NeuroImage: Clinical 4 (2014) 526–530

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

Data on the safety of repeated MRI in healthy children

Scott K. Holland [a,b], Mekibib Altaye [a,c,*], Sara Robertson [a], Anna W. Byars [a,d], Elena Plante [a,e], Jerzy P. Szaflarski [a,f]

[a] Pediatric Neuroimaging Research Consortium, Cincinnati Children's Research Foundation, Cincinnati, OH, USA
[b] Department of Radiology, University of Cincinnati College of Medicine, Cincinnati, OH, USA
[c] Department of Pediatrics, Division of Biostatistics and Epidemiology, University of Cincinnati College of Medicine, Cincinnati, OH, USA
[d] Department of Pediatrics, Division of Neurology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA
[e] Department of Speech and Hearing Sciences and Disorders, University of Arizona, Tucson, AZ, USA
[f] Department of Neurology, University of Alabama at Birmingham, Birmingham, AL, USA

ARTICLE INFO

Article history:
Received 9 December 2013
Received in revised form 15 January 2014
Accepted 17 January 2014
Available online 9 February 2014

Keywords:
MRI
Safety healthy children repeated
IQ
BMI
OWLS
BRIEF
Longitudinal

ABSTRACT

Purpose: To address the question of the safety of MRI for research in normal, healthy children. We examined MRI, neurocognitive and biometric data collected in a group of healthy, normally developing children who have participated in a 10 year longitudinal fMRI study.
Materials and methods: Thirty-one healthy children ranging in age from 5 to 7 years were enrolled between 2000 and 2002 and were tested yearly as part of a longitudinal study of normal language development. Twenty-eight of these children have completed multiple neuroimaging, neurocognitive and biometric exams. These children ranged in age from 5 to 18 years during the course of the study and were exposed to up to 10 annual MRI scans. Linear regression of the IQ (WISC-III) (Wechsler, 1991), executive function (BRIEF) (Gioia et al., 2002), and language (OWLS) (Carrow-Woolfolk, 1995) measures was performed against the number of years of exposure to MRI in the study. Body mass index (BMI) (Ogden et al., 2006) was also examined as a function of years and compared with normative values.
Results: The WISC-III Full Scale (FSIQ) in our longitudinal cohort was higher than the average at baseline. There was no significant change over time in mean FSIQ $p = 0.80$, OWLS $p = 0.16$, or BRIEF $p = 0.67$. Similarly, over 10 years there were no significant changes in the Coding subtest of WISC III and height and body mass index did not deviate from norms (50th percentile).
Conclusions: Examination of neurocognitive and biometric data from a decade-long, longitudinal fMRI study of normal language development in this small, longitudinal sample of healthy children in the age range of 5 to 18 years, who received up to 10 MRI scans, provides scientific evidence to support the belief that MRI poses minimal risk for use in research with healthy children.

1. Introduction

Examining the current literature on magnetic resonance imaging (MRI) for keywords relating to biological effects of MRI turns up primarily articles relating to the operational hazards associated with MRI (Gangarosa et al., 1987) and protecting patients and radiology personnel from risks associated with ferromagnetic objects becoming projectiles in close proximity to MRI magnets (Gallauresi and Woods, 2008; Shellock and Crues, 2004). There is no question that the benefits outweigh the risks of MRI for clinical diagnostic purposes. However, for research in vulnerable populations such as children and minors who are dependent on parents or guardians for consent to participate in research protocols, it is the responsibility of the research community to insure that the risk is minimal if there is no direct benefit to the participant. Most Institutional Review Boards (IRBs) classify MRI as a minimal risk procedure and therefore the risk/benefit ratio works in favor of approval for many research protocols involving children as human subjects. According to the NIH-sanctioned Collaborative Institutional Training Initiative (CITI) program (Braunschweiger and Goodman, 2007), minimal risk means "The probability (of occurrence) and magnitude (seriousness) of harm or discomfort (e.g., psychological, social, legal, economic) associated with the research are not greater than those ordinarily encountered in daily life (of the average person in the general population) or during the performance of routine physical or psychological examinations or tests." Minimal risk, therefore, is used to define a threshold of anticipated harm or discomfort associated with the research that is low. This classification is based on a lack of evidence to the contrary.

