	**<0.001**	31 (22–42)	32 (24–43)	0.346
Number of unimanual taps non-dominant hand	23 (17–33)	29 (21–43)	**<0.001**	25 (18–35)	27 (20–38)	0.117
Number of valid alternating taps	8 (2–18)	13 (5–30)	**<0.001**	8 (2–20)	11 (5–26)	**0.039**
Number of valid synchronous taps	16 (8–24)	21 (12–31)	**<0.001**	18 (11–27)	19 (11–27)	0.768
Inhibition and flexibility*						
ΔReaction time (inhibition), ms	313 (120–536)	200 (79–485)	**0.008**	261 (98.7–438)	258 (94–500)	0.876
ΔNumber of errors (inhibition)	2 (0–3)	1 (0–2)	**0.002**	1.5 (0–3)	1 (0–2)	**0.038**
ΔReaction time (flexibility), ms	637 (367–878)	550 (283–798)	**0.043**	603 (330–848)	623 (345–869)	0.726
ΔNumber of errors (flexibility)	2 (0–4)	1 (0–3)	0.243	2 (0–4)	1 (0–3)	0.104
Memory†						
Verbal-auditory						
Numbers (range 1–19): memory span (forward)	8 (5–9)	9 (7–11)	**<0.001**	8 (6–11)	9 (7–11)	0.149
Numbers (range 1–19): working memory (backward)	9 (6–11)	10 (8–13)	**<0.001**	9.5 (6.2–12)	10.5 (9–13)	**0.003**
Word pairs (proportion of correct responses): learning	0.45 (0.33–0.53)	0.5 (0.38–0.66)	**0.001**	0.46 (0.36–0.57)	0.46 (0.35–0.6)	0.627
Word pairs (proportion of correct responses): immediate memory	0.4 (0.2–0.5)	0.5 (0.35–0.64)	**<0.001**	0.4 (0.23–0.5)	0.5 (0.3–0.6)	**0.012**
Word pairs (proportion of correct responses): delayed memory	0.3 (0.2–0.4)	0.4 (0.3–0.5)	**<0.001**	0.35 (0.21–0.42)	0.4 (0.28–0.5)	0.068
Word pairs (proportion of correct responses): recognition	0.96 (0.9–1)	1 (0.95–1)	**0.011**	0.97 (0.9–1)	0.97 (0.93–1)	0.404
Non-verbal, visual-spatial (proportion of correct responses)						
Pictures: memory span	0.83 (0.71–0.89)	0.89 (0.8–0.93)	**<0.001**	0.84 (0.73–0.9)	0.86 (0.76–0.93)	0.114
Dots: learning	0.83 (0.66–0.88)	0.88 (0.83–0.94)	**<0.001**	0.87 (0.70–0.94)	0.88 (0.80–0.94)	0.061
Dots: immediate memory	0.83 (0.62–1)	1 (0.75–1)	**<0.001**	0.87 (0.66–1)	1 (0.68–1)	0.224
Dots: delayed memory	0.83 (0.5–1)	1 (0.75–1)	**<0.001**	0.83 (0.66–1)	0.87 (0.68–1)	0.268
Learning index (range 50–150)	93 (84–103)	101 (90–109)	**<0.001**	96 (87–103)	99 (87–109)	0.102

Continued

Table 2 Continued

	Tested population			Propensity score-matched population		
	Tested—post heart surgery population (N=361)	Healthy controls (N=215)	p Value	Tested—post heart surgery population (N=167)	Healthy controls (N=167)	p Value
Behaviour (by proxy), T-score 1						
CBCL-internalising problems (range 29–100)	52 (45–61)	48 (41–57)	**<0.001**	51 (43–59)	49 (41–58)	**0.046**
CBCL-externalising problems (range 28–100)	50 (43–56)	46 (40–55)	**0.006**	49 (41–55)	47 (40–56)	0.221
CBCL-total problems (range 24–100)	52 (45–59)	47 (40–55)	**<0.001**	51 (43–56)	48 (40–56)	**0.049**

Median (IQR) are reported.
*Higher scores reflect worse performance.
†Higher scores reflect better performance.
CBCL, Child Behavior Checklist.

the meta-analysis and LGC trial are of similar magnitude to those in prematurely born children,[35] and are in line with the meta-analysis of Karsdorp *et al*.[19] Despite the worse intelligence scores for the CHD group compared with healthy controls, the mean IQ scores of the children with CHD were still within the 'normal' average range, when compared with population norms. Nevertheless, these IQ differences can have consequences in a learning environment.[36] Population norms might insufficiently capture the dynamic nature of development. A well-matched healthy control group of typically developing children without CHD may offer a more representative reflection of normal variation between and within typically developing children[37 38] who grow up in the same time period as children with CHD.

Alertness, inhibition and memory

Similarly to intelligence, alertness was also consistently impaired across the available studies in the meta-analysis. The alerting system is one of the attentional networks in addition to the orienting and executive attention network.[39] It is responsible for achieving and maintaining a state of high sensitivity to incoming information. An efficient alerting system is pivotal for other more complex cognitive functions, such as EF. Additionally, detecting problems of alertness has an important clinical consequence because improving alertness is considered essential for cognitive rehabilitation.

Recently, more consideration has been given to impairments in EF in children with CHD.[2 16 17] The LGC trial exactly showed that, in critically ill children, the more complex EF, flexibility, was improved by tight glucose control, almost to the levels of healthy controls[20] However, from this meta-analysis, we have to conclude that studies examining these specific neurocognitive skills are scarce. The few studies included in this meta-analysis involved several specific subgroups of CHD with only a limited number of patients in each subgroup. This resulted in low statistical power and precision. Also, the use of different tests for EF, adding to the ascertained heterogeneity, may have had an impact on low precision.

Nevertheless, the SMD (0.57) for EF reaction time was in line with the deficit in intelligence (0.53). It indicates that children with CHD react slower when performing specific EF tasks measuring inhibition. Children with CHD from the LGC trial showed the smallest SMD in EF reaction time.[20] This may be attributed to the larger size of the study and to the fact that the EF reaction time was corrected for baseline speed and the propensity score matching. The CHD group of the LGC trial also made more errors during an inhibition test. This meta-analysis confirmed the findings from previous reviews,[2 16 17] contending that the risk of memory deficits in children with CHD may be lower. However, results from the LGC trial showed worse performance for immediate memory and working memory of verbal information in the CHD group. The use of pooled memory scores in the

