

Influence of Surgeon Experience, Hospital Volume, and Specialty Designation on Outcomes in Pediatric Surgery

A Systematic Review

Jarod P. McAteer, MD; Cabrini A. LaRiviere, MD, MPH; George T. Drugas, MD; Fizan Abdullah, MD, PhD; Keith T. Oldham, MD; Adam B. Goldin, MD, MPH

Importance: Analyses of volume-outcome relationships in adult surgery have found that hospital and physician characteristics affect patient outcomes, such as length of stay, hospital charges, complications, and mortality. Similar investigations in children's surgical specialties are fewer in number, and their conclusions are less clear.

Objective: To review the evidence regarding surgeon or hospital experience and their influence on outcomes in children's surgery.

Evidence Review: A MEDLINE and EMBASE search was conducted for English-language studies published from January 1, 1980, through April 13, 2012. Titles and abstracts were screened in a standardized manner by 2 reviewers. Studies selected for inclusion had to use a measure of hospital or surgeon experience as a predictor variable and had to report postoperative outcomes as dependent response variables. Included studies were reviewed with regard to methodologic quality, and study results were extracted.

Findings: Sixty-three studies were reviewed. Significant heterogeneity was detected in exposure definitions, outcome measures, and risk adjustment, with the greatest heterogeneity seen in appendectomy studies. Various ex-

posure levels were examined: hospital level in 48 (68%) studies, surgeon level in 11 (17%), and both in 9 (14%). Nineteen percent of studies did not adjust for confounding, and 57% did not adjust for sample clustering. The most consistent methods and reproducible results were seen in the pediatric cardiac surgical literature. Forty-nine studies (78%) showed positive correlation between experience and most primary outcomes, but differences in outcomes and exposure definitions made comparisons between studies difficult. In general, hospital-level factors tended to correlate with outcomes for high-complexity procedures, whereas surgeon-level factors tended to correlate with outcomes for more common procedures.

Conclusions and Relevance: Data on experience-related outcomes in children's surgery are limited in number and vary widely in methodologic quality. Future studies should seek both to standardize definitions, making results more applicable, and to differentiate procedures affected by surgeon experience from those more affected by hospital resources and system-level variables.

JAMA Pediatr. 2013;167(5):468-475.

Published online March 25, 2013.

doi:10.1001/jamapediatrics.2013.25

Author Affiliations: Division of Pediatric General and Thoracic Surgery, Seattle Children's Hospital, Seattle, Washington (Drs McAteer, Drugas, and Goldin); Department of Surgery, University of Washington School of Medicine, Seattle (Drs McAteer, Drugas, and Goldin); Department of Surgery, Louisiana State University, New Orleans (Dr LaRiviere); Department of Surgery, The Johns Hopkins University, Baltimore, Maryland (Dr Abdullah); and Division of Pediatric Surgery, Children's Hospital of Wisconsin, Milwaukee (Dr Oldham).

HOSPITAL AND SURGEON characteristics (eg, operative volume, institution designation, and fellowship training) have often been implicated in influencing outcomes. These measures are generally thought to serve as proxy measures of provider experience and resources. Numerous investigations have examined the outcomes associated with surgical experience, with most studies focusing on operative volume. Although the conclusions of these studies do not necessarily establish a causal relationship, several have established a strong enough association to guide practice and shape policy.¹⁻⁵ Studies in adults have been

numerous and generally consistent in methodology, but the quality and quantity of similar data in children are less consistent.

Although reviews of the volume-outcome relationship in adults have shown some methodologic shortcomings, the

*For editorial comment
see page 485*

larger numbers of studies and of investigators engaged in adult research have made these data fairly reliable.⁶ Similar data in the children's surgical literature are sparse, despite the need for such data to help guide practice. Children's surgical practices, however, unlike those for adults, gener-

ally do not subspecialize in the performance of rare, technically demanding procedures. Moreover, the total numbers of rare pediatric procedures (eg, esophageal atresia repair and portoenterostomy) are far fewer than similar high-risk procedures in adults (eg, esophagectomy and pancreaticoduodenectomy). The relationship between surgical frequency and complexity as they affect outcomes in children must be understood better to provide the best care to these patients. Although guidelines have been developed by expert opinion highlighting the importance of involving specialists in children's surgical care, more data are needed to define the specific situations for which referral is indicated.⁷

To date, to our knowledge, no comprehensive review has examined the data on the correlation between hospital and surgeon characteristics and outcomes in children's surgery. To examine the association between such characteristics and outcomes in pediatric surgical patients, we reviewed observational studies in children assessing the relationship between prespecified measures of surgical experience and clinical outcomes.

METHODS

STUDY SELECTION AND DATA SOURCES

We systematically reviewed English-language studies published since 1980 that focused on pediatric patients undergoing operative procedures. Included studies were required to have a measure of hospital or surgeon experience (operative volume, hospital designation, or surgeon subspecialty) as a predictor variable, and any clinical outcome (eg, mortality, complications, length of stay, and readmission) as a response variable. Studies that evaluated only patient characteristics at presentation rather than outcomes of care were excluded.

Studies were identified by searching the MEDLINE and EMBASE databases, with the most recent search performed on April 13, 2012. Our search process included the terms *volume*, *regionalization*, *designation*, *specialty*, *pediatric*, *children*, *surgery*, *outcomes*, *outcome assessment*, and *volume-outcome*. Bibliographies were reviewed, and experts were consulted about missed studies. Eligibility assessment was performed independently in a standardized manner by 2 of us (J.P.M. and A.B.G.) screening titles and abstracts. Articles were selected according to the aforementioned eligibility criteria, and disagreements were resolved through consensus.

