

## Effect of Hospital Volume of Thyroidectomies on Outcomes Following Substernal Thyroidectomy

Fredric M. Pieracci · Thomas J. Fahey III

Published online: 15 January 2008  
© Société Internationale de Chirurgie 2007

### Abstract

**Background** Substernal thyroidectomy (ST), as compared to conventional, cervical thyroidectomy, is a technically demanding procedure that is associated with increased morbidity and mortality. We tested the hypothesis that outcomes following ST are improved at centers that perform a high volume of thyroidectomies.

**Methods** Patients who underwent ST from 1998 to 2004 were extracted from the New York State Statewide Planning and Research Cooperative System database. Hospital volume of thyroidectomies was divided into low (<33 per year), middle (33–99 per year), and high ( $\geq 100$  per year) volumes. Outcome variables included hospital length of stay (LOS), recurrent laryngeal nerve (RLN) injury, hypoparathyroidism, postoperative bleeding, respiratory failure, blood transfusion, and mortality.

**Results** A total of 1153 STs were analyzed; 372 (32.2%) were performed at low-volume centers, 388 (33.7%) at middle-volume centers, and 393 (34.0%) at high-volume centers. Linear associations were observed between increasing hospital volume of thyroidectomies and decreasing age ( $p = 0.003$ ), increasing co-morbidity ( $p < 0.0001$ ), increased likelihood of total versus subtotal thyroidectomy ( $p < 0.0001$ ), and increased likelihood of thyroid malignancy ( $p < 0.0001$ ). Despite this, increasing hospital volume of thyroidectomies predicted a decreased likelihood of overall complications ( $p = 0.005$ ), postoperative bleeding ( $p = 0.01$ ), blood transfusion ( $p = 0.04$ ), respiratory failure ( $p = 0.04$ ) and mortality ( $p = 0.004$ ), as well as a trend

toward a decreased LOS ( $p = 0.06$ ). The overall complication rate and the mortality rate remained significantly associated with volume group by multivariate analysis.

**Conclusion** Despite more extensive surgery on patients with greater co-morbidity, LOS, morbidity, and mortality were all decreased when ST occurred at hospitals that perform a high volume of thyroidectomies.

A robust and consistent association has been observed between outcomes following surgical procedures and the hospital volume of that procedure, with patients who undergo surgery at high-volume centers having better outcomes than those who undergo surgery at low-volume centers [1–13]. Importantly, this relation appears constant over varying levels of surgical complexity and specialties, having been observed following procedures ranging from appendectomy [10] to pancreaticoduodenectomy [2, 14].

As a result of these data, several authors have argued for the regionalization of technically demanding procedures that are performed relatively infrequently to high-volume centers. Such a model has enjoyed success in the field of trauma surgery, where severely injured patients have significantly better outcomes when treated at high-volume trauma centers [15]. Accordingly, credentialing agencies as well as third-party payers have incorporated volume standards into hospital referral criteria for myriad surgical procedures that have documented volume-outcome relations [16–18].

The relation between hospital volume of thyroidectomies and outcomes following substernal thyroidectomy (ST) has not been studied. Substernal thyroidectomy is indicated for removal of a substernal goiter (defined most commonly as extension of  $> 50\%$  of the thyroid gland into the mediastinum [19]), which is present in approximately

F. M. Pieracci (✉) · T. J. Fahey III  
Department of Surgery, Department of Public Health, Weill  
Medical College of Cornell University, 411 East 69th Street,  
KB-220, New York, NY 10021, USA  
e-mail: frp9005@med.cornell.edu

6% of patients who present for thyroidectomy [19–28]. Substernal thyroidectomy has been advocated for the treatment of all substernal goiters for several reasons—including the natural history of substernal goiter, which is one of progressive expansion with eventual onset of compressive symptoms, the fact that medical therapy is nearly always unsuccessful [29], and because surgery often relieves compressive symptoms [23]. Additionally, approximately 10% of substernal goiters harbor an underlying malignancy [20–22, 25, 27].

Compared to conventional cervical thyroidectomy (CT), ST is associated with a increased hospital length of stay (LOS), increased postoperative morbidity, and increased mortality [30, 31]. Using the New York State Statewide Planning and Cooperative Systems (SPARCS) Database, we recently documented a more than eightfold risk-adjusted increase in the likelihood of mortality following ST compared to CT [31]. Worse outcomes following ST compared to CT are likely due to both the technical difficulty encountered during excision and challenges faced during perioperative management of patients with a substernal goiter (e.g., tracheal compression necessitating fiberoptic bronchoscopy for endotracheal intubation).

The purpose of this study was to determine if a volume–outcome relation exists between hospital volume of thyroidectomies and outcomes following ST. Such an association would have potential health care policy ramifications regarding regionalization. We therefore tested the hypothesis that LOS, morbidity, and mortality are improved when ST is performed at a high-volume hospital.