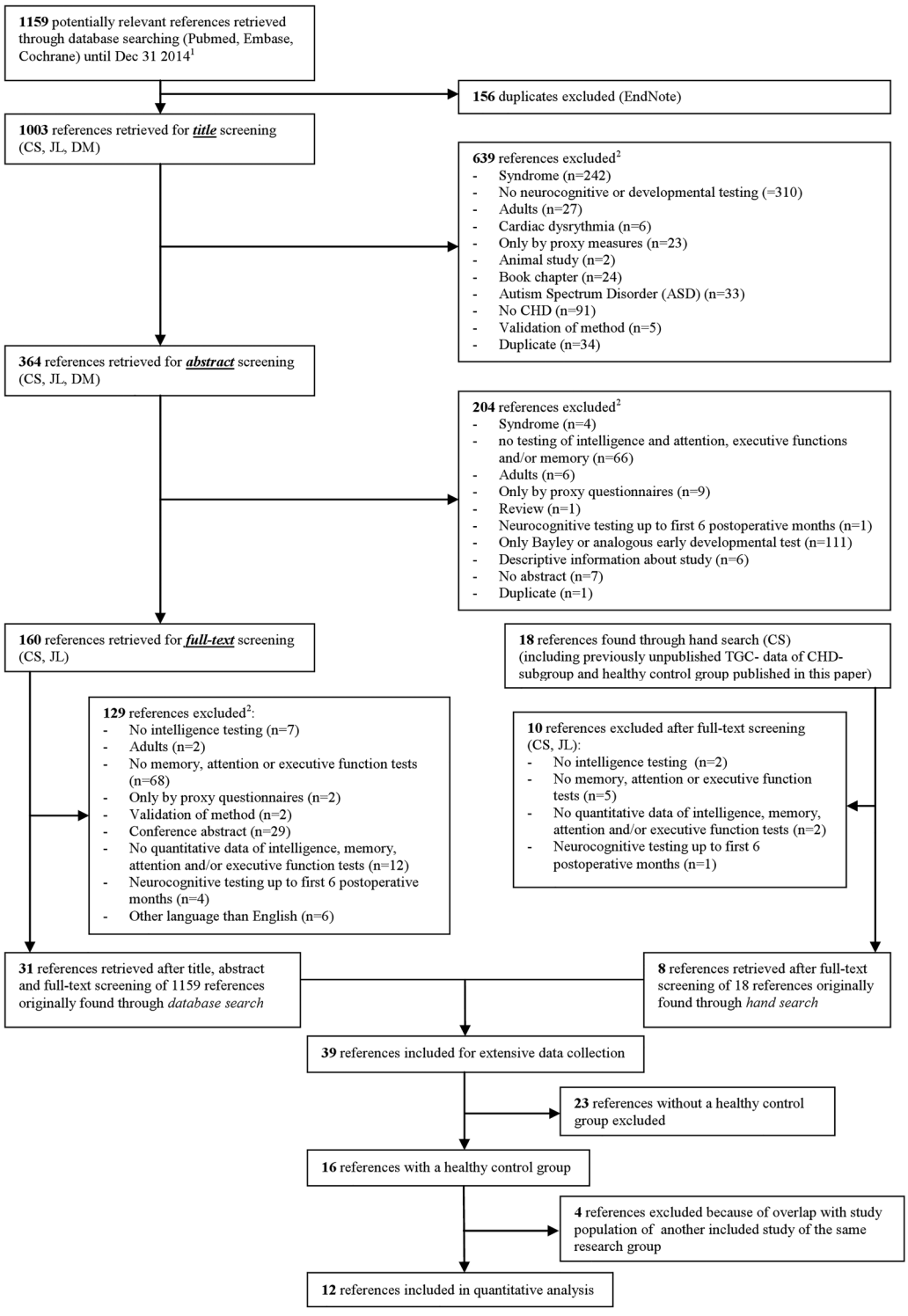


Figure 1 Flow diagram of Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA). [1]Search strategy in eMethods 1. [2]Manuscripts could be excluded for more than one reason (CHD, congenital heart disease).

meta-analysis may have hidden specific memory deficits. Children may also have been too young to detect differences in tests, which examine functions that are continuously developing through childhood.

The aforementioned indicates that the difficulties that children with CHD experience when performing IQ tests and complex tasks requiring EFs may be explained by a basic alertness deficit. This has also been found in other paediatric populations and may be linked with white matter changes.[40] Alternatively, other EFs, such as flexibility[20] or working memory, may preferentially be affected.

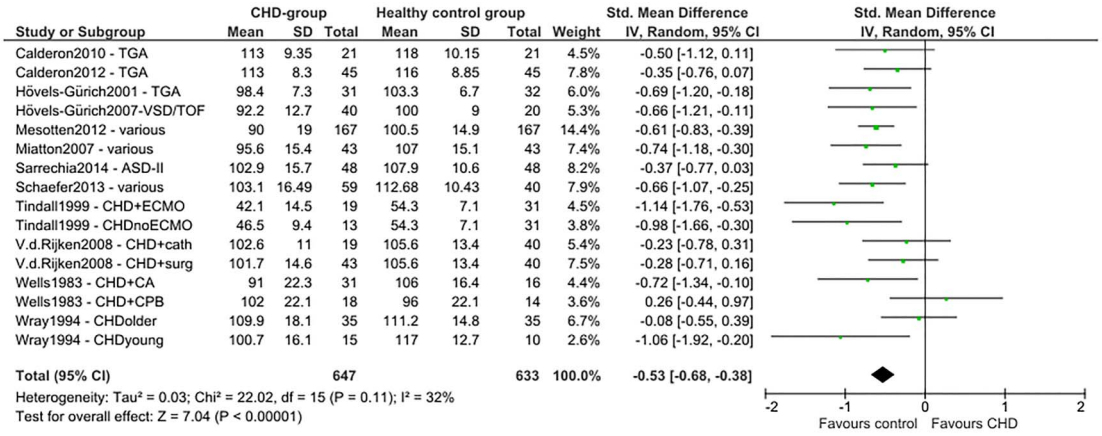


Figure 2 Forest plot of intelligence (author/year/journal—CHD type). TGA, d-transposition of the great arteries; VSD, ventricular septal defect; TOF, tetralogy of Fallot; ASD, atrial septal defect; CHD, congenital heart disease; ECMO, extracorporeal membrane oxygenation; V.d.Rijken, Van der Rijken; cath, catheterisation; surg, surgery; CA, circulatory arrest; CPB, cardiopulmonary bypass.

Strengths

This systematic review and meta-analysis has several strengths. Exploring the effects of CHD on not only intelligence but also on EFs, attention and memory in the same patients has been a new avenue in the assessment of neurocognitive function in children with CHD. This allowed us to compare the impact of CHD on intelligence with the impact of CHD on specific neurocognitive skills. The studies that included a healthy control group could be combined for the analysis of a SMD, despite the use of different tests. Therefore, the findings are fairly robust. The analysis of the LGC trial data offers a valuable contribution of neurocognitive data from a large CHD and healthy control group, thereby increasing power and precision in the meta-analysis. Certainly, propensity score matching for relevant factors in neurocognitive development improved the stringency and reliability of the analyses.

Limitations

The meta-analysis has a few inherent limitations though. First, attention and EFs involve several functions, which may interact.[9] Although the lower aggregated score on intelligence is clinically relevant, it is not clear whether the poorer results for attention and EF have any clinical impact, because universal definitions and test protocols are lacking. Unfortunately, the number of studies assessing different aspects of attention and EF was low. As a result, only one attentional function and one EF could be examined in this meta-analysis. Therefore, no conclusions can be drawn for other attention components and EFs. Future studies thus ought to use comparable test batteries for neurocognitive function assessments. Second, the very small sample sizes of the CHD and control groups, partially due to the separate reporting of a high variation in CHD diagnoses, reduced statistical power and precision both at the individual study level and at meta-analysis level. Because of the separate reporting of CHD subgroups, neurocognitive data of healthy controls were sometimes included twice in the forest plots. The slightly asymmetrical funnel plot of intelligence reflects a possible publication bias and may have overestimated the effect size. Future research should pay more attention to the statistical powering of

