cancer surgery: A retrospective analysis. Anesthesiology 2016; 124:69–79
2. Brenner H: Long-term survival rates of cancer patients achieved by the end of the 20th century: A period analysis. Lancet 2002; 360:1131–5
3. Brenner H, Hakulinen T: Very-long-term survival rates of patients with cancer. J Clin Oncol 2002; 20:4405–9
4. Winkelmayer WC, Kurth T: Propensity scores: help or hype? Nephrol Dial Transplant 2004; 19:1671–3
5. Buckley A, Quaid MS, Johnson P, Buggy D: Serum from women undergoing breast cancer surgery, randomized to propofol-paravertebral anaesthetic technique, maintain natural killer cell anti-tumour activity compared with sevoflurane/opioid technique: ESAPC1-5. Eur J Anaesthesiol 2014; 31:2
6. Melamed R, Bar-Yosef S, Shakhar G, Shakhar K, Ben-Eliyahu S: Suppression of natural killer cell activity and promotion of tumor metastasis by ketamine, thiopental, and halothane, but not by propofol: Mediating mechanisms and prophylactic measures. Anesth Analg 2003; 97:1331–9

(Accepted for publication March 28, 2016.)

In Reply:

Many thanks for your comments on our recent retrospective study.[1]

With regard to the first point posed by Drs. Ali and Ghori, concerning the use of the term "long-term cancer survival," we agree that together with 1-yr survival, 5- and 10-yr survival rates are commonly used when reporting cancer survival. However, contextually, mortality rates for perioperative interventions are commonly reported as either 30 days or length of stay, and as such the reported follow-up of between 18 months and 4.5 yr would qualify as long term.

With regard to the use of the propensity model and all-cause mortality data, we agree that a better approach would have been to consider cancer-attributable mortality. However, these data are not reliably available in the United Kingdom. National cancer registries do not cover the broad span of cancers we considered, and in addition often have incomplete data for the early years covered by the study.

We agree that Kaplan–Meier curves for the propensity-matched groups should have been included in the study. These are now included in figure 1, A–C, and as you can see are very similar to those for the nonmatched groups.

Drs. Doleman, Lun, and Williams raise interesting queries regarding our analysis, much of which we are in agreement with.

There is little doubt that cancer type and stage have a profound effect on outcome, and the lack of accurate data within our study for the latter in particular is a potential major confounder. However, as we state in the discussion, the very lack of availability of staging data to the practitioners administering the anesthesia lessens this fact since it could not have been a deciding factor in the choice of anesthetic.

James C. Eisenach, M.D., served as Editor-in-Chief for this exchange.

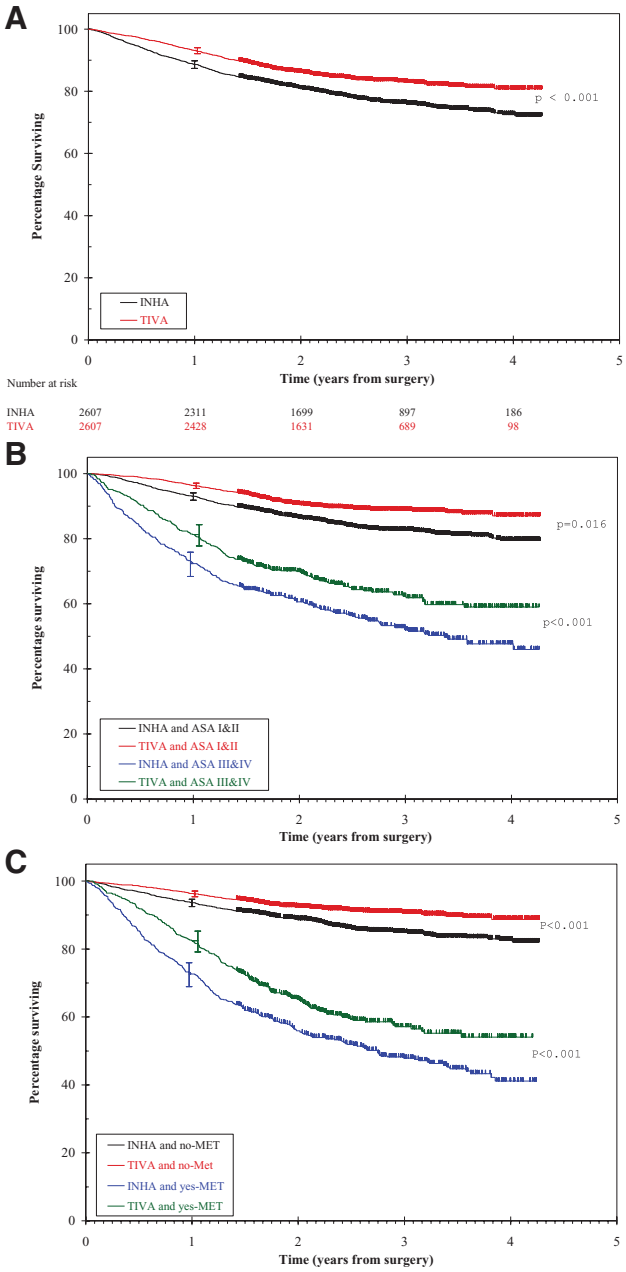

Fig. 1. (*A*) Kaplan–Meier (KM) plot for the propensity-matched patients. (*B*) KM plot for the propensity-matched patients by anesthesia type and American Society of Anesthesiologists (ASA) groups. (*C*) KM plot for the propensity-matched patients by anesthesia type and metastasis status. INHA = volatile inhalational; no-MET = no detected metastases; TIVA = total IV anesthesia; yes-MET = known metastases at the time of surgery.