DATA EXTRACTION AND METHODOLOGIC QUALITY

Information was extracted from each study on (1) procedure of interest, (2) exposure level (hospital, surgeon, or both), (3) exposure definition, (4) primary outcome measure(s), (5) association between exposure and primary outcome(s), (6) database used (administrative, clinical, or none), (7) confounding adjustment, and (8) sample sizes. As in other systematic reviews organizing results by procedure, we elected to look at each separate procedure of interest in a single article as a stand-alone study.⁶

Validity within studies was evaluated by examining whether studies adjusted results for confounding. The adjustment method was recorded (eg, logistic regression), as well as whether authors reported model diagnostics (eg, goodness-of-fit tests). Authors' approach to risk adjustment was also recorded if the adjustment method accounted for disease severity or patient

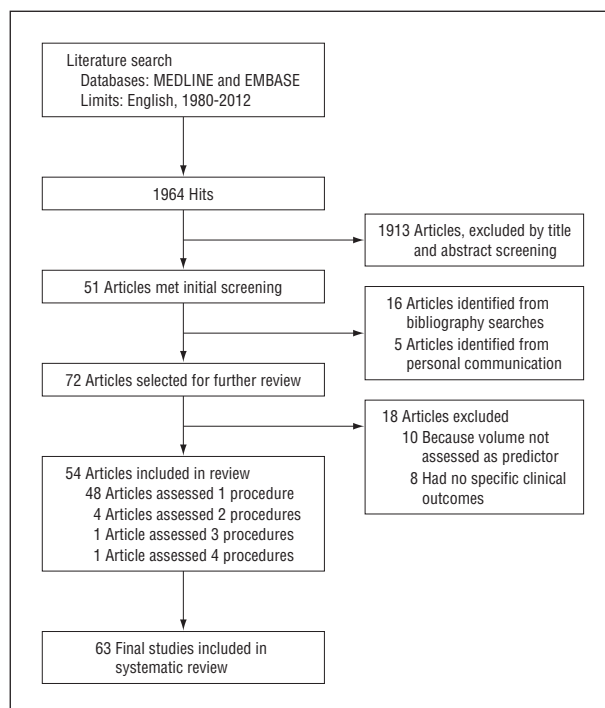


Figure. Study selection flowchart.

complexity to control for differences in case mix across hospitals and surgeons. Adjustment for demographic factors alone was not considered risk adjustment. Data were also gathered regarding whether authors adjusted for clustering and non-independence of sampling. If the study found a statistically significant advantage ($P < .05$) in the primary outcomes for more "experienced" institutions or providers, results were considered positive. If only some outcomes were improved or if positive results were found only in certain subgroups, results were considered mixed. If no association was seen, results were considered negative. A uniform summary measure could not be applied due to the variation in outcome measures.

RESULTS

STUDY SELECTION AND METHODOLOGIC QUALITY ASSESSMENT

Our systematic review generated 1964 articles, of which 51 were selected after initial screening. After addition of 21 articles identified by bibliography review and personal communication, 72 articles were selected for further review. Of these, 10 articles that did not examine a measure of hospital or surgeon experience as a predictor and 8 articles that did not explicitly examine clinical outcomes were excluded. Of the remaining 54 articles, 6 examined more than 1 procedure, generating a total of 63 individual studies (Figure).

Studies were quite heterogeneous in methodology. We found variability in the exposure level analyzed, exposure definition, outcome measures, and, for volume studies, categorization of volume variables (eTable; <http://www.jamapeds.com>). Sample sizes were generally reasonable, with 62% (39 studies) using study populations larger than 1000 patients (Table 1). Only 2 studies^{8,9} evaluated sample sizes smaller than 100, and both

Table 1. Summary of Study Attributes

Study Attribute	63 Studies in Review, No. (%)
Sample size >1000	39 (62)
No. of hospitals or physicians in each study	
<25 hospitals and/or <50 surgeons	12 (19)
≥25 hospitals and/or ≥50 surgeons	40 (63)
Data not published	11 (17)
Exposure level examined	
Surgeon only	11 (17)
Hospital only	43 (68)
Surgeon and hospital	9 (14)
Exposure definitions	
Hospital or surgeon	
Annual case volume	36 (57)
Specialty designation	11 (17)
Specialty designation and caseload	5 (8)
Case volume	
Total case volume for study period	6 (10)
Highest volume year	1 (2)
Historical total No. of procedures performed	1 (2)
Weekly or monthly case volume	1 (2)
Total admissions per year, condition specific	1 (2)
Total pediatric admissions per year	1 (2)
Primary outcome	
Mortality vs survival	35 (56)
Outcome combinations not including mortality	28 (44)
Multivariate adjustment method	
Logistic, linear, and/or Cox proportional-hazard regression	46 (73)
Stratified analysis	3 (5)
Observed-to-expected mortality ratios	2 (3)
No adjustment (only crude associations reported)	12 (19)
Multivariate studies with adjustment for disease severity risk (n = 51)	45 (88)
Adjustment for clustering at hospital or surgeon level	27 (43)
Type of database used	
Administrative	37 (59)
Clinical	16 (25)
Not a database study	10 (16)

reviewed rare conditions (congenital diaphragmatic hernia [CDH] and biliary atresia). Most studies (68% or 43 studies) focused on hospital-level characteristics. Studies focusing on surgeon-level characteristics tended to investigate more common, less resource-intensive procedures, such as appendectomy and pyloromyotomy.¹⁰⁻¹⁶ Outcome measures varied according to diagnosis or procedure of interest, with most studies of high-acuity conditions (eg, congenital heart disease) considering primarily mortality and most studies of lower-acuity conditions considering combinations of other outcomes (eg, length of stay, readmission, charges, and complications).

Of the studies reviewed, 57% (36 studies) used a single specific exposure definition (annual hospital or surgeon operative volume for the procedure of interest), and the rest used a variety of definitions; 72% (31 of 43) of hospital-level studies used the annual volume definition, compared with only 45% (5 of 11) of surgeon-level studies. We also noted differences in definitions between specialties and procedures. Whereas 93% (13 of

14) of included studies on congenital heart disease used annual hospital volume as their volume definition, only 33% (3 of 9) of appendicitis studies used annual surgeon or hospital volume as a definition, the remainder considering primarily surgeon subspecialty or hospital designation.

