OPEN

PRS GLOBAL OPEN

VIEWPOINT
Craniofacial/Pediatric

IQ Assessment in Craniofacial Neurocognitive Studies: Interpreting Results Relative to Evidence-based Systematic Analysis

Aaron S. Long, BS*; Sacha C. Hauc, MPH*; Adam H. Junn, BS*; Caitlin Meyer, MLIS†; Linda Mayes, MD‡; John A. Persing, MD*; Michael Alperovich MD, MSc, FACS*

Recent studies have evaluated the relationship of timing of surgery, type of surgery, and genetic factors with long-term cognitive function in craniosynostosis.[1,2] Intelligence quotient (IQ), which is standardized to a mean score of 100 and standard deviation (SD) of 15, evaluates overall cognitive ability and is a commonly used outcome measure in craniosynostosis research. Concerns have been raised about the generalizability of studies where the reported mean IQ was greater than 100.[3,4] Central to this issue is whether high average scores reflect selective recruitment at specific academic centers or occur generally across academic institutions and disciplines.

In conjunction with an experienced research librarian and with input from a neuropsychologist, we systematically identified all North American studies conducted at an academic center that used the Wechsler Abbreviated Scale of Intelligence to assess IQ. (**See table, Supplemental Digital Content 1**, which describes the search and screening strategy, **http://links.lww.com/PRSGO/C79**.) To allow for investigation across heterogeneous study populations, inclusion was limited to studies including a healthy control group in addition to the condition being evaluated (eg, autism). Details of recruitment, study group characteristics, and mean of full-scale IQ (FSIQ), verbal IQ (VIQ), and performance IQ (PIQ) were recorded.

In all, 31 studies met inclusion criteria. Twelve studies were in a pediatric population, 16 were in an adult population and three had both adult and pediatric patients. (**See table, Supplemental Digital Content 2**, which shows all recorded variables, **http://links.lww.com/PRSGO/C80**.) Average FSIQ values reported for healthy control subjects had a median of 112.0 (interquartile range [IQR]: 108.5–15.5), and the median scores for average VIQ and PIQ were 116.2 (IQR: 112.9–119.1) and 113.5 (IQR: 109.7–115.3), respectively (Table 1; Fig. 1).

*From the *Department of Surgery, Section of Plastic Surgery, Yale School of Medicine, New Haven, Conn.; †Harvey Cushing/John Hay Whitney Medical Library, Yale University, New Haven, Conn.; and ‡Yale Child Study Center, Yale School of Medicine, New Haven, Conn.*

Received for publication May 2, 2022; accepted May 11, 2022.

Plast Reconstr Surg Glob Open 2022;10:e4405; doi: 10.1097/GOX.0000000000004405; Published online 20 July 2022.

Table 1. Mean Full-scale IQ, Verbal IQ, and Performance IQ Values of Healthy Control Groups

IQ Values (N = 31)	Median (IQR)
Full-scale IQ	
Pediatric (N = 10)	112.3 (104.9–116.2)
Adult (N = 15)	112.1 (108.97–115.3)
Total (N = 28)	112.0 (108.5–115.5)
Verbal IQ	
Pediatric (N = 5)	119.1 (116.2–123.6)
Adult (N = 3)	114.5 (108.2–116.1)
Total (N = 9)	116.2 (112.9–119.1)
Performance IQ	
Pediatric (N = 5)	114.3 (113.1–115.3)
Adult (N = 3)	113.5 (106.3–115.9)
Total (N = 9)	113.5 (109.7–115.3)

Pediatric: ≤18 years old; Adult: > 18 years old.

Mean IQ scores greater than 100 were common across North American academic-recruitment studies regardless of discipline (eg, psychiatry, neurology, etc.). The utility of these studies lies in the relative differences, or "delta," of IQ found in between-group comparisons. For instance, an average FSIQ score of 109.4 for patients with alcohol dependence would seem high compared to the normative mean but is lower than the 118.7 observed among controls in the same study.

High IQ values observed in research studies may reflect a natural phenomenon of increasing IQ in the general population termed the Flynn effect.[5] Selection bias likely also plays a role as subjects choosing to participate in research may not reflect the general population.

Future neurocognitive studies within craniofacial surgery should be designed with at least two cohorts for comparison as there is limited utility in comparing IQs to a national mean.

The normative IQ distribution with a mean of 100 is useful for educational placement or evaluation for disability; it is not as well-suited for university-based outcome studies. The results of this systematic investigation would caution against interpreting high IQ values as indicative of errors in methodology or recruitment within an individual study.

Michael Alperovich, MD, MSc, FACS
Department of Surgery
Section of Plastic and Reconstructive Surgery
Yale School of Medicine
330 Cedar Street, Boardman Building
New Haven, CT 06510
E-mail: Michael.alperovich@yale.edu

Related Digital Media are available in the full-text version of the article on www.PRSGlobalOpen.com.

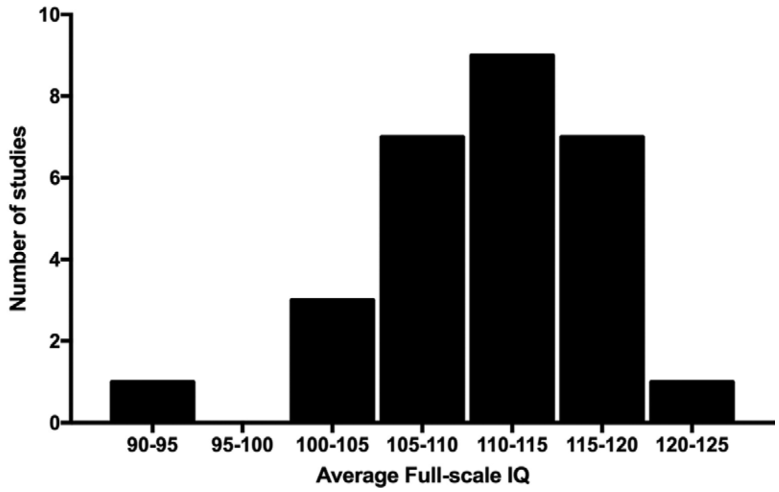

Fig. 1. Mean full-scale IQ values of healthy control groups in each study.

DISCLOSURE

The authors have no financial interest to declare in relation to the content of this article. Presented as a display poster at the American Cleft Palate-Craniofacial Association Annual Meeting, April 2022 (Fort Worth, TX).

REFERENCES

1. Alperovich M, Runyan CM, Gabrick KS, et al. Long-term neurocognitive outcomes of spring-assisted surgery versus cranial vault remodeling for sagittal synostosis. *Plast Reconstr Surg.* 2021;147:661–671.
2. Wu RT, Timberlake AT, Abraham PF, et al. SMAD6 genotype predicts neurodevelopment in nonsyndromic craniosynostosis. *Plast Reconstr Surg.* 2020;145:117e–125e.
3. Purnell CA, Kapp-Simon KA. Discussion: long-term neurocognitive outcomes of spring-assisted surgery versus cranial vault remodeling for sagittal synostosis. *Plast Reconstr Surg.* 2021;147:674–675.
4. Rogers GF. Discussion: long-term neurocognitive outcomes of spring-assisted surgery versus cranial vault remodeling for sagittal synostosis. *Plast Reconstr Surg.* 2021;147:672–673.
5. Trahan LH, Stuebing KK, Fletcher JM, et al. The Flynn effect: a meta-analysis. *Psychol Bull.* 2014;140:1332–1360.