## Materials and methods

Data were extracted from the SPARCS database for the years 1998–2004. The SPARCS database contains information on all patients discharged from acute-care, nonfederal hospitals in New York State (NYS). A total of 15 diagnostic fields and 15 procedures fields, which are based on the International Classification for Disease, 9th revision (ICD-9) classification system, are available for each patient discharge. For the present analysis, a data set containing patients who underwent either partial (06.52) or total (06.51) ST was created. Substernal thyroidectomies in which the extent of resection was unclear (06.53) were excluded.

The primary independent variable was the average hospital volume of thyroidectomies performed over the 7-year study period. Although hospital volume of STs was considered originally as the variable on which to base the volume calculation, too few STs occurred to make this a valid metric. Specifically, among the hospitals that performed at least one ST over the 7-year study period, the

mean and median annual hospital volume of STs was 1.0 and 0.4, respectively.

The volume variable was thus based on the average number of thyroidectomies performed. Thyroidectomies were captured using the aforementioned ST codes, as well as the following codes for CT: unilateral cervical thyroid lobectomy (06.2), partial CT (06.3), and total CT (06.4). Patients who underwent thyroid biopsy (06.1) only, as well as cases in which the extent of thyroid resection was not specified (06.53 and 06.98) were not included in the hospital volume calculation.

Because the hospital volume of thyroidectomies was not distributed normally, it was analyzed as an ordinal categorical variable, with cutoff points determined by division into tertiles. Tertiles were chosen to assess for linear trends between hospital volume groups (in contrast to analysis of a dichotomous variable) while retaining sufficient sample sizes within groups for multivariable analysis. Low-volume hospitals were defined as centers that averaged <33 thyroidectomies per year, middle-volume as between 33 and 99 thyroidectomies per year, and high-volume as  $\geq 100$  thyroidectomies per year. Division into tertiles resulted in approximately equal sample sizes within each volume group.

Additional covariates abstracted included age, sex, comorbidity, race (white vs. nonwhite), insurance status (private insurance vs. other), extent of thyroidectomy (total vs. subtotal), and thyroid pathology [malignancy (19.3) vs. other]. Patient co-morbidities were captured using the Deyo co-morbidity index [32], which assigns points based on the ICD-9 coding system for 19 preexisting co-morbid conditions, ranging from 1 to 6 for each condition, for a total possible score of 37. However, owing to a high degree of co-linearity between the thyroid malignancy variable and the total co-morbidity score (which incorporates thyroid malignancy), a modified Deyo co-morbidity score, which excluded thyroid malignancy (19.3), was calculated. The co-morbidity variable was analyzed as continuous.

Outcome variables included hospital LOS, hospital mortality, and surgical complications. Complications following thyroidectomy were captured using ICD-9 diagnostic coding as outlined by Sosa et al. [30]. Recurrent laryngeal nerve (RLN) injury was defined as the presence of a diagnostic code in any field for either vocal cord paralysis (47.83) or a surgical complication involving the nervous system (99.700, 997.09). Endocrine complications were captured using the diagnostic codes for either hypoparathyroidism (25.21) or hypocalcemia (27.541) (grouped into the variable hypoparathyroidism for the present analysis). Postoperative bleeding was captured using the diagnostic codes for either hematoma formation (998.12) or hemorrhage related to a procedure (998.11). Acute respiratory failure (518.81) and red blood cell (RBC) transfusion (99.04) were also queried.

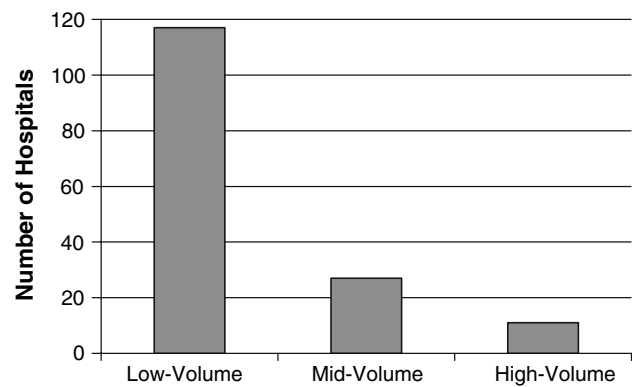
Statistical analyses were computed using SAS Version 9.1 (SAS Institute, Carey, NC, USA). All  $p$  values were two-sided, with statistical significance evaluated at the 0.05  $\alpha$  level. The assessment of differences in continuous variables between the three volume groups was performed using analysis of variance (ANOVA). The  $p$  values derived from this test are listed as  $p_{ANOVA}$ . In the case of overall significance, the assessment of significance between individual volume groups was performed using Tukey's studentized range (HSD) test [33]. The assessment of differences in nominal categorical variables between volume groups was performed using the omnibus  $\chi^2$  test. The  $p$  values derived from this test are listed as  $p_{\chi^2 omnibus}$ . Furthermore, the assessment of a linear trend between increasing volume groups and both covariates and outcomes was performed by calculating the  $\chi^2$  trend statistic. The  $p$  values derived from this test are listed as  $p_{\chi^2 trend}$ .

