Association between surgical case volume and survival in T1 bladder cancer: A need for centralization of care?

CUA PRIZE ESSAY

Marian S. Wettstein, MD[1,2,3]; Song Pham, MSc[2]; Syed R. Qadri, MHI[1]; Kathy Li, MPH[1]; Christian D. Fankhauser, MD[3]; Benedikt Kranzbühler, MD[3]; Ning Liu, PhD[2]; Theodorus van der Kwast, MD, PhD[4]; Thomas Hermanns, MD[3]; Girish S. Kulkarni, MD, PhD[1,2]

[1]Division of Urology, Department of Surgery, Princess Margaret Cancer Centre, University Health Network, University of Toronto, Toronto, ON, Canada; [2]ICES, Toronto, ON, Canada; [3]Department of Urology, University Hospital of Zurich, University of Zurich, Zurich, Switzerland; [4]Department of Pathology, Princess Margaret Cancer Centre, University Health Network, University of Toronto, Toronto, ON, Canada

Cite as: Wettstein MS, Pham S, Qadri SR, et al. Association between surgical case volume and survival in T1 bladder cancer: A need for centralization of care? *Can Urol Assoc J* 2020;14(9):E394-401. http://dx.doi.org/10.5489/cuaj.6812

Appendix available at *cuaj.ca*

Abstract

Introduction: Prior research demonstrated an association between surgeon case volume and survival in muscle-invasive bladder cancer (BC). This relationship, however, has not been investigated in the setting of high-risk, non-muscle-invasive BC (NMIBC). Hence, we investigated whether a higher surgeon case volume of T1 BC translates into improved survival outcomes.

Methods: Province-wide pathology reports (January 2002 to December 2015) were linked with health administrative data to identify patients diagnosed with T1 BC. For each patient, we determined the T1 case volume of the involved surgeon by benchmarking (percentile) her/him against his/her colleagues during a lookback period of one year. The volume-outcome (overall survival) relationship was then investigated by Cox proportional hazards regression (unadjusted and adjusted for a wide range of assumed confounders) that incorporated volume in three different ways (≥80th percentile vs. below, ≥ median vs. below, continuous [quintiles]). Effect sizes were presented as hazard ratios (95% confidence interval).

Results: We identified 7426 patients who were diagnosed with T1 BC and followed for a median of 4.8 years. A third of all patients (n=1895, 25.5%) received surgery by a high-volume surgeon (80th percentile and higher). Higher T1 case volume was associated with improved survival both in unadjusted (80th percentile: 0.93 [0.86–0.99]; median: 0.93 [0.87–0.99]; continuous: 0.97 [0.94–0.99]) and adjusted analysis (80th percentile: 0.94 [0.88–1.01]; median: 0.93 [0.87–0.99]; continuous: 0.97 [0.95–0.99]) regardless of the method by which volume was analyzed.

Conclusions: This population-based cohort study demonstrated a volume-outcome relationship in T1 BC and raises questions regarding the centralization of care in high-risk NMIBC.

Introduction

T1 bladder cancer (BC) is a highly aggressive malignancy with up to 34% of all patients progressing to muscle-invasion and a cancer-specific mortality of up to 20% by year 5.[1,2] Experienced urologic surgeons are not only required to perform a high-quality initial transurethral resection of the bladder tumor (TURBT) but also for downstream care, namely to provide re-resection, adequately prescribe intravesical immunotherapy, provide rigorous followup, and identify patients who benefit from immediate or early radical cystectomy.[3,4]

There is prior evidence that higher surgeon volumes improve long-term survival in patients undergoing radical cystectomy[5] and centralization of radical cystectomies is often discussed.[6,7] While the need for high-volume, expert care for muscle-invasive BC is clear because of the multidisciplinary nature of the disease and the complexity of cystectomy, the need for high volumes in non-muscle-invasive BC (NMIBC) is probably underappreciated. Yet, optimal decision-making and surgical skill in T1 BC is paramount, as these patients have the most to lose from both over- and under-treatment. Specifically, surgeons with the highest expertise in managing T1 BC are most likely to be best positioned to navigate complex oncological decisions pertaining to re-resection, intravesical therapies, identifying intravesical therapy failures, and providing optimal radical cystectomy, both from a timing and technical perspective. Thus, we hypothesize that higher surgeon volumes in T1 BC will translate into improved survival outcomes relative to lower volumes. To our knowledge, this association has not been investigated in the setting of T1 BC.