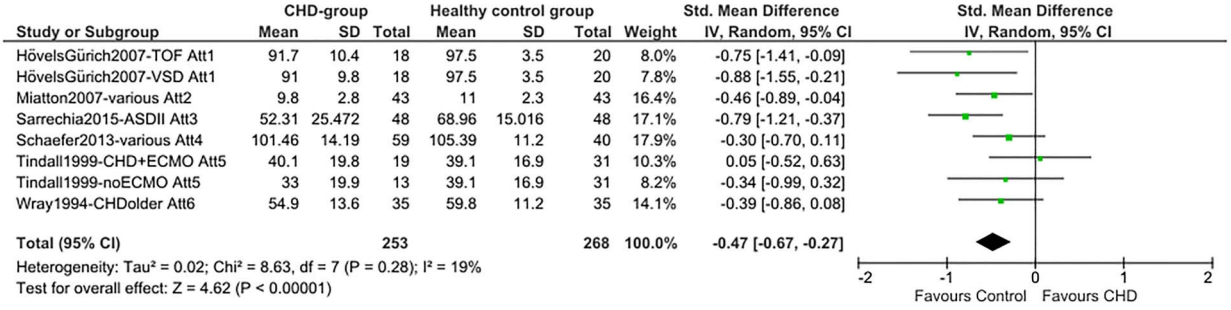


Figure 3 Forest plot of attention (Att; alertness) non-reaction time (author/year/journal—CHD-type/attention measure). Att1, mean accuracy (%) Attention Network Test; Att2, visual attention (NEuroPSYchological Assessment, NEPSY); Att3, auditory attention (NEPSY); Att4, processing speed (Wechsler Intelligence Scale for Children, WISC-IV); Att5, delay task, vigilance hits; Att6, speed of information (BAS, British Ability Scales); VSD, ventricular septal defect; TOF, tetralogy of Fallot; ASD, atrial septal defect; CHD, congenital heart disease; ECMO, extracorporeal membrane oxygenation.

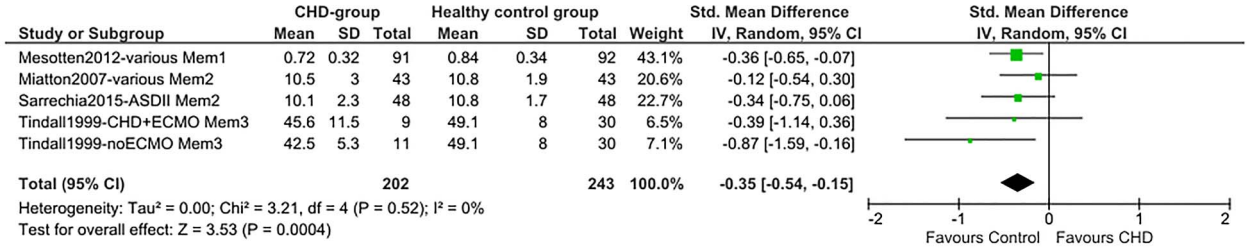


Figure 4 Forest plot of verbal memory (Mem; author/year/journal—CHD-type/memory measure). Mem1, sum of immediate and delayed verbal memory (Children's Memory Scale, Word Pairs, proportion correct responses); Mem2, narrative memory (NEuroPSYchological Assessment, NEPSY); Mem3, verbal memory (Wide Range Assessment of Memory and Learning, WRAML); ASD, atrial septal defect; CHD, congenital heart disease; ECMO, extracorporeal membrane oxygenation.

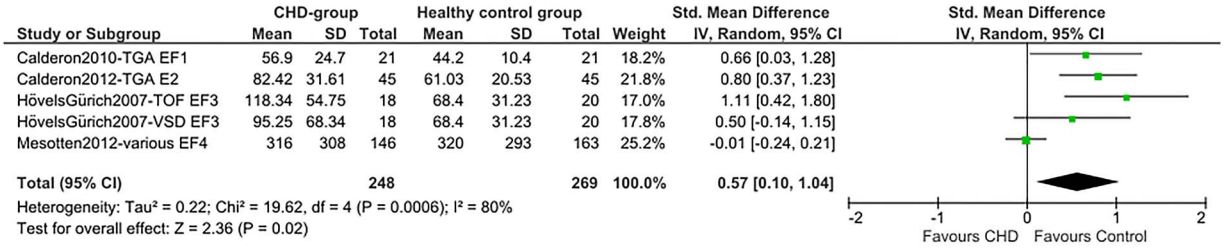


Figure 5 Forest plot of EF (inhibition) RT (author/year/journal—CHD-type/EF measure). EF, executive function; RT, reaction time; EF-test1, incongruent stroop; EF-test2, stroop; EF-test3, conflict (Attention Network Test); EF-test4, inhibition (Amsterdam Neuropsychological Tasks); TGA, d-transposition of the great arteries; VSD, ventricular septal defect; TOF, tetralogy of Fallot; CHD, congenital heart disease.

effects on neurocognitive outcome. Third, the exclusion of six non-English studies (one Italian, one Chinese, one French, one Spanish and two German) might have raised a language bias. However, research has shown no evidence of systematic bias from the use of language restrictions in systematic review-based meta-analyses.[41] Fourth, this meta-analysis cannot draw any conclusion on the interaction between effect of IQ and alertness, EF and memory. At least to exclude the fact that effects on EFs and memory are not entirely explained by a generalised or more basic impairment, one ought to examine IQ and alertness in addition to the other functions. Finally, studies on EFs, attention and memory outcomes in adolescence are lacking, indicating the need for long-term longitudinal follow-up studies to the level of secondary education,[18] when academic difficulties can appear. Adolescence is particularly essential for the late maturation of the prefrontal cortex and EFs.[37] Understanding the impact of these prefrontal cortex functions on daily life and school functioning is clinically relevant, because a better understanding and early detection of deficits may improve the daily functioning of children with CHD.[2]

CONCLUSIONS

Children with CHD who have undergone heart surgery have consistently worse performance of intelligence and alertness. In this meta-analysis, memory appears to be less affected. The effect of CHD on EF in young children cannot be reliably assessed due to poor standardisation of the testing methodology. Larger, more standardised, longitudinal long-term follow-up studies of specific neurocognitive skills in a large group of children with CHD and a matched healthy control group are necessary for a better understanding of neurocognitive deficits, and their impact on daily life and school functioning.

Acknowledgements The authors thank Jens De Groot (Biomedical Library, Catholic University Leuven) for his help with the search strategies.

Contributors CS and JL conceptualised and designed the study, carried out analyses, drafted the initial manuscript and approved the final manuscript as submitted. IV carried out analyses, and reviewed and approved the final manuscript as submitted. GVdB conceptualised and designed the study, carried out analyses, and reviewed and approved the final manuscript as submitted. DM conceptualised and designed the study, reviewed and revised the manuscript, and approved the final manuscript as submitted.

Funding This study was funded by the Flemish Agency for Innovation by Science and Technology (IWT/070695/TBM). DM holds a FWO-Flanders Senior Clinical Investigator Fellowship. GVdB via the University of Leuven receives structural research financing via the Methusalem programme, funded by the Flemish government, and holds an ERC advanced grant from the EU-FP7 (2012-321670).

Competing interests None declared.

Patient consent Obtained.