Multivariate logistic regression models were created to evaluate the independent effect of hospital volume of thyroidectomies on both the occurrence of at least one complication and mortality following ST. Variables associated with hospital volume at the  $p < 0.10$  level by univariate analysis were added to the model using a forward selection method. The overall contribution of the fitted model to predicting variability in the outcome of interest was assessed using the likelihood ratio  $\chi^2$  test. The independent contribution of individual variables was assessed using the Wald  $\chi^2$  test. Model fit was assessed using the Hosmer-Lemeshow goodness-of-fit  $\chi^2$  statistic, with  $p > 0.05$  indicating acceptable model calibration.

## Results

Of the 217 New York State hospitals that performed at least one thyroidectomy over the study period, 155 (71.4%) performed at least one ST. Figure 1 depicts the distribution of these 155 hospitals according to annual volume of thyroidectomies. Whereas most hospitals averaged  $< 33$  thyroidectomies per year ( $n = 117$ , 75.5%) (low volume), only 11 hospitals (7.1%) averaged  $\geq 100$  thyroidectomies per year (high volume). The mean number of thyroidectomies performed annually was 10.7 (range 0.3–31.2) for low-volume centers, 53.0 (range 35.7–89.3) for middle-volume centers, and 170.3 (range 100.9–287.4) for high-volume centers ( $p_{ANOVA} < 0.0001$ ).

A total of 1153 patients underwent ST over the 7-year study period. Sample characteristics are shown in Table 1. The sample consisted of a relatively young, healthy population with a mean age of 56.7 years and a mean co-morbidity score of 0.7. Malignancy was present in 261 patients (24.3%); 680 (59.0%) patients underwent total thyroidectomy; and 988 (85.7%) patients underwent



**Fig. 1** Distribution of the 155 New York state hospitals that performed at least one substernal thyroidectomy according to the volume of thyroidectomies performed over the 7-year study period. Most ( $n = 117$ , 75.5%) of the hospitals performed fewer than 33 thyroidectomies per year (low-volume), whereas only 11 hospitals (7.1%) performed  $\geq 100$  thyroidectomies per year (high-volume). The remaining hospitals ( $n = 27$ , 17.4%) performed between 33 and 99 thyroidectomies per year (middle-volume)

surgery at a teaching hospital. The volume groups were distributed evenly, with 372 patients (32.2%) undergoing ST at a low-volume center, 388 (33.7%) at a middle-volume center, and 393 (34.0%) at a high-volume center.

Analysis of patient covariates revealed significant differences according to volume groups (Table 2). Age decreased slightly with increasing volume; the mean age of patients who underwent surgery at a high-volume facility was 4 years younger than that of patients who underwent surgery at a low-volume center (58.5 vs. 54.5, respectively,  $p_{ANOVA} = 0.003$ ). The co-morbidity score also varied significantly by volume group ( $p_{ANOVA} < 0.0001$ ), with high-volume patients having a mean score that was twice that of low-volume patients (1.2 vs. 0.6, respectively). Furthermore, the likelihood of private insurance

**Table 1** Sample demographics

| Parameter                    | Value       |
|------------------------------|-------------|
| Age (years)*                 | 56.7 (16.5) |
| Modified co-morbidity score* | 0.7 (1.9)   |
| Female                       | 854 (74.0%) |
| White                        | 663 (57.5%) |
| Private insurance            | 636 (55.1%) |
| Malignancy                   | 261 (24.3%) |
| Total thyroidectomy          | 680 (59.0%) |
| Teaching hospital            | 988 (85.7%) |
| High-volume                  | 393 (34.0%) |
| Middle-volume                | 388 (33.7%) |
| Low-volume                   | 372 (32.2%) |

Continuous variables (\*) are expressed as the mean (SD). Categorical variables are expressed as the number