© 2020 Canadian Urological Association

Methods

Design and setting

We performed a retrospective, observational, population-based cohort study in the province of Ontario, Canada (population 14.7 million).[8] Manually abstracted, province-wide pathology reports were linked with health administrative data: 1) to identify patients diagnosed with incident T1 BC; (2) to ascertain their survival outcomes; 3) to determine the case volume of the surgeon; and 4) and to measure potential confounders. These datasets were linked using unique encoded identifiers and analyzed at ICES (Toronto, ON, Canada). Detailed data sources, codes, and definitions can be found in the Appendix (available at *cuaj.ca*). Institutional research ethics board approval was obtained and reporting is in accordance with the REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement.[9]

Patients

Cancer Care Ontario (Toronto, ON, Canada) provided province-wide BC pathology reports (collected between January 1, 2001 and December 31, 2015) that were manually abstracted by trained data abstractors. We linked reports classified as T1 BC to health administrative data sources. We then identified the first occurrence of a T1 tumor for each patient to restrict our analyses to incident cases. We further excluded individuals: 1) not covered by the Ontario Health Insurance Plan during the three years preceding the diagnosis; 2) older than 105 years or younger than 18 years; 3) diagnosed by a procedure other than TURBT; and 4) with a non-urothelial histology. In accordance with our expert uropathologist (T.v.d.K), we further excluded T1LG/G1 tumors to mitigate potential staging errors.

Ascertainment of survival outcomes

Overall survival was verified by the Registered Persons Database (RPDB). Patients were censored at their date of last contact with the healthcare system (maximal followup until September 30, 2019). Cancer-specific survival was ascertained by the Office of the Registrar General-Deaths (ORGD) database. For each deceased patient, this database provides the leading cause of death according to the methodology described by Becker.[10] The main input source of the ORGD database is death certificates that are translated into leading causes of death by professional data abstractors. Patients were censored if they died due to another leading cause than BC or at their last contact with the healthcare system (maximal followup in OGRD database to December 31, 2016).

Ascertainment of surgeon case volume

For each patient, we determined the T1 BC case volume of the involved surgeon during a lookback window of one year. To provide such a lookback window for each patient, we started our study cohort on January 1, 2002 (BC pathology reports were available from 2001 to enable the lookback). Next, the case volume during the lookback window of a specific surgeon was benchmarked (percentile rank) against the case volumes of her/his colleagues who were surgically active during the exact same lookback window. We used this standardization approach with a one-year lookback window to prevent the potential for a spurious volume-outcome association caused by the increasing incidence of BC, and thus increased number of incident pathology reports, over time. In the final analysis, surgeon case volume was investigated in three different ways, specifically ≥80th percentile vs. below, ≥median vs. below, and continuous (quintiles).

Ascertainment of assumed confounders

Besides age and sex, we hypothesized that any volume-outcome relationship would be confounded by the domains "tumor," "comorbidity/healthcare system utilization," and "socio-economics." The domain "tumor" was accounted for by histology, grade, sufficiently sampled muscularis propria, concurrent carcinoma in situ, lymphovascular invasion, tumor size/multiplicity (billing claim-based), and prior urinary tract cancer. The domain "comorbidity/healthcare system utilization" was captured by the Charlson comorbidity index (lookback window of three years),[11] The Johns Hopkins Adjusted Clinical Groups® (version 10; hereinafter called ACG®; lookback window of one year)[12] resource utilization band, ACG® frailty indicator, prevalence (year before diagnosis) of asthma, congestive heart failure, chronic obstructive pulmonary disease, inflammatory bowel disease, dementia, diabetes mellitus, hypertension, psychiatric comorbidity, or rheumatoid arthritis; prior cancer diagnosis, myocardial infarction, and number of home care claims/drug claims/physician billing claims during year preceding the diagnosis. Rurality index, postal code-based socio-economic status, marginalization (captured by the postal code-based Ontario Marginalization Index measuring residential instability/material deprivation/dependency/ethnic concentration[13]), and ethnicity[14] accounted for the domain "socio-economics."

Analysis

We used the absolute standardized difference of the mean to compare baseline characteristics between patients who were operated on by high-volume (80th percentile and higher) and low-volume (below 80th percentile) surgeons.[15] The crude association between surgeon case volume and

overall/cancer-specific survival was investigated by Kaplan-Meier curves/log-rank tests and univariable Cox proportional hazards regression, while the adjusted association was estimated by multivariable Cox proportional hazards regression (adjusted for the above-mentioned confounding variables). Effect sizes were presented as hazard ratios (95% confidence interval) and $p<0.05$ was considered statistically significant (two-sided). All analyses were performed in R 3.5.1 (R Foundation for Statistical Computing, Vienna, Austria).

Results

Fig. 1 shows the derivation of our study cohort. Out of 15 262 linked T1 BC pathology reports, we identified 7426 patients who fulfilled the inclusion criteria. After benchmarking each surgeon against her/his colleagues who were surgically active during the same lookback window, patients were classified as follows: using the cutoff at the 80th percentile, a quarter of all patients (n=1895; 25.5%) received surgery by a high-volume surgeon. Using the cutoff at the median, 59.5% of all patients (n=4417) received surgery by a high-volume surgeon.