In addition to the variability in exposure definition in general, studies specifically using volume as an exposure also varied in their treatment of the volume variable itself. Of studies using volume measures, 17% (9 of 52) treated volume as a continuous variable, and the rest categorized the variable in a myriad of ways that showed little consistency, even across studies of similar procedures (eTable). Similarly, the methods for delineating volume categories were not consistent. Some authors defined cut points a priori, but many used data-driven categories based on case number equality within groups or percentiles.

The quality of risk-adjustment methods varied widely (eTable). Eighty-one percent (51 of 63) of the studies used some form of adjustment, but 12% (6 of 51) of those studies failed to adjust for disease severity (Table 1). Only 43% (27 of 63) of studies adjusted for sample clustering. Although multivariate regression models were common across studies, only 4 studies reported model diagnostics, and 3 of them were in the congenital heart literature.

Regarding the type of database used, most studies (n = 37) used administrative data. Of 10 nondatabase studies, 5 (50%) used surgeon subspecialty or hospital designation rather than a volume measure as the exposure of interest, and 5 (50%) were appendicitis studies.

OVERVIEW OF STUDY FINDINGS

Cardiac Surgery

Although methodologic quality and results varied across studies, the most consistent literature was seen in congenital heart surgery (**Table 2**). Fourteen studies were evaluated, all of which uniformly used a volume measure as exposure. Twelve studies found significant correlation between increasing hospital volume and positive outcomes. All studies evaluated in-hospital mortality except 1 study¹⁷ that evaluated postdischarge mortality and found no association. Nearly all studies used rigorous risk-adjustment models, and most adjusted for clustering. These studies generally focused purely on hospital volume, although 1 study considered both surgeon and hospital volume and found significantly improved outcomes for both.¹⁸ All studies used well-established databases.

Neurosurgery

Four neurosurgery studies^{19,30-32} were included, all using volume measures as exposures. Three studies focused on ventriculoperitoneal shunt placement, and the other evaluated craniotomies for tumor resection. Three studies used hospital or surgeon annual caseload as their volume definition, while 1 used the annual number of procedure-specific admissions. Two studies considered both sur-

Table 2. Summary of Study Results

Specialty or Condition of Interest ^a	Results, No. (%)			Database Studies, No. (%)	Adjustment, No. (%)			Mortality as Primary Outcome, No. (%)
	Positive	Mixed	Negative		Any Multivariate	Disease Severity	Clustering	
Congenital heart disease (n = 14) ¹⁷⁻²⁹	12 (86)	0	2 (14)	14 (100)	14 (100)	12 (86)	8 (57)	14 (100)
Neurosurgery (n = 4) ^{19,30-32}	3 (75)	0	1 (25)	4 (100)	4 (100)	4 (100)	3 (75)	3 (75)
Otolaryngology or craniofacial surgery (n = 3) ^{19,33,34}	1 (33)	1 (33)	1 (33)	3 (100)	3 (100)	3 (100)	3 (100)	2 (66)
Orthopedics (n = 2) ^{19,35}	0	1 (50)	1 (50)	2 (100)	2 (100)	2 (100)	1 (50)	2 (100)
Transplantation (n = 2) ^{36,37}	2 (100)	0	0	2 (100)	2 (100)	2 (100)	0	1 (50)
Urology (n = 3) ³⁸⁻⁴⁰	1 (33)	2 (67)	0	3 (100)	3 (100)	3 (100)	2 (66)	1 (33)
General surgery								
Appendicitis (n = 9) ^{10-12,16,41-45}	4 (44)	4 (44)	1 (11)	4 (44)	5 (56)	3 (33)	2 (22)	0
Pyloric stenosis (n = 7) ^{13-15,44,46-48}	2 (29)	5 (71)	0	6 (86)	4 (57)	3 (43)	3 (43)	0
Congenital diaphragmatic hernia (n = 4) ^{8,49-51}	3 (75)	0	1 (25)	4 (100)	2 (50)	2 (50)	1 (25)	4 (100)
Solid tumor (Wilms tumor or neuroblastoma) (n = 4) ^{52,53}	0	0	4 (100)	4 (100)	4 (100)	4 (100)	4 (100)	4 (100)
Biliary atresia (n = 4) ^{9,54-56}	3 (75)	0	1 (25)	0	3 (75)	3 (75)	0	4 (100)
Other general surgery (n = 7) ^{44,49,57-61}	1 (14)	4 (57)	2 (29)	7 (100)	5 (71)	4 (57)	0	1 (14)

^aSome of the referenced articles included studies on more than 1 procedure or population of interest.

geon and hospital volume, 1 surgeon volume alone, and 1 hospital volume alone. Risk-adjustment models were rigorous and the studies were based on large administrative samples. Three studies found a strong positive association between high surgeon and hospital volume and improved outcomes, but the study focusing on hospital volume alone found no association.

Otolaryngology or Craniofacial Surgery

Three studies^{19,33,34} were included in otolaryngology, 2 focusing on tracheotomy and 1 on cleft lip repair. Both tracheotomy articles considered only hospital volume, and the study on cleft lip repair evaluated both surgeon and hospital volume. All 3 studies measured volume by annual caseload and used robust risk-adjustment models. Results were inconsistent across studies. One tracheotomy study found positive associations across all outcomes, and the other found no positive associations. Although the cleft lip repair study found generally positive results, the associations for various outcomes differed depending on whether the focus was on hospital or surgeon volume. Only high surgeon volume was correlated with a decreased rate of complications.