**Table 2** Sample characteristics according to volume group

| Variable                                   | Low-volume<br>(n = 372) | Middle-volume<br>(n = 388) | High-volume<br>(n = 393) | $P_{ANOVA}^a$ | $P_{\chi^2}^b$ omnibus | $P_{\chi^2}^c$ trend |
|--|-------------------------|----------------------------|--------------------------|---------------|------------------------|----------------------|
| Age (years) <sup>d,e</sup>                 | 58.5 (16.3)             | 57.1 (16.5)                | 54.5 (16.5)              | 0.003         | —                      | —                    |
| Modified co-morbidity score <sup>d,f</sup> | 0.6 (1.6)               | 0.5 (1.4)                  | 1.2 (2.5)                | < 0.0001      | —                      | —                    |
| Female                                     | 276 (74.2%)             | 285 (73.5%)                | 293 (74.6%)              | —             | 0.94                   | 0.91                 |
| White                                      | 208 (55.9%)             | 246 (63.4%)                | 209 (53.2%)              | —             | 0.01                   | 0.42                 |
| Private insurance                          | 169 (45.4%)             | 205 (52.8%)                | 262 (66.7%)              | —             | < 0.0001               | < 0.0001             |
| Malignancy                                 | 64 (17.2%)              | 52 (13.4%)                 | 145 (36.9%)              | —             | < 0.0001               | < 0.0001             |
| Total thyroidectomy                        | 165 (44.4%)             | 253 (65.2%)                | 262 (66.7%)              | —             | < 0.0001               | < 0.0001             |
| Teaching hospital                          | 270 (72.6%)             | 325 (83.8%)                | 393 (100%)               | —             | < 0.0001               | < 0.0001             |

<sup>a</sup> The  $p$  value obtained from analysis of variance (ANOVA)

<sup>b</sup> The  $p$  value obtained from the omnibus  $\chi^2$  test

<sup>c</sup> The  $p$  value obtained from the  $\chi^2$  trend statistic

<sup>d</sup> Continuous variables ( $d$ ) are expressed as the mean (SD). Categorical variables are expressed as the number

<sup>e</sup> Significant differences in age existed only between the high-volume and low-volume groups

<sup>f</sup> Significant differences in the co-morbidity score existed between the high-volume and low-volume groups as well as between the high-volume and middle-volume groups, but not between the mid-volume and low-volume groups

( $p_{\chi^2\text{trend}} < 0.0001$ ), total thyroidectomy ( $p_{\chi^2\text{trend}} < 0.0001$ ), malignancy ( $p_{\chi^2\text{trend}} < 0.0001$ ), and presentation to a teaching hospital ( $p_{\chi^2\text{trend}} < 0.0001$ ) all increased with increasing volume group. Whereas only 72.6% of low-volume patients ( $n = 270$ ) underwent surgery at a teaching hospital, all of the 11 high-volume hospitals were teaching facilities. Although race varied significantly by volume group ( $p_{\chi^2\text{omnibus}} = 0.01$ ), no evidence of a linear trend was observed ( $p_{\chi^2\text{trend}} = 0.42$ ). Finally, sex did not vary significantly by volume group ( $p_{\chi^2\text{omnibus}} = 0.94$ ).

Outcomes following ST by volume group are summarized in Table 3 and Figures 2 and 3. The likelihood of at least one postoperative complication increased significantly with decreasing hospital volume ( $p_{\chi^2\text{trend}} = 0.005$ ). Neither RLN injury ( $p_{\chi^2\text{omnibus}} = 0.74$ ) nor hypoparathyroidism ( $p_{\chi^2\text{omnibus}} = 0.14$ ) varied in incidence according to volume group. However, a linear trend was observed between increasing hospital volume of thyroidectomies and

decreasing incidence of postoperative bleeding ( $p_{\chi^2\text{trend}} = 0.01$ ), red blood cell (RBC) transfusion ( $p_{\chi^2\text{trend}} = 0.04$ ), and respiratory failure ( $p_{\chi^2\text{trend}} = 0.04$ ) following ST (Fig. 2). Furthermore, a trend toward a decreased hospital LOS with increasing volume group was observed ( $p_{ANOVA} = 0.06$ ). Finally, mortality was nearly 10-fold higher in the low-volume group than in the high-volume group ( $p_{\chi^2\text{trend}} = 0.004$ ) (Fig. 3).

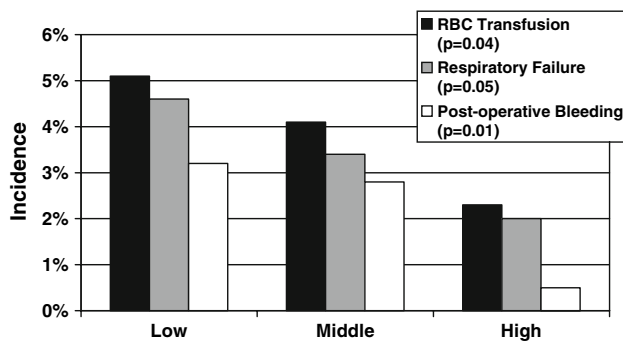
Results from multivariate logistic regression analysis, using both the incidence of at least one complication and mortality as the dependent variables, are shown in Tables 4 and 5, respectively. Age, race, insurance status, co-morbidity, extent of thyroidectomy, and presence of thyroid malignancy were added to the models in addition to hospital volume. Controlling for these covariates, patients who underwent ST at either a low-volume hospital or a middle-volume hospital were more than two times as likely to incur a complication as compared to patients who underwent ST at a