Ethics approval Institutional Review Board.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. Moons P, Sluysmans T, De Wolf D, *et al*. Congenital heart disease in 111 225 births in Belgium: birth prevalence, treatment and survival in the 21st century. *Acta Paediatr* 2009;98:472–7.
2. Marino BS, Lipkin PH, Newburger JW, *et al*. Neurodevelopmental outcomes in children with congenital heart disease: evaluation and management: a scientific statement from the American Heart Association. *Circulation* 2012;126:1143–72.
3. McCrindle BW, Williams RV, Mitchell PD, *et al*. Relationship of patient and medical characteristics to health status in children and adolescents after the Fontan procedure. *Circulation* 2006;113:1123–9.
4. Shillingford AJ, Glanzman MM, Ittenbach RF, *et al*. Inattention, hyperactivity, and school performance in a population of school-age children with complex congenital heart disease. *Pediatrics* 2008;121:e759–67.
5. Kamphuis M, Ottenkamp J, Vliegen HW, *et al*. Health related quality of life and health status in adult survivors with previously operated complex congenital heart disease. *Heart* 2002;87:356–62.
6. Hansen BM, Dinesen J, Hoff B, *et al*. Intelligence in preterm children at four years of age as a predictor of school function: a longitudinal controlled study. *Dev Med Child Neurol* 2002;44:517–21.
7. Forbess JM, Visconti KJ, Bellinger DC, *et al*. Neurodevelopmental outcomes in children after the fontan operation. *Circulation* 2001;104 (12 Suppl 1):I127–132.
8. Chapman SB. Neurocognitive stall: a paradox in long term recovery from pediatric brain injury. *Brain Inj Prof* 2007;3:10–13.
9. Diamond A. Executive functions. *Annu Rev Psychol* 2013;64:135–68.
10. Bull R, Espy KA, Wiebe SA. Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Dev Neuropsychol* 2008;33:205–28.
11. Alloway TP, Alloway RG. Investigating the predictive roles of working memory and IQ in academic attainment. *J Exp Child Psychol* 2010;106:20–9.
12. Limperopoulos C, Tworetzky W, McElhinney DB, *et al*. Brain volume and metabolism in fetuses with congenital heart disease: evaluation with quantitative magnetic resonance imaging and spectroscopy. *Circulation* 2010;121:26–33.
13. Miller SP, McQuillen PS, Hamrick S, *et al*. Abnormal brain development in newborns with congenital heart disease. *N Engl J Med* 2007;357:1928–38.
14. Mahle WT, Tavani F, Zimmerman RA, *et al*. An MRI study of neurological injury before and after congenital heart surgery. *Circulation* 2002;106(12 Suppl 1):I109–14.
15. Rivkin MJ, Watson CG, Scoppettuolo LA, *et al*. Adolescents with D-transposition of the great arteries repaired in early infancy demonstrate reduced white matter microstructure associated with clinical risk factors. *J Thorac Cardiovasc Surg* 2013;146:543–9.e1.
16. Miatton M, De Wolf D, Francois K, *et al*. Neurocognitive consequences of surgically corrected congenital heart defects: a review. *Neuropsychol Rev* 2006;16:65–85.
17. Bellinger DC, Newburger JW. Neuropsychological, psychosocial, and quality-of-life outcomes in children and adolescents with congenital heart disease. *Prog Pediatr Cardiol* 2010;29:87–92.
18. Snookes SH, Gunn JK, Eldridge BJ, *et al*. A systematic review of motor and cognitive outcomes after early surgery for congenital heart disease. *Pediatrics* 2010;125:e818–27.
19. Karsdorp PA, Everaerd W, Kindt M, *et al*. Psychological and cognitive functioning in children and adolescents with congenital heart disease: a meta-analysis. *J Pediatr Psychol* 2007;32:527–41.
20. Mesotten D, Gielen M, Sterken C, *et al*. Neurocognitive development of children 4 years after critical illness and treatment with tight glucose control: a randomized controlled trial. *JAMA* 2012;308:1641–50.
21. Calderon J, Bonnet D, Courtin C, *et al*. Executive function and theory of mind in school-aged children after neonatal corrective cardiac surgery for transposition of the great arteries. *Dev Med Child Neurol* 2010;52:1139–44.
22. Calderon J, Angeard N, Moutier S, *et al*. Impact of prenatal diagnosis on neurocognitive outcomes in children with transposition of the great arteries. *J Pediatr* 2012;161:94–8.e1.
23. Hovels-Gurich HH, Seghaye MC, Sigler M, *et al*. Neurodevelopmental outcome related to cerebral risk factors in children after neonatal arterial switch operation. *Ann Thorac Surg* 2001;71:881–8.
24. Hovels-Gurich HH, Konrad K, Skorzenski D, *et al*. Attentional dysfunction in children after corrective cardiac surgery in infancy. *Ann Thorac Surg* 2007;83:1425–30.
25. Miatton M, De Wolf D, Francois K, *et al*. Neuropsychological performance in school-aged children with surgically corrected congenital heart disease. *J Pediatr* 2007;151:73–8, 78 e71.
26. Sarrechia I, De Wolf D, Miatton M, *et al*. Neurodevelopment and behavior after transcatheter versus surgical closure of secundum type atrial septal defect. *J Pediatr* 2015;166:31–38 e31.
27. Schaefer C, von Rhein M, Knirsch W, *et al*. Neurodevelopmental outcome, psychological adjustment, and quality of life in adolescents with congenital heart disease. *Dev Med Child Neurol* 2013;55:1143–9.
28. Tindall S, Rothermel RR, Delamater A, *et al*. Neuropsychological abilities of children with cardiac disease treated with extracorporeal membrane oxygenation. *Dev Neuropsychol* 1999;16:101–15.
29. Van Der Rijken R, Hulstijn-Dirkmaat G, Kraaimaat F, *et al*. Open-heart surgery at school age does not affect neurocognitive functioning. *Eur Heart J* 2008;29:2681–8.
30. Wells FC, Coghill S, Caplan HL, *et al*. Duration of circulatory arrest does influence the psychological development of children after cardiac operation in early life. *J Thorac Cardiovasc Surg* 1983;86:823–31.
31. Wray J, Pot-Mees C, Zeitlin H, *et al*. Cognitive function and behavioural status in paediatric heart and heart-lung transplant recipients: the Harefield experience. *BMJ* 1994;309:837–41.
32. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health* 1998;52:377–84.
33. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd edn. Hillsdale, NJ: L. Erlbaum Associates, 1988.
34. Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.
35. Bhutta AT, Cleves MA, Casey PH, *et al*. Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. *JAMA* 2002;288:728–37.
36. Prentice DA Miller D. When small effects are impressive. *Psychol Bull* 1992;112:160–4.
37. Casey BJ, Tottenham N, Liston C, *et al*. Imaging the developing brain: what have we learned about cognitive development? *Trends Cogn Sci* 2005;9:104–10.
38. Waber DP, De Moor C, Forbes PW, *et al*. The NIH MRI study of normal brain development: performance of a population based sample of healthy children aged 6 to 18 years on a neuropsychological battery. *J Int Neuropsychol Soc* 2007;13:729–46.
39. Posner MI. Measuring alertness. *Ann N Y Acad Sci* 2008;1129:193–9.
40. Potharst ES, van Wassenaer-Leemhuis AG, Houtzager BA, *et al*. Perinatal risk factors for neurocognitive impairments in preschool children born very preterm. *Dev Med Child Neurol* 2013;55:178–84.
41. Morrison A, Polisena J, Husereau D, *et al*. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. *Int J Technol Assess Health Care* 2012;28:138–44.
42. Jenkins KJ, Gauvreau K, Newburger JW, *et al*. Consensus-based method for risk adjustment for surgery for congenital heart disease. *J Thorac Cardiovasc Surg* 2002;123:110–18.
43. Leteurtre S, Martinot A, Duhamel A, *et al*. Validation of the paediatric logistic organ dysfunction (PELOD) score: prospective, observational, multicentre study. *Lancet* 2003;362:192–7.

APPENDIX

eMethods 1: More details on data analysis and study selection

eMethods 2: Search strategy Pubmed & Embase

eResults: eTable 1 and eFigures 1-4

eReferences

eMethods 1: More details on study selection and data analysis

Data analysis LGC-trial data

Propensity scores were estimated using logistic regression with syndrome, origin (exclusive European), language (exclusively Dutch language) and gender, as dichotomous covariates and age at follow-up and socio-economic status, as continuous covariates. One-to-one nearest neighbor matching without replacement was performed, using a caliper of 0.07. Satisfactory matching was obtained, as indicated by absolute standardized difference in means less than or equal to 0.1 for all variables.[1] Matching was successful over the full range of propensity scores. 167 CHD-patients were matched to 167 controls.

