

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

Caroline Sterken,<sup>1</sup> Jurgen Lemiere,<sup>2,3</sup> Ilse Vanhorebeek,<sup>1</sup> Greet Van den Berghe,<sup>1</sup> Dieter Mesotten<sup>1</sup>

**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



CrossMark

<sup>1</sup>Division of Cellular and Molecular Medicine, Clinical Department and Laboratory of Intensive Care Medicine, KU Leuven, Leuven, Belgium

<sup>2</sup>Child and Adolescent Psychiatry, University Hospitals Leuven, Leuven, Belgium

<sup>3</sup>Paediatric Haemato-Oncology, University Hospitals Leuven, Leuven, Belgium

**Correspondence to**  
Dr Dieter Mesotten; [dieter.mesotten@med.kuleuven.be](mailto: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](http://clinicaltrials.gov) Identifier NCT00214916.

## INTRODUCTION

As the survival rate of children with congenital heart disease (CHD) continues to improve thanks to medical advances,<sup>1</sup> public interest and research have been focusing more on how children with CHD survive.<sup>2</sup>

Children with CHD may experience difficulties in daily life<sup>3</sup> and academic functioning,<sup>4</sup> which may persist into adulthood.<sup>5</sup> 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.<sup>6</sup> However, children with CHD often have intelligence scores within average range compared with population norms.<sup>7</sup> 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.<sup>8</sup> 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<sup>9</sup> and contribute to academic performance.<sup>10 11</sup> Other neurocognitive skills, such as memory and attention, which are inter-related with EFs, also contribute to academic performance.<sup>9</sup>

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.<sup>12</sup> Postnatally, infants and older children with CHD may have preoperative<sup>13</sup> and postoperative white matter abnormalities,<sup>14</sup> which may relate to worse neurocognitive outcome in children with CHD.<sup>15</sup> Surgery seems to impact neurocognitive outcome in CHD as well.<sup>15</sup> 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.<sup>2 16 17</sup>

However, to our knowledge, the outcome data of specific neurocognitive functions have not been analysed systematically in children with CHD. One systematic review<sup>18</sup> and one meta-analysis<sup>19</sup> 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.<sup>20</sup> 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.<sup>20</sup> 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, socio-economic 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  $\chi^2$  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.<sup>20</sup>

### 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<sup>20–31</sup> 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.<sup>32</sup> 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.<sup>33</sup> For clinical interpretation, the overall SMD for intelligence scores was multiplied by the typical SD of the normal IQ distribution (mean  $100 \pm SD$  15). The  $I^2$  statistic was used to evaluate heterogeneity.  $I^2 \geq 50\%$  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  $\Delta N$  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 ( $\Delta N$  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).<sup>34</sup>

An overview of the 12 included studies is presented in [table 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 0.07–3.85) 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  $0.42$ ),  $p = 0.004$ ,  $I^2 = 0\%$ ; [figure 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 ([figure 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  $-0.29$ ),  $p < 0.00001$ ,  $I^2 = 0\%$ ; [figure 3](#)). The funnel plot for intelligence was slightly asymmetrical ([figure 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;<sup>43</sup> RACHS, risk adjustment for congenital heart surgery.<sup>42</sup>

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

Median (IQR) are reported.

\*Higher scores reflect worse performance.

†Higher scores reflect better performance.

CBCL, Child Behavior Checklist.