**Table 3** Outcomes according to volume group

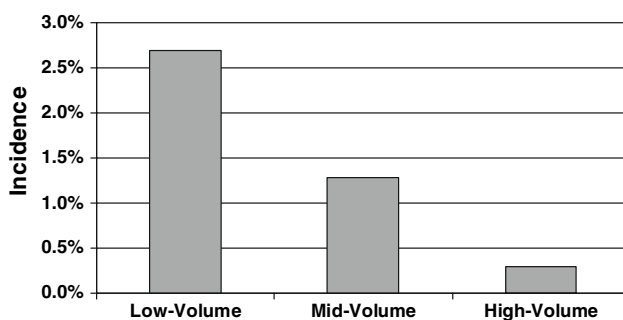
| Variable               | Low-volume (n = 372) | Middle-volume (n = 388) | High-volume (n = 393) | $P_{ANOVA}$ | $P_{\chi^2\text{omnibus}}$ | $P_{\chi^2\text{trend}}$ |
|------------------------|----------------------|-------------------------|-----------------------|-------------|----------------------------|--------------------------|
| Any complication       | 60 (16.1%)           | 65 (16.8%)              | 36 (9.2%)             | —           | 0.003                      | 0.005                    |
| RLN injury             | 6 (1.6%)             | 9 (2.3%)                | 9 (2.3%)              | —           | 0.74                       | 0.52                     |
| Hypoparathyroidism     | 16 (4.1%)            | 28 (7.2%)               | 16 (4.1%)             | —           | 0.14                       | 0.51                     |
| Postoperative bleeding | 12 (3.2%)            | 11 (2.8%)               | 2 (0.5%)              | —           | 0.02                       | 0.01                     |
| RBC transfusion        | 19 (5.1%)            | 16 (4.1%)               | 9 (2.3%)              | —           | 0.12                       | 0.04                     |
| Respiratory failure    | 17 (4.6%)            | 13 (3.4%)               | 8 (2.0%)              | —           | 0.15                       | 0.04                     |
| Length of stay (days)* | 5.6 (11.1)           | 5.3 (19.6)              | 3.4 (7.2%)            | 0.06        | —                          | —                        |
| Mortality              | 10 (2.7%)            | 5 (1.3%)                | 1 (0.3%)              | —           | 0.02                       | 0.004                    |

RLN: recurrent laryngeal nerve; RBC: red blood cell

Continuous variables (\*) are expressed as the mean (SD). Categorical variables are expressed as the number



**Fig. 2** Postoperative complications following substernal thyroidectomy that varied significantly according to volume group. The incidences of postoperative bleeding, red blood cell (RBC) transfusion, and respiratory failure following substernal thyroidectomy were inversely related to hospital volume of thyroidectomies



**Fig. 3** Mortality following substernal thyroidectomy according to volume group. There was an inverse, linear relation between hospital volume of thyroidectomies and the likelihood of mortality following substernal thyroidectomy ( $\chi^2$  trend = 8.3,  $p = 0.004$ )

high-volume center: OR = 2.11, 95% CI [1.24–3.58,  $p = 0.006$  and OR = 2.23, 95% CI 1.34–3.69,  $p = 0.002$ , respectively) (Table 4). Patients who underwent ST at a low-volume hospital were also more than 10 times as likely to die as patients who underwent ST at a high-volume hospital (OR = 10.5, 95% CI 1.08–102.4,  $p = 0.04$ ) (Table 5). Patients who underwent surgery at a middle-volume hospital also demonstrated an increased likelihood of mortality following ST compared to high-volume patients, although this value did not reach statistical significance: OR = 7.8, 95% CI 0.76–80.9,  $p = 0.08$ ).

## Discussion

Using NYS administrative data, we found that patients who underwent ST at a hospital that averaged a relatively high volume of thyroidectomies enjoyed better outcomes than patients who underwent ST at a hospital that averaged fewer thyroidectomies. These associations were observed despite that fact that patients who presented to high-volume centers had greater co-morbidity, were more likely to have a thyroid malignancy, and were more likely to require a