Systematic review

Study selection

We excluded studies that only used the Bayley Infant Developmental Test[2] or an equivalent early developmental test, because such general developmental tests do not offer additional data on the specific neurocognitive skills.

Data analysis

If neurocognitive data of all tested children with CHD together were available in a particular study, the sample size of the total CHD-group was used in the sample size calculation. Otherwise the sample sizes of the tested subgroups were used. If age data were missing or only median ages or ranges were mentioned, these were not included in the age calculation. In case of overlap of study population between studies of the same research group, the study with the largest sample sizes was included in the meta-analysis. In case of equal sample sizes, the study with neurocognitive data that were most related to other included studies, was chosen.

Because of a different interpretation of reaction time and non-reaction time measures of executive

functions (EFs) and attention, the SMD-calculation was done separately. For attention, reaction and non-reaction time data related to the same attentional function, namely alertness, were included. For EFs, reaction and non-reaction time data related to the same EF, namely inhibition, were included.

With regard to memory, the SMD-calculation for verbal and non-verbal measures was done separately. If neurocognitive data of all tested children with CHD together were available in a particular study, these neurocognitive data were used in the effect size size calculation. Otherwise the data of CHD-subgroups were used.

The I^2-statistic was used to evaluate heterogeneity. It describes the percentage of the variability in effect estimates that is due to heterogeneity rather than chance (http://handbook.cochrane.org/). A funnel plot is a simple scatter plot of the effect estimates (here SMD's) from individual studies against some measure of each study's size or precision (here standard error, SE) (http://handbook.cochrane.org/). The precision of the estimated effect increases as the standard error decreases (or study size increases), hereby creating a funnel.

eMethods 2

Search strategy Pubmed

("Reaction Time"[Mesh] OR "Arousal"[Mesh] OR "Impulsive Behavior"[Mesh:noexp] OR "Mental Processes"[Majr] OR "Cognition Disorders"[Majr] OR "Neurobehavioral Manifestations"[Majr] OR "Mental Disorders Diagnosed In Childhood"[Majr] OR "Intelligence"[Mesh] OR "Psychomotor Performance"[Mesh] OR "Inhibition (Psychology)"[Mesh] OR neurocognitive function*[tiab] OR neuropsychological function*[tiab] OR neurodevelopmental function*[tiab] OR developmental function*[tiab] OR intellectual function*[tiab] OR mental function*[tiab] OR "neurocognitive development"[tiab] OR "neuropsychological development"[tiab] OR intellectual development*[tiab] OR “mental development”[tiab] OR "neurocognitive follow-up"[tiab] OR "neuropsychological follow-up"[tiab] OR "neurodevelopmental follow-up"[tiab] OR "developmental follow-up"[tiab] OR neurocognitive outcome*[tiab] OR neuropsychological outcome*[tiab] OR neurodevelopmental outcome*[tiab] OR developmental outcome*[tiab] OR academic outcome*[tiab] OR intellectual outcome*[tiab] OR mental outcome*[tiab] OR neurocognitive deficit*[tiab] OR neuropsychological deficit*[tiab] OR neurodevelopmental deficit*[tiab] OR developmental deficit*[tiab] OR academic deficit*[tiab] OR intellectual deficit*[tiab] OR mental deficit*[tiab] OR neurocognitive disabilit*[tiab] OR neuropsychological disabilit*[tiab] OR neurodevelopmental disabilit*[tiab] OR developmental disabilit*[tiab] OR academic disabilit*[tiab] OR intellectual disabilit*[tiab] OR mental disabilit*[tiab] OR neurocognitive disorder*[tiab] OR neuropsychological disorder*[tiab] OR neurodevelopmental disorder*[tiab] OR developmental disorder*[tiab] OR intellectual disorder*[tiab] OR mental disorder*[tiab] OR neurocognitive delay*[tiab] OR neuropsychological delay*[tiab] OR neurodevelopmental delay*[tiab] OR developmental

delay*[tiab] OR academic delay*[tiab] OR intellectual delay*[tiab] OR mental delay*[tiab] OR neurocognitive impairment*[tiab] OR neuropsychological impairment*[tiab] OR neurodevelopmental impairment*[tiab] OR developmental impairment*[tiab] OR academic impairment*[tiab] OR intellectual impairment*[tiab] OR mental impairment*[tiab] OR neurocognitive difficult*[tiab] OR neuropsychological difficult*[tiab] OR neurodevelopmental difficult*[tiab] OR developmental difficult*[tiab] OR academic difficult*[tiab] OR intellectual difficult*[tiab] OR mental difficult*[tiab] OR neurocognitive disturbance*[tiab] OR neuropsychological disturbance*[tiab] OR neurodevelopmental disturbance*[tiab] OR developmental disturbance*[tiab] OR academic disturbance*[tiab] OR intellectual disturbance*[tiab] OR mental disturbance*[tiab] OR intelligence[tiab] OR "attention"[tiab] OR ADHD[tiab] OR executive function*[tiab] OR "working memory"[tiab] OR inhibition[tiab] OR inhibitory[tiab] OR inhibiting[tiab] OR inhibits[tiab] OR inhibit[tiab] OR cognit*[tiab] OR learning[tiab] OR memory[tiab] OR reading[tiab] OR reaction time*[tiab] OR response time*[tiab] OR "processing speed"[tiab] OR vigilance[tiab] OR “executive control” [tiab] OR “problem solving”[tiab] OR “solving problems”[tiab] OR dyslexia[tiab] OR dyscalculia[tiab] OR impulsiveness[tiab] OR impulsivity[tiab])

AND

("Heart Defects, Congenital"[Majr] OR "Cardiac Surgical Procedures"[Majr] OR "Cardiac Catheterization"[Majr] OR cardiac malformation*[tiab] OR ((heart defect*[tiab] OR heart disease*[tiab] OR cardiac disease*[tiab] OR cardiac defect*[tiab]) AND (congenital[tiab] OR inherited[tiab] OR hereditary[tiab] OR inborn[tiab] OR genetic[tiab])) OR "cardiac surgery"[tiab] OR heart surger*[tiab] OR heart operation*[tiab] OR catheterization*[tiab])

AND

("Psychological Tests"[Mesh] OR neurocognitive assessment*[tiab] OR neuropsychological assessment*[tiab] OR neurodevelopmental assessment*[tiab] OR developmental

assessment*[tiab] OR academic assessment*[tiab] OR intellectual assessment*[tiab] psychological assessment*[tiab] OR mental assessment*[tiab] OR neurocognitive test*[tiab] OR psychological test*[tiab] OR neuropsychological test*[tiab] OR neurodevelopmental test*[tiab] OR developmental test*[tiab] OR academic test*[tiab] OR mental test*[tiab] OR neurocognitive evaluation*[tiab] OR neuropsychological evaluation*[tiab] OR neurodevelopmental evaluation*[tiab] OR developmental evaluation*[tiab] OR intellectual evaluation*[tiab] OR psychological evaluation*[tiab] OR mental evaluation*[tiab] OR test[tiab] OR tests[tiab] OR testing[tiab] OR scale*[tiab] OR scaling[tiab])