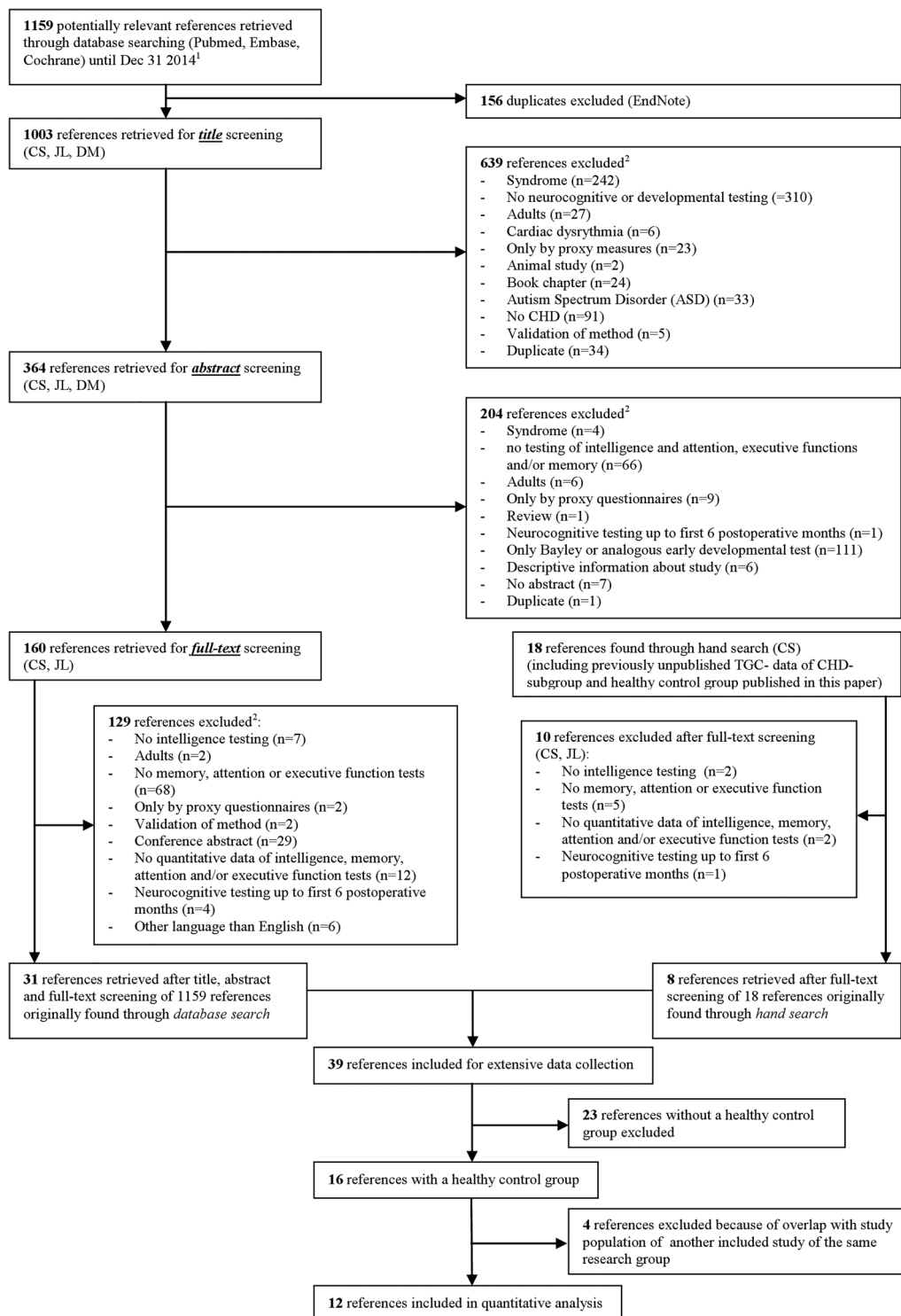
the meta-analysis and LGC trial are of similar magnitude to those in prematurely born children,<sup>35</sup> and are in line with the meta-analysis of Karsdorp *et al.*<sup>19</sup> Despite the worse intelligence scores for the CHD group compared with healthy controls, the mean IQ scores of the children 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.<sup>36</sup> 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<sup>37 38</sup> 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.<sup>39</sup> 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.<sup>2 16 17</sup> 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<sup>20</sup> 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.<sup>20</sup> 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,<sup>2 16 17</sup> 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

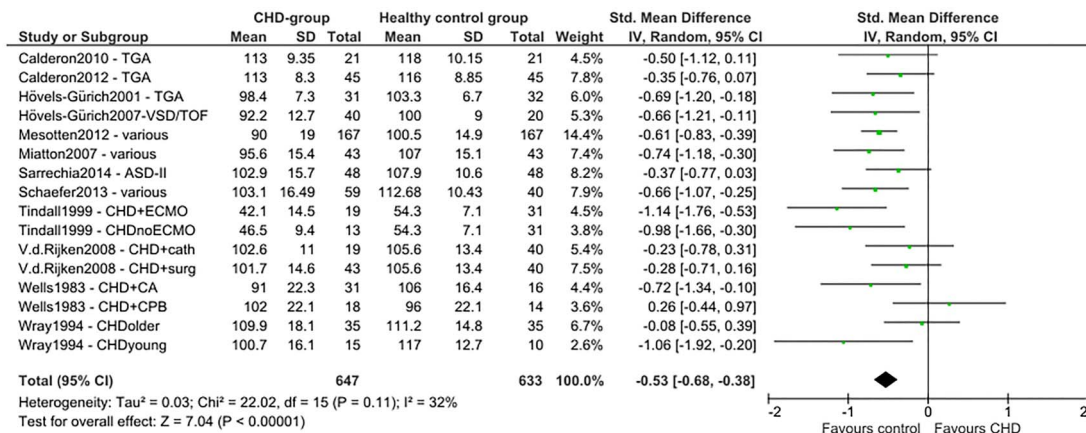


**Figure 1** Flow diagram of Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA). <sup>1</sup>Search strategy in eMethods 1. <sup>2</sup>Manuscripts could be excluded for more than one reason (CHD, congenital heart disease).

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.<sup>40</sup> Alternatively, other EFs, such as flexibility<sup>20</sup> 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.

