Hospital characteristics, rather than surgical volume, predict length of stay following colorectal cancer surgery

Don Vicendese,[1,2] Luc Te Marvelde,[1,2] Peter D. McNair,[3,4] Kathryn Whitfield,[2] Dallas R. English,[1,5] Souhaib Ben Taieb,[6] Rob J. Hyndman,[6] Robert Thomas[7]

Length of hospital stay (LOS) following colorectal surgery is considered an important proxy for successful treatment.[1,2] LOS is influenced by neoplasm (e.g. site and stage) and patient characteristics including comorbidities, social and geographic factors and discharge destination.[3] While a short hospital stay may negatively affect quality of care and health outcomes for patients,[2] a protracted LOS may increase the risk of – or signal complications including, but not limited to – cardiopulmonary, gastrointestinal and muscular morbidities, thromboembolic disease and hospital acquired infection.[2,4,5] Hospital-specific factors also influence LOS, including pre- and post-operative processes, the incorporation of multi-disciplinary teams into patient management practices and the degree of implementation of emerging evidence-based best practice.[6,7] In addition to the impact on patients, an unnecessarily prolonged LOS is an inefficient use of finite health system resources that could be allocated to other areas of patient need. This problem is not insignificant. In 2016, the Victorian Auditor General identified $125 million per year in potential statewide savings from efficient LOS performance.[2] LOS is an important component in the delivery of efficient, high-quality patient care.[8]

Abstract

Objective: Length of hospital stay (LOS) is considered a vital component for successful colorectal surgery treatment. Evidence of an association between hospital surgery volume and LOS has been mixed. Data modelling techniques may give inconsistent results that adversely impact conclusions. This study applied techniques to overcome possible modelling drawbacks.

Method: An additive quantile regression model formulated to isolate hospital contextual effects was applied to every colorectal surgery for cancer conducted in Victoria, Australia, between 2005 and 2015, involving 28,343 admissions in 90 Victorian hospitals. The model compared hospitals' operational efficiencies regarding LOS.

Results: Hospital LOS operational efficiencies for colorectal cancer surgery varied markedly between the 90 hospitals and were independent of volume. This result was adjusted for pertinent patient and hospital characteristics.

Conclusion: No evidence was found that higher annual surgery volume was associated with lower LOS for patients undergoing colorectal cancer surgery. Our model showed strong evidence that differences in LOS efficiency between hospitals was driven by hospital contextual effects that were not predicted by provider volume. Further study is required to elucidate these inherent differences between hospitals.

Implications for public health: Our model indicated improved efficiency would benefit the patient and medical system by lowering LOS and reducing expenditure by more than $3 million per year.

Key words: surgery, colorectal, cancer, length of stay, quantile regression

Since Adams et al.[9] in 1973 and Lufte et al.[10] in 1979 reported a link between provider volume and successful surgical outcomes, provider volume has been used as a surrogate measure of hospital quality. Evidence of an association between hospital or surgeon cancer surgery volume and better patient outcomes has been mixed;[3,7,8,11-31] however, previous analyses have had important limitations. A 2002 review that examined seven statistical modelling techniques used to assess association between patient factors and LOS in a cardiovascular setting found that choice of model influenced the conclusion of

1. *Cancer Epidemiology and Intelligence Division, Cancer Council Victoria*
2. *Cancer Strategy and Development, Department of Health and Human Services, Victoria*
3. *The Victorian Agency for Health Information, The Department of Health and Human Services, Victoria*
4. *The Melbourne School of Population and Global Health, The University of Melbourne, Victoria*
5. *Centre for Epidemiology and Biostatistics, The University of Melbourne, Victoria*
6. *Department of Econometrics and Business Statistics, Monash University, Victoria*
7. *Faculty of Medicine, Dentistry and Health Sciences, The University of Melbourne, Victoria*

Correspondence to: Dr Don Vicendese, College of Science, Health and Engineering, Plenty Road, Bundoora, VIC 3086; e-mail: vicendese.don@gmail.com

Submitted: March 2019; Revision requested: June 2019; Accepted: July 2019

The authors have stated they have no conflict of interest.

Aust NZ J Public Health. 2020; 44:73-82; doi: 10.1111/1753-6405.12932

the analyses. Model results were inconsistent due to LOS being a complex phenomenon and unmet assumptions regarding distributional fit. A small proportion of patients with very long hospital stays made it inherently difficult for simpler parametric methods to effectively model the data.[32] These models have been employed also in colorectal cancer (CRC) studies of association between LOS and provider volume and hence some doubt is cast on both negative and positive conclusions.[7,13,14,20,26,30,31] Some studies have used arbitrary thresholds for the categorisation of LOS and volumes (low, medium and high),[12-14,18-22,25,31,33] which may reduce statistical power.[34] If the skewness of LOS changes as provider volume changes, then arbitrary categorisations of LOS may be problematic.[34-36] Additionally, heterogeneity due to arbitrary categorisations has prevented synthesis of findings.[29,37]

A weakness of some hospital patient outcome studies is not accounting for the non-independence of the data if it is important to do so. Analyses may be biased if there are correlations between outcomes within a hospital that are unaccounted for – sometimes referred to as random effects.[38] Furthermore, surgeries are performed in the context of a hospital with a distinct infrastructure and management that affect their outcome.[19] Modelling this contextual effect may yield important information regarding its association with patient outcomes such as LOS.[39-41] The systematic differences in patients' outcomes across hospitals that persist after differences in patients' risk profiles have been accounted for reflect differences in hospitals' quality of care.[42] In this study, we examined the relationship between LOS following CRC surgery and provider volume by using a quantile regression model that makes no distributional assumption about LOS or error terms[43,44] and avoids arbitrarily categorising LOS or provider volume.[34] The model takes the individual patient as the unit of analysis but uses the data structure to analyse the association between the hospital context and LOS.[39-41]

Methods

The Victorian Admitted Episode Dataset (VAED) includes all separations (discharges and transfers) undertaken within all Victorian hospitals. All separations between 1 July 2005 and 30 June 2015 recorded in the VAED, which included one of 30 ICD-10-AM Australian Classification of Health Interventions procedure codes for colorectal surgery as the primary reason for admission, were identified (Supplementary Table 1). There were 62,774 admissions for 57,446 patients. Analysis was restricted to admissions whose principal diagnosis was for CRC, ICD-10-AM codes C18.x to C21.x, which resulted in a final data set of 28,343 admissions for 27,633 patients.

Length of stay was defined as the number of days from admission to discharge for the episode of care including transfers to other hospitals and geriatric and rehabilitation centres.

As it was conceivable that LOS and provider volume were not necessarily linearly related, we used an additive quantile regression (AQR) model that does not require a predetermined functional fit but instead determines the best fit from the data.[34,44-46] Provider volume was defined as the number of colorectal surgical procedures performed by a hospital within a fiscal year (1 July to 30 June), whether patients had a principal diagnosis of CRC or not; that is, annual volume (AV). The model we used was based on a formulation by Mundlak[47] and is in a class commonly referred to as a within and between effects model.[39,41,48] It required that we enter both AV and mean annual volume (MAV). Mean annual volume (MAV) was defined for each hospital as the mean of all AV over the number of years the hospital operated within the 10-year study period. Not all hospitals performed colorectal surgical procedures in every study year.[19] The within effect was modelled by AV. It estimated the effect on LOS within hospitals as AV varied and its interpretation is equivalent to any fixed effect estimator.[48] The between effect was modelled by MAV. Due to the model formulation used, it estimated the effect on LOS if a patient were to attend another hospital with a different MAV, that is, the hospital contextual effect.[39,40,42] This method draws comparisons across hospitals and estimates the effect of hospital choice on patient LOS or, in other words, hospitals' quality of care or efficiency regarding LOS.[42]

The model was adjusted for various patient and hospital factors that may confound the association between provider volume and LOS.[3,8,38,49] These included patient factors such as: sex; age; Elixhauser Comorbidity Index; the American Society of Anesthesiologists' physical status classification system (ASA); cancer site (colon, rectum or anus); whether the cancer was metastatic or not based on ICD-10-AM codes; admission type; separation mode;[3,50] and the colorectal procedure performed. Hospital factors were: number of daily colorectal surgical admissions; type of hospital (private for profit, private non-profit or public); and whether the hospital was co-located or not.[51] Possible seasonal variation in LOS was modelled by month of surgery. The Elixhauser comorbidity index was calculated based on the diagnoses codes (up to 40) present in the surgical episode. All other patient and hospital factors are recorded in the VAED except for colocation and private for profit hospitals. These were determined by checking hospital websites or The Victorian Department of Health and Human Services internal records.