AND

("infant"[MeSH Terms] OR "child"[MeSH Terms] OR "adolescent"[MeSH Terms] OR "Young Adult"[Mesh] OR infant*[tw] OR child*[tw] OR adolescen*[tw] OR young adult*[tw])

NOT

((animal NOT human) OR “Review”[Publication Type] OR "Case Reports" [Publication Type])

Search strategy Embase

#84 #78 NOT #83

#83 #77 OR #79 OR #80 OR #81 OR #82

#82 'editorial'/exp

#81 'conference paper'/exp

#80 'case report'/exp

#79 'case study'/exp

#78 #74 NOT (#75 NOT #76)

#77 'review'/exp OR 'systematic review'/exp OR 'systematic review (topic)'/exp

#76 human AND [embase]/lim

#75 animal AND [embase]/lim

#74 #69 AND #73

#73 #70 OR #71 OR #72

#72 infant* OR child* OR adolescen* OR 'young adult' OR 'young adults' AND [embase]/lim

#71 'adolescent'/exp

#70 'child'/exp

#69 #54 AND #62 AND #68

#68 #63 OR #64 OR #65 OR #66 OR #67

#67 scale*:ab,ti OR scaling:ab,ti AND [embase]/lim

#66 ((neurocognitive OR neuropsychological OR neurodevelopmental OR psychological OR developmental OR mental OR academic) NEXT/1 evaluation*):ab,ti AND [embase]/lim

#65 ((neurocognitive OR neuropsychological OR psychological OR neurodevelopmental OR developmental OR mental OR academic) NEXT/1 test*):ab,ti AND [embase]/lim

#64 ((neurocognitive OR neuropsychological OR neurodevelopmental OR psychological OR developmental OR intellectual OR mental OR academic) NEXT/1 assessment*):ab,ti AND [embase]/lim

#63 'psychologic test'/exp

#62 #55 OR #56 OR #57 OR #58 OR #59 OR #60 OR #61

#61 catheterization* AND [embase]/lim

#60 'cardiac surgery':ab,ti OR 'heart surgery':ab,ti OR 'heart operation':ab,ti OR 'heart operations':ab,ti AND [embase]/lim

#59 (('heart defect' OR 'heart defects' OR 'heart disease' OR 'heart diseases' OR 'cardiac defect' OR 'cardiac defects' OR 'cardiac disease' OR 'cardiac diseases') NEAR/1 (congenital OR inherited OR hereditary OR inborn OR genetic)):ab,ti AND [embase]/lim

#58 'cardiac malformation':ab,ti OR 'cardiac malformations':ab,ti AND [embase]/lim

#57 'heart catheterization'/exp/mj

#56 'heart surgery'/exp/mj

#55 'congenital heart malformation'/exp/mj

#54 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40 OR #41 OR #42 OR #43 OR #44 OR #45 OR #46 OR #47 OR #48 OR #49 OR #50 OR #51 OR #52 OR #53

#53 impulsiv*:ab,ti AND [embase]/lim

#52 dyslexia:ab,ti OR dyscalculia:ab,ti AND [embase]/lim

#51 (problem* NEAR/1 solving):ab,ti AND [embase]/lim

#50 vigilance:ab,ti AND [embase]/lim

#49 'processing speed':ab,ti AND [embase]/lim

#48 ((react* OR response) NEAR/2 time*):ab,ti AND [embase]/lim

#47 learning:ab,ti OR memory:ab,ti OR reading:ab,ti AND [embase]/lim

#46 cognit*:ab,ti AND [embase]/lim

#45 inhibition:ab,ti OR inhibitory:ab,ti OR inhibit:ab,ti OR inhibits:ab,ti OR inhibiting:ab,ti AND [embase]/lim

#44 'working memory':ab,ti AND [embase]/lim

#43 (executive NEXT/1 (control* OR function*)):ab,ti AND [embase]/lim

#42 intelligence:ab,ti AND [embase]/lim

#41 ((neurocognitive OR neuropsychological OR neurodevelopmental OR developmental OR intellectual OR mental OR academic) NEXT/1 disturbance*):ab,ti AND [embase]/lim

#40 ((neurocognitive OR neuropsychological OR neurodevelopmental OR developmental OR intellectual OR mental OR academic) NEXT/1 difficult*):ab,ti AND [embase]/lim

#39 ((neurocognitive OR neuropsychological OR neurodevelopmental OR developmental OR intellectual OR mental OR academic) NEXT/1 impairment*):ab,ti AND [embase]/lim

#38 ((neurocognitive OR neuropsychological OR neurodevelopmental OR developmental OR intellectual OR mental OR academic) NEXT/1 delay*):ab,ti AND [embase]/lim

#37 ((neurocognitive OR neuropsychological OR neurodevelopmental OR developmental OR intellectual OR mental) NEXT/1 disorder*):ab,ti AND [embase]/lim

#36 ((neurocognitive OR neuropsychological OR neurodevelopmental OR developmental OR intellectual OR mental OR academic) NEXT/1 disabilit*):ab,ti AND [embase]/lim

#35 ((neurocognitive OR neuropsychological OR neurodevelopmental OR developmental OR intellectual OR mental OR academic) NEXT/1 deficit*):ab,ti AND [embase]/lim

#34 ((neurocognitive OR neuropsychological OR neurodevelopmental OR developmental OR intellectual OR mental OR academic) NEXT/1 outcome*):ab,ti AND [embase]/lim

#33 ((neurocognitive OR neuropsychological OR neurodevelopmental OR developmental) NEXT/1 'follow-up'):ab,ti AND [embase]/lim

#32 ((neurocognitive OR neuropsychological OR intellectual OR mental) NEXT/1 development*):ab,ti AND [embase]/lim

#31 ((neurocognitive OR neuropsychological OR neurodevelopmental OR developmental OR intellectual OR mental) NEXT/1 function*):ab,ti AND [embase]/lim

#30 'mental deficiency'/exp

#29 'learning disorder'/exp

#28 'perception disorder'/exp OR 'psychomotor disorder'/exp

#27 'impulse control disorder'/de

#26 'communication disorder'/de

#25 'attention deficit disorder'/exp

#24 'autism'/exp

#23 'memory disorder'/de

#22 'language disability'/exp

#21 'intellectual impairment'/de

#20 'developmental coordination disorder'/exp

#19 'cognitive defect'/exp

#18 'attention disturbance'/exp

#17 'apraxia'/exp

#16 'thinking'/exp

#15 'perception'/exp

#14 'orientation'/exp

#13 'mental capacity'/exp OR 'mental development'/exp OR 'mental performance'/exp

#12 'learning'/exp OR 'memory'/exp

#11 'executive function'/exp

#10 'cognition'/de

#9 'impulsiveness'/exp

#8 'wakefulness'/exp

#7 'sensorimotor function'/exp

#6 'psychophysics'/exp

#5 'response time'/exp

#4 'psychological refractory period'/exp

#3 'attention'/exp

#2 'arousal'/exp

#1 'psychophysiology'/de

eResults

eTable 1

Reference	N CHD	CHD-group	Control group	N Control	Age surgery	Age testing	IQ test	Attention test	Memory test	Executive function test	Risk of bias
Calderon J et al. J Pediatr. 2012;161(1):94-8.	45	TGA-prenatal diagnosis (N=29) TGA-postnatal diagnosis (N=16)	Recruited and examined in the same period and same geographic area	45	TGA- prenatal Mean 7.5d (SD 3.9) TGA-postnatal Mean 6.6d (SD 2.9)	4-6y	Columbia Mental Maturity Scale	/	/	Stroop Test, knock and tap subtest (NEPSY), digit span test (WISC-IV) and spatial span task (BEM-144 blocks), Dimensional Card Sorting test	18
Calderon J et al. Developmental Medicine and Child Neurology. 2010;52(12):1139-44.	21	TGA	Children randomly recruited in two Parisian primary schools over the same time period and matched to the study group according to age and sex	21	Neonatal arterial switch operation	TGA: mean 7y4m control: mean 7y6m	Columbia Mental Maturity Scale	/	/	Incongruent Stroop, Statue (NEPSY), Digit Span Backwards, spatial span task (BEM-144 blocks), Tower of London	17