Orthopedics

Two studies^{19,35} in orthopedics were identified, both examining spinal fusion. Both studies analyzed hospital volume and used administrative databases. Results were largely negative, because the only improvement seen at high-volume institutions was a lower reoperation rate.

Transplantation

Two abdominal transplantation studies^{36,37} were identified, 1 on liver and 1 on renal transplantation. Both studies considered hospital volume and used a United Network for Organ Sharing (UNOS) clinical database. Each investigation found a strong positive association between center volume and improved outcomes after adjustment.

Urology

Three urology studies met criteria. One study³⁸ examined hospital volume as an exposure in patients undergoing bladder exstrophy repair, and 2 studies^{39,40} examined both hospital and surgeon volume in patients undergoing ureteral reimplantation. Strong risk-adjustment models were used. Results were mixed for hospital volume, but a positive association between surgeon volume and outcomes was observed for ureteral reimplantation.

General Surgery

Studies in pediatric general surgery varied widely in quality and results. The most consistent results were found in studies on CDH and biliary atresia. Similar to cardiac studies, hospital volume was consistently defined as annual caseload and mortality was the primary outcome. Three CDH studies^{8,49,50} used a well-established Canadian database, and the fourth study⁵¹ used a national US administrative database. These studies generally reported significantly decreased mortality at higher-volume hospitals. One study⁴⁹ that did not report a significant association did not adjust for risk.

The biliary atresia studies describe the French and United Kingdom experience and used either national health registries or provider surveys as data sources. Both United Kingdom studies^{9,54} showed a strong association between increasing hospital volume and survival, with national outcomes improving after regionalization. The earlier French study revealed a similar association, but this association became statistically insignificant after implementation of national practice standards without formal regionalization.^{55,56}

Most studies in the general surgical literature focused on appendectomy and pyloromyotomy, and these studies displayed the most significant variability in methods. The appendectomy literature in particular presented few common threads between studies. Only 3 studies analyzed volume measures,^{11,41,42} and only 4 used databases.^{16,41,42,44} Studies that did not use volume mea-

tures used hospital designation and/or surgeon subspecialty as exposures. Four studies reported positive results, 4 mixed, and 1 negative. The most consistent result was a lower negative appendectomy rate for more experienced surgeons and hospitals. Results for other outcomes were inconsistent. Rigorous risk adjustment was uncommon. Although surgeon-level characteristics generally showed stronger associations with outcomes than hospital-level characteristics, the heterogeneity in definitions and lack of risk adjustment made this finding difficult to interpret.

Pyloromyotomy studies generally used better risk-adjustment methods than appendectomy studies and typically demonstrated a strong association between surgeon experience and improved outcomes. Specifically, postoperative complications (eg, duodenal perforation) were noted to be lower and length of stay shorter for high-volume and pediatric surgeons.^{13-15,46,47} Less association was reported for hospital-level characteristics, although 3 studies^{44,46,48} reported significantly decreased lengths of stay at high-volume centers and designated children's hospitals. In studies that looked at both surgeon volume and surgeon subspecialty, the effect of subspecialty on complications disappeared when surgeon volume was controlled for.¹⁵ Similar results illustrating the greater importance of surgeon volume vs subspecialty were noted in large-database studies of inguinal herniorrhaphy, cholecystectomy, and thyroid surgery.⁵⁷⁻⁶⁰

The literature for other procedures was sparse. Two studies^{44,61} of intussusception reduction were included and showed generally negative results, although neither study adjusted adequately for disease severity or reduction modality. Only 2 cancer articles^{52,53} were reviewed, both of which assessed survival for Wilms tumor and neuroblastoma and found no association with hospital volume. Rather, these studies found hospital characteristics, such as Children's Oncology Group membership and use of specific chemotherapy regimens, to be strongly associated with outcomes.

COMMENT

To our knowledge, this systematic review is the first critical analysis of the available evidence for the relationships between outcome and volume or experience in children's surgery. Our review included 63 studies evaluating 25 distinct procedures. There is significant heterogeneity in study methodology and definitions. The limitations of these data relate to both heterogeneous definitions that impair comparisons across studies as well as limited risk adjustment that impairs the internal validity of certain studies. Although this heterogeneity precludes a formal meta-analysis, some important findings can be extracted from the cumulative results of these articles once the limitations are understood.

First, volume-related research is limited by the types of outcomes that can be reliably identified. This research relies on large samples and is therefore limited to existing large databases, which have limitations.⁶² In noting the concerns regarding the validity of clinical and administrative databases, combined approaches using both

data types may prove useful, as long as standard definitions and methods are used.^{63,64} Identifying appropriate outcomes is a complex process and is a limiting factor in study design and in the validity of conclusions. In-hospital mortality is frequently used as a primary outcome measure in adult studies because it is easily and reliably identified, but this outcome is too rare for most pediatric procedures (eg, appendectomy). The single field in which mortality has been useful in children's surgery is congenital heart surgery research. There is strong evidence for a positive association between hospital volume and improved survival in this area. These articles highlight the importance of quality and consistency in study design, definitions, data sources, and risk adjustment in outcomes research. Although the exact mechanism of improved outcomes across the congenital cardiac literature remains unclear, these data have led to policy changes that have improved outcomes.⁶⁵ The CDH literature shows a similar association, but the number and quality of studies have not been sufficient to effect such policy changes. Further research with larger numbers and more thorough adjustment models will help to solidify these findings.

Another limitation in the current literature is the variability in risk adjustment. Nineteen percent of studies did not adjust for any covariates, and of those that did, several adjusted only for demographic factors. It is essential to adjust for case mix because such factors have a strong influence on outcome disparities. Ideally, models should adjust for factors specifically related to the procedures and outcomes of interest. An excellent example is the Risk Adjustment for Congenital Heart Surgery score used in several cardiac studies.⁶⁶ Other risk measures, such as diagnosis-related group scores, are somewhat less specific, and study results for such adjustment modalities in our review are largely negative.^{19,35} Adjustment for clustering is another important consideration, and only 43% of studies ($n = 27$) accounted for this. Outcomes investigations must use statistical methods that account for nonindependence of sampling because such clustering can affect risk estimates and standard errors.