We did not know *a priori* whether all variables in the model were confounders and hence may have over-specified the model.[52] We tested confounding for the continuous variables by running separate additive quantile regressions (AQR) with them as sole covariate and LOS and AV separately as outcomes. This was done for the 1st, 5th, 25th, 50th, 75th, 95th and 99th percentiles of LOS. The results indicated that the continuous covariates were all associated with both AV and LOS and hence could be accepted as potential confounders. We tested the categorical variables for confounding by running Kruskal-Wallis rank sum tests for both AV and LOS. The results indicated that, except for sex, all the categorical variables were associated with both AV and LOS and therefore could also be accepted as potential confounders. Sex was associated with LOS but not with AV; however, we included it with the full model in the boosting process to assess its influence. We did not include ICU use or ICU hours in the model as they are directly on the causal pathway to LOS (use of ICU necessarily means a lengthening in LOS) and hence an over-adjustment bias.[52] We assessed possible random effects[33] by including a random intercept for hospital into the boosting process, where hospital was represented by a dummy variable.

Model building was aided with boosting, a statistical technique in the class of machine learning methods.[34,43,53-56] Boosting assessed which variables and functional fit better predicted LOS. These results were used to build a second AQR model that did not use boosting but used shrinkage and penalisation for model estimates.[44,45] We tested how well the models represented the data by using them to predict the empirical cumulative distribution (CDF) of LOS and assessed its fit. We refer to this as the recovered distribution (see Supplementary File: Statistical Details).

As the model formulation we used estimated the effect on LOS if a patient were to attend another hospital,[39,40,42] we used the model to simulate the change in LOS if CRC patients were to have counterfactually attended a hospital that the model indicated to be more LOS efficient. We predicted each percentile 1 to 99; these were combined to obtain the predicted CDF of LOS.[56] This was termed the counterfactual CDF. The area under the counterfactual CDF was calculated and compared to the area under the recovered CDF. As the area under each CDF directly related to the total sum of LOS days, the difference in areas estimated the change in total sum of LOS due to this hypothetical experiment. A 95% confidence interval (CI) was computed for the estimated change.

We calculated a dollar value for the change in total LOS by assuming an average of $1,000 per bed day. This was based on the 2016 report by the Victorian Auditor General who estimated an average cost of $864 per bed day due to direct costs such as daily labour, primarily for medical and nursing staff, and indirect costs such as lighting, heating and cleaning. The cost estimate excluded theatre costs, pharmacy and pathology costs, depreciation, capital and maintenance costs, as these costs are unlikely to be saved by reduced LOS.[2] The Auditor General relied on data up to 2014 and so – allowing for CPI increases of 2% per annum to 2019 – $864 converted to approximately $1,000 but, when applied to total LOS, the estimate was rounded down to the nearest million.

Statistical significance was set at the 0.05 level. All statistical analysis was performed using R, version 3.3.3[57] using the following packages: quantreg for the AQR[58]; mboost for the boosting[59]; and flux for computing the area under the graph using the auc function.[60] (Further details are provided in the Supplementary File).

Results

LOS ranged between 1 and 258 days with median 10 days and 25th and 75th percentiles (interquartile range, IQR) 7 and 15 days, respectively. The 28,343 episodes of care generated a total LOS of 386,647 days over the 10-year study period between fiscal years 2006–2015. The episodes of care in the highest 25 percentiles of LOS generated 55% of the total LOS, illustrating the long tail of the LOS distribution.

All surgeries were performed at 90 Victorian hospitals. AV ranged between 1 and 608 with median 44 (IQR, 14-139). MAV ranged between 1 and 566.7 with mean and standard deviation (SD) of 70.4 and 91.9, respectively, and median 30.5 (IQR, 8.3-118.0). The hospital with highest volumes was a considerable outlier as the next highest MAV and AV were 279.6 and 345, respectively. More than half of total LOS was generated in the 13 hospitals with highest MAV. They had MAV greater than 191 and accounted for 52% of admissions. Approximately 1.5% of total LOS was generated in 30 hospitals with MAV less than 13 (at most one per month on average), accounting for 1.9% of admissions. LOS had increasing skewness with increasing volume (results not shown), further indicating difficulties for models with distributional assumptions and suggesting it may not be advisable to categorise LOS in order to use a logit model.[34-36]

Model Fit

Boosting indicated all entered variables may be important for predicting LOS except for Elixhauser score and co-location status. Separation mode and month were the strongest and most consistent predictors across the percentiles. Boosting suggested that non-linear fits for MAV, AV, age, and month better predicted LOS than linear, but a linear fit for year and number of daily colorectal surgery admissions better predicted LOS. Boosting also indicated that allowing for hospital random effects was not useful for predicting LOS. These results were used to build the second AQR model. We compared how well the models represented the data by using them to predict the empirical cumulative distribution (CDF) of LOS and assessed their fits. We chose the second model as a better fit to the data (see Supplementary File: Statistical Details).

Mean Annual Volume Association with LOS

The model graphs (Figure 1) indicate that hospitals' performances (contextual effect) regarding patient LOS varied greatly for all percentiles and that this variation was not systematically associated with MAV. The graphs display initial falls in LOS for MAV up to approximately 33, which are then followed by marked variation, with low points in LOS at MAV of 122.1 and 245.8 and a high point at 105.6 MAV. The former two MAV had the lowest LOS over all percentiles 40 and 20 times, respectively, while the latter had the highest 89 times. In a sensitivity analysis, the admissions to the three hospitals with these three MAV were removed from the analysis. This resulted in very much flattened graphs (not shown here). For all percentiles, the *p*-values from an F statistic for model fit were less than 1×10^{-6}.

Annual Volume Association with LOS

The model graphs (Figure 2) indicated that the association between AV and LOS varied with percentile of LOS. For percentiles 10 to 95, rising AV was generally associated with rising LOS. For percentiles 10–60 this association was generally attenuated or reversed for AV greater than about 100. This was similar for percentiles 65–95, except the attenuation occurred later at AV of approximately 200. The magnitude of these variations in association generally ranged between 0.5 and 4 days and increased with increasing LOS percentile. For all percentiles, the *p*-values from an F statistic for model fit were less than 1×10^{-6}.

Counterfactual prediction of change in LOS contingent on change in MAV

To carry out the counterfactual prediction, we selected the hospital with MAV of 122.1 as an LOS-efficient hospital due to its consistent association with reduced LOS over many percentiles. There were 68 hospitals that had AV of 122.1 or less and they generated 106,488 LOS days (27.5%) from 7,979 episodes of care (28.1%). We carried out the counterfactual experiment in two ways. Firstly, simulating change in LOS limited to patients that had attended hospitals with MAV lower than 122.1, and secondly for all patients.

The first counterfactual prediction estimated a fall of 13.0% in total LOS for all patients who attended hospitals with MAV less than 122.1, $p<0.009$. The 95%CI for this percentage fall was (3.8%, 22.2%). This percentage change equated to an estimated reduction of 13,822 total LOS days with 95%CI (4,033, 23,612 days) or a predicted saving of about $13 million in present day terms over the 10-year study period.[2] The second counterfactual prediction estimated a fall of 8.5% in total LOS over all patients and hospitals, $p<0.007$, with 95%CI (2.9%, 14.6%). This equated to a reduction of

32,842 total LOS days, 95%CI (11,225, 56,434 days) or a predicted saving of about $32 million in present day terms over the 10-year study period.[2] Figure 3 demonstrates that, for the second counterfactual experiment, the predicted savings mainly came from reduced LOS for patients who had LOS between percentiles 14 and 94.