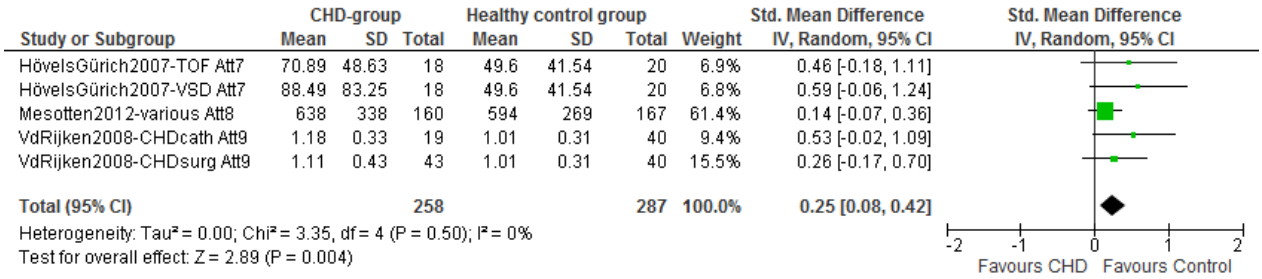
Mesotten D et al. JAMA. 2012;308(16):1641-50.	361 (167 PSM)	various CHD's	healthy controls	215 (167 PSM)	Matched CHD-group: median 2.11 (IQR 0.34-4.77),	Matched CHD-group: median 6.02 (IQR 4.21-8.78), Matched control group: median 5.91 (IQR 4.58 - 9.07)	Wechsler Intelligence Scales (depending on age)	Baseline Speed (Amsterdam Neuropsychological Tests)	Numbers, Word Pairs, Pictures, Dots (Children's Memory Scale)	Response Organization Objects (Amsterdam Neuropsychological Tests), Numbers Backwards (Children's Memory Scale)	16
Hovels-Gurich HH et al. Ann Thorac Surg. 2007;83(4):1425-30.	40	20 TOF 20 VSD	healthy controls	20	TOF: Mean 0.7y (SD 0.03y) VSD: Mean 0.7y (SD 0.2y)	TOF: Mean 7.4y (SD 1.4y) VSD: Mean 7.4y (SD 1.9y) controls Mean 8.1y (SD 1.7y)	Kaufman Battery of Children	Attention Network Test (ANT)	/	Attention Network Test (ANT)	15
Hovels-Gurich HH et al. Ann Thorac Surg. 2001;71(3):881-8.	33	24 TGA 4 TGA + unimportant VSD 4 with TGA + VSD closed 1 with TGA + coarctatio aortae	32 age-matched healthy children without perinatal complications and with a normal neurodevelopmental status according to their parents'	32	Mean 7.0 (SD 4.7d)	Study group: Mean 3.6 (SD 0.5y) Control group: Mean 3.8y (SD 0.6)	Vienna Developmental Test	/	Vienna Developmental Test: subscale learning and memory	/	17

Miatton M et al. J Pediatr. 2007;151(1):73-8, 8 e1.	43	cyanotic group (N=26), acyanotic group (N=17)	healthy controls group matched for sex, age and educational-level	43	not mentioned (all children underwent open-heart surgery)	Total CHD-group: Mean 8y8m (SD 1y6m) Control group: Mean 8y11m (SD 1y7m)	WISC-III:	attention and executive functioning (NEPSY)	memory domain (NEPSY)	Attention and Executive Functioning (NEPSY)	15
Sarrechia I et al. The Journal of pediatrics. 2015;166(1):31-38 e31.	48	ASD-II: surgical repair group (N=18), transcatheter repair group (N=30)	Healthy controls recruited through approval of primary school boards and matched with the patients on sex, age, and parental education	48	Surgical repair group Mean 2 y9m (SD 1y8m) Transcatheter repair group Mean 4y2m (SD 1y7m)	Surgical repair group: age Mean 9y2m (SD 2y2m) Transcatheter repair group: Mean 9y3m (SD 1y7m)	Shortened version of WISC-III	Auditory attention (NEPSY-II-NL)	Memory and Learning: (NEPSY-II-NL)	Executive domain (NEPSY-II-NL)	16
Schaefer C et al. Dev Med Child Neurol. 2013;55(12):1143-1149.	59	various CHD's (49% acyanotic CHD's, 51% cyanotic CHD's)	Healthy children, similar to CHD-group in terms of age, sex and socio-economic status	40	median 0.9y (range 0-5.6y)	median age: 13y8m (range 11y5m - 16y11m)	WISC-IV	processing speed (WISC-IV)	ROCF: memory	working memory (WISC-IV), ROCF-Test	17

Tindall S et al. Developmental neuropsychology. 1999;16(1):101-15.	22 (ECMO / CHD)	9 CHD + ECMO 13 CHD + No ECMO various CHD's	31 children obtained from daycare settings and private referrals	31	ECMO: mean 11.7m (SD 9.8) CHD: Mean 13.4m (SD 8.7)	ECMO: mean 63.5m (SD 9.8) CHD: mean 65.3 (SD 9.8) Control: mean 64.1 (SD 8.4)	McCarthy Scales of Children's Abilities (MSCA)	Gordon Diagnostic Systemat Delay Task (GDS)	Wide Range Assessment of Memory and Learning (WRAML)	Gordon Diagnostic System Delay Task (GDS): efficiency ratio (measure of impulsivity)	11
Van Der Rijken R et al. European Heart Journal. 2008;29(21):2681-8.	62	43 various CHD's awaiting surgery 19 various CHD's awaiting catheterization	healthy children from regional mainstream primary and secondary schools, matched for age, sex, educational level, general intelligence, and parental educational level with both patient groups	40	not mentioned exactly; assessment took place in the period prior to surgery or cath. and again 1 year afterwards	CHD-surg Mean 11.6 (SD 3.1) CHD-cath Mean 11.0 (SD 2.6) Control group 11.7 (SD 2.8)	WISC-III	Complex Reaction Time Task, Bourdon-Vos, Letter Detection Task (Amsterdam Neuropsychological Test), Computerized drawing task	/	/	18