**Table 4** Results from multivariable logistic regression on the occurrence of at least one complication following substernal thyroidectomy

| Parameter                        | Odds ratio        | 95% CI     | p     |
|----------------------------------|-------------------|------------|-------|
| Age (years)                      | 1.02 <sup>a</sup> | 1.00, 1.03 | 0.03  |
| Modified co-morbidity score      | 1.13 <sup>b</sup> | 1.02, 1.24 | 0.02  |
| White race                       | 1.09              | 0.74, 1.62 | 0.66  |
| Private insurance                | 0.53              | 0.34, 0.82 | 0.005 |
| Teaching hospital                | 0.75              | 0.46, 1.22 | 0.25  |
| Total thyroidectomy              | 1.59              | 1.07, 2.37 | 0.02  |
| Thyroid malignancy               | 1.28              | 0.77, 2.12 | 0.34  |
| High-volume hospital (reference) | 1.00              | —          | —     |
| Middle-volume hospital           | 2.23              | 1.34, 3.69 | 0.002 |
| Low-volume hospital              | 2.11              | 1.24, 3.58 | 0.006 |

CI: confidence interval

<sup>a</sup> Refers to the increase in odds of mortality for a 1-year increase in age

<sup>b</sup> Refers to the increase in odds of mortality for a one-point increase in the co-morbidity score

Maximum likelihood  $\chi^2 = 60.1$ ,  $p < 0.0001$ , Hosmer-Lemeshow  $p = 0.08$

total thyroidectomy than patients who underwent surgery at a low-volume center. Using multivariate logistic regression, we found an independent association between the hospital volume of thyroidectomies and both mortality and complications following ST, with patients who underwent ST at a low-volume center incurring a more than twofold increased likelihood of at least one postoperative complication and a more than 10-fold increased likelihood of mortality compared to patients who underwent ST at a

**Table 5** Results from multivariable logistic regression on mortality following substernal thyroidectomy

| Parameter                        | Odds ratio        | 95% CI      | p        |
|----------------------------------|-------------------|-------------|----------|
| Age (years)                      | 1.06 <sup>a</sup> | 1.01, 1.11  | 0.02     |
| Modified co-morbidity score      | 1.76 <sup>b</sup> | 1.43, 2.17  | < 0.0001 |
| White race                       | 0.28              | 0.07, 1.11  | 0.07     |
| Private insurance                | 0.10              | 0.01, 1.00  | 0.05     |
| Teaching hospital                | 0.39              | 0.11, 1.43  | 0.16     |
| Total thyroidectomy              | 0.52              | 0.16, 1.73  | 0.28     |
| Thyroid malignancy               | 0.77              | 0.09, 6.88  | 0.81     |
| High-volume hospital (reference) | 1.00              | —           | —        |
| Middle-volume hospital           | 7.84              | 0.76, 80.95 | 0.08     |
| Low-volume hospital              | 10.52             | 1.08, 102.4 | 0.04     |

<sup>a</sup> Refers to the increase in odds of mortality for a 1-year increase in age

<sup>b</sup> Refers to the increase in odds of mortality for a one-point increase in the co-morbidity score. Maximum likelihood  $\chi^2 = 64.0$ ,  $p < 0.0001$ , Hosmer-Lemeshow  $p = 0.91$



high-volume center. These findings add to the sizable literature documenting a volume–outcome relation in both endocrine surgery and surgery in general. Although these data appear to favor regionalization of STs to high-volume centers, several methodological considerations warrant elaboration.

The most commonly invoked explanation for the volume–outcome relation in surgery involves the theory that both surgeon and hospital experience results in improved patient care—put simply, “practice makes perfect.” Due to the nature of the SPARCS database, we were unable to distinguish the relative contributions of surgeon volume from hospital volume. Independent of hospital volume, surgeon expertise has been associated with improved outcomes following a wide range of procedures, including thyroidectomy [30, 34–38]. Whereas some investigators have found that most of the volume–outcome relation is explained by surgeon experience [30, 38], others have noted an independent effect of hospital volume [34]. Harmon et al. reported that the outcomes of low-volume surgeons following colorectal resection are improved when surgery is performed at high-volume centers [39].

The relation between surgeon and hospital volumes of thyroidectomy and patient outcomes was originally studied by Sosa et al. [30]. Overall, a significant association was observed between increased surgeon volume and improved outcomes, including postoperative complications, LOS, and hospital charges. In a subgroup analysis of surgeons with operating privileges at more than one hospital, LOS was found to be significantly associated with surgeon, but not hospital, volume of thyroidectomies. However, results for additional outcomes by hospital volume groups, including mortality, were not reported.

Outcomes following ST appear intuitively dependent on both surgeon and ancillary staff expertise. The increased risk of RLN injury associated with ST as compared to CT [31] argues for the importance of surgeon experience. However, the incidence of RLN injury did not vary by volume group in the present study, although the relative rarity of this complication precluded a meaningful multivariate analysis. In contrast, because perioperative management of patients with a substernal goiter frequently mandates use of specialized procedures, such as fiberoptic bronchoscopy, sternotomy, and postoperative mechanical ventilation [25, 27], ancillary staff expertise (e.g., anesthesiologists, surgical intensivists) may also have a substantial impact on patient outcome. This hypothesis is consistent with the decreased likelihood of respiratory failure observed in patients who undergo ST at a high-volume center.