Conclusions within the literature are also limited by the heterogeneity of exposure definitions. Exposures in these studies aim to capture some measure of hospital and surgeon experience, measured by operative volume, surgeon subspecialty, or hospital designation. The first conceptual question addresses why there should be a relationship between volume and outcome or between surgeon subspecialty or hospital designation and outcome and what the nature of that relationship is. Other authors have addressed theories behind this,⁶⁷ but here we will simply address the issue as to whether volume and specialty or designation are simply associated with improved outcome, or if they are causally related to improved outcome. It is important to understand these underlying theories because exposure definition affects conclusions. In general, the surgeon-volume definition of experience argues for a causal relationship—the “practice makes perfect” theory. Hospital volume is the definition that has been most consistently applied in the literature and is often assumed to be a proxy for surgeon volume, and therefore again assumed to be describing a causal relationship. Specialty or design-

nation definitions also imply a causal relationship: expertise and experience lead to improved outcomes. With different exposure definitions, one may not necessarily be superior to the other, but some of the most informative studies considered both volume and specialty or designation measures.^{11,15,46,57,58} Similarly, studies assessing both surgeon- and hospital-level exposures are able to assess the relative contribution of each. Regarding volume measures specifically, investigators should try to define volume categories a priori and standardize volume cutoffs for specific procedures. This enhances generalizability and limits bias. Another option is to treat volume as a continuous variable, but this can impair interpretability and limit the ability to make clinical or policy recommendations.

A fourth limitation in these studies is the a priori temporal relationship between volume and outcome defined in the study design. If the assumption is that increased volume is causally related to improved outcome, then the exposure must precede the outcome. All of the volume measure studies in our review define volume within the same time period as the outcomes being measured. In this case, there is no time for a “practice makes perfect” relationship to be established. Preferably, researchers should define the volume-outcome relationship using hospital or surgeon caseloads from prior time periods to define volume strata for subsequent time periods. For example, 1 study (C.A.L., J.P.M., J. Huaco, MD, MPH, M. Garrison, PhD, J. Avansino, MD, T. Koepsell, MD, MPH, K.T.O., and A.B.G., June 2011, unpublished data) used the average annual hospital volume during a 2-year period to define the volume category to be associated with outcomes observed in the third year. This can help to decrease exposure misclassification. By the same token, only outcomes that temporally follow the exposure of interest should be assessed. For example, several studies assessed perforation rate as an outcome in appendectomy studies, even though this is a factor determined at presentation.^{68,69}

The most important point manifested by the results of these studies is that hospital-level characteristics are often strongly associated with improved outcomes in less common, more complex problems (eg, CDH and congenital heart surgery), whereas surgeon-level factors appear to be more important in more common, less resource-demanding procedures (eg, appendectomy, pyloromyotomy, ureteral reimplantation, and cleft lip repair) as well as in procedures commonly encountered in adult surgery (thyroidectomy, inguinal herniorrhaphy, and cholecystectomy). These results highlight the importance of surgeon- vs system-level factors, depending on the condition of interest. Although surgeon volume and subspecialty probably represent measures of practitioner experience and competence, hospital volume and designation may serve as proxies for other characteristics of large centers of care. This point is highlighted by the findings of studies on solid tumors,⁵³ showing improved care at Children’s Oncology Group–designated centers, regardless of volume, and of the French studies in patients with biliary atresia,⁵⁶ showing improved outcomes after nationwide standardization but no formal regionalization policy. Ideally, future research should evalu-

ate surgeon- and hospital-level measures simultaneously, which will provide more information than evaluating either independently.

The system-level factors at play in high-complexity procedures include numerous multidisciplinary resources. Infants and children with complex surgical diagnoses require high-intensity therapies from many sources. Nurses, neonatologists, intensivists, anesthesiologists, and subspecialty consultants are essential components of the system necessary to care for the sickest children. High hospital volume, therefore, may not be the key exposure that is being captured in complex procedures. Rather, the system of care associated with high volume (or a children’s hospital designation) is the key. Therefore, policies designed to ensure that children requiring resource-demanding procedures are cared for in fully equipped environments may improve outcomes.

One possible way to produce this improvement would be to define and standardize the hospital-level resources (eg, pediatric surgeons, anesthesiologists, and nurses) necessary to care for pediatric patients with low-, medium-, or high-complexity surgical conditions. Several specialty guideline articles⁷⁰⁻⁷³ have been written to suggest system-level requirements needed to create the safest environment for pediatric surgical patients, much like the United States trauma system that has led to clear improvements in trauma outcomes.⁷⁴ With the adoption of national guidelines for hospital-level resources needed to care for low-, medium-, and high-acuity patients, many hospital-level confounders will be removed and comparisons between hospitals and physicians will become more robust and effective. This will address many of the limitations of research using existing databases and may facilitate the creation and use of national disease-specific outcome registries.

Although conclusions regarding experience and outcomes in pediatric surgery can be drawn in some areas, as already highlighted, a great deal more work must be done to fully elucidate this relationship and the circumstances under which it is most important. Children’s surgery presents special challenges in assessing the influence of hospital- and surgeon-level factors on patient outcomes. This fact, coupled with the ever-increasing societal focus on outcomes in medicine, makes it essential that 2 things occur in parallel. First, high-quality clinical research must be conducted to identify the key hospital and surgeon factors that affect patient outcomes, with an initial focus on high-acuity resource-intensive procedures, such as intracavitary neonatal procedures. Second, given the amount of time that it will take to complete such research, we should define and consider implementing resource standards for specific conditions that reflect current data and expert consensus opinion. Our review of the literature highlights a number of strengths and limitations in the present evidence base, and clarifying these may help inform the design of future studies that would be more amenable to comparison and meta-analysis.

