Explaining the Variation in Surgical Practice for Differentiated Thyroid Cancer in Ontario, Canada

Stephen F. Hall, MSc, MD, FRCSC; Jonathan C. Irish, MSc, MD, FRCSC; Rebecca J. Griffiths, BSc; Marlo Whitehead, MSc, PSAT

IMPORTANCE In the management of differentiated thyroid cancer (DTC), the extent of surgical treatment required for most patients remains controversial and varies widely. This variation may be associated with the Enthusiasm Hypothesis, the notion that geographic differences in use of health care services are driven by the prevalence of physicians with a preference for particular services.

OBJECTIVE To evaluate the Enthusiasm Hypothesis and its applicability to the variation in the surgical treatment of thyroid cancer in Ontario, Canada.

DESIGN, SETTING, AND PARTICIPANTS This population-based study obtained electronic data on all health care–associated events for the complete population of insured residents of Ontario, Canada. Patients (n = 28 754) who were 18 years of age or older and underwent an initial therapeutic thyroid surgical procedure for a papillary or follicular carcinoma diagnosis between January 1, 2000, and December 31, 2015, were included. The final data analysis was performed January 31, 2019.

EXPOSURES Therapeutic thyroid cancer surgical procedure.

MAIN OUTCOMES AND MEASURES Numbers of surgeons, extent of surgical treatment, and case volumes in 14 geographic regions in Ontario.

RESULTS In total, 28 754 patients with DTC were included, of whom 22 600 (78.6%) were female, with a mean (SD) age of 49 (14) years. The use of total thyroidectomy for DTC varied widely across geographic regions of Ontario, from 45.8% to 77.1% of all cancer cases. More than 90% of the overall variation in practice (total thyroidectomy vs less-than-total thyroidectomy) could be explained by the practices of the high-volume surgeons (enthusiasts) in regions with the highest rates of thyroid cancer diagnosis.

CONCLUSIONS AND RELEVANCE The Enthusiasm Hypothesis appears to be consistent with the findings of this study. The practices of the high-volume surgeons who were enthusiastic for total thyroidectomy in the geographic regions with the highest rates of thyroid cancer diagnosis may account for the variation in practice across Ontario between 2000 and 2015.
the extent of surgical procedure for thyroid disease and specifically for differentiated thyroid cancer (DTC) has been debated for decades and has been associated with marked differences in treatment philosophy; however, to our knowledge, how those differences evolved has not been explored. Franci et al recently published a study on the 2014 rates of thyroidectomy for all causes across geographic US regions among Medicare beneficiaries 65 years or older. They reported a 6.2-fold difference in rates unrelated to health care availability, regional socioeconomic status, or surgeons per capita. Variations in surgical treatment rates have been reported for many conditions, and Francis et al found that the treatment of thyroid disease had a higher variation compared with prostate cancer treatment, suggesting the “widely divergent local beliefs and practice patterns” as the reason. Similar geographic variations in all aspects of thyroid cancer treatment, and specifically in the surgical management of DTC, across the regions of Ontario, Canada, have also been reported. These differences exist despite the publication by numerous organizations of management guidelines for thyroid disease, including cancer. As Wennberg stated, “When there is no evidence, treatment will vary,” and the lack of quality evidence plus the interpretation of the poor evidence by surgeons has created these differences or variations.

The Enthusiasm Hypothesis was proposed by Dr Mark Chassin as an explanation for the geographic variations in the utilization of specific health care services. Chassin compared the use of carotid endarterectomy for stroke prevention in US regions in 1981. At that time, carotid endarterectomy was controversial, with no good evidence and with known practice variation. Chassin found that the higher-use geographic regions had 8.2 times the number of high-volume surgeons (enthusiasts) compared with the lower-use regions, and he noted that “geographic differences in use of health care services are caused by differences in the prevalence of physicians who are Enthusiasts for particular services.” Enthusiasts, according to Chassin et al, were “believers,” “not uncertain,” and had “faith in the utility of the procedure.” Park et al suggested that variations in utilization rates were associated with the inappropriate overuse of procedures in the regions with the highest use of those procedures, and Chassin et al warned that inappropriate use can follow practitioner enthusiasm when “Enthusiasm can seduce a practitioner into believing in the effectiveness of a procedure where no data exists to support it.”