* Corresponding author at: Division of Biostatistics and Epidemiology, Department of Pediatrics, Cincinnati Childrens's Hospital Medical Center.

http://dx.doi.org/10.1016/j.nicl.2014.01.013

Over the course of three decades of MRI use in humans, there have not been any acute or long-term deleterious biological effects attributed to MRI exposure, aside from the obvious physical injuries that occur because of ferromagnetic projectiles colliding with people on their path along the flux lines of the superconducting magnets that power the MRI machines. Still, there is a dearth of literature describing systematic studies of MRI biological effects using scientific or epidemiological methods to produce evidence upon which to base a conclusion or even make an estimate of how large such effects could be. This study aims to provide scientific evidence to test the hypothesis that MRI produces measureable adverse effects on cognitive and physical development in children who are exposed to repeated MRI scans between the ages of 5 and 18 years. While there is no existing data to support this hypothesis that we are aware of and we do not expect our data to allow us to validate this claim, we are forced to test this positive hypothesis because it is not possible to reject the null hypothesis with any degree of certainty based on one, small scale study such as the one reported here. Conversely, we expect to be able to reject the hypothesis that adverse effects will be found in our sample and to use our data to set an upper bound on the magnitude of such effects if they exist. Further we expect our result to provide justification for the classification of research using MRI as minimal risk.

Much of the research involving the use of MRI in pediatric populations is aimed at understanding development and disorders of cognitive functions such as language and attention. Functional MRI of the developing brain exposes the brain and the entire human body to a static magnetic field, gradient magnetic field changes, and radio frequency (RF) electromagnetic fields (Haake et al., 1999). FDA guidelines and manufacturer limits prevent acute biological effects from RF heating and peripheral and vestibular nerve stimulation (Zaremba, 2003, 2008). While acute effects of MRI below these limits have not been reported, researchers must question whether MRI exposure of the cerebral cortex, brain stem, thalamus, and neuroendocrine glands that moderate growth and development could possibly produce long-term effects, even though mechanisms underlying such effects have not been described (Chou, 2007; Dini and Abbro, 2005; Robertson et al., 2009; Weiss et al., 1992). Continued vigilance for such effects is incumbent upon us as medical researchers. While we aim to improve child health through scientific investigations, harm to human research subjects and particularly to a vulnerable population of children, is not an acceptable cost for such scientific advances.

Here we examine the question of the safety of MRI from the point of view of its impact on physical and cognitive growth and development in healthy children. We address this question using MRI, cognitive, and biometric data that we have collected in a group of healthy, normally-developing children who have participated in a longitudinal study of language development using fMRI for the past 10 years (Szaflarski et al., 2006). Admittedly our data set is limited and the lack of significant MRI related effects on cognitive and biometric measures does not preclude discovery of biological effects from repeated MRI in the future. However, the data permit us to establish an upper limit for how large an effect could be and still avoid detection using the gross biometric and cognitive assessments that we have obtained in this longitudinal sample of healthy children. Controlling for relevant growth variables we are also able to estimate the sample size needed to detect measureable effects at specified levels. A verifiable positive finding would have implications for research in children and could allow us to estimate the scale of the potential impact that MRI exposure might have on the selected biomarkers. Results of this study establish a baseline for MRI bioeffects and gauge the necessity and scale for prospective studies of MRI bioeffects in the future.

2. Materials and methods

A longitudinal cohort of 31 healthy children was enrolled between 2000 and 2002 at age 5 (n = 9), 6 (n = 7) or 7 (n = 15) years. Twenty-eight (13 girls, 15 boys) of these children have completed multiple years of annual neuroimaging, biometric, neurological exams, and cognitive testing as listed in Table 1.

Table 1
Number of scan per subject.