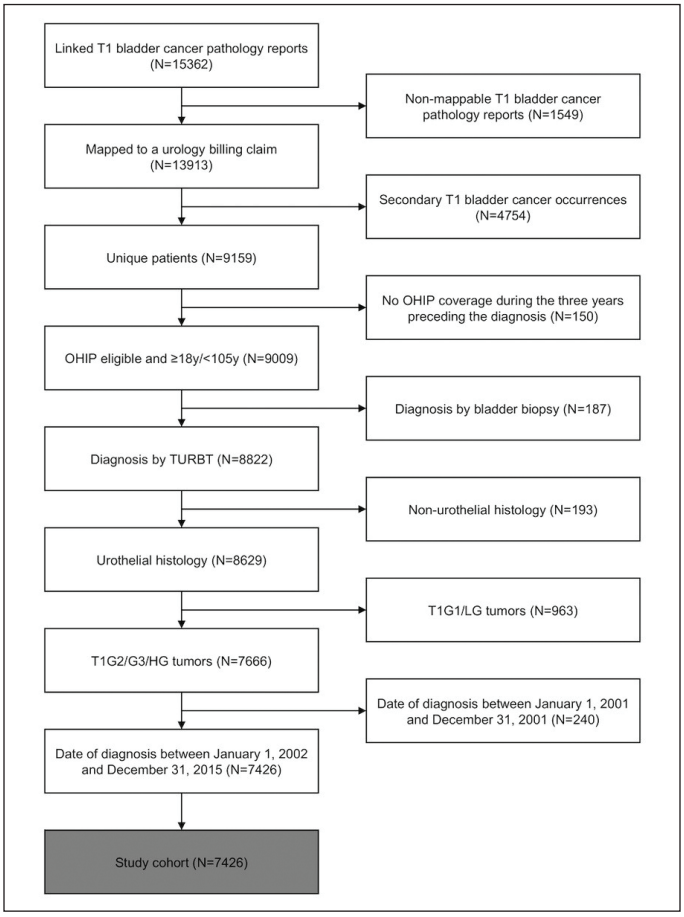

Fig. 1. Derivation of study cohort. OHIP: Ontario Health Insurance Plan; TURBT: transurethral resection of bladder tumor; y: years.

Baseline characteristics of the study cohort are presented in Table 1. Median age at diagnosis was 74 years (interquartile range [IQR] 66–81) and 21.9% of all patients were female. Pure urothelial histology was found in 87.3%, followed by urothelial carcinoma with squamous differentiation (4.8%) and glandular differentiation (3.8%). In close to three-quarters of all patients (72.9%), tumor grade was reported by the World Health Organization (WHO) 2004/2016 classification. Of the remaining patients reported by the WHO 1973 classification, 59.6% were diagnosed with a poorly differentiated (G3) tumor, while 40.4% had a moderately differentiated (G2) T1 BC. Sufficiently sampled muscularis propria, concurrent carcinoma in situ, and lymphovascular invasion were reported in 65.1%, 15.5%, and 6.0% of all patients, respectively. When we stratified the baseline characteristics by surgeon case volume (cutoff at 80th percentile), we observed a mostly well-balanced cohort with regards to the absolute standardized difference of means. Only grade (0.103), sufficiently sampled muscularis propria (0.106), number of tumors (0.135), and rurality index (0.123) showed marginal imbalances.

Followup information was available for 7424 of 7426 included patients (>99.9%). Median (IQR) followup for overall survival was 4.8 years (2.2–8.0). Out of 7424 patients, 4346 patients (58.5%) died by the end of the followup period. Standardized to 10 000 person years of followup, the mortality rate was lower in the group of patients who received surgery by a high-volume surgeon (2.61 deaths; 80th percentile and higher) in comparison to patients who received surgery by a low-volume surgeon (2.83 deaths; below 80th percentile). The corresponding Kaplan-Meier curves are presented in Fig. 2 and show a statistically significant log-rank test (p=0.04). Overall survival at five, 10, and 15 years in the high-volume vs. the low-volume group was 60.0%, 40.8%, and 26.3% vs. 57.9%, 38.3%, and 24.8%, respectively. The corresponding univariable Cox proportional hazards regression shows (Table 2) that patients who received surgery by a high-volume (80th percentile and higher) in comparison to a low-volume (below 80th percentile) surgeon were 7% less likely to die (hazard ratio [HR] 0.93 [0.85–0.99], p=0.04). This effect could be confirmed when we used the median as the cutoff (HR 0.93 [0.87–0.99], p=0.02) and when we incorporated surgeon case volume in a continuous fashion (quintiles; HR 0.97 [0.95–0.99], p=0.04).

For overall survival, the adjusted analysis could confirm the effect sizes estimated during univariable modelling, although the dichotomization at the 80th percentile did not reach statistical significance whereas it did for the other two volume methods (80th percentile and higher vs. below: HR 0.94 [0.89–1.01], p=0.12; median and higher vs. below: HR 0.93 [0.87–0.99], p=0.02; continuous (quintiles): HR 0.97 [0.95–0.99], p=0.009) (Table 2).