A second explanation for the observed volume–outcome relation involves reverse causality. Specifically, providers may selectively refer patients to hospitals that are known to

produce favorable patient outcomes, thereby increasing patient volume at these hospitals. Owing to the retrospective nature of this study, these two possibilities cannot be distinguished.

Irrespective of the relative contributions of surgeon and ancillary staff expertise, the independent association between hospital volume and outcome following ST remains germane; regardless of the mechanism, patients who presented to high-volume hospitals incurred better outcomes than patients who presented to low-volume hospitals. This association must be interpreted with caution as unmeasured patient characteristics that differ by volume group may independently affect outcomes. However, several parameters associated with adverse outcomes following thyroidectomy, including co-morbidity, thyroid malignancy, and extent of surgery [30], were in fact more common in patients who presented to high-volume centers. Furthermore, an independent effect of hospital volume on both complications and mortality following ST was observed after controlling for the aforementioned variables using multivariate logistic regression.

Although the use of administrative data is advantageous for analysis of outcomes following relatively infrequent procedures such as ST, important limitations are also introduced. Certain variables, such as goiter size, degree of substernal extension, and use of a sternotomy incision during ST, are impossible to capture. Series of substernal goiters have reported a mean goiter mass ranging from 140 g [20] to 264 g [23], and increased goiter mass has been associated with postoperative complications following ST [27]. Goiter size may thus confound the relation between ST and adverse outcomes, although there are no data to suggest that goiter size varies by volume group. Analysis of institutional databases that capture goiter size could help to elucidate these relations, although such studies would likely suffer from inadequate power.

Additional limitations of administrative data include undercoding demographic variables such as patient co-morbidity. Furthermore, additional unmeasured patient characteristics may differ based on hospital volume groupings and introduce biases. This phenomenon, known as “clustering,” may in turn exaggerate measures of association in volume–outcome studies [40]. Finally, because hospital administrative data are limited to the inpatient period, the durations of both RLN injury and hypoparathyroidism are impossible to ascertain.

## Conclusions

The literature concerning the hospital volume of surgical procedures and patient outcomes continues to evolve. In the present study, we provide evidence that the hospital

volume of thyroidectomies is directly related to both morbidity and mortality following ST. This effect appears to be independent of differences in patient age, socioeconomic status, co-morbidity, or extent of surgery among volume groups. Although these results suggest potential value to the regionalization of ST to high-volume centers, they must be examined in light of the limitations common to studies of this kind. Further research is needed prior to definitive recommendations.