Subject ID	Age	# of scan
05F003	5	7
05F004	5	9
05F008	5	6
05M002	5	7
05M003	5	10
05M005	5	7
05M008	5	7
05M019	5	7
05M024	5	6
06F001	6	5
06F011	6	8
06F018	6	6
06M001	6	10
06M005	6	10
06M012	6	8
07F002	7	10
07F007	7	9
07F009	7	9
07F010	7	9
07F015	7	8
07F021	7	9
07F024	7	8
07M001	7	10
07M004	7	8
07M005	7	8
07M006	7	10
07M009	7	10
07M012	7	7

Biometric data reported here include height, weight, and Body Mass Index (BMI) (Ogden et al., 2006). For each visit, MRI scanning was completed, if possible, given the child's status (e.g. orthodontic braces, and medical status). Cognitive, developmental, and biological measures were recorded according to the schedule in Table 2 for the longitudinal cohort. IRB approval was obtained for the study and informed consent was obtained from parents as well as assent of minor participants.

We examined the longitudinal change in the Wechsler Intelligence Scale for Children, Third Edition (Wechsler, 1991) (WISC-III) administered to children prior to the first MRI and after the 3rd and 5th scans. Data from years 1, 3, and 5 for the FSIQ from WISC-III are reported. In addition, the Coding subtest from the WISC-III was administered to all participants again in year 10 and is used to model the longitudinal trend across all scan years (1st, 3rd, 5th and 10th). We computed the linear regression of the Coding subtest scores for WISC-III, accounting for the repeated nature of the data. The resulting line for the test with

Table 2
List and administration time of relevant neuroimaging, cognitive and biometric measurements for the longitudinal cohort.

	Years									
Measurements	1	2	3	4	5	6	7	8	9	10
Neuroimaging: MRI	X	X	X	X	X	X	X	X	X	X
Cognitive:										
WISC-III/WPPSI-III (Wechsler, 1991)	X		X		X					
Coding	X		X		X					X
WASI										X
OWLS (Carrow-Woolfolk, 1995)	X		X		X					
Listening comprehension	X		X		X					
Oral expression	X		X		X					
Oral comprehension	X		X		X					
BRIEF (Gioia et al., 2000, 2002) — parent						X	X	X	X	X
BRI						X	X	X	X	X
GEC						X	X	X	X	X
MI						X	X	X	X	X
Weight	X	X	X	X	X	X	X	X	X	X
Height	X	X	X	X	X	X	X	X	X	X

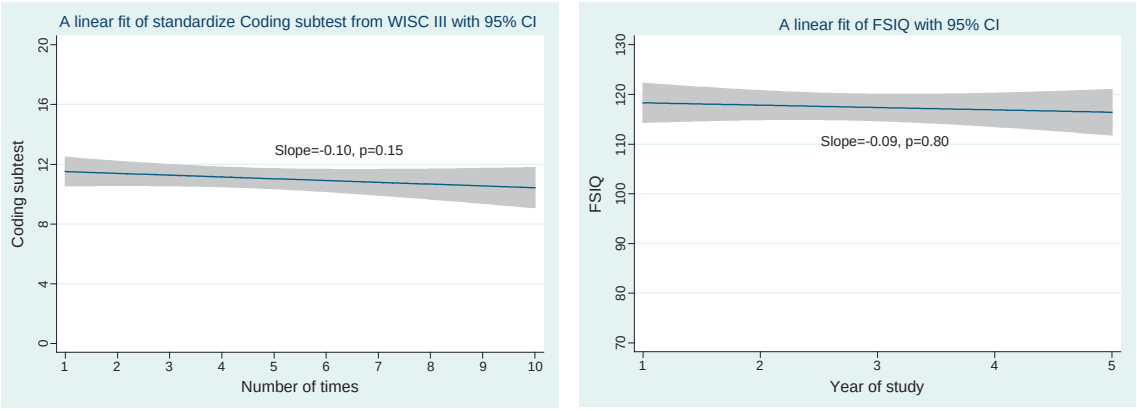

Fig. 1. Linear regression of 10-year longitudinal data from WISC III Coding subtest (left) and 5 year WISC III Full Scale IQ scores (FSIQ) (right) against the number of annual exposure with 95% CI.

the corresponding 95% confidence interval is shown in Fig. 1. We also fit a linear regression for the FSIQ obtained from WISC-III across the 1st, 3rd and 5th scan times as displayed in Fig. 1 (right).