Median (IQR) followup for cancer-specific survival was 3.4 years (1.6–6.6). Out of 7426 patients, 1236 (16.1%) died because of BC, 2319 (30.3%) died due to other causes, and

Table 1. Baseline characteristics of study cohort (n=7246), overall and stratified by surgeon case volume (80th percentile and higher vs. below 80th percentile)

Baseline characteristics	Overall	High-volume (≥80th percentile)	Low-volume (<80th percentile)	ASDM
	(n=7426)	(n=1895)	(n=5529)	
Demographics				
Age in years, median (IQR)	74 (66–81)	74 (66–81)	74 (66–81)	0.006
Female, n (%)	1623 (21.9%)	414 (21.8%)	1209 (21.9%)	<0.001
Tumor				
Histology, n (%)				0.015
Pure urothelial carcinoma	6484 (87.3%)	1657 (87.4%)	4825 (87.3%)	
Squamous differentiation	355 (4.8%)	93 (4.9%)	262 (4.7%)	
Glandular differentiation	279 (3.8%)	69 (3.6%)	210 (3.8%)	
Other urothelial histology	308 (4.1%)	76 (4.0%)	232 (4.2%)	
Grade, n (%)				0.103*
Poorly differentiated/G3 (WHO 1973)	1199 (16.1%)	291 (15.4%)	907 (16.4%)	
Moderately differentiated/G2 (WHO 1973)	813 (10.9%)	254 (13.4%)	559 (10.1%)	
High-grade (WHO 2004/2016)	5414 (72.9%)	1350 (71.2%)	4063 (73.5%)	
Muscularis propria, n (%)				0.106*
Sampled	4836 (65.1%)	1210 (63.9%)	3624 (65.5%)	
Not sampled	2218 (29.9%)	615 (32.5%)	1603 (29.0%)	
Not reported	372 (5.0%)	70 (3.7%)	302 (5.5%)	
Concurrent CIS, n (%)				0.026
Present	1150 (15.5%)	280 (14.8%)	869 (15.7%)	
Absent/not reported	6276 (84.5%)	1615 (85.2%)	4660 (84.3%)	
LVI, n (%)				0.011
Present	449 (6.0%)	111 (5.9%)	338 (6.1%)	
Absent/not reported	6977 (94.0%)	1784 (94.1%)	5191 (93.9%)	
Tumor size/multiplicity (billing claim-based), n (%)				0.081
Multiple tumors	3865 (52.0%)	930 (49.1%)	2934 (53.1%)	
Single tumor >2 cm	3307 (44.5%)	892 (47.1%)	2414 (43.7%)	
Single tumor up to 2 cm	254 (3.4%)	73 (3.9%)	181 (3.3%)	
Number of tumors (pathology report-based), median (IQR)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	1.00 (1.00–1.00)	0.135*
Missing	1467 (19.8%)	470 (24.8%)	995 (18.0%)	
Maximal tumor dimension in cm (pathology report-based), median (IQR)	2.00 (1.00–3.00)	1.70 (1.00–3.00)	2.00 (1.00–3.00)	0.050
Missing	3914 (52.7%)	963 (50.8%)	2949 (53.3%)	
Tumor weight in g (pathology report-based), median (IQR)	4.00 (1.80–9.00)	4.00 (2.00–8.73)	4.00 (1.60–9.00)	0.007
Missing	4455 (60.0%)	1159 (61.2%)	3295 (59.6%)	
Tumor volume in ml (pathology report-based), median (IQR)	2.50 (1.00–5.00)	3.00 (1.00–5.00)	2.50 (1.00–5.00)	0.071
Missing	6501 (87.5%)	1649 (87.0%)	4851 (87.7%)	
Prior urinary tract cancer, n (%)	1135 (15.3%)	279 (14.7%)	856 (15.5%)	0.021
Comorbidity/healthcare system utilization				
Charlson comorbidity index, mean (SD)[a]	1.48 (3.72)	1.44 (3.66)	1.31 (3.46)	0.046
ACG® resource utilization band, n (%)				0.075
Low morbidity	45 (0.6%)	15 (0.8%)	30 (0.5%)	
Moderate morbidity	3251 (43.8%)	874 (46.1%)	2377 (43.0%)	
High morbidity	2482 (33.4%)	614 (32.4%)	1867 (33.8%)	
Very high morbidity	1648 (22.2%)	392 (20.7%)	1255 (22.7%)	

*Absolute standardized difference of means ≥0.1 (indicative of unbalanced cohort). [a]Mean (SD) was chosen instead of median (IQR) due to more informative numerical appearance. [b]Postal code-based income quintiles (first quintile: lowest income). [c]Postal code-based marginalization quintiles (first quintile: least marginalized patients). ACG: Adjusted Clinical Groups®; ASDM: absolute standardized difference of means; CIS: carcinoma in situ; IQR: interquartile range; LVI: lymphovascular invasion; SD: standard deviation; WHO: World Health Organization .