Please see the Supplementary File for an in-depth display of the results for the associations between all other patient and hospital factors in our model and LOS**.**

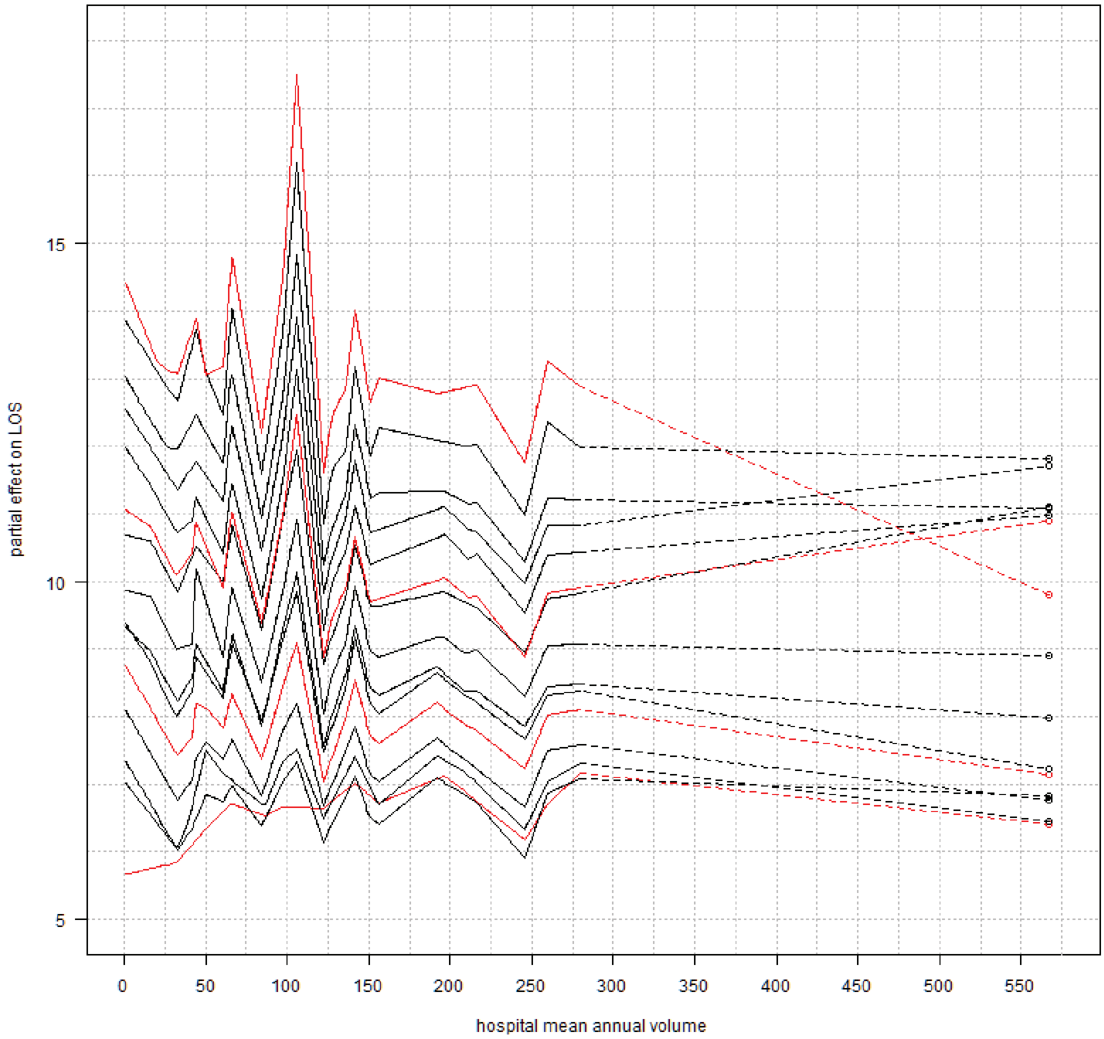
Discussion

Our model indicated that there was marked variation across hospitals' performances regarding LOS and that it was not contingent on MAV. This contextual effect in our model reflects hospitals' quality of care or efficiency regarding LOS.[42] It estimated change in patient LOS if patients were admitted, counterfactually, to another hospital.[39,42] If the hospital contextual effect was negatively associated with LOS, for example, the graphs in Figure 1 would have shown consistent falls in LOS with increasing MAV. Instead they showed increased or decreased LOS was largely independent of change in MAV. Therefore, the hospital contextual effect on LOS was associated with factors other than MAV.

The substantial size of the predicted change produced by the counterfactual experiment and its statistical significance indicated that the hospital contextual effect was important in explaining LOS. Both counterfactual predictions implied a statistically significant and substantial reduction in total LOS if less-efficient hospitals were to function at the level of a more-efficient hospital identified by the model. This represents an improved outcome for CRC patients due to quicker discharge times, and for the health system due to substantial savings of approximately $3 million per year in present day terms. As the model predicted statistically significant savings in LOS among (but not necessarily all) hospitals with provider volumes both lower and greater than the reference hospital with MAV of 122.1, this further illustrates that LOS efficiency is not necessarily based on provider volume.

Ash et al., in their review of statistical issues in assessing hospital performance that was commissioned by the Committee of Presidents of Statistical Societies, strongly

Figure 1A: Model estimates for between hospital differences (contextual effect) in the association between annual volumes and LOS for percentiles 5-75, in intervals of 5.



Notes:

Red lines are percentiles 5, 25, 50 and 75.

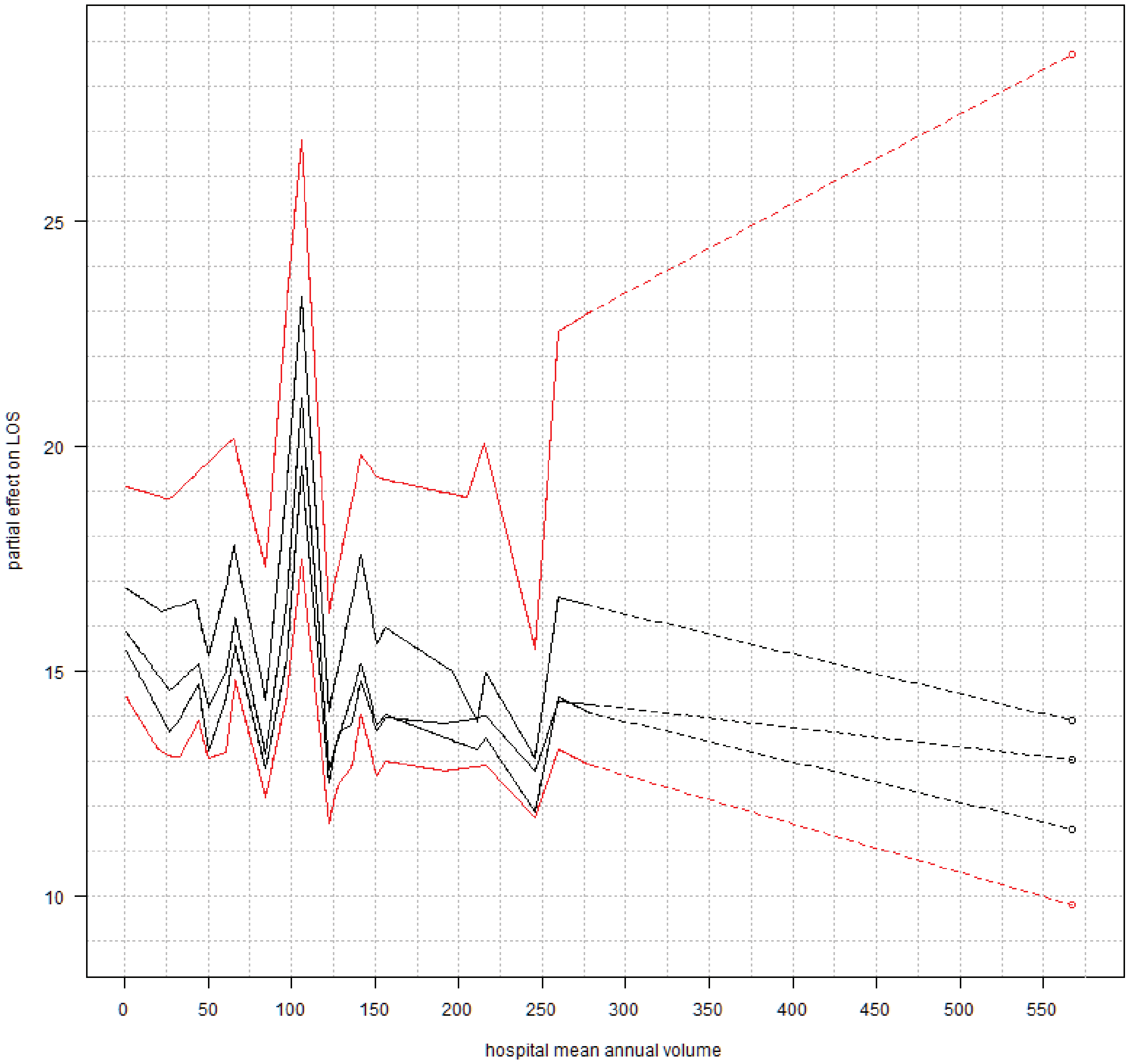
The lines are dotted between 279.6 and 566.7 as there were no hospitals with MAV between these values. A few of the percentile fits display quantile crossing – see Supplementary for statistical details.