In years 6–10, executive functioning was assessed annually by administering the parent form of the Behavior Rating Inventory of Executive Function (BRIEF) (Gioia et al., 2000). In this analysis we use the Global Executive Composite (GEC) score from the BRIEF as an overarching summary T-score with a mean of 50 and standard deviation of 10. As with the Wechsler scales above, we fit a regression model that accounts for the repeated measures nature of the data to examine the relationship between the number of MRI scans and these scores. We plotted the fitted line with the corresponding 95% confidence interval as shown in Fig. 2 (right). In addition, we examined the Oral and Written Language Scales (OWLS) (Carrow-Woolfolk, 1995) administered to children prior to the first MRI and after the third and fifth annual scans. These results are also shown graphically in Fig. 2 (left).

Finally, we also evaluated collected biometric data for weight, height, and Body Mass Index (BMI) in this cohort and compared it to the corresponding norms, using age- and sex-adjusted data from the National Center for Health Statistics (NCHS) of the Center for Disease Control and Prevention (CDC). We used the 5th, 50th and 95th percentiles for body mass index (BMI) to illustrate the corresponding norm for our longitudinal cohort as shown in Fig. 3.

3. Results

Based on the initial five years of data using only WISC-III FSIQ for the longitudinal cohort, the mean and standard deviation at baseline, year 3, and year 5 was 117.9(13.5), 118.2(11.1), and 115.8(10.8), respectively. The resulting slope from the linear fit was −0.09 with $p = 0.80$, a non-significant result as shown in Fig. 1 (right). Similar non-significant trends were observed for OWLS in years 1, 3, and 5 in mean Listening Comprehension scores (106.2 (18.6), 105.5(13.1), 110.4(15.28.1); $p = 0.33$), Oral Expression scores (110.6(10.9), 109.9(1), 116.1(14.2); $p = 0.435$), and Oral Comprehension scores (109.0(14.49), 108.2(11.9), 114.6(14.5); $p = 0.17$) The linear regression plot of the average OWLS composite data is given in Fig. 2 (left).

The mean and standard deviation of the Coding subtests obtained from the WISC-III at baseline and at year 10 was 11.4(3.08) and 10.4(3.5), with a p-value of 0.35. The plot and fit of the data across years 1, 3, 5 and 10 have a non-significant slope ($p = 0.15$) as illustrated in Fig. 1 (left). The mean BRIEF GEC scores in years 6–10 were (49.5(9.3), 45.6(7.5), 47.3(11.2), 47.0(10.0) and 46.1(9.3); $p = 0.67$) respectively; again the trend in the linear regression with the number of annual MRI scans does not reach significance (Fig. 2, right).

Similarly, the body mass index did not deviate from norms (50th percentile) and most of the measurements are within the 5th and 95th percentile of the CDC BMI chart over 10 years (Fig. 3).