Table 1 (cont'd). Baseline characteristics of study cohort (n=7246), overall and stratified by surgeon case volume (80th percentile and higher vs. below 80th percentile)

Baseline characteristics	Overall	High-volume (≥80th percentile)	Low-volume (<80th percentile)	ASDM
	(n=7426)	(n=1895)	(n=5529)	
Frailty (ACG®), n (%)	409 (5.5%)	96 (5.1%)	313 (5.7%)	0.026
Asthma, n (%)	869 (11.7%)	214 (11.3%)	655 (11.8%)	0.017
Congestive heart failure, n (%)	840 (11.3%)	232 (12.2%)	608 (11.0%)	0.039
Chronic obstructive pulmonary disease, n (%)	1997 (26.9%)	488 (25.8%)	1508 (27.3%)	0.034
Inflammatory bowel disease, n (%)	62 (0.8%)	15 (0.8%)	47 (0.9%)	0.006
Dementia, n (%)	373 (5.0%)	83 (4.4%)	290 (5.2%)	0.040
Diabetes mellitus, n (%)	2113 (28.5%)	541 (28.5%)	1572 (28.4%)	0.003
Hypertension, n (%)	4987 (67.2%)	1260 (66.5%)	3726 (67.4%)	0.019
Prior cancer diagnosis, n (%)	1242 (16.7%)	322 (17.0%)	920 (16.6%)	0.009
Prior myocardial infarction, n (%)	306 (4.1%)	85 (4.5%)	220 (4.0%)	0.025
Psychiatric comorbidity, n (%)	1590 (21.4%)	403 (21.3%)	1187 (21.5%)	0.005
Rheumatoid arthritis, n (%)	136 (1.8%)	38 (2.0%)	97 (1.8%)	0.018
Home care claims during preceding year, mean (SD)a	6.42 (31.1)	5.77 (29.3)	6.64 (31.7)	0.028
Drug claims during preceding year, median (IQR)	18.0 (3.00–35.0)	17.0 (3.00–35.5)	18.0 (3.00–35.0)	0.021
Physician billing claims during preceding year, median (IQR)	67.0 (43.0–102)	66.0 (42.0–100)	67.0 (44.0–103)	0.036
Socio-economics				
Rurality index, mean (SD)[a]	13.0 (18.0)	14.6 (19.0)	12.4 (17.6)	0.123*
Missing	56 (0.8%)	12 (0.6%)	44 (0.8%)	
Socio-economic status as quintile, mean (SD)[a,b]	3.04 (1.40)	3.07 (1.39)	3.03 (1.40)	0.025
Missing	27 (0.4%)	7 (0.4%)	20 (0.4%)	
Residential instability as quintile, mean (SD)[a,c]	3.29 (1.37)	3.29 (1.35)	3.29 (1.38)	0.006
Missing	83 (1.1%)	16 (0.8%)	67 (1.2%)	
Material deprivation as quintile, mean (SD)[a,c]	2.99 (1.40)	2.92 (1.39)	3.02 (1.40)	0.072
Missing	83 (1.1%)	16 (0.8%)	67 (1.2%)	
Dependency as quintile, mean (SD)[a,c]	3.47 (1.39)	3.54 (1.39)	3.45 (1.39)	0.061
Missing	83 (1.1%)	16 (0.8%)	67 (1.2%)	
Ethnic concentration as quintile, mean (SD)[a,c]	2.81 (1.40)	2.64 (1.35)	2.86 (1.41)	0.084
Missing	83 (1.1%)	16 (0.8%)	67 (1.2%)	
Ethnicity, n (%)				0.084
General population	7245 (97.6%)	1866 (98.5%)	5377 (97.3%)	
Chinese	121 (1.6%)	19 (1.0%)	102 (1.8%)	
South Asian	60 (0.8%)	10 (0.5%)	50 (0.9%)	

*Absolute standardized difference of means ≥0.1 (indicative of unbalanced cohort). [a]Mean (SD) was chosen instead of median (IQR) due to more informative numerical appearance. [b]Postal code-based income quintiles (first quintile: lowest income). [c]Postal code-based marginalization quintiles (first quintile: least marginalized patients). ACG: Adjusted Clinical Groups®; ASDM: absolute standardized difference of means; CIS: carcinoma in situ; IQR: interquartile range; LVI: lymphovascular invasion; SD: standard deviation; WHO: World Health Organization .

3871 (53.6%) were censored. As observed for overall survival, the cancer-specific mortality rate was lower in the group of patients who received surgery by a high-volume surgeon (0.91 deaths per 10 000 person years of followup; 80th percentile and higher) in comparison to patients who received surgery by a low-volume surgeon (1.03 deaths per 10 000 followup years; below 80th percentile). The corresponding Kaplan-Meier curves (Fig. 3) show a clear separation that was not statistically significant (p=0.12). Cancer-specific survival at five and 10 years in the high-volume vs. the low-volume group was 82.8% and 78.8% vs. 81.2% and 76.2%, respectively. Univariable Cox proportional hazards regression analysis evaluating the association between cancer-specific survival and surgeon case volume consistently confirmed the effect sizes observed with overall survival analyses but without reaching statistical significance (80th percentile and higher vs. below: HR 0.90 [0.79–1.03], p=0.12; median and higher vs. below: HR 0.99 [0.88–1.12], p=0.91; continuous (quintiles): HR 0.97 [0.93–1.01], p=0.20) (Table 2).