As far as including cancer types in either the propensity model or the multivariate analysis (beyond the broad groups that have already been included in the analysis, see tables 1 and 2), the major issue is the numbers of different types and subtypes, with consequent substantial implications for outcome. A look through the data reveals more than 20 broad cancer types. Within those types are further subdivisions, for example, triple-negative breast cancer has a very different

Copyright © 2016, the American Society of Anesthesiologists, Inc. Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

Downloaded from http://asa2.silverchair.com/anesthesiology/article-pdf/125/2/420/487620/20160800_0-00036.pdf by guest on 13 March 2024

Table 1. Univariate Analysis for Cancer Type

	Overall Patients			Matched Pairs		
Variables	Events/Total n	HR (95% CI for HR)	*P* Value	Events/Total n	HR (95% CI for HR)	*P* Value
Breast						
TIVA	103/1560	1	0.264	70/876	1	0.541
INHA	52/603	1.210 (0.866–1.690)		42/546	0.887 (0.605–1.302)	
Gastrointestinal tract						
TIVA	137/418	1	0.004	116/363	1	< 0.001
INHA	223/504	1.372 (1.109–1.698)		183/365	1.681 (1.332–2.122)	
Gynecologic						
TIVA	81/331	1	0.221	69/274	1	0.731
INHA	133/428	1.189 (0.901–1.569)		100/347	1.056 (0.776–1.436)	
Sarcoma						
TIVA	77/491	1	0.043	62/344	1	0.673
INHA	128/625	1.340 (1.010–1.778)		85/502	0.932 (0.672–1.293)	
Urology						
TIVA	41/670	1	< 0.001	38/560	1	< 0.001
INHA	81/432	3.094 (2.125–4.506)		51/294	2.545 (1.672–3.875)	

HR = hazard ratio; INHA = volatile inhalational; TIVA = total intravenous anesthesia.

Table 2. Multivariate Analysis for Cancer Type

	Overall Patients		Matched Pairs	
Variables	HR (95% CI for HR)	*P* Value	HR (95% CI for HR)	*P* Value
Gastrointestinal tract				
TIVA	1	0.010	1	0.009
INHA	1.327 (1.069–1.646)		1.379 (1.083–1.756)	
Sarcoma				
TIVA	1	0.697	Not significant in the univariate model	
INHA	1.058 (0.795–1.408)			
Urology				
TIVA	1	0.752	1	0.129
INHA	1.064 (0.723–1.568)		1.405 (0.906–2.178)	

HR = hazard ratio; INHA = volatile inhalational; TIVA = total intravenous anesthesia.

outcome from estrogen receptor–positive breast cancer, as do the different types of thyroid cancer. The numbers of individual cancers are often small, and as a result, statistical power would be lost if we were to subdivide beyond the point that has already been undertaken.

Dr. Dawson makes specific reference to the effects of transfusion, and we agree that there are a significantly greater number of patients who underwent transfusion in the group receiving volatile anesthesia. The fundamental reason for this relates to the allowance margin in the propensity scoring and to the numbers of factors included in the propensity model. While statistically significant, the actual difference in numbers of transfused patients is low at only 40. Given that the unadjusted morality rate for transfused patients is 51%, this would have little impact on the overall mortality difference between the 2 groups (190 patients), but we would agree that it is a confounding factor. Eliminating the difference between groups completely would have resulted in reducing the number of patients in the propensity-matched model further.

The strength of our analysis lies in the fact that it considered a large number of unselected cancer patients admitted for elective surgery. The weakness is that it is a retrospective study with all the inherent problems and unaccounted for confounders that come with that. The lack of data around staging and the small numbers of individual cancer subtypes prohibit further analysis without loss of meaning. As stated in our conclusion, the only assertion we make is that our study found an association between mode of anesthesia administration and mortality. Further adequately powered prospective studies of specific cancer types with comprehensive staging data now need to be undertaken to confirm or refute our findings.

Copyright © 2016, the American Society of Anesthesiologists, Inc. Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

Downloaded from http://asa2.silverchair.com/anesthesiology/article-pdf/125/2/420/487620/20160800_0-00036.pdf by guest on 13 March 2024

Competing Interests

The authors declare no competing interests.