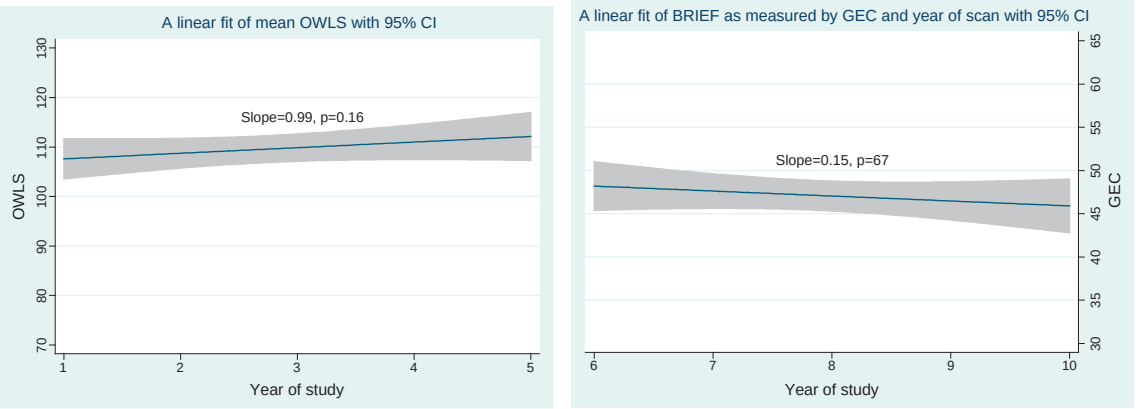

Fig. 2. Average OWLS Composite scores (left) with 95% CI for the longitudinal cohort and BRIEF score (right), against the number of annual MRI exposure with 95% CI for the longitudinal cohort.

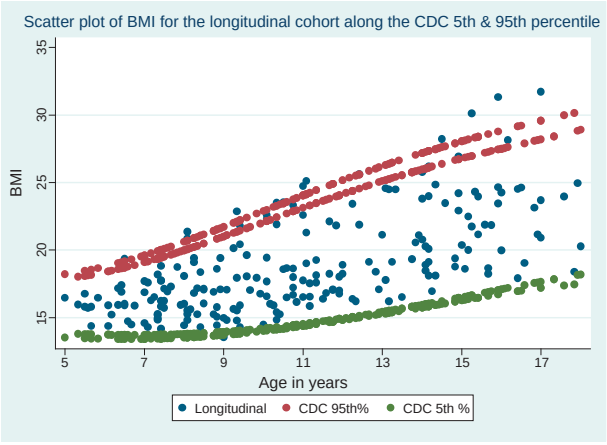

Fig. 3. Scatter plot of body mass index (BMI) against age in years for longitudinal cohort data (bottom-right).

Note that the elevated scores for the cognitive measures in our cohort at baseline render comparisons with the population norms for the tests irrelevant. For example, the mean and standard deviation WISC-III FSIQ at baseline was 117.9 ± 13.5. Comparing our cohort directly with norms (100 ± 15) might suggest that only higher scoring children participate in MRI brain imaging research studies, which is not a relevant point to this study. Consequently we focus primarily on analysis of the trends in biometric and cognitive scores over time, relative to normative trends.

4. Discussion

Adverse cognitive or biological effects from repeated MRI scans are not evident in the data from this longitudinal sample of children in the age range of 5 to 18 years using the gross cognitive and growth rate measures administered during the course of 10 years of exposure to annual MRI scans. The effect size estimated as the least square mean difference between the scores at the last and first time points is small (effect size for WISC FSIQ = 0.17, BRIEF = 0.31 and OWLS = 0.38) and without consistent positive or negative trends. This suggests that any changes due to repeated MRI scanning are likely to be very subtle and not clinically significant. Based on these effect sizes, estimated sample sizes of 280, 83, and 55 would be needed to detect significant positive or negative changes over time in FSIQ, BRIEF, and OWLS, respectively. These estimates are based on five-year average exposure of MRI scans and 80% power.

It is not possible to prove conclusively that deleterious effects do not occur with repeated MRI in children and the present data can only be properly interpreted as an upper limit on how safe repeated MRI can be for children in the specific age range of 5 to 18 years. In the present study, change over time for cognitive measures was less than the standard error for these measures. This magnitude of change is within the range that would be expected for retesting children on these measures, regardless of whether they had received repeated MRI. Likewise, the distribution of BMI is not distinguishable from normal trends.

We attempted to minimize practice effects on the WISC-III and WAIS-III by not repeating the tests every year. The tests were administered every other year in most cases and the WASI test was administered 5 years after the last exposure to WISC for most children. The "Flynn Effect" is also known to result in increasing IQ scores in populations, related to increases in fluid and crystallized intelligence over time (Flynn, 1994). If practice effects or Flynn Effect are present in our dataset, it would tend to inflate the cognitive test scores over time. Such an effect could be offset in our data by decreasing cognitive ability due to the repeated MRI exposures. There is no way to disambiguate these factors based on our retrospective study design and data we have collected and this is a limitation of the study. Other limitations of this retrospective study on the potential impact of repeated MRI exposure on physical and cognitive growth and development in healthy children include the small sample size, inconsistent cognitive testing due to the wide age range and a lack of cognitive measures specifically designed to be sensitive to longitudinal trends.