All p values, from an F statistic that assessed model fit for all percentiles 1-75, were less than 1 x 10-6.

recommended analytical approaches that respect the multilevel structure of hospital data and allow for both between and within hospital variation.[38] Our first model was multilevel and allowed for between and within effects via the Mundlak formulation. The second model allowed for between and within effects via the Mundlak formulation but was not multilevel in the sense of allowing for correlation between outcomes within a hospital – referred to as random effects. However, our model building process indicated that, for our data, random effects were not important for predicting LOS and the second model (without random effects) that we used for the final analysis was superior. The weakness in not allowing for random effects is that, although coefficient estimation is still consistent and close to the estimate when random effects is employed, coefficient standard errors may be underestimated.[61] That is, statistical significance is over optimistic. However, even if our model building process was inaccurate, the *p*-values for model fit for MAV and AV were so low – less than 1×10^{-6} for all percentiles – it would be unlikely that standard errors would be increased to the extent that statistical significance at the 0.05 level would be lost. As our modelling respected the multilevel structure in the data by incorporating a Mundlak formulation and checking the importance of random effects, we compared our findings to other findings from multilevel models.

Other studies using multilevel models have made similar findings to our study. Zheng et al. used a multilevel model based on Poisson regression to analyse outcomes following laparoscopic colectomy for 4,617 elderly Stage I-III patients in 465 US hospitals. Similar to our results, they also found a significant hospital contextual effect for LOS and that LOS was not associated with hospital volumes. In that study, hospital volume was dichotomised based on data derived results.[20] Liu et al. used a multilevel model based on logistic regression to analyse 61,728 CRC surgeries in 218 Taiwanese hospitals and found that LOS was associated with surgeon volume but not hospital volume. In that study, LOS was dichotomised based on

Figure 1B: Model estimates for between hospital differences (contextual effect) in the association between annual volumes and LOS for percentiles 75-95 in intervals of 5.



Notes:

Red lines are percentiles 75 and 95.

The lines are dotted between 279.6 and 566.7 as there were no hospitals with MAV between these values. A few of the percentile fits display quantile crossing – see Supplementary for statistical details.

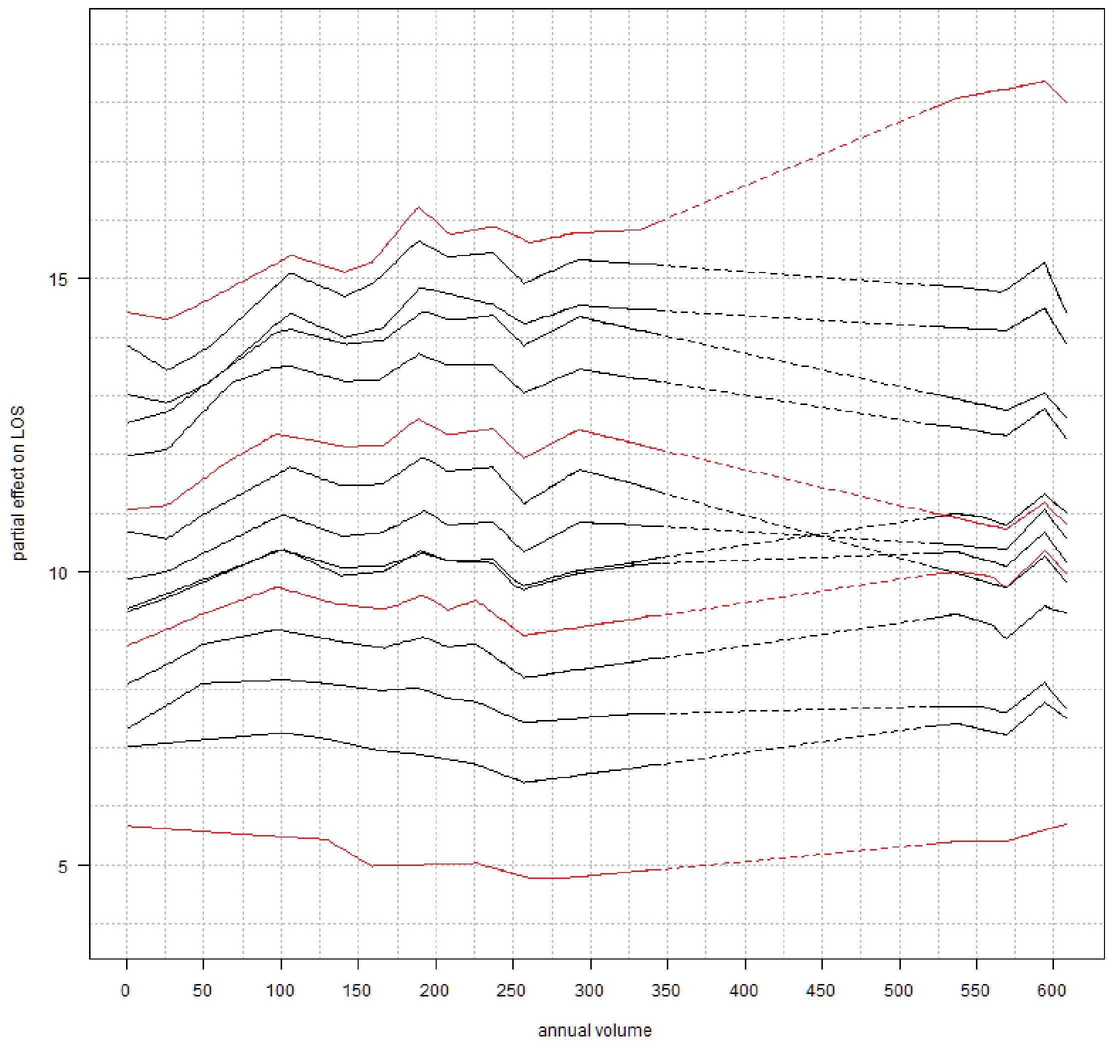
All p values, from an F statistic that assessed model fit for all percentiles 75-99, were less than 1 x 10-6.

median LOS and volumes categorised into four levels.[12]

Burns et al. used a multilevel model based on logistic regression to analyse outcomes following 109,261 elective colorectal surgeries in England.[19] In that study, volumes were entered in separate regressions as either a three-level categorical (low, medium, high) or numerical variable. The basis for dichotomising LOS was not disclosed. The authors found no association between numerical hospital and surgeon volumes and LOS (*p*=0.484 and 0.448, respectively) but found a negative association between categorical hospital and surgeon volumes and LOS. Curiously, even though these results were conflicting, the authors still concluded that LOS was associated with hospital and surgeon volumes. The adjusted odds ratios comparing the higher volume categories to the lowest were quite modest, between 0.95 and 0.98. Although statistically significant due to the very large sample size (at the 0.05 level, at least, as the *p*-value was not disclosed), there is some doubt about this conclusion on statistical and practical grounds. The conflicting result, dependent on outcome characterisation, is sufficient to be cautious about the conclusion on statistical grounds. On the other hand, it may indicate a non-linear association with continuous volume on the logit scale in the logistic regression. Similarly to our study, the authors noted that even among the highest-volume providers, wide variation in outcome was observed, which implies that the arbitrary categorisation of LOS may have been statistically unsound.[34-36] A 5% reduction, at best, in the *odds* of high LOS, if treated by a high volume hospital or surgeon, equates to a risk *difference* of 1.28% at most (calculations not shown). Due to the wide variation in outcome even among the highest-volume providers, a risk difference of this size, although welcome, is not of practical assistance to enable the patient to confidently discern a hospital with more favourable LOS outcomes. However, the authors tempered their conclusion with the salient remark that centralisation may be ineffective in improving results unless the higher volume is directed towards high-quality providers. This is, in effect, what was demonstrated with our counterfactual experiment.