Timothy J. Wigmore, M.A., F.R.C.A., F.F.I.C.M., F.C.I.C.M., Shaman Jhanji, Ph.D., M.R.C.P., F.R.C.A., F.F.I.C.M., Kabir Mohammed, M.Sc. Royal Marsden Foundation Trust, London, United Kingdom (T.J.M.). timothy.wigmore@rmh.nhs.uk

Reference

1. Wigmore TJ, Mohammed K, Jhanji S: Long-term survival for patients undergoing volatile *versus* IV anesthesia for cancer surgery: A retrospective analysis. Anesthesiology 2016; 124:69–79

(Accepted for publication April 6, 2016.)

Current Quality Registries Lack the Accurate Data Needed to Perform Adequate Reliability Adjustments

To the Editor:

We would like to thank Drs. Wakeam and Hyder[1] for their excellent discussion and description of reliability adjustment in the recent issue of Anesthesiology. The authors correctly highlight the important role that the statistical analysis of data submitted to the various registries can play in the ranking of institutions. This is particularly important now that the Centers for Medicare and Medicaid Services requires providers to participate in a Physician Quality Reporting System[2] using a Qualified Clinical Data Registry. These requirements are a precursor to altering physician payments based upon measures of care quality.

We would like to raise the issue of another area of "reliability": the reproducibility of the underlying data themselves. While some registries such as National Surgical Quality Improvement Program do periodic data audits and have well-described accuracy thresholds,[3] many do not. In fact, some registries, including the Anesthesia Quality Institute and the American Society of Anesthesiologists Perioperative Surgical Home initiative, allow for widely divergent methods of data collection, yet lump these data together assuming they are comparable. For example, one group might define postoperative nausea and vomiting based on postanesthesia care unit antiemetic administration, while another bases it on direct patient interviews. Other registries, such as some maintained by the National Quality Foundation, utilize administrative claims data, which have been shown to be discordant with data collected by other methods.[4–8] Despite these very different methods of data collection, all of these examples are considered equally valid national quality registries.

We find the idea that the underlying data used in these registries may be inconsistent to be worrisome. Ideally, the data on patients in various registries should be identical regardless of the method by which they were collected. At the very least, even if the data are not identical between registries, it is critical that within a registry, the data from various sites be of equal quality and have the same definitions, something the major registries in our own specialty lack. If the data inputs are not consistent, we are left with the question of which data to believe, and the conclusion is that the risk adjustment models used may be unable to control for patient-specific risk factors the way they are intended.

It seems inevitable that in the near future, providers will be compared to each other and paid partially based on these comparisons. This concept is based upon the unverified supposition that we can effectively compare patients across institutions. On the basis of the current landscape, we find this supposition unlikely, and we are concerned that using these inadequate tools may lead to incorrect choices in the near future. Drs. Wakeam and Hyder are absolutely correct that "big data" require more than assembling a large sample size and assuming that the "N" will solve the problem, but rather a thorough understanding of statistics and attention to detail. Unfortunately, it seems that the goals of some of the quality registries are outpacing the science behind them.

Competing Interests

The authors declare no competing interests.

Ira S. Hofer, M.D., Yannick Le Manach, M.D., Ph.D., Maxime Cannesson, M.D., Ph.D. David Geffen School of Medicine at UCLA, Los Angeles, California (I.S.H.). ihofer@mednet.ucla.edu

References

1. Wakeam E, Hyder JA: Reliability of reliability adjustment for quality improvement and value-based payment. Anesthesiology 2016; 124:16–8
2. Manchikanti L, Hirsch JA: Regulatory burdens of the Affordable Care Act. Harvard Health Policy Rev 2012; 13:9–12
3. Shiloach M, Frencher SK Jr, Steeger JE, Rowell KS, Bartzokis K, Tomeh MG, Richards KE, Ko CY, Hall BL: Toward robust information: Data quality and inter-rater reliability in the American College of Surgeons National Surgical Quality Improvement Program. J Am Coll Surg 2010; 210:6–16
4. McIsaac DI, Gershon A, Wijeysundera D, Bryson GL, Badner N, van Walraven C: Identifying obstructive sleep apnea in administrative data: A study of diagnostic accuracy. Anesthesiology 2015; 123:253–63
5. Quach S, Blais C, Quan H: Administrative data have high variation in validity for recording heart failure. Can J Cardiol 2010; 26:306–12
6. Quan H, Parsons GA, Ghali WA: Validity of information on comorbidity derived from ICD-9-CCM administrative data. Med Care 2002; 40:675–85
7. Pasquali SK, He X, Jacobs JP, Jacobs ML, Gaies MG, Shah SS, Hall M, Gaynor JW, Peterson ED, Mayer JE, Hirsch-Romano JC: Measuring hospital performance in congenital heart surgery: Administrative *versus* Clinical Registry Data. Ann Thorac Surg 2015; 99:932–8

Copyright © 2016, the American Society of Anesthesiologists, Inc. Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

Downloaded from http://asa2.silverchair.com/anesthesiology/article-pdf/125/2/420/487620/20160800_0-00036.pdf by guest on 13 March 2024