While each individual in the group may have a different trend for a specific measure collected at different time points, positive or negative variations in individual trends are expected. In most cases the variations in individual trends, upward or downward with time, fall within the standard deviations of the measures. To make the association between MRI exposure and neurobehavioral or biometric measures, we can only make statistical inferences from the group data. In this case we are able to estimate the significance of the trends relative to norms from the general population and generalize our findings to the larger population. At the group level none of these trends is statistically significant.

Despite the limitations described above, we are able to reject the first part of our initial hypothesis, that repeated exposure to MRI produces measureable adverse effects on neurocognitive development in children who are exposed to repeated MRI scan between the ages of 5 and 18 years. The biometric data in Fig. 3, although limited to BMI trends, also points to the rejection of the hypothesis that repeated exposure to MRI produces measureable adverse effects on physical development, though admittedly BMI is a very gross biometric measure and does not allow us to explore impact of MRI on specific areas of growth and development.

If direct evidence for an adverse interaction of magnetic fields or MRI with biological systems is identified, then researchers using MRI to study human development must pause to consider the implications. Until such a mechanism is discovered we can only examine the relationships between MRI exposure and biological and behavioral measures of development using an epidemiological approach. Recent discussion of the safety of MRI for research in healthy children (Holland et al., 2010; Jiao, 2010; Prato et al., 2010) motivates us to use this approach to examine data from our longitudinal cohort of pediatric subjects as they grow into adulthood.

Future studies should be conducted comparing participants who have had repeated MRI scans to a normative control group without exposure. Ideally, a prospective study from birth to adulthood would be conducted in a large cohort of participants in a longitudinal study with repeated exposures to MRI along with consistent cognitive assessments using instruments designed to be administered repeatedly without influence from practice effects. There are a number of such instruments available such as the ANAM (Kabat et al., 2001) and Cogstate (Falleti et al., 2006). Generally these tests are designed to test for subtle decline in cognitive ability due to brain injury or neurodegenerative diseases in adults. However, there are few such instruments with norms for children. By our own estimates for effect sizes described above, cohorts of 100 to 300 children would be needed to detect significant changes over time using the gross measures we had available for this retrospective study. Using modern computer-based cognitive assessments designed to avoid practice effects in repeated administrations of the tests should improve sensitivity and might reduce sample size requirements. However, this type of study will take decades to complete and there are many disincentives to perform it, including cost, perceived risk to subjects, and reluctance in the medical community and corporate interests to turn up any adverse effects from MRI in children. Given the lack of evidence for acute adverse effects from MRI scanning during its long history and widespread clinical use, it appears unlikely that such effects exist. The benefit of MRI for clinical diagnosis is unequivocal and the medical-legal system in the United States weighs heavily in favor of using MRI in children to avoid missing a diagnosis or subjecting children to more invasive or risky procedures such as biopsies or X-rays. Consequently it is unlikely that

the ideal prospective, longitudinal MRI bioeffects study will ever be funded or conducted in children. Meanwhile the data reported here provide some level of assurance that up to 10 MRI scans do not produce observable deleterious bioeffects in children and the results can be used to define a framework for the design of a larger scale study.

5. Conclusion

Examination of cognitive and biometric data from a decade-long longitudinal fMRI study of normal language development in this small, longitudinal sample of healthy children in the age range of 5 to 18 years, who received up to 10 MRI scans, provides evidence to support the belief that MRI poses minimal (if any) risk for use in research with healthy children.

References

Braunschweiger, P., Goodman, K.W., 2007. The CITI program: an international online resource for education in human subjects protection and the responsible conduct of research. Acad. Med. 82 (9), 861–864 (Sep).

Carrow-Woolfolk, E., 1995. Oral and Written Language Scales. American Guidance Service, Circle Pines, MN.

Chou, C.K., 2007. Thirty-five years in bioelectromagnetics research. Bioelectromagnetics 28 (1), 3–15 (Jan).