Aravani et al. used two multilevel models based on logistic regression with either

Figure 2A: Model estimates for within hospital differences in the association between annual volumes and LOS for percentiles 5-75, in intervals of 5.



Notes:

Red lines are percentiles 5, 25, 50 and 75.

The lines are dotted between 345 and 520 as there were no hospitals with AV between these values. A few of the percentile fits display quantile crossing –see Supplementary for statistical details.

All p values, from an F statistic that assessed model fit for all percentiles 1-75, were less than 1 x 10-6.

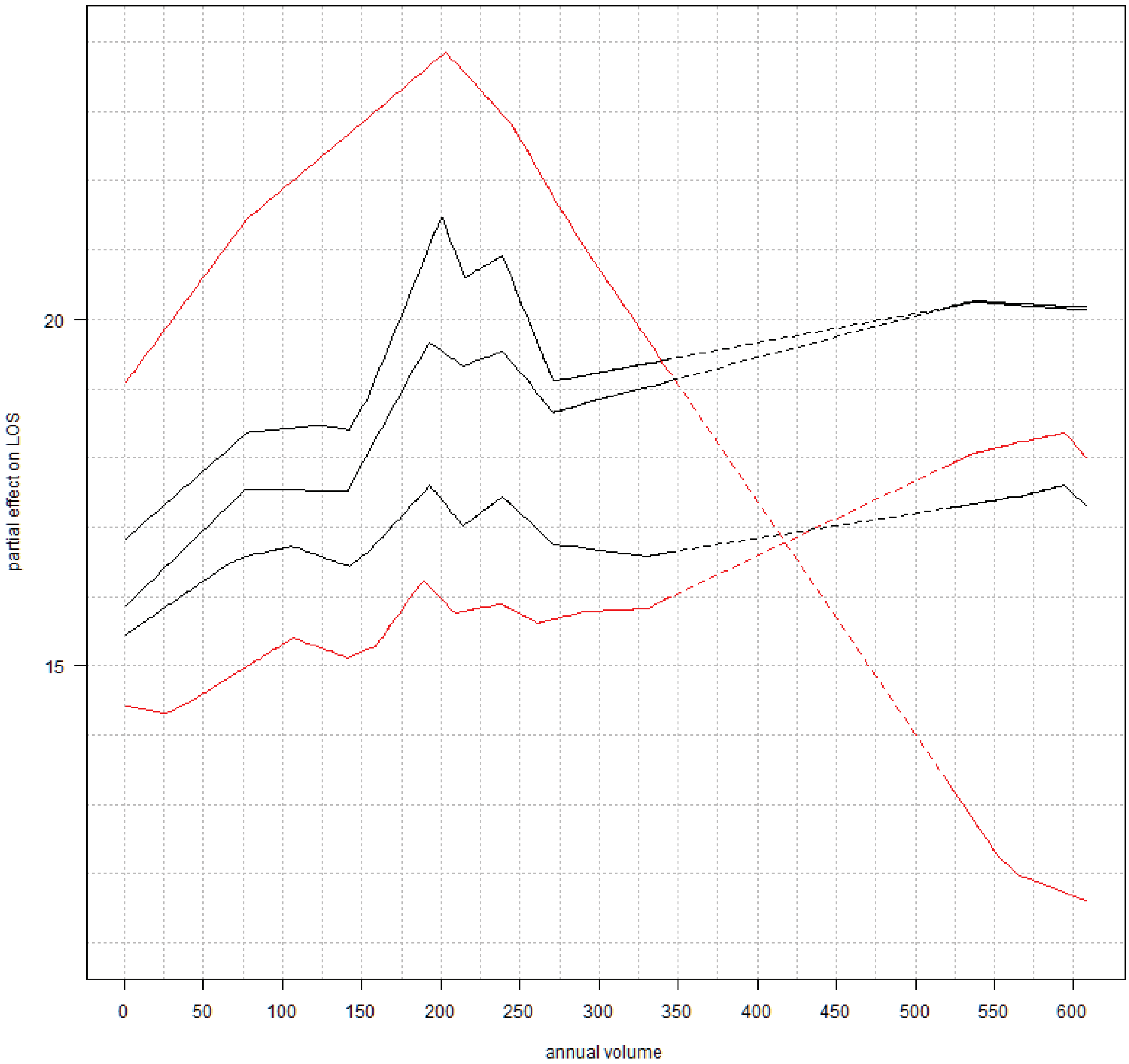
ideal (≤5 days postoperative) or prolonged (≥90th percentile, 21 days postoperative) as outcomes following CRC surgery for 240,873 patients.[8] This study produced funnel plots based on their analyses that, similarly to our study, clearly displayed marked variation between providers in both outcomes but with no systematic association with provider volume.

We found no evidence that higher AV was associated with lower LOS. AV was the within hospital effect in our model. The model indicated that over the 10-year study period, increased volume in a single year was generally associated with higher LOS. This association may be explained by the demands on management and service staff resources that may increase incommensurately compared to increases in AV and which eventually effect operational efficiencies in the short term.[62] We also found the association between AV and LOS was non-linear. Increase in LOS was attenuated after AV of around 100 for percentiles 60 and lower of LOS and AV of around 200 for higher percentiles of LOS. Increasing AV beyond these threshold points generally showed little change in LOS. This implies that, in the face of growth in provider volume, lower-volume hospitals would be more subject to increased LOS, but higher-volume hospitals would show little change. This may be due to *relative* capacities to absorb the demands on management and service staff resources. These results concord with Faiz, who noted that there was little evidence that increasing volumes of colorectal surgery hospitals was a warranted strategy for improving quality of care.[63]

In our model, the association between AV and LOS was independent of the hospital contextual effect, as these factors have been mutually adjusted for in our model.[39] Therefore, the effect of increased AV could be augmented or attenuated depending on the hospital's contextual effect. This adds further evidence to the conclusion by Burns et al.[19] mentioned above – that centralisation may be ineffective in improving results unless the higher volume is directed towards high-quality providers. This recommendation has been echoed by others.[50,64]

There have been conflicting results from many studies regarding the relationship between colorectal surgery hospital volumes

Figure 2B: Model estimates for within hospital differences in the association between annual volumes and for percentiles 75-95 in intervals of 5.