The effect sizes observed in unadjusted analysis of cancer-specific survival could be confirmed in the multivariable models (80th percentile and higher vs. below: HR 0.92 [0.80–1.05], p=0.23; median and higher vs. below: HR 0.97 [0.86–1.09], p=0.60; continuous (quintiles): HR 0.97 [0.93–1.02], p=0.22), although these analyses were not statistically significant (Table 2).

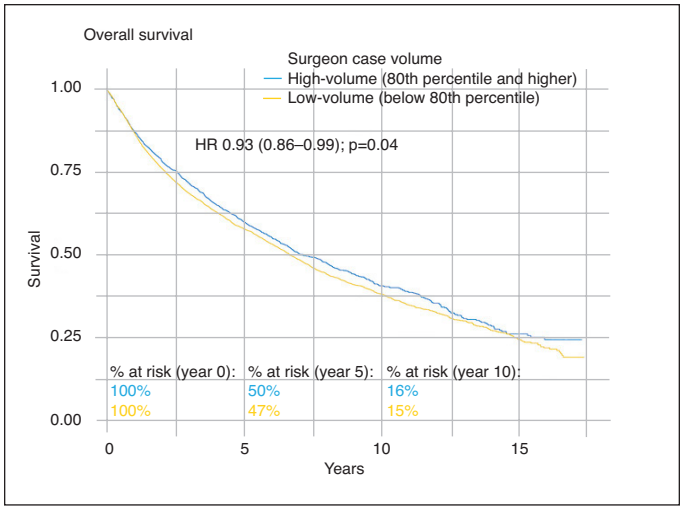

Fig. 2. Kaplan-Meier curve and log-rank test comparing overall survival between patients who received surgery by a high-volume surgeon (≥80th percentile and higher; blue curve) and patients who received surgery by a low-volume surgeon (<80th percentile; yellow curve). HR: hazard ratio.

Discussion

This population-based cohort study demonstrated an independent association between surgeon case volume and overall survival in T1 BC. The minimal discrepancy between unadjusted and adjusted analysis is not surprising in light of a well-balanced cohort. Although the analysis of cancer-specific survival mostly reproduced the effect sizes observed for overall survival, we were not able to demonstrate statistical significance. This is probably caused by limited power due to two effects: first, the event rate for cancer-specific mortality is much lower than the event rate for overall mortality; second, followup for cancer-specific mortality was only available until December 31, 2016 since the manual abstraction of death certificates lags behind overall death ascertainment by two to three years in ICES databases.At any time of followup, patients who received surgery by a high-volume in comparison to a low-volume surgeon were between 1% and 10% less likely to die from any cause or from BC. It is generally known that population-based, high-powered studies like ours bear the risk to detect statistically significant but clinically irrelevant differences. Hence, the clinical relevance of absolute differences of 0.22 deaths (2.61 vs. 2.83) and 0.12 cancer-specific deaths (0.91 vs. 1.03) per 10 000 person years of followup is unknown. Our effect sizes are comparable to the volume-outcome results reported by Kulkarni et al in patients undergoing radical cystectomy in the province of Ontario (unadjusted HR [continuous] 0.98 [0.97–0.99], p<0.001; adjusted HR [continuous] 0.98 [0.98–0.99], p=0.002).[5] The comparability between this study and ours might be questioned since radical cystectomy is a procedure with a high peri-/postoperative morbidity and mortality discrepant to the safety profile of a TURBT, which may have theoretically driven the observed cystectomy volume-outcome results.[16] However, the authors performed sensitivity analyses that were robust to the exclusion of peri-/postoperative mortality.[5] Thus, data from both of these studies point out the importance of initial, upfront clinical decision-making regarding long-term outcomes in patients with high-risk BC.

The main strengths of our study are its province-wide nature involving more than 7000 patients, its long follow-up duration, and its rich set of confounders (ranging from tumor-specific parameters to patient comorbidity/healthcare system utilization and socio-economic indices) that could be controlled for. Our work, however, has two important limitations: first, as with all observational studies, our work is at risk for residual, unmeasured confounding; second, the performance of the initial resection by a high-volume surgeon does not necessarily translate into downstream care by the same surgeon. Therefore, we could not verify if a patient initially resected by a low-volume surgeon was eventually referred to a high-volume surgeon for re-resection and followup care. However, in routine clinical care, the first steps of therapy and many initial oncological decisions are usually made by the first resecting surgeon, with subsequent refer-

Table 2. Cox proportional hazards regression quantifying the association between surgeon case volume and overall/cancer-specific survival