This current study examined the rates of total thyroidectomy by region in Ontario, Canada. Given the known regional variation in the extent of required surgical interventions for thyroid cancer, we defined enthusiasts as high-volume surgeons who prefer total thyroidectomy to less-than-total thyroidectomy.

The universal health care (Ontario Health Insurance Plan [OHIP]), the legislated health records collection, and the data set available from ICES (Institute of Clinical Evaluative Sciences) for the Province of Ontario create opportunities to conduct unique studies of large, complete, unselected patient populations. The practice patterns for the treatment of DTC across the health authority regions of Ontario have been previously studied. This study was designed according to the Chassin method to better understand practice patterns and variations.

**Key Points**

**Question** What theory or concept may explain the wide variation in the surgical management of differentiated thyroid cancer?

**Findings** In a population-based study of 28,754 patients with differentiated thyroid cancer in Ontario, Canada, the variation in the extent of required surgical treatment was associated with the practices of high-volume surgeons in regions with the highest rates of new cases of thyroid cancer. This finding is consistent with the Enthusiasm Hypothesis, which posits that the prevalence of physicians with a preference for certain health care services leads to geographic differences in the use of those services.

**Meaning** This study found that high-volume surgeons with enthusiasm for a particular treatment appeared to have a role in creating the variation in surgical practice across a large population of patients with thyroid cancer in Ontario.

**Methods**

**Data Set and Data Sources** This research was conducted at ICES Queen’s in Ontario, Canada. It received approval from the Queen’s University Research Ethics Board and the institutional review board at Sunnybrook Hospital. No informed consent was required by these review boards because the study used anonymized patient data.

The electronic data holdings of ICES include all health care-associated events for the complete population of insured Ontario residents, whose data are linked through anonymous unique identifiers. Linkages include all hospital encounters and all physician billing data for treatments, investigations, and procedures, thus presenting a clinical story for each anonymized patient. Approximately 3% of the population is not covered by the provincially funded OHIP, and we have no access to their information. These individuals include transients; tourists; and those covered by the Canadian federal health insurance such as active members of the armed forces, indigenous persons living on reservations, inmates of federal prisons, and some refugees.

We obtained data from several sources. First, the Ontario Cancer Registry (OCR) is a population-based cancer registry that captures information on all incident cases of cancer in Ontario. The OCR contains abstracted pathological reports from all hospitals and laboratories in Ontario, electronic records submitted by regional cancer centers, hospital discharge records, and death reports from the Office of the Registrar General in Ontario. The OCR also provides patient demographics, date of diagnosis, and vital status.

Second, the OHIP database contains data on all fee-for-service claims submitted by and paid to physicians, including office visits, consultations, laboratory tests, imaging tests, and biopsies. The OHIP data include the dates of all health care encounters, the referring physician information, and the physician type or specialty.

Third, the Registered Persons Database provides demographic information on all Ontario residents eligible for OHIP coverage.
Fourth, the Canadian Institute for Health Information includes the Discharge Abstract Database, which contains discharge data on all hospital admissions, including those for surgical procedures, and the National Ambulatory Care Reporting System, which is a data set of outpatient hospital visits.

**Study Setting, Population, and Variables**

For the purpose of funding and providing health care for 14 million people, Ontario was divided into 14 regions or Local Health Integration Networks (LHINs) according to health care utilization and political boundaries. Given Ontario's diverse geography, the LHINs vary in size, population (covering from 250,000 to >2 million residents), and urban-rural mix. At least 1 Canadian teaching hospital with an affiliated multidisciplinary cancer treatment center is located within all but 1 LHIN.

