Therapeutics and Clinical Risk Management

Dovepress
open access to scientific and medical research

Open Access Full Text Article

REVIEW

Is There a Surgeons' Effect on Patients' Physical Health, Beyond the Intervention, That Requires Further Investigation? A Systematic Review

Christoph Schnelle [1], Justin Clark [1], Rachel Mascord [2], Mark A Jones [1]

[1]Institute for Evidence-Based Healthcare, Bond University, Robina, Queensland, Australia; [2]General Dentist, BMA House, Sydney, New South Wales, Australia

Correspondence: Christoph Schnelle, Institute for Evidence-Based Healthcare, Bond University, Robina, Queensland, Australia, Email christoph.schnelle@student.bond.edu.au

Video Abstract

Point your SmartPhone at the code above. If you have a QR code reader the video abstract will appear. Or use: https://youtu.be/pL-eGyAGhSk

Objective: To find and review published papers researching surgeons' effects on patients' physical health. Clinical outcomes of surgery patients with similar prognoses cannot be fully explained by surgeon skill or experience. Just as there are "hospital" and "psychotherapist" effects, there may be "surgeons" effects that persist after controlling for known variables like patient health and operation riskiness.
Methods: Cohort studies and randomized controlled trials (RCTs) of any surgical intervention, which, after multivariate adjustment, either showed proportion of variance in patients' physical health outcomes due to surgeons (random effects) or graded surgeons from best to worst (fixed effects). Studies with <15 surgeons or only ascribing surgeons' effects to known variables excluded. Medline, PubMed, Embase, and PsycINFO were used for search until June 2020. Manual search for papers referring/referred by resulting studies. Risk of bias assessed by Cochrane risk-of-bias tool and Newcastle–Ottawa Scale.
Results: Included studies: 52 cohort studies and three RCTs of 52,436+ surgeons covering 102 outcomes (33 unique). Studies either graded surgeons from best to worst or calculated the intra-class correlation coefficient (ICC), the percentage of patients' variation due to surgeons, in diverse ways. Sixteen studies showed exceptionally good and/or bad performers with confidence intervals wholly above or below the average performance. ICCs ranged from 0 to 47%, median 4.0%. There are no well-established reporting standards; highly heterogeneous reporting, therefore no meta-analysis.
Discussion: Interpretation: There is a surgeons' effect on patients' physical health for many types of surgeries and outcomes, ranging from small to substantial. Surgeons with exceptional patient outcomes appear regularly even after accounting for all known confounding variables. Many existing cohort studies and RCTs could be reanalyzed for surgeons' effects especially after methodological reporting guidelines are published.
Conclusion: In terms of patient outcomes, it can matter which surgeon is chosen. Surgeons with exceptional patient outcomes are worth studying further.
Keywords: physicians, physicians' effect, doctors' effect, therapists' effect, practice effect, clinical competence, professional practice gap, surgeons' practice pattern, quality of health care, delivery of health care

Introduction

What is already known on this topic: Previous research has shown associations between characteristics of surgeons, such as their level of surgical experience, and patient health outcomes. It is unclear whether surgeons have an influence on patients' physical health that has not been captured by known variables and how large that influence is.

What this study adds: This study is the first systematic review of unexplained surgeons' influence on patients' physical health. Findings are highly variable, depending on the type of outcome and surgery that can result in substantial differences in patient health outcomes between surgeons.

Received: 20 January 2022
Accepted: 4 April 2022
Published: 26 April 2022

Rationale

If you want to find a good surgeon, an internet query will provide advice from many sources.[3–5] There are also databases of what the database provider considers to be the best surgeons, calculated from raw death and complication rates plus other doctors' recommendations.[6] Surgeons themselves have given their opinion on what makes a good or outstanding surgeon,[7–11] with Barry Jackson's essay perhaps being the most comprehensive.[12] However, this information mostly relies on personal experiences, although Jackson's essay does mention "First-class outcomes after allowing for case-mix". Existing evidence suggests that some surgeons are more effective at applying interventions than others as there is, for example, a substantial volume effect, ie case volume, and years of practice effect in a number of surgical specialties.[13,14] In fact, there are few studies where author-selected outstanding practitioners are investigated,[15–17] with only Schenck et al mentioning surgeons.