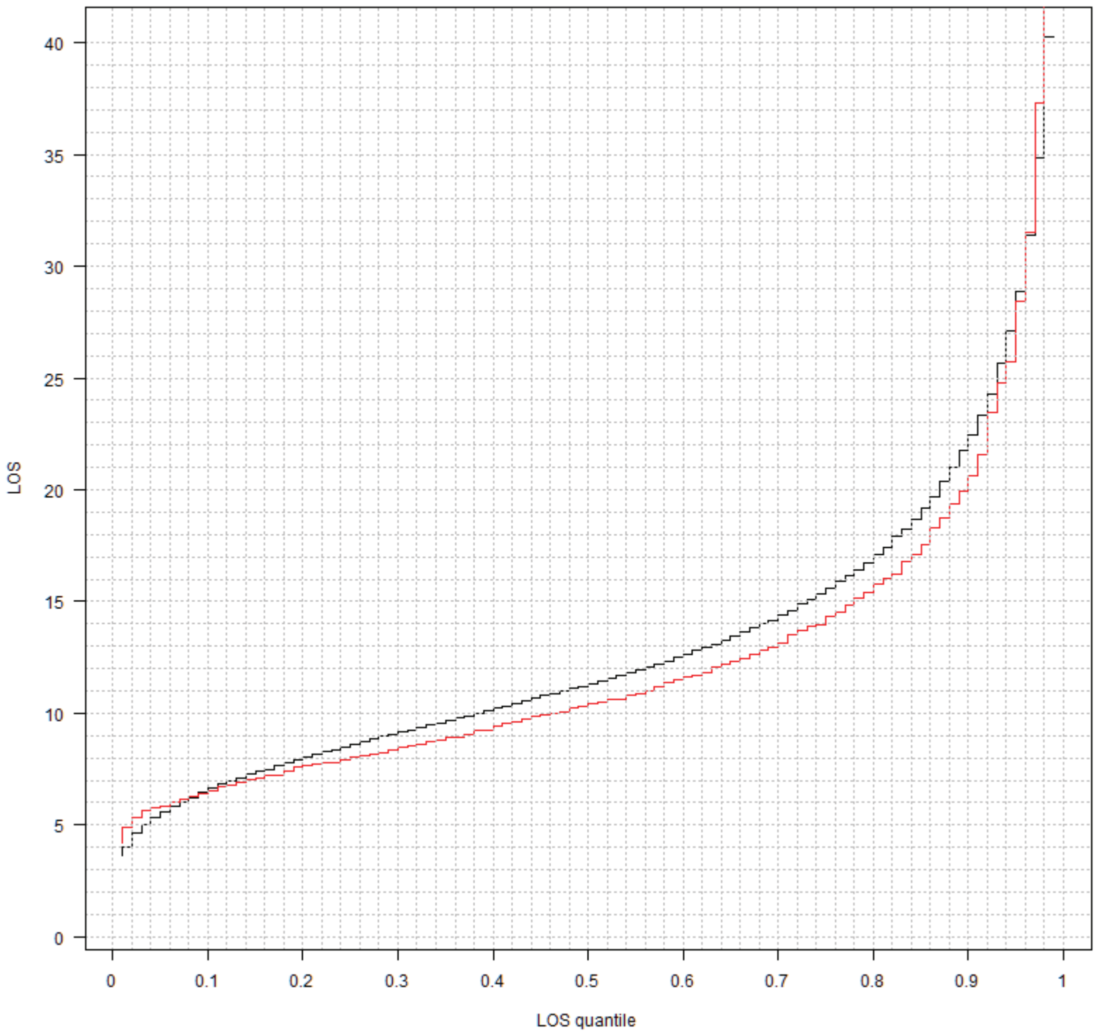
Notes:
Red lines are percentiles 75 and 95.
The lines are dotted between 345 and 520 as there were no hospitals with AV between these values. A few of the percentile fits display quantile crossing –see Supplementary.
All p values, from an F statistic that assessed model fit for all percentiles 75-99, were less than 1 x 10-6.

and LOS.[3,7,8,11-31] Chowdhury et al., in their comprehensive systematic review, analysed 127 studies that examined surgical outcomes including LOS. They included 29 involving oncological surgeries of which four were for colorectal cancer. They concluded that, although the initial impression is that the literature overwhelmingly substantiates a benefit from high hospital volume, analysis of the quality of the data suggests that this conclusion is unsafe.[29] They suggest that studies had been biased by not including the extra LOS if a patient was transferred to another hospital to recover. We included this contribution to LOS in our study. It has also been suggested that there may be publication bias that may contribute to the impression of the importance of hospital volume.[50]

Some of the uncertainty regarding the association between provider volume and LOS may be partially explained by reverse causality. Market forces may have directed high volume referrals to hospitals with known quality outcomes, or perceived as so, rather than the high volumes causing the high quality outcomes as such.[63] That is, hospital volume has an endogenous and exogenous quality. It can be both a predictor and outcome of hospital quality. Ash et al. pointed out this endogeneity and exemplified it from the other end of the volume spectrum. The current *low* volume of a hospital may be a result of poor quality and hence in the causal pathway between the hospital and the outcome being considered.[38] Therefore, because of the combined exogeneity and endogeneity of provider volume, careful consideration needs to be given to its inclusion as a way of characterising hospital performance.

As with all observational studies, a weakness of this study is the risk of bias due to inadequate adjustment for included confounders or non-adjustment due to non-included confounders, especially for confounders that may be differentially distributed across patients and hospitals.[38,42] These could be such factors as: hospital waiting list pressure[62]; tumour stage; socioeconomic status (SES); patient health behaviour; multidisciplinary teams; a broader range of specialist and technology-based services at high-volume hospitals; better coordination of personnel or a higher level of collaboration between physicians and nurses; the presence of more or better-trained nurses[65]; use of specific clinical pathways; the method by which laboratory tests are reported to physicians; policies and operations of hospitals[20]; and surgeon volume and specialisation.[29,66] We had no data for any of these factors. Omitted variable bias has been controlled to a certain extent due to the Mundlak model formulation; however, it is not a panacea.[38,41,42] Surgeon

Figure 3: In black we have the recovered distribution. Red is the counterfactual distribution obtained by setting all MAV to 122.1 and all AV to the same AV in each year as generated by the hospital with MAV of 122.1.



volume and specialisation are important confounders for the association between hospital volume and CRC surgical outcomes, internationally and in Australia.[29,66-68] There is evidence from a systematic review and meta-analysis that indicated surgeon volume and experience may be more important than hospital volume.[29,37] It may be possible that high volume can be caused by a large group of low-volume surgeons in one hospital, while a hospital volume may be low if a single – perhaps more experienced – surgeon performs all the procedures. This infers that interpreting the association between hospital volume and outcomes is difficult without controlling for surgeon volume.[29] Our model would have benefited from including surgeon as another level.[38]

As the hospital contextual effect seemed largely due to factors other than provider volumes, it would be beneficial for patients and the health system if some hospitals were directly investigated to determine how they are functioning more efficiently regarding LOS and conversely why some hospitals are functioning less efficiently.[37,42,69,70] This would provide valuable information that may greatly benefit the patient and health system and be a check on our model. These could be such factors as systemised protocols and procedures, multi-disciplinary teams and ERAS (Enhanced Recovery After Surgery), for which data was not available. As LOS efficiency was not dependent on provider volume, this implies that it may be possible for all hospitals, regardless of volume, to operate more efficiently in regard to LOS and hence realise the important benefits to the patient and healthcare system alike.[71] However, it may be unlikely that hospitals with very low volumes could justify the allocation of resources to achieve operational efficiencies.

An inconsistent association between MAV and LOS, the general positive association between AV and LOS and possible endogeneity, suggest it is inadvisable to use provider volume as a marker of quality regarding LOS following CRC surgery in Victoria. The Victorian Auditor General also found provider volume was not a reliable indicator of LOS efficiency following surgery for many conditions other than CRC.[2] We found it was more pertinent to use the hospital contextual effect to analyse the performance of hospitals in regard to LOS.

Using the hospital contextual effect, our model presents itself as a useful method for assessing provider performance regarding LOS, adjusted for pertinent patient and hospital factors.[38] The model can produce graphs such as that in Figure 1 to guide analysis. More importantly, by using the model's counterfactual prediction capacity, an index of performance with a 95% CI can be obtained for any hospital. The prediction would immediately indicate if the hospital was performing better, worse or at about the same level as all other hospitals. If performing the counterfactual experiment – using the hospital being assessed as the basis of comparison – resulted in a 5% drop in total LOS, then that hospital would have an index of 0.95 with an associated confidence interval. An index lower than 1 indicates superior efficiency, higher than 1 indicates inferior and 1 shows no difference. This index is independent of any distributional or model fit assumptions or arbitrary categorisation. For the analysis of LOS, the Victorian Auditor General's report resorted to using trimmed data in a linear regression. This method may be subject to statistical objections if a mean is not representative of the whole data and because information in the tail(s) of the distribution, that may represent patients who may not fit the profile of a mean patient, is discarded[2,44,72,73] (see Supplementary File).