Accepted for Publication: December 21, 2012.

Published Online: March 25, 2013. doi:10.1001/jamapediatrics.2013.25

Correspondence: Jarod P. McAteer, MD, Division of Pediatric General and Thoracic Surgery, Seattle Children's Hospital, 4800 Sand Point Way NE, Seattle, WA 98105 (jarodmc@u.washington.edu).

Author Contributions: Dr McAteer had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. *Study concept and design:* McAteer, LaRiviere, Oldham, and Goldin. *Acquisition of data:* McAteer and Goldin. *Analysis and interpretation of data:* All authors. *Drafting of the manuscript:* McAteer and Goldin. *Critical revision of the manuscript for important intellectual content:* LaRiviere, Drugas, Abdullah, Oldham, and Goldin. *Statistical analysis:* McAteer and Goldin. *Administrative, technical, and material support:* Drugas and Goldin. *Study supervision:* Abdullah, Oldham, and Goldin.

Conflict of Interest Disclosures: None reported.

Online-Only Material: An eTable is available at <http://www.jamapeds.com>.

REFERENCES

1. Birkmeyer JD, Siewiers AE, Finlayson EV, et al. Hospital volume and surgical mortality in the United States. *N Engl J Med*. 2002;346(15):1128-1137.
2. Birkmeyer JD, Stukel TA, Siewiers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. *N Engl J Med*. 2003;349(22):2117-2127.
3. Khuri SF, Henderson WG. The case against volume as a measure of quality of surgical care. *World J Surg*. 2005;29(10):1222-1229.
4. Khuri SF, Henderson WG, Daley J, et al; Principal Investigators of the Patient Safety in Surgery Study. Successful implementation of the Department of Veterans Affairs' National Surgical Quality Improvement Program in the private sector: the Patient Safety in Surgery study. *Ann Surg*. 2008;248(2):329-336.
5. Boudourakis LD, Wang TS, Roman SA, Desai R, Sosa JA. Evolution of the surgeon-volume, patient-outcome relationship. *Ann Surg*. 2009;250(1):159-165.
6. Halm EA, Lee C, Chassin MR. Is volume related to outcome in health care? a systematic review and methodologic critique of the literature. *Ann Intern Med*. 2002;137(6):511-520.
7. Surgical Advisory Panel, American Academy of Pediatrics. Guidelines for referral to pediatric surgical specialists. *Pediatrics*. 2002;110(1, pt 1):187-191.
8. Javid PJ, Jaksic T, Skarsgard ED, Lee S; Canadian Neonatal Network. Survival rate in congenital diaphragmatic hernia: the experience of the Canadian Neonatal Network. *J Pediatr Surg*. 2004;39(5):657-660.
9. McKiernan PJ, Baker AJ, Kelly DA. The frequency and outcome of biliary atresia in the UK and Ireland. *Lancet*. 2000;355(9197):25-29.
10. Emil SG, Taylor MB. Appendicitis in children treated by pediatric versus general surgeons. *J Am Coll Surg*. 2004;39(1):34-39.
11. Somme S, To T, Langer JC. Effect of subspecialty training on outcome after pediatric appendectomy. *J Pediatr Surg*. 2007;42(1):221-226.
12. Alexander F, Magnuson D, DiFiore J, Jirousek K, Sedic M. Specialty versus generalist care of children with appendicitis: an outcome comparison. *J Pediatr Surg*. 2001;36(10):1510-1513.
13. Brain AJ, Roberts DS. Who should treat pyloric stenosis: the general or specialist pediatric surgeon? *J Pediatr Surg*. 1996;31(11):1535-1537.
14. Prankoff T, Campbell BT, Travis J, Hirschl RB. Differences in outcome with subspecialty care: pyloromyotomy in North Carolina. *J Pediatr Surg*. 2002;37(3):352-356.
15. Langer JC, To T. Does pediatric surgical specialty training affect outcome after Ramstedt pyloromyotomy? a population-based study. *Pediatrics*. 2004;113(5):1342-1347.
16. Kokoska ER, Minkes RK, Silen ML, et al. Effect of pediatric surgical practice on the treatment of children with appendicitis. *Pediatrics*. 2001;107(6):1298-1301.
17. Chang RK, Rodriguez S, Lee M, Klitzner TS. Risk factors for deaths occurring within 30 days and 1 year after hospital discharge for cardiac surgery among pediatric patients. *Am Heart J*. 2006;152(2):386-393.
18. Hannan EL, Raczyk M, Kavey RE, Quaegebeur JM, Williams R. Pediatric cardiac surgery: the effect of hospital and surgeon volume on in-hospital mortality. *Pediatrics*. 1998;101(6):963-969.
19. Berry JG, Lieu TA, Forbes PW, Goldmann DA. Hospital volumes for common pediatric specialty operations. *Arch Pediatr Adolesc Med*. 2007;161(1):38-43.
20. Chang RK, Klitzner TS. Can regionalization decrease the number of deaths for children who undergo cardiac surgery? a theoretical analysis. *Pediatrics*. 2002;109(2):173-181.
21. Bazzani LG, Marcin JP. Case volume and mortality in pediatric cardiac surgery patients in California, 1998-2003. *Circulation*. 2007;115(20):2652-2659.