Surgeon case volume	Overall survival				Cancer-specific survival			
	Unadjusted		Adjusted[a]		Unadjusted		Adjusted[a]	
	HR (95% CI)	p	HR (95% CI)	p	HR (95% CI)	p	HR (95% CI)	p
80th percentile and higher vs. below	0.93 (0.86–0.99)	0.04*	0.94 (0.89–1.01)	0.12	0.90 (0.79–1.03)	0.12	0.92 (0.80–1.05)	0.23
Median and higher vs. below	0.93 (0.87–0.99)	0.02*	0.93 (0.87–0.99)	0.02*	0.99 (0.88–1.12)	0.91	0.97 (0.86–1.09)	0.60
Continuous (quintiles)	0.97 (0.95–0.99)	0.004*	0.97 (0.95–0.99)	0.009*	0.97 (0.93–1.01)	0.20	0.97 (0.93–1.02)	0.22

*Statistically significant. [a]Adjusted for age, sex, histology, grade, sufficiently sampled muscularis propria, concurrent carcinoma in situ, lymphovascular invasion, tumor size/multiplicity (billing claim-based), prior urinary tract cancer, Charlson comorbidity, ACG® resource utilization band, ACG® frailty indicator, prevalence of asthma, congestive heart failure, chronic obstructive pulmonary disease, inflammatory bowel disease, dementia, diabetes mellitus, hypertension, psychiatric comorbidity, or rheumatoid arthritis; prior cancer diagnosis, myocardial infarction, number of home care claims/drug claims/physician billing claims during year preceding the diagnosis, rurality index, postal code-based socio-economic status, postal code-based components of Ontario Marginalization Index (residential instability, material deprivation, dependency, ethnic concentration), and surname-based ethnicity. Number of tumors, maximal dimension, weight, and volume (pathology report-based) were excluded as confounding variables from the multivariable models because of the high proportion of missing values. Out of 7424 patients with available followup data, 7287 individuals (98.2%) could be considered for multivariable modelling after restricting the cohort to patients with complete observations of all assumed confounding variables. ACG: Adjusted Clinical Groups; CI: confidence interval; HR: hazard ratio.

ral occurring after many of these initial attempts at therapy. Thus, we feel this limitation is unlikely to have had a material impact on the results.

In the light of the inherent limitations of our study and the relatively small absolute effect size, our work should not be regarded as a plea for centralization in T1 BC but rather as a hypothesis-generating perspective. The signal we detected at the population-level might be a surrogate for improved care provided by high-volume surgeons that certainly warrants further investigation. Further research can not only explore specific differences in practice patterns between high-volume and low-volume surgeons (such as TURBT quality, utilization of re-resection, instillation of adjuvant immunotherapy, or use of immediate/early radical cystectomy) but might also help to define high-volume/low-volume surgeons in the event centralization is introduced by policymakers. On this occasion, it should be highlighted that the sampling of muscularis propria, an important indicator of TURBT quality,[17] was comparable between high-volume surgeons (63.9%) and low-volume surgeons (65.9%).

Conclusions

This population-based cohort study demonstrated a volume-outcome relationship in T1 BC and raises questions regarding the centralization of care. It further points to the importance of early, optimal decision-making and care in the setting of high-risk, NMIBC. The generalizability of our findings is not only limited by the fact that the performance of the initial resection by a high-volume surgeon does not necessarily translate into downstream care by the same surgeon but also by the relatively small absolute effect size. Further studies are ultimately warranted to shed more light on these important findings.

Acknowledgements/funding: This study was supported by ICES, which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC). The opinions, results, and conclusions reported in this paper are those of the authors and are independent from the funding sources. No endorsement by ICES or the Ontario MOHLTC is intended or should be inferred. Parts of this material are based on data and/or information compiled and provided by the Canadian Institute for Health Information (CIHI), and Cancer Care Ontario (CCO). However, the analyses, conclusions, opinions, and statements expressed in the material are those of the authors, and do not reflect those of the funding or data sources; no endorsement is intended or should be inferred. Furthermore, parts of this material are based on Ontario Registrar General (ORG) information on deaths, the original source of which is ServiceOntario. The views expressed therein are those of the author and do not necessarily reflect those of ORG or Ministry of Government Services. The authors thank all data abstractors for their essential work and IMS Brogan Inc. for the use of their Drug Information Database.

Dr. Wettstein is funded by the Swiss Cancer Research Foundation (BIL KFS-4009-08-2016), the Professor Dr. Max Cloetta Foundation & Uniscientia Foundation (Vaduz), the Swiss Association of Urology, and the Kurt & Senta Herrmann Foundation. The remaining authors report no competing personal or financial interests related to this work.

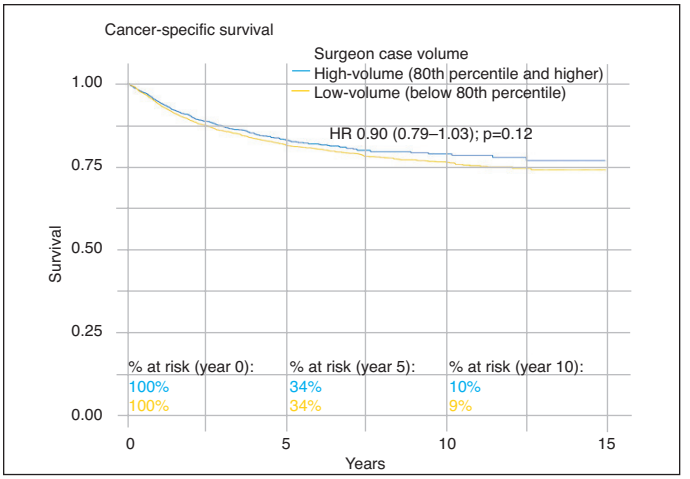

Fig. 3. Kaplan-Meier curve and log-rank test comparing cancer-specific survival between patients who received surgery by a high-volume surgeon (≥80th percentile; blue curve) and patients who received surgery by a low-volume surgeon (<80th percentile; yellow curve). HR: hazard ratio.