Our model could be used in national or international settings as it can allow for nesting in those levels, and so help assess hospital efficiency in regard to LOS in broader contexts. This would assist with synthesis of future international studies when, in the past, diverse categorisation methods had impeded synthesis. It can be extended to analyse variation between hospitals regarding other outcomes such as mortality and readmission following CRC surgery. However, our model requires very large data sets, especially if many patient factors were entered into the analysis.

Our model was further complemented by shrinkage and penalisation to obtain more accurate estimates and reduce statistical error[38,45] (see Supplementary File). Shrinkage may not generally be robust to outlier values of the outcome,[42] but quantile regression naturally copes with them.[74]

Please see the Supplementary File for results regarding all other model variables such as laparoscope use, cancer site, seasonality and type of admission and discharge.

Conclusion and implications for public health

Our model showed marked variation between Victorian hospitals regarding LOS efficiency following CRC surgery between 2005 and 2015. This efficiency was not predicted by provider volume. Patients and the health system would benefit from an analysis of the function of identified efficient hospitals to clarify factors contributing to their relative efficiency regarding LOS and conversely to rectify factors that cause some hospitals to function less efficiently. As volume was not a determining factor for LOS efficiency, all providers could benefit – assuming capacity to allocate required resources. This has the potential to improve convalescent times for CRC surgical patients and save the Victorian health system more than $3 million per year in present day terms.

Due to its endogeneity, we recommend not using raw provider volume as an indicator of LOS operational efficiency. Further study is required to determine if LOS efficiency may be associated with increased risk of post-operative complications.

References

1. Boyle EM, Walsh SR. 'Predictors of length of stay in patients having elective colorectal surgery within an enhanced recovery protocol' Int J Surg 2010; 8: 628-32. Commentary on: *Int J Surg.* 2011;9:355-6.
2. Victorian Auditor-General's and Office. *Hospital Performance: Length of Stay* [Report]. Mlebourne (AUST): State Government of Victoria; 2016.
3. Cologne KG, Byers S, Rosen DR, et al. Factors Associated with a Short (<2 Days) or Long (>10 Days) Length of Stay after Colectomy: A Multivariate Analysis of over 400 Patients. *Am Surg.* 2016;82:960-3.
4. Jakobsen DH, Sonne E, Andreasen J, Kehlet H. Convalescence after colonic surgery with fast-track vs conventional care. *Colorectal Dis.* 2006;8:683-7.
5. Kehlet H. Fast-track colorectal surgery. *Lancet.* 2008;371:791-3.
6. Richardson B, Preskitt J, Lichliter W, et al. The effect of multidisciplinary teams for rectal cancer on delivery of care and patient outcome: Has the use of multidisciplinary teams for rectal cancer affected the utilization of available resources, proportion of patients meeting the standard of care, and does this translate into changes in patient outcome? *Am J Surg.* 2016;211:46-52.
7. Gatt M, Anderson AD, Reddy BS, Hayward-Sampson P, Tring IC, MacFie J. Randomized clinical trial of multimodal optimization of surgical care in patients undergoing major colonic resection. *Br J Surg.* 2005;92:1354-62.
8. Aravani A, Samy EF, Thomas JD, Quirke P, Morris EJ, Finan PJ. A retrospective observational study of length of stay in hospital after colorectal cancer surgery in England (1998-2010). *Medicine (Baltimore).* 2016;95:e5064.
9. Adams DF, Fraser DB, Abrams HL. The complications of coronary arteriography. *Circulation.* 1973;48:609-18.
10. Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. *N Engl J Med.* 1979;301:1364-9.
11. Ahmed Ali U, Dunne T, Gurland B, Vogel JD, Kiran RP. Actual versus estimated length of stay after colectomy surgery: Which factors influence a deviation? *Am J Surg.* 2014;208:663-9.