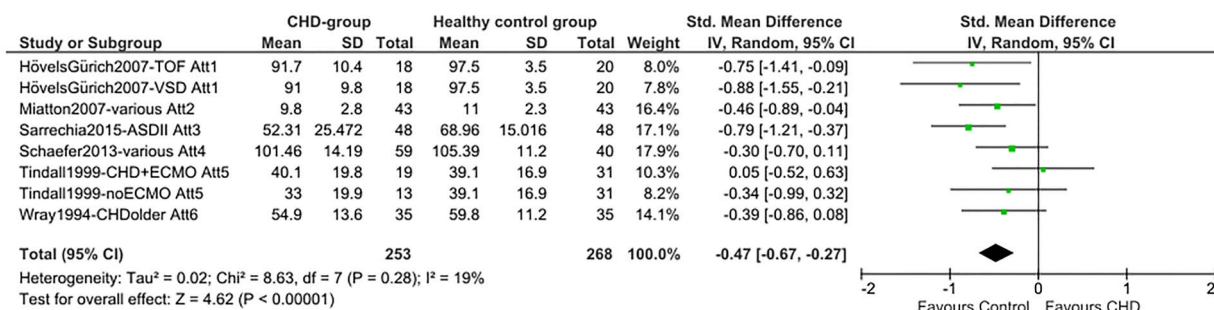
**Strengths**

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**

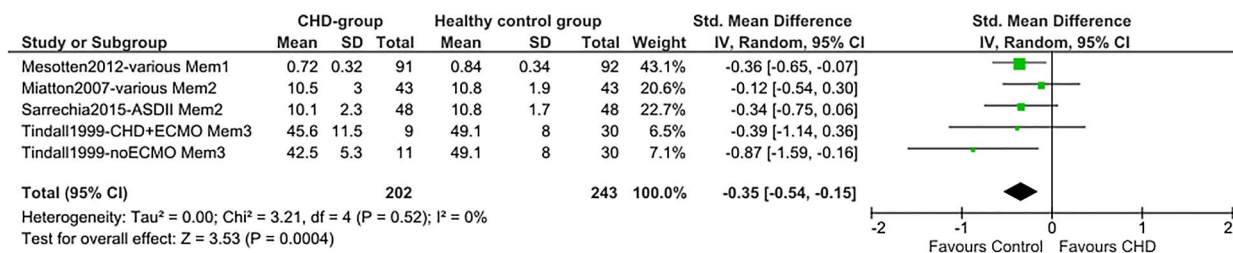
The meta-analysis has a few inherent limitations though. First, attention and EFs involve several functions, which

may interact.<sup>9</sup> 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

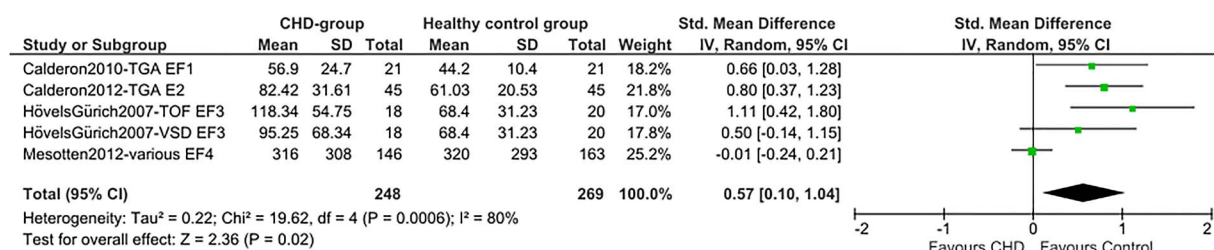


**Figure 3** Forest plot of attention (Att; alertness) non-reaction time (author/year/journal—CHD-type/attention measure). Att1, mean accuracy (%) Attention Network Test; Att2, visual attention (NEuroPSYchological Assessment, NEPSY); Att3, auditory attention (NEPSY); Att4, processing speed (Wechsler 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.<sup>41</sup> 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,<sup>18</sup> when academic difficulties can appear. Adolescence is particularly essential for the late maturation of the prefrontal cortex and EFs.<sup>37</sup> 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.<sup>2</sup>

## 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.

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license,

which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

## REFERENCES

- 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.
- 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.
- McCordle 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.
- 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.
- 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.
- 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.
- Forbess JM, Visconti KJ, Bellinger DC, *et al*. Neurodevelopmental outcomes in children after the fontan operation. *Circulation* 2001;104(12 Suppl 1):1127–132.
- Chapman SB. Neurocognitive stall: a paradox in long term recovery from pediatric brain injury. *Brain Inj Prof* 2007;3:10–13.
- Diamond A. Executive functions. *Annu Rev Psychol* 2013;64:135–68.
- 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.
- Alloway TP, Alloway RG. Investigating the predictive roles of working memory and IQ in academic attainment. *J Exp Child Psychol* 2010;106:20–9.
- 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.
- 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.
- 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):1109–14.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- Van Der Rijken R, Hulstijn-Dirkmaat G, Kraaijaat F, *et al*. Open-heart surgery at school age does not affect neurocognitive functioning. *Eur Heart J* 2008;29:2681–8.
- 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.
- 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.
- 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.
- Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd edn. Hillsdale, NJ: L. Erlbaum Associates, 1988.
- Moher D, Liberati A, Tetzlaff J, *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.
- 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.
- Prentice DA, Miller D. When small effects are impressive. *Psychol Bull* 1992;112:160–4.
- 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.
- 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.
- Posner MI. Measuring alertness. *Ann N Y Acad Sci* 2008;1129:193–9.
- 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.
- 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.
- 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.
- 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.