Using International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) code 193, Canadian Classification of Diagnostic, Therapeutic, and Surgical Procedures codes (CCP 19.1, 19.3, 19.21, and 19.29), and Canadian Classification of Health Interventions codes (CCI 1FU87 and 1FU89) indicated on hospital discharge data and OCR records, we identified 30,441 patients 18 years of age or older who underwent an initial therapeutic thyroid surgical procedure (hemi, subtotal, or total thyroidectomy) for papillary or follicular carcinoma between January 1, 2000, and December 31, 2015 (study period). We excluded patients who had only an open biopsy or nodulectomy and 1,687 patients who had only a completion thyroidectomy during the study period, leaving 28,754 patients for inclusion in the study.

The date of diagnosis was the date of the first therapeutic surgical procedure (index surgery). The index surgery was classified as either less- than-total thyroidectomy or total thyroidectomy. We did not include in the total thyroidectomy cohort any patients who had completion thyroidectomy within 1 month of less-than-total thyroidectomy.

The assignment of a patient’s LHIN and diagnosis was based on patient postal code. Treatment (total thyroidectomy as initial treatment) was reported according to the LHIN of the treatment facility, but because of LHIN boundaries, physician locations, and referral patterns, some patient migration was inevitable especially in urban regions.

**Statistical Analysis**

The final analysis was performed January 31, 2019. We calculated the mean diagnosis rates of thyroid cancer per 100,000 population (the total number of cases among the total population) for each LHIN during the 16-year study period. The 95% CIs were based on the mean diagnosis rates, assuming a Poisson distribution. Age and sex adjustments were not practical because of the small number of cases per year in some LHINs and because the rates were based on the mean across 16 years.

Using surgeon billing data, we calculated the total thyroidectomy case volume for all surgeons who performed at least 1 thyroid cancer operation during the study period. We also created thyroid cancer case-volume categories (similar to the Chassin method) for the surgeons: occasional (<10), low (10–49), moderate (50–99), high (100–149), and highest (≥150). For this study, we defined enthusiast surgeons as those with more than 100 cases.

The number of thyroid cancer surgeons per 100,000 population was calculated for each LHIN and for each LHIN’s case-volume categories (<10, 10–49, 50–99, 100–149, and ≥150). The LHINs were grouped into high or low thyroid cancer rates, and we followed the method described in the Chassin article to identify and describe the extent of variance in practice between regions. We calculated (1) the number of surgeons per 100,000 population by case volume for the combined thyroid cancer rate regions and (2) the rates of total thyroidectomy per 100,000 population for each of the case-volume categories for the combined thyroid cancer rate regions. In this way, the practices of similar-volume surgeons across regions could be compared, and the ratio could represent the variation in required surgical practice.

**Results**

We identified 28,754 patients with DTC (papillary or follicular) who underwent a therapeutic surgical procedure in Ontario between January 1, 2000, and December 31, 2015. Of these included patients, 22,600 (78.6%) were female, and the mean (SD) age was 49 (14) years.

**Figure 1** presents (in order from largest to smallest) the mean DTC case rate or incidence per 100,000 population over the 16 years by LHIN and with 95% CIs. The LHINs were divided into high-rate (LHIN 5, 6, 7, 8, and 9) and low-rate (LHIN 1, 2, 3, 4, 10, 11, 12, 13, and 14) groups. The mean rate of thyroid cancer diagnosis for the low-rate group was 11.97 per 100,000 people (95% CI, 11.73–12.21) and was 24.81 per 100,000 population (95% CI, 24.45–25.16) for the high-rate group, which was 2.07 times that of the low-rate group.

We identified 343 surgeons who performed at least 1 total thyroidectomy for thyroid cancer in the 16-year period. **Figure 2** presents the number of surgeons performing total thyroidectomy for cancer in each of 5 case-volume categories. Of these surgeons, 165 (48.1%) performed fewer than 10 total thyroidectomy procedures, and 32 (9.3%) performed more than 150 procedures. Overall, 257 (74.9%) of 343 surgeons performed fewer than 50 procedures, and 47 (13.7%) performed more than 100 procedures.