12. Liu CJ, Chou YJ, Teng CJ, et al. Association of surgeon volume and hospital volume with the outcome of patients receiving definitive surgery for colorectal cancer: A nationwide population-based study. *Cancer*. 2015;121:2782-90.
13. Thompson BS, Coory MD, Gordon LG, Lumley JW. Cost savings for elective laparoscopic resection compared with open resection for colorectal cancer in a region of high uptake. *Surg Endosc*. 2014;28:1515-21.
14. Yasunaga H, Matsuyama Y, Ohe K. Effects of hospital and surgeon volumes on operating times, postoperative complications, and length of stay following laparoscopic colectomy. *Surg Today*. 2009;39:955-61.
15. Shida D, Tagawa K, Inada K, et al. Enhanced recovery after surgery (ERAS) protocols for colorectal cancer in Japan. *BMC Surg*. 2015;15:90.
16. Nygren J, Soop M, Thorell A, Hausel J, Ljungqvist O. An enhanced-recovery protocol improves outcome after colorectal resection already during the first year: A single-center experience in 168 consecutive patients. *Dis Colon Rectum*. 2009;52:978-85.
17. Boudourakis LD, Wang TS, Roman SA, Desai R, Sosa JA. Evolution of the surgeon-volume, patient-outcome relationship. *Ann Surg*. 2009;250:159-65.
18. Kelly M, Sharp L, Dwane F, Kelleher T, Comber H. Factors predicting hospital length-of-stay and readmission after colorectal resection: A population-based study of elective and emergency admissions. *BMC Health Serv Res*. 2012;12:77.
19. Burns EM, Bottle A, Almoudaris AM, et al. Hierarchical multilevel analysis of increased caseload volume and postoperative outcome after elective colorectal surgery. *Br J Surg*. 2013;100:1531-8.
20. Zheng Z, Hanna N, Onukwugha E, Bikov KA, Mullins CD. Hospital center effect for laparoscopic colectomy among elderly stage I-III colon cancer patients. *Ann Surg*. 2014;259:924-9.
21. Faiz O, Haji A, Burns E, Bottle A, Kennedy R, Aylin P. Hospital stay amongst patients undergoing major elective colorectal surgery: Predicting prolonged stay and readmissions in NHS hospitals. *Colorectal Dis*. 2011;13:816-22.
22. Harmon JW, Tang DG, Gordon TA, et al. Hospital volume can serve as a surrogate for surgeon volume for achieving excellent outcomes in colorectal resection. *Ann Surg*. 1999;230:404-11; discussion 11-3.
23. Patel SS, Patel MS, Mahanti S, et al. Laparoscopic versus open colon resections in California: A cross-sectional analysis. *Am Surg*. 2012;78:1063-5.
24. Burns EM, Mamidanna R, Currie A, et al. The role of caseload in determining outcome following laparoscopic colorectal cancer resection: An observational study. *Surg Endosc*. 2014;28:134-42.
25. Keller DS, Hashemi L, Lu M, Delaney CP. Short-term outcomes for robotic colorectal surgery by provider volume. *J Am Coll Surg*. 2013;217:1063-9. e1.
26. Rea JD, Lu KC, Diggs BS, Cone MM, Hardiman KM, Herzig DO. Specialized practice reduces inpatient mortality, length of stay, and cost in the care of colorectal patients. *Dis Colon Rectum*. 2011;54:780-6.
27. Damle RN, Macomber CW, Flahive JM, et al. Surgeon volume and elective resection for colon cancer: An analysis of outcomes and use of laparoscopy. *J Am Coll Surg*. 2014;218:1223-30.
28. Larson DW, Marcello PW, Larach SW, et al. Surgeon volume does not predict outcomes in the setting of technical credentialing: Results from a randomized trial in colon cancer. *Ann Surg*. 2008;248:746-50.
29. Chowdhury MM, Dagash H, Pierro A. A systematic review of the impact of volume of surgery and specialization on patient outcome. *Br J Surg*. 2007;94:145-61.
30. Huebner M, Hubner M, Cima RR, Larson DW. Timing of complications and length of stay after rectal cancer surgery. *J Am Coll Surg*. 2014;218:914-9.
31. Borowski DW, Bradburn DM, Mills SJ, et al. Volume-outcome analysis of colorectal cancer-related outcomes. *Br J Surg*. 2010;97:1416-30.
32. Austin PC, Rothwell DM, Tu JV. A comparison of statistical modeling strategies for analyzing length of stay after CABG surgery. *Health Serv Outcomes Res Methodol*. 2002;3:107-33.
33. Gruen RL, Pitt V, Green S, Parkhill A, Campbell D, Jolley D. The effect of provider case volume on cancer mortality: Systematic review and meta-analysis. *CA Cancer J Clin*. 2009;59:192-211.
34. Fenske N, Fahrmeir L, Hothorn T, Rzehak P, Hohle M. Boosting structured additive quantile regression for longitudinal childhood obesity data. *Int J Biostat*. 2013;9(1). DOI: 10.1515/ijb-2012-0035.
35. Borghi E, de Onis M, Garza C, et al. Construction of the World Health Organization child growth standards: Selection of methods for attained growth curves. *Stat Med*. 2006;25:247-65.
36. Wang X, Dey DK. Generalized extreme value regression for binary response data: An application to B2B electronic payments system adoption. *Ann Appl Stat*. 2010;4:2000-23.
37. Archampong D, Borowski D, Willejrgensen P, Iversen LH. Workload and surgeons specialty for outcome after colorectal cancer surgery. *Cochrane Database Syst*. 2012;3(3):CD005391.
38. Arlene SA, Stephen EF, Thomas AL, Sharon-lise TN, Thérèse AS, Utts J. *Statistical Issues In Assessing Hospital Performance*. Commissioned by the Committee of Presidents of Statistical Societies. Baltimore (MD): U.S. Centers for Medicare and Medicaid Services; 2011.
39. Bell A, Fairbrother M, Jones K. Fixed and Random effects models: Making an informed choice. *Qual Quant*. 2019;53(2):1051–74.
40. Feaster D, Brincks A, Robbins M, Szapocznik J. Multilevel models to identify contextual effects on individual group member outcomes: A family example [report]. *Fam Process*. 2011;50:167.
41. Dieleman JL, Templin T. Random-effects, fixed-effects and the within-between specification for clustered data in observational health studies: A simulation study. *PloS One*. 2014;9:e110257.
42. Danks L, Duckett S. *All Complications Should Count: Using Our Data to Make Hospitals Safer*. Carlton (AUST): Grattan Institute; 2018.
43. Hofner B, Mayr A, Robinzonov N, Schmid M. Model-based boosting in R: A hands-on tutorial using the R package mboost. *Comput Stat*. 2014;29:3-35.
44. Koenker R. *Quantile Regression*. Cambridge (UK): Cambridge University Press; 2005.
45. Koenker R. Additive models for quantile regression: Model selection and confidence bandaids. *Braz J Probab Stat*. 2011;25:239-62.
46. Koenker R. (2005). Quantile Regression in R: A Vignette. In: *Quantile Regression (Econometric Society Monographs)*. Cambridge (UK): Cambridge University Press; 2005. p. 295-316.
47. Mundlak Y. On The Pooling Of Time Series And Cross Section Data. *Econometrica*. 1978;46:69.
48. van de Pol M, Wright J. A simple method for distinguishing within- versus between-subject effects using mixed models. *Anim Behav*. 2009;77:753-8.
49. Field K, Shapiro J, Wong HL, et al. Treatment and outcomes of metastatic colorectal cancer in Australia: Defining differences between public and private practice. *Intern Med J*. 2015;45:267-74.
50. Burns EM, Faiz O. Evolution of the surgeon-volume, patient outcome relationship. *Ann Surg*. 2010;251:991-2. Author Reply 2.
51. Gibson S, McConigley R. Unplanned oncology admissions within 14 days of non-surgical discharge: A retrospective study. *Support Care Cancer*. 2016;24:311-17.
52. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology*. 2009;20:488-95.
53. Efron B, Hastie T. *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science*. Cambridge (UK): Cambridge University Press; 2016.
54. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. New York (NY): Springer; 2009.
55. Hofner B, Hothorn T, Kneib T, Schmid M. A framework for unbiased model selection based on boosting. *J Comput Graph Stat*. 2011;20:956-71.
56. Taieb SB, Huser R, Hyndman RJ, Genton MG. Forecasting uncertainty in electricity smart meter data by boosting additive quantile regression. *IEEE Trans Smart Grid*. 2016;7:2448-55.
57. R Core Team. *The R Project for Statistical Computing. A Language and Environment for Statistical Computing*. Vienna (AU): R Foundation for Statistical Computing; 2017.
58. Koenker R. *quantreg: Quantile Regression*. Vienna (AU): Institute for Statistics and Mathematics The R Project for Statistical Computing; 2018.
59. Hothorn T, Buehlmann P, Kneib T, Schmid M, Hofner B. *mboost: Model-Based Boosting*. R Package Version 2.9-1. Vienna (AU): R Foundation for Statistical Computing; 2017
60. Jurasinski G, Koebsch F, Guenther A, Beetz S. *flux: Flux Rate Calculation from Dynamic Closed Chamber Measurements*. Vienna (AU): R Foundation for Statistical Computing; 2014.
61. Skrondal A, Rabe-Hesketh S. *Multilevel Modelling*. Thousand Oaks (CA): SAGE Publications; 2010.
62. Kc DS, Terwiesch C. Impact of workload on service time and patient safety: an econometric analysis of hospital operations. *Manage Sci*. 2009;55:1486-98.
63. Faiz O. The volume–outcome relationship in colorectal surgery. *Tech Coloproctol*. 2014;18(10):961–2.
64. Killeen SD, O'Sullivan MJ, Coffey JC, Kirwan WO, Redmond HP. Provider volume and outcomes for oncological procedures. *Br J Surg*. 2005;92:389-402.
65. Kizer KW. The volume–outcome conundrum. *N Engl J Med*. 2003;349:2159-61.
66. McGrath DR, Leong DC, Gibberd R, Armstrong B, Spigelman AD. Surgeon and hospital volume and the management of colorectal cancer patients in Australia. *ANZ J Surg*. 2005;75:901-10.
67. Jones IT. More evidence of the benefits of case volume and specialization in colorectal surgery. *ANZ J Surg*. 2005;75:839.
68. Morris M, Platell CFE. Surgical volume influences survival in patients undergoing resections for stage ii colon cancers. *ANZ J Surg*. 2007;77:902-6.
69. Byrne BE, Pinto A, Aylin P, Bottle A, Faiz OD, Vincent CA. Understanding how colorectal units achieve short length of stay: An interview survey among representative hospitals in England. *Patient Saf Surg*. 2015;9:2.
70. Cornish JA, Tekkis PP, Tan E, Tilney HS, Thompson MR, Smith JJ. The national bowel cancer audit project: The impact of organisational structure on outcome in operative bowel cancer within the United Kingdom. *Surg Oncol*. 2011;20:e72-7.
71. Van Dorp DR, Boston A, Berri RN. Establishing a complex surgical oncology program with low morbidity and mortality at a community hospital. *Am J Surg*. 2015;209:536-41.
72. Hohl K. Beyond the average case: The mean focus fallacy of standard linear regression and the use of quantile regression for the social sciences. *SSRN*. 2009. DOI: 10.2139/ssrn.1434418.
73. Tukey J. More honest foundations for data analysis. *J Stat Plan Inference*. 1997;57:21-8.
74. John OO. Robustness of quantile regression to outliers. *Am J Appl Math Stat*. 2015;3:86-8.

Supporting Information

Additional supporting information may be found in the online version of this article:

Supplementary File: A table of the 30 surgical procedures that were included in the study; further analytical results not presented in the main paper regarding patient and hospital factors and their association with length of stay following colorectal cancer surgery; and statistical methodology details to be read in conjunction with the methods section in the main body of the article.