## References

- Begg CB, Cramer LD, Hoskins WJ, et al. (1998) Impact of hospital volume on operative mortality for major cancer surgery. *JAMA* 280:1747–1751
- Birkmeyer JD, Finlayson SR, Tosteson AN, et al. (1999) Effect of hospital volume on in-hospital mortality with pancreaticoduodenectomy. *Surgery* 125:250–256
- Nguyen NT, Paya M, Stevens CM, et al. (2004) The relationship between hospital volume and outcome in bariatric surgery at academic medical centers. *Ann Surg* 240:586–593; discussion 593–584
- Patti MG, Corvera CU, Glasgow RE, et al. (1998) A hospital's annual rate of esophagectomy influences the operative mortality rate. *J Gastrointest Surg* 2:186–192
- Peck C, Peck J, Peck A (2001) Comparison of carotid endarterectomy at high- and low-volume hospitals. *Am J Surg* 181:450–453
- van Lanschot JJ, Hulscher JB, Buskens CJ, et al. (2001) Hospital volume and hospital mortality for esophagectomy. *Cancer* 91:1574–1578
- Weller WE, Hannan EL (2006) Relationship between provider volume and postoperative complications for bariatric procedures in New York State. *J Am Coll Surg* 202:753–761
- Solomon RA, Mayer SA, Tarmey JJ (1996) Relationship between the volume of craniotomies for cerebral aneurysm performed at New York State hospitals and in-hospital mortality. *Stroke* 1996;27:13–17
- Hannan EL, Popp AJ, Tranmer B, et al. (1998) Relationship between provider volume and mortality for carotid endarterectomies in New York State. *Stroke* 29:2292–2297
- Farber BF, Kaiser DL, Wenzel RP (1981) Relation between surgical volume and incidence of postoperative wound infection. *N Engl J Med* 305:200–204
- Birkmeyer JD, Siewers AE, Finlayson EV, et al. (2002) Hospital volume and surgical mortality in the United States. *N Engl J Med* 346:1128–1137
- Finlayson EV, Goodney PP, Birkmeyer JD (2003) Hospital volume and operative mortality in cancer surgery: a national study. *Arch Surg* 138:721–725
- Lin HC, Xirasagar S, Lee HC, et al. (2006) Hospital volume and inpatient mortality after cancer-related gastrointestinal resections: the experience of an Asian country. *Ann Surg Oncol* 13:1182–1188
- Lieberman MD, Kilburn H, Lindsey M, et al. (1995) Relation of perioperative deaths to hospital volume among patients undergoing pancreatic resection for malignancy. *Ann Surg* 222:638–645
- Nathens AB, Jurkovich GJ, Maier RV, et al. (2001) Relationship between trauma center volume and outcomes. *JAMA* 285:1164–1171
- <http://www.leapfroggroup.org>. Accessed October 3, 2007
- <http://www.facs.org/cqi/bscn/levelsdefined.html>. Cited October 3, 2007
- Epstein AM (2002) Volume and outcome—it is time to move ahead. *N Engl J Med* 346:1161–1164
- Wax MK, Briant TD (1992) Management of substernal goitre. *J Otolaryngol* 21:165–170
- Arici C, Dertsiz L, Altunbas H, et al. (2001) Operative management of substernal goiter: analysis of 52 patients. *Int Surg* 86:220–224
- Erbil Y, Bozboru A, Barbaros U, et al. (2004) Surgical management of substernal goiters: clinical experience of 170 cases. *Surg Today* 34:732–736
- Hedayati N, McHenry CR (2002) The clinical presentation and operative management of nodular and diffuse substernal thyroid disease. *Am Surg* 68:245–251
- Moron JC, Singer JA, Sardi A (1998) Retrosternal goiter: a six-year institutional review. *Am Surg* 64:889–893
- Ozdemir A, Hasbahceci M, Hamaloglu E, et al. (2000) Surgical treatment of substernal goiter. *Int Surg* 85:194–197
- Parra-Membrives P, Sanchez-Blanco JM, Gomez-Rubio D, et al. (2003) Retrosternal goiters: safety of surgical treatment. *Int Surg* 88:205–210
- Rodriguez JM, Hernandez Q, Pinero A, et al. (1999) Substernal goiter: clinical experience of 72 cases. *Ann Otol Rhinol Laryngol* 108:501–504
- Shen WT, Kebebew E, Duh QY, et al. (2004) Predictors of airway complications after thyroidectomy for substernal goiter. *Arch Surg* 139:656–659
- Sancho JJ, Kraimps JL, Sanchez-Blanco JM, et al. (2006) Increased mortality and morbidity associated with thyroidectomy for intrathoracic goiters reaching the carina tracheae. *Arch Surg* 141:82–85
- Mack E (1995) Management of patients with substernal goiters. *Surg Clin North Am* 75:377–394
- Sosa JA, Bowman HM, Tielsch JM, et al. (1998) The importance of surgeon experience for clinical and economic outcomes from thyroidectomy. *Ann Surg* 228:320–330
- Pieracci F, Fahey TJ (2007) Substernal thyroidectomy, as compared to conventional, cervical thyroidectomy, is associated with increased morbidity and mortality. *J Am Coll Surg* 205:1–7
- Deyo RA, Cherkin DC, Ciol MA (1992) Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 45:613–619
- Kleinbaum DG, Kupper LL, Muller KE (1988) *Applied Regression Analysis and Other Multivariable Methods*. 2nd edition. PWS-Kent, Boston
- Callahan MA, Christos PJ, Gold HT, et al. (2003) Influence of surgical subspecialty training on in-hospital mortality for gastrectomy and colectomy patients. *Ann Surg* 238:629–636
- Hermanek P, Hohenberger W (1996) The importance of volume in colorectal cancer surgery. *Eur J Surg Oncol* 22:213–215
- Lavernia CJ, Guzman JF (1995) Relationship of surgical volume to short-term mortality, morbidity, and hospital charges in arthroplasty. *J Arthroplasty* 10:133–140
- Porter GA, Soskolne CL, Yakimets WW, et al. (1998) Surgeon-related factors and outcome in rectal cancer. *Ann Surg* 227:157–167
- Lien YC, Huang MT, Lin HC (2007) Association between surgeon and hospital volume and in-hospital fatalities after lung cancer resections: the experience of an Asian country. *Ann Thorac Surg* 83:1837–1843
- Harmon JW, Tang DG, Gordon TA, et al. (1999) Hospital volume can serve as a surrogate for surgeon volume for achieving excellent outcomes in colorectal resection. *Ann Surg* 230:404–411
- Panageas KS, Schrag D, Riedel E, et al. (2003) The effect of clustering of outcomes on the association of procedure volume and surgical outcomes. *Ann Intern Med* 139:658–665