Dini, L., Abbro, L., 2005. Bioeffects of moderate-intensity static magnetic fields on cell cultures. Micron 36 (3), 195–217.

Falleti, M.G., Maruff, P., Collie, A., Darby, D.G., 2006. Practice effects associated with the repeated assessment of cognitive function using the CogState battery at 10-minute, one week and one month test-retest intervals. J. Clin. Exp. Neuropsychol. 28 (7), 1095–1112 (Oct).

Flynn, J.R., 1994. IQ gains over time. In: Sternberg, R.J. (Ed.), Encyclopedia of Human Intelligence. Macmillan, New York, pp. 617–623.

Gallauresi, B.A., Woods, T., 2008. Danger: "sandbag" in the MRI room. Nursing 38 (12), 60 (Dec).

Gangarosa, R.E., Minnis, J.E., Nobbe, J., Praschan, D., Genberg, R.W., 1987. Operational safety issues in MRI. Magn. Reson. Imaging 5 (4), 287–292.

Gioia, G.A., Isquith, P.K., Guy, S.C., Kenworthy, L., 2000. Behavior rating inventory of executive function. Child Neuropsychol. 6 (3), 235–238 (Sep).

Gioia, G.A., Isquith, P.K., Retzlaff, P.D., Espy, K.A., 2002. Confirmatory factor analysis of the Behavior Rating Inventory of Executive Function (BRIEF) in a clinical sample. Child Neuropsychol. 8 (4), 249–257 (Dec).

Haake, E.M., Brown, R.W., Thompson, M.R., Venkatesan, R., 1999. Magnetic Resonance Imaging: Physical Principles and Sequence Design. John Wiley & Sons, Inc., New York.

Holland, S.K., Byars, A.W., Plante, E., Szaflarski, J.P., Dietrich, K., Altaye, M., 2010. Studies support probable long-term safety of MRI. Science 329 (5991), 512–513.

Jiao, L., 2010. Fear of MRI scans trips up brain researchers. Science, vol. 327. AAAS, p. 931.

Kabat, M.H., Kane, R.L., Jefferson, A.L., DiPino, R.K., 2001. Construct validity of selected Automated Neuropsychological Assessment Metrics (ANAM) battery measures. Clin. Neuropsychol. 15 (4), 498–507 (Dec).

Ogden, C.L., Carroll, M.D., Curtin, L.R., McDowell, M.A., Tabak, C.J., Flegal, K.M., 2006. Prevalence of overweight and obesity in the United States, 1999–2004. JAMA 295 (13), 1549–1555.

Prato, S.F., Thomas, A.W., Legros, A., et al., 2010. MRI safety not scientifically proven. Science 328, 568.

Robertson, J.A., Theberge, J., Weller, J., Drost, D.J., Prato, F.S., Thomas, A.W., 2009. Low-frequency pulsed electromagnetic field exposure can alter neuroprocessing in humans. J. R. Soc. Interface 7 (44), 467–473.

Shellock, F.G., Crues, J.V., 2004. MR procedures: biologic effects, safety, and patient care. Radiology 232 (3), 635–652 (Sep).

Szaflarski, J.P., Schmithorst, V.J., Altaye, M., et al., 2006. A longitudinal functional magnetic resonance imaging study of language development in children 5 to 11 years old. Ann. Neurol. 59 (5), 796–807 (Feb 23).

Wechsler, D., 1991. Wechsler Intelligence Scale for Children, Third edition. The Psychological Corporation, San Antonio, TX.

Weiss, J., Herrick, R.C., Taber, K.H., Contant, C., Plishker, G.A., 1992. Bio-effects of high magnetic fields: a study using a simple animal model. Magn. Reson. Imaging 10 (4), 689–694.

Zaremba, L.A., 2003. Criteria for Significant Risk Investigations of Magnetic Resonance Diagnostic Devices. U.S. Department of Health and Human Services; Food and Drug Administration Guidelines 793 (July 14).

Zaremba, L.A., 2008. In: CfDaR, Health, FaD, Administration (Eds.), FDA Guidelines for Magnetic Resonance Equipment Safety.