22. Benavidez OJ, Gauvreau K, Del Nido P, Bacha E, Jenkins KJ. Complications and risk factors for mortality during congenital heart surgery admissions. *Ann Thorac Surg*. 2007;84(1):147-155.
23. Hirsch JC, Gurney JG, Donohue JE, Gebremariam A, Bove EL, Ohye RG. Hospital mortality for Norwood and arterial switch operations as a function of institutional volume. *Pediatr Cardiol*. 2008;29(4):713-717.
24. Welke KF, Diggs BS, Karamlou T, Ungerleider RM. The relationship between hospital surgical case volumes and mortality rates in pediatric cardiac surgery: a national sample, 1988-2005. *Ann Thorac Surg*. 2008;86(3):889-896.
25. Welke KF, O'Brien SM, Peterson ED, Ungerleider RM, Jacobs ML, Jacobs JP. The complex relationship between pediatric cardiac surgical case volumes and mortality rates in a national clinical database. *J Thorac Cardiovasc Surg*. 2009;137(5):1133-1140.
26. Morales DL, Zafar F, Rossano JW, et al. Use of ventricular assist devices in children across the United States: analysis of 7.5 million pediatric hospitalizations. *Ann Thorac Surg*. 2010;90(4):1313-1319.
27. Davies RR, Russo MJ, Hong KN, et al. Increased short- and long-term mortality at low-volume pediatric heart transplant centers: should minimum standards be set? retrospective data analysis. *Ann Surg*. 2011;253(2):393-401.
28. Pasquali SK, Jacobs JP, He X, et al. The complex relationship between center volume and outcome in patients undergoing the Norwood operation. *Ann Thorac Surg*. 2012;93(5):1556-1562.
29. Pasquali SK, Li JS, Burstein DS, et al. Association of center volume with mortality and complications in pediatric heart surgery. *Pediatrics*. 2012;129(2):e370-e376. doi:10.1542/peds.2011-1188.
30. Cochrane DD, Kestle JR. The influence of surgical operative experience on the duration of first ventriculoperitoneal shunt function and infection. *Pediatr Neurosurg*. 2003;38(6):295-301.
31. Smith ER, Butler WE, Barker FG II. In-hospital mortality rates after ventriculoperitoneal shunt procedures in the United States, 1998 to 2000: relation to hospital and surgeon volume of care. *J Neurosurg*. 2004;100(2)(Suppl Pediatrics):90-97.
32. Smith ER, Butler WE, Barker FG II. Craniotomy for resection of pediatric brain tumors in the United States, 1988 to 2000: effects of provider caseloads and progressive centralization and specialization of care. *Neurosurgery*. 2004;54(3):553-565.
33. Lewis CW, Carron JD, Perkins JA, Sie KC, Feudtner C. Tracheotomy in pediatric patients: a national perspective. *Arch Otolaryngol Head Neck Surg*. 2003;129(5):523-529.
34. Hopper RA, Lewis C, Umbdenstock R, Garrison MM, Starr JR. Discharge practices, readmission, and serious medical complications following primary cleft lip repair in 23 U.S. children's hospitals. *Plast Reconstr Surg*. 2009;123(5):1553-1559.
35. Vitale MA, Arons RR, Hyman JE, Skaggs DL, Roye DP, Vitale MG. The contribution of hospital volume, payer status, and other factors on the surgical outcomes of scoliosis patients: a review of 3,606 cases in the state of California. *J Pediatr Orthop*. 2005;25(3):393-399.
36. Schurman SJ, Stablein DM, Perlman SA, Warady BA. Center volume effects in pediatric renal transplantation: a report of the North American Pediatric Renal Transplant Cooperative Study. *Pediatr Nephrol*. 1999;13(5):373-378.
37. Tracy ET, Bennett KM, Danko ME, et al. Low volume is associated with worse patient outcomes for pediatric liver transplant centers. *J Pediatr Surg*. 2010;45(1):108-113.
38. Nelson CP, Dunn RL, Wei JT, Gearhart JP. Surgical repair of bladder exstrophy in the modern era: contemporary practice patterns and the role of hospital case volume. *J Urol*. 2005;174(3):1099-1102.
39. Dickstein RJ, Barone JG, Liao JG, Burd RS. The effect of surgeon volume and hospital characteristics on in-hospital outcome after ureteral reimplantation in children. *Pediatr Surg Int*. 2006;22(5):417-421.
40. Nguyen HG, Chamie K, Nguyen KG, Durbin-Johnson B, Kurzrock EA. Outcomes after pediatric ureteral reimplantation: a population based analysis. *J Urol*. 2011;185(6):2292-2297.
41. Ponsky TA, Huang ZJ, Kittle K, et al. Hospital- and patient-level characteristics and the risk of appendiceal rupture and negative appendectomy in children. *JAMA*. 2004;292(16):1977-1982.
42. Smink DS, Finkelstein JA, Kleinman K, Fishman SJ. The effect of hospital volume of pediatric appendectomies on the misdiagnosis of appendicitis in children. *Pediatrics*. 2004;113(1 pt 1):18-23.
43. Whisker L, Luke D, Hendrickse C, Bowley DM, Lander A. Appendicitis in children: a comparative study between a specialist paediatric centre and a district general hospital. *J Pediatr Surg*. 2009;44(2):362-367.