**Figure 3** presents and compares the extent of treatment across the LHINs of Ontario. **Table 1** presents the number of surgeons per 100,000 population by regions with high and low thyroid cancer rates. Overall, 3.49 surgeons per 100,000 population were providing thyroid cancer surgical procedures across Ontario, but 5.81 times as many highest-volume surgeons (enthusiasts) were operating in the regions with the higher rates of thyroid cancer.

**Table 2** presents the number of total thyroidectomy cases per 100,000 population for all of the surgeons in each of the case-volume categories by regions with high and low thyroid cancer rates. The total thyroidectomy rate for the region with the highest thyroid cancer rate was 187.26, and the difference from the lower-rate region was 164 per 100,000 population (95% CI, 187–23). The overall difference for all cases was 182 per 100,000 population (95% CI, 274–92), and the ratio of 164 per 182 represented the proportion of the overall variance associated with the highest-volume surgeons. However, the low-rate region had 2.07 fewer thyroid cancer cases, and when the
rate of total thyroidectomy was adjusted for the low-rate region, the proportion of the overall variance associated with the high-volume surgeons was more than 100%. The practice of the high-volume surgeons in the regions with the highest thyroid cancer rate may explain the more than 90% variation in surgical treatment for thyroid cancer across the province.

Discussion

The objective of this study was to use the Chassin method to investigate the implication of enthusiast surgeons for the practice patterns for DTC across Ontario. We found wide variations in DTC diagnosis rates and in the extent of surgical procedures for DTC that were similar to findings in previous publications. We also found that most surgeons treating thyroid cancer performed low volumes of total thyroidectomy, but 13.7% of surgeons had high volumes of total thyroidectomy. The overall numbers of surgeons per 100 000 population did not vary across Ontario, but the number of high-volume surgeons in the regions with the highest thyroid cancer rate may explain the more than 90% variation in surgical treatment for thyroid cancer across the province.

Chassin proposed that geographic variations in health care utilization could be explained by the differences in the prevalence of physicians who were enthusiasts for particular services, as demonstrated by a large study of patients who underwent carotid endarterectomy in the United States in 1981. At that time, carotid endarterectomy was a common operation for the prevention of stroke but was not supported by evidence. Chassin found that the US regions with the highest rates of carotid endarterectomy also had a larger number of high-volume surgeons who performed the procedure and that 77% of the variation in practice was associated with the number of these region-based surgeons. From 2000 to 2015, no rigorous evidence existed to support total thyroidectomy over less-than-total thyroidectomy for most cases of thyroid cancer; instead, practice was influenced by entrenched beliefs, institutional biases, consensus-based guidelines of the day, and physician factors such as those described by Wennberg, Birkmeyer et al, and Saini et al.
Other factors may explain practice variation and our findings. One explanation could be referral bias, which assumes that complex cases requiring extensive surgical interventions are referred to more experienced surgeons in larger centers. Referral bias, however, is unlikely given that experienced thyroid surgeons (otolaryngologist or head and neck surgeons, general surgeons, and surgical oncologists) were on staff at all of the teaching centers in all LHINs except for 1. Furthermore, based on findings from previous Ontario-based work that 61% of tumors were less than 2 cm, the increasing use of diagnostic ultrasonography, articles on tumor size evolution over time, and the principles of overdiagnosis, most of the patients in this study likely had tumors that were less than 2 cm and not complex. We also found that tumor size was not a factor in the rates of total thyroidectomy.

Patient treatment migration could also explain these results. In a previous report based on a subset of this study’s population, the LHIN of residence was compared with the LHIN of index surgery to assess patient migration. Some migration of patients occurred to and from most LHINs, and the regions with low thyroid cancer rate treated more than 80% of their residents. The LHINs with high cancer rates had more patient migration to and from adjacent LHINs, and the LHINs with the greatest number of major hospitals treated almost twice their resident case volumes from the adjacent LHINs. Patient migration in this study can be explained by geographic-based LHIN boundaries, especially in urban centers, as well as established referral practices and access, but it is an unlikely factor in this study’s results.