The dataset from this study is held securely in coded form at ICES. While data-sharing agreements prohibit ICES from making the dataset publicly available, access may be granted to those who meet pre-specified criteria for confidential access, available at *www.ices.on.ca/DAS*. The full dataset creation plan and underlying analytic code are available from the authors upon request, understanding that the computer programs may rely upon coding templates or macros that are unique to ICES and are therefore either inaccessible or may require modification.

This paper has been peer-reviewed

References

1. Cancer Research UK. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bladder-cancer/survival. Accessed Feb. 29, 2020.
2. Fernandez-Gomez J, Madero R, Solsona E, et al. Predicting non-muscle-invasive bladder cancer recurrence and progression in patients treated with bacillus Calmette-Guerin: The CUETO scoring model. *J Urol* 2009;182:2195-203. https://doi.org/10.1016/j.juro.2009.07.016
3. Babjuk M, Burger M, Compérat E, et al. EAU guidelines on non-muscle-invasive bladder cancer (TaT1 and CIS). Arnhem, The Netherlands: EAU Guidelines Office; 2019.
4. Chang SS, Boorjian SA, Chou R, et al. Diagnosis and treatment of non-muscle-invasive bladder cancer: AUA/SUO guideline. *J Urol* 2016;196:1021-9. https://doi.org/10.1016/j.juro.2016.06.049
5. Kulkarni GS, Urbach DR, Austin PC, et al. Higher surgeon and hospital volume improves long-term survival after radical cystectomy. *Cancer* 2013;119:3546-54. https://doi.org/10.1002/cncr.28235
6. Kulkarni GS. Improving outcomes of radical cystectomy: A call for passive overactive regionalization. *Can Urol Assoc J* 2020;14:97-8. https://doi.org/10.5489/cuaj.6505
7. Williams SB, Ray-Zack MD, Hudgins HK, et al. Impact of centralizing care for genitourinary malignancies to high-volume providers: A systematic review. *Eur Urol Oncol* 2019;2:265-73. https://doi.org/10.1016/j.euo.2018.10.006
8. Statistics Canada. Table 17-10-0009-01 Population estimates, quarterly. Available at: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000901. Accessed March 31, 2020.
9. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely collected health Data (RECORD) statement. *PLoS Med* 2015;12:e1001885. https://doi.org/10.1371/journal.pmed.1001885
10. Becker R, Silvi J, Ma Fat D, et al. A method for deriving leading causes of death. *Bull World Health Organ* 2006;84:297-304. https://doi.org/10.2471/BLT.05.028670
11. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *J Chronic Dis* 1987;40:373-83. https://doi.org/10.1016/0021-9681(87)90171-8

12. Reid RJ, MacWilliam L, Verhulst L, et al. Performance of the ACG case-mix system in two Canadian provinces. *Med Care* 2001;39:86-99. https://doi.org/10.1097/00005650-200101000-00010
13. Matheson FI, Dunn JR, Smith KL, et al. Development of the Canadian Marginalization Index: A new tool for the study of inequality. *Can J Public Health* 2012;103:S12-6. https://doi.org/10.1007/BF03403823
14. Shah BR, Chiu M, Amin S, et al. Surname lists to identify South Asian and Chinese ethnicity from secondary data in Ontario, Canada: A validation study. *BMC Med Res Methodol* 2010;10:42. https://doi.org/10.1186/1471-2288-10-42
15. Austin PC. Using the standardized difference to compare the prevalence of a binary variable between two groups in observational research. *Comm Stats Simul Comp* 2009;38:1228-34. https://doi.org/10.1080/03610910902859574
16. Shabsigh A, Korets R, Vora KC, et al. Defining early morbidity of radical cystectomy for patients with bladder cancer using a standardized reporting methodology. *Eur Urol* 2009;55:164-74. https://doi.org/10.1016/j.eururo.2008.07.031
17. Mariappan P, Zachou A, Grigor KM; Edinburgh Uro-Oncology Group. Detrusor muscle in the first, apparently complete transurethral resection of bladder tumor specimen is a surrogate marker of resection quality, predicts risk of early recurrence, and is dependent on operator experience. *Eur Urol* 2010;57:843-9. https://doi.org/10.1016/j.eururo.2009.05.047

Correspondence: Dr. Girish S. Kulkarni, Princess Margaret Cancer Centre, Toronto, ON, Canada; girish.kulkarni@uhn.ca