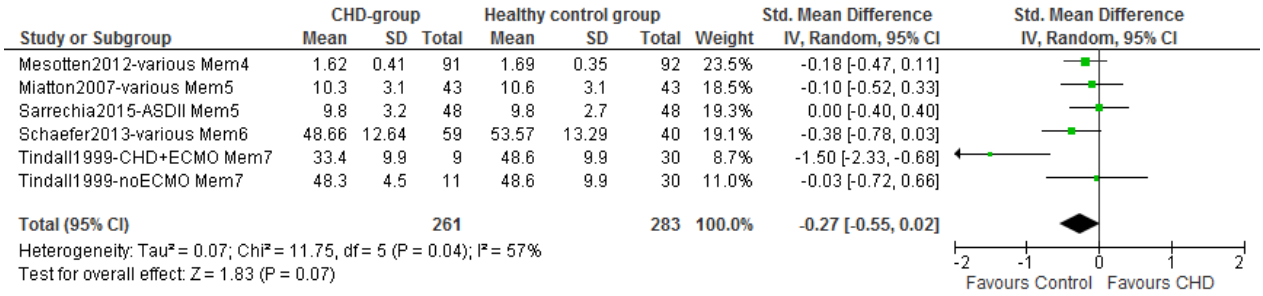
Wells FC et al. J Thorac Cardiovasc Surg. 1983;86(6):823-31.	49	Various CHD's: TCA-group (hypothermia and circulatory arrest) (N=31) CPB-group (continuous cardiopulmonary bypass) (N=18)	16 siblings of TCA patients & 13 siblings of CPB patients	29	TCA-group: 15m ± 2.0 CPB-group: 16m ± 2.0	TCA-group: Mean 68m (SD 1.1) CPB-group: Mean 70m (SD 1.7) Siblings TCA-group: Mean 86m (SD 4.8) Siblings CPB-group: Mean 84m (SD 5.9)	McCarthy Scales of Children's Abilities (MSCA) or Wechsler Intelligence Scales (depending on age)	/	Memory subscale (MSCA)	/	10
Wray J et al. BMJ 1994;309(6958):837-41.	50	Conventional heart surgery Heart or heart-lung-transplantation	healthy children with no medical problems	45	not exactly mentioned; Tx-group: mean time after Tx at the time of assessment = 10m (range 3-25m); Cardiac group: cardiac surgery over 24m	mean age Tx-group: 9.4y (range 0.6-16.6y) Cardiac group: 6.2y Control group: 8.2y	Ruth Griffiths mental developmental scales or Short Form IQ from British Ability scales (BAS), depending on age	speed of information (BAS):	short term memory, retrieval of knowledge (BAS)	/	15

eFigure 1. Forest plot of attention (Att) (alertness) reaction time (author/year/journal – CHD-type / attention measure)



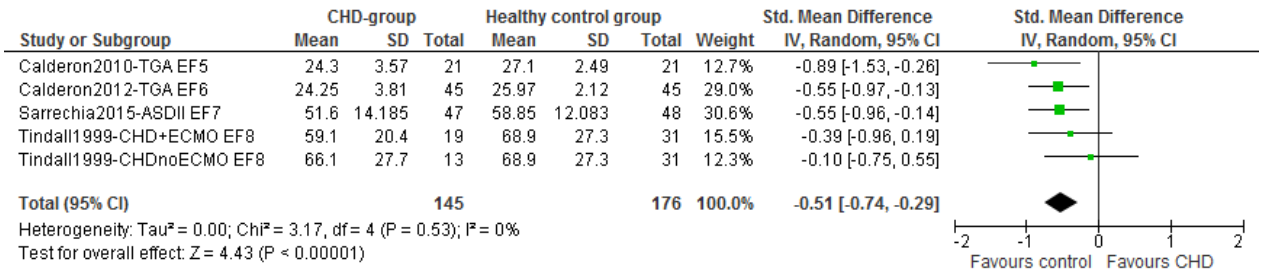
Abbreviations: Att7 = Alerting (msec) (Attention Network Test); Att8 = Mean reaction time (msec), alertness task (Amsterdam Neuropsychological Tasks); VdRijken = Van der Rijken; cath = catheterization; Att9 = Amsterdam Neuropsychological Task, reaction time; surg = surgery

eFigure 2. Forest plot of non-verbal memory (author/year/journal – CHD-type / Memory measure)



Legend: Mem4 = Sum of Immediate and Delayed Non-Verbal Memory (Children's Memory Scale, Dots, proportion correct responses); Mem5 = Memory for Faces (NEuroPSYchological Assessment, NEPSY);Mem6 = Rey-Osterrieth Complex Figure Test (ROCF), memory; Mem7 = Visual Memory (Wide Range Assessment of Memory and Learning, WRAML)

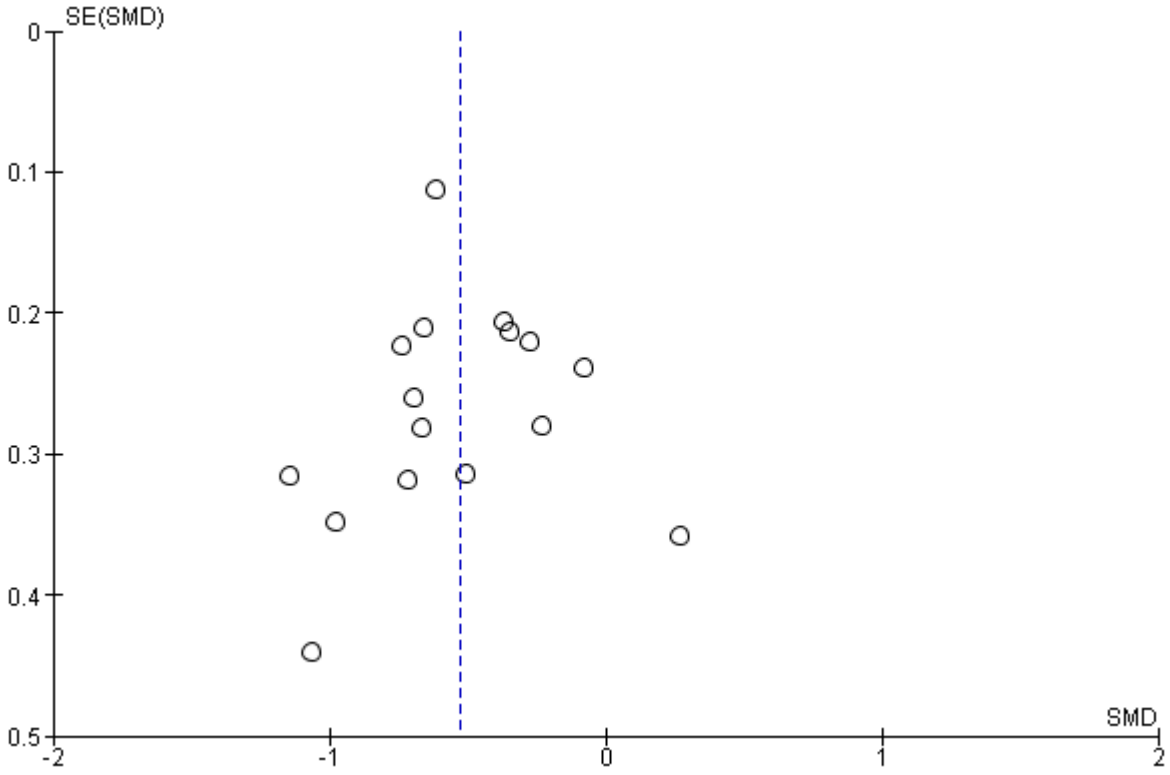
eFigure 3. Forest plot of executive function (EF) (inhibition) non-reaction time (author/year/journal – CHD-type / EF measure)



Legend: EF-Test5 = Statue; EF-Test6 = Knock & Tap; EF-Test7 = Inhibition (NEPSY); EF-Test8 = Delay Task, Efficiency ratio

eFigure 4. Slightly asymmetrical funnel plot of intelligence (N=16)

12 original studies; Abbreviations: SE = standard error; SMD = standardized mean difference



eReferences

1. Gayat E, Pirracchio R, Resche-Rigon M, Mebazaa A, Mary JY, Porcher R. Propensity scores in intensive care and anaesthesiology literature: a systematic review. *Intensive care medicine.* 2010;36(12):1993-2003.
2. Bayley N. *Manual for the Bayley Scales of Infant and Toddler Development. 3rd ed.* . San Antonio, TX: Harcourt Assessment; 2006.