Uncertainty about the evidence and about the various consensus-based guidelines of the day is more likely to help create the wide variations in practice. Eddy’s uncertainty theory was based on the concept of a large gray area between appropriateness and inappropriateness that left physicians uncertain about indications for interventions. Uncertain physicians would then follow the enthusiasts, who often became opinion leaders, thus generating geographic variations in practice such as those found in this study of Ontario populations and in Francis et al’s study of Medicare beneficiaries in the United States. The 2015 version of the American Thyroid Association Guidelines on the management of thyroid cancer has proposed de-escalation in the extent of surgical procedure for DTC, suggesting that many patients may be (or may have been) candidates for less-than-total thyroidectomy. However, Chassin hypothesized that enthusiasts’ faith may be difficult to shake even when good evidence emerged. Because most of the recommendations of the new guidelines may be based on lower-level evidence, surgical practice needs to be monitored for compliance with these guidelines. Furthermore, by recognizing enthusiasts’ role as opinion leaders in practice variation and as possible experts on uncertainty, current health care leaders, funders, and policy makers have the potential to reduce variation.

Strengths and Limitations
The major strength of this study was it was made possible by the universal health care in Ontario, the legislated collection of data from multiple sources on all patients, and the availability of centralized administrative data sets at ICES. The major weakness of this study was that electronic data of neither tumor size nor extent of disease were available in Ontario for the study period. A second limitation was that we did not account for new, retiring, migrated, or multi-center surgeons, and therefore the case volumes of some surgeons were low. The missing numbers were likely small but would mean that we may have underestimated the numbers of higher-volume surgeons and therefore the variance. A third limitation was that we have not reported outcomes. According to a previous report, however, the outcomes of overall survival and disease-specific survival in Ontario did not differ on the basis of regional variations in the extent of treatment for DTC.

Conclusions
This study found that, in the absence of high-quality evidence on the extent of surgical procedure in DTC, the variation in practice across Ontario, Canada, between 2000 and 2015 may be associated with the practices of the high-volume surgeons who were enthusiastic for total thyroidectomy in geographic regions with the highest rates of thyroid cancer. High-quality evidence is needed by patients, physicians, and policy makers on the outcomes of total thyroidectomy and less-than-total thyroidectomy for DTC, especially as higher numbers of cases are increasing demand for treatment.
Variation in Surgical Practice for Differentiated Thyroid Cancer

Redirecing, Without Dampening, the Enthusiasm of Surgeons

Amy Y. Chen, MD, MPH

It was J. Alison Glover who first noted variation in rates of tonsillectomy across Britain more than 75 years ago.1 He noted that the variation was because of the school health officers who would refer (or not refer) students for tonsillectomy. Fast forward to Wennberg and Gittelsohn’s article2 35 years later that again documented variation in rates of tonsillectomy, this time across Vermont. In this issue of JAMA Otolaryngology-Head & Neck Surgery, Hall and colleagues3 report on high variability in rates of thyroidectomy within the province of Ontario, Canada. They propose that Chassin’s enthusiasm hypothesis4 may help explain high rates of thyroidectomy in certain geographic regions of Ontario.

Wenbenberg and Gittelsohn,2 and later Birkmeyer,5 have published extensively on the potential causes of regional variation in rates of surgical procedures. In short, different attitudes and beliefs about the indications for surgery are the most important reasons for regional variation. But how can this be in thyroid surgery? We have numerous guidelines published regarding indications for surgery (hemithyroidectomy, total thyroidectomy) or not (active surveillance) from national and international organizations.

REFERENCES

4. Hall SF, Irish JC, Groome P, Hurlbut D. Do lower-risk thyroid cancer patients who lived in regions with more aggressive treatments have better outcomes? Thyroid. 2017;27(10):1246-1257. doi:10.1089/thy.2017.0103