44. Cosper GH, Hamann MS, Stiles A, Nakayama DK. Hospital characteristics affect outcomes for common pediatric surgical conditions. *Am Surg.* 2006;72(8):739-745.
45. Collins HL, Almond SL, Thompson B, et al. Comparison of childhood appendicitis management in the regional paediatric surgery unit and the district general hospital. *J Pediatr Surg.* 2010;45(2):300-302.
46. Ly DP, Liao JG, Burd RS. Effect of surgeon and hospital characteristics on outcome after pyloromyotomy. *Arch Surg.* 2005;140(12):1191-1197.
47. Safford SD, Pietrobon R, Safford KM, Martins H, Skinner MA, Rice HE. A study of 11,003 patients with hypertrophic pyloric stenosis and the association between surgeon and hospital volume and outcomes. *J Pediatr Surg.* 2005;40(6):967-973.
48. Raval MV, Cohen ME, Barsness KA, Bentrem DJ, Phillips JD, Reynolds M. Does hospital type affect pyloromyotomy outcomes? analysis of the Kids' Inpatient Database. *Surgery.* 2010;148(2):411-419.
49. Baird R, Eeson G, Safavi A, Puligandla P, Laberge JM, Skarsgard ED; Canadian Pediatric Surgery Network. Institutional practice and outcome variation in the management of congenital diaphragmatic hernia and gastroschisis in Canada: a report from the Canadian Pediatric Surgery Network. *J Pediatr Surg.* 2011;46(5):801-807.
50. Grushka JR, Laberge JM, Puligandla P, Skarsgard ED; Canadian Pediatric Surgery Network. Effect of hospital case volume on outcome in congenital diaphragmatic hernia: the experience of the Canadian Pediatric Surgery Network. *J Pediatr Surg.* 2009;44(5):873-876.
51. Bucher BT, Guth RM, Saito JM, Najaf T, Warner BW. Impact of hospital volume on in-hospital mortality of infants undergoing repair of congenital diaphragmatic hernia. *Ann Surg.* 2010;252(4):635-642.
52. Gutierrez JC, Koniaris LG, Cheung MC, Byrne MM, Fischer AC, Sola JE. Cancer care in the pediatric surgical patient: a paradigm to abolish volume-outcome disparities in surgery. *Surgery.* 2009;145(1):76-85.
53. Gutierrez JC, Cheung MC, Zhuge Y, Koniaris LG, Sola JE. Does Children's Oncology Group hospital membership improve survival for patients with neuroblastoma or Wilms tumor? *Pediatr Blood Cancer.* 2010;55(4):621-628.
54. McClement JW, Howard ER, Mowat AP. Results of surgical treatment for extrahepatic biliary atresia in United Kingdom 1980-2: survey conducted on behalf of the British Paediatric Association Gastroenterology Group and the British Association of Paediatric Surgeons. *Br Med J (Clin Res Ed).* 1985;290(6465):345-347.
55. Chardot C, Carton M, Spire-Bendelac N, Le Pommelet C, Golmard JL, Auvvert B. Prognosis of biliary atresia in the era of liver transplantation: French national study from 1986 to 1996. *Hepatology.* 1999;30(3):606-611.
56. Serinet MO, Broué P, Jacquemin E, et al. Management of patients with biliary atresia in France: results of a decentralized policy 1986-2002. *Hepatology.* 2006;44(1):75-84.
57. Borenstein SH, To T, Wajja A, Langer JC. Effect of subspecialty training and volume on outcome after pediatric inguinal hernia repair. *J Pediatr Surg.* 2005;40(1):75-80.
58. Sosa JA, Tuggle CT, Wang TS, et al. Clinical and economic outcomes of thyroid and parathyroid surgery in children. *J Clin Endocrinol Metab.* 2008;93(8):3058-3065.
59. Tuggle CT, Roman SA, Wang TS, et al. Pediatric endocrine surgery: who is operating on our children? *Surgery.* 2008;144(6):869-877.
60. Chen K, Cheung K, Sosa JA. Surgeon volume trumps specialty: outcomes from 3596 pediatric cholecystectomies. *J Pediatr Surg.* 2012;47(4):673-680.
61. Bratton SL, Haberkern CM, Waldhausen JH, Sawin RS, Allison JW. Intussusception: hospital size and risk of surgery. *Pediatrics.* 2001;107(2):299-303.
62. Aylin P, Bottle A, Majeed A. Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models. *BMJ.* 2007;334(7602):1044. doi:10.1136/bmj.39168.496366.55.
63. Welke KF, Karamlou T, Diggs BS. Databases for assessing the outcomes of the treatment of patients with congenital and paediatric cardiac disease—a comparison of administrative and clinical data. *Cardiol Young.* 2008;18(suppl 2):137-144.
64. Welke KF, Diggs BS, Karamlou T, Ungerleider RM. Comparison of pediatric cardiac surgical mortality rates from national administrative data to contemporary clinical standards. *Ann Thorac Surg.* 2009;87(1):216-222, discussion 222-223.
65. Allen SW, Gauvreau K, Bloom BT, Jenkins KJ. Evidence-based referral results in significantly reduced mortality after congenital heart surgery. *Pediatrics.* 2003;112(1, pt 1):24-28.
66. Jenkins KJ, Gauvreau K, Newburger JW, Spray TL, Moller JH, Iezzoni LI. Consensus-based method for risk adjustment for surgery for congenital heart disease. *J Thorac Cardiovasc Surg.* 2002;123(1):110-118.
67. Luft HS, Hunt SS, Maerki SC. The volume-outcome relationship: practice-makes-perfect or selective-referral patterns? *Health Serv Res.* 1987;22(2):157-182.
68. Bratton SL, Haberkern CM, Waldhausen JH. Acute appendicitis risks of complications: age and Medicaid insurance. *Pediatrics.* 2000;106(1, pt 1):75-78.
69. Smink DS, Fishman SJ, Kleinman K, Finkelstein JA. Effects of race, insurance status, and hospital volume on perforated appendicitis in children. *Pediatrics.* 2005;115(4):920-925.
70. Arul GS, Spicer RD. Where should paediatric surgery be performed? *Arch Dis Child.* 1998;79(1):65-72.
71. Stark AR; American Academy of Pediatrics Committee on Fetus and Newborn. Levels of neonatal care. *Pediatrics.* 2004;114(5):1341-1347.
72. Hackel A, Badgwell JM, Binding RR, et al. Guidelines for the pediatric perioperative anesthesia environment. *Pediatrics.* 1999;103(2):512-515.
73. Stolar CJ; Workforce Committee American Pediatric Surgical Association. Best practice for infant surgery—a position statement from the American Pediatric Surgical Association. *J Pediatr Surg.* 2008;43(9):1585-1586.
74. MacKenzie EJ, Rivara FP, Jurkovich GJ, et al. A national evaluation of the effect of trauma-center care on mortality. *N Engl J Med.* 2006;354(4):366-378.