It is well established that there is a hospital effect, ie that hospitals have a substantial influence on patients' health outcomes and that there are wide variations in patients' health outcomes between hospitals.[18–23] There is also substantial research on a therapist effect in psychotherapy with wide variations among therapists, so much so that this finding has made it into training material for psychotherapists.[24,25] Recent research also suggests that provider expectations could have a causal role in treatment effectiveness.[26] At the same time, the placebo effect, which can be substantial,[27] including in surgery[28–31] with some dissent for orthopedic surgery,[32] is suggestive of a surgeon's effect. The placebo effect shows that even with an inert or inactive intervention, there is an effect on patients. It is possible that part of this effect is due to the surgeon administering the placebo, usually a type of sham surgery.[28] However, there are currently no well-established standards on how to assess surgeons as an intervention in their own right or as an effect modifier of a given intervention. Recent research has endeavored to analyze the effect size of surgeons[33] by investigating 10 surgical trials for surgeon intra-cluster correlation coefficients (ICCs) and found that surgeons have a range of effects on patient health that differ between surgical specialties.

In the study by Udyavar et al[34] of 2149 surgeons performing 569,767 emergency surgeries it was shown that in five out of seven types of surgery, surgeons were responsible for 23% to 47% of the variability in patient mortality. This difference in outcome could not be explained by the choice of treatment, prognostic or diagnostic factors, patient clinical or demographic factors, hospital-level factors, or surgeon volume. To date studies such as Udyavar et al have not been synthesized. In this systematic review we have addressed this gap in the literature.

Objectives

This systematic review aims to identify and evaluate all the research to date examining the effect of surgeons on patient physical health outcomes after known variables have been accounted for. It is part of a larger research project that includes a systematic review of non-surgical practitioners, and a methodological study on how to report practitioners' effects on patients' physical health.

Methods

Eligibility Criteria

A systematic review was conducted following Synthesis without meta-analysis (SWiM) guidelines.[35] This review limits itself to studies that investigated actual patients' physical outcomes and excluded studies that focused on patients' opinions or satisfaction levels, with the rationale that these outcomes are often a more ambiguous way to measure surgeons' effects.[36]

The PICO is as follows:

Population	**P**	**Surgeons**
Intervention	I	
Comparison	C	
Outcome	O	Surgeons' effect on patients' physical health outcome

Information Sources and Search Strategy

We initially searched three databases: Medline via its PubMed interface, Embase, and PsycINFO from inception to June 2020 to identify relevant studies that investigate the influence of surgeons on patients' physical health outcomes. The search strategy used for each database is reported in Supplemental File 1 and was designed by JMC, a specialist in this area. In addition to the electronic search of databases, we further manually searched the references lists of the eligible articles and previous systematic reviews to identify potentially relevant studies that did not appear in the literature search. The following systematic review registries were searched for similar reviews: PROSPERO and Cochrane's CENTRAL register. One study was suggested by a reviewer.[37]

Selection Process and Further Eligibility Criteria

Two reviewers independently screened titles/abstracts for inclusion. Any discrepancies were resolved by discussion or in consultation with a third reviewer.

Study designs considered for inclusion were retrospective and prospective observational studies, case-control studies, and randomized controlled studies, where either the proportion of variance in patient outcomes explained by differences between practitioners, ie practitioners' random effects, are measured, or the difference between the individual practitioners is highlighted, ranging from best to worst, ie practitioners' fixed effects are measured. Any medical practitioner except psychotherapists were included. At this stage both surgeons and doctors who were not surgeons were included, however this paper only includes studies of surgeons. All other medical doctors are reviewed in a separate paper.

Any patient's physical health-related outcome was eligible, examples of which are repair reoperations, readmission rate, survival/mortality rate, embryo transfer rate, length of hospital stay, infection rate, estimated blood loss, recurrence rates, pain, and other post-operative complications. There were no date or language restrictions.

We excluded studies that only ascribed a surgeons' effect to particular surgeon-related variables, such as volume of procedures performed or specialty of surgeon; studies with fewer than 15 surgeons; cross-sectional studies, ie surveys of doctors or patients, as they had an increased risk of bias; and two studies that mentioned fixed or random effects but did not actually list the effects either graphically or in numerical form.[38,39]

The authors could not find a recommendation for the minimum number of clusters in a study for a systematic review – in this case the minimum number of practitioners. We took 15 practitioners as the smallest cluster size but appreciate that this is an arbitrary number. (Figure 1).

Data Collection Process and Data Items

Titles and abstracts were collected using Endnote 9 and uploaded into Rayyan for inclusion or exclusion where the two reviewers independently screened titles and abstracts. The resulting eligible studies were marked as members of a group in the original Endnote library and their full text documents were added to the library.

CS and a second extractor independently and in duplicate extracted the relevant data from each eligible study and collected the following variables using Excel:

- Unique publication identifier consisting of first author and year
- Surgical specialty
- Type of study (RCT, Cohort)
- Type of intervention (can be multiple)
- Outcome type (multiple)
- Significant surgeons' effect as per authors' evaluation Y/N
- Number of surgeons
- Number of patients or procedures
- Number of hospitals/institutions
- ICC (intra-class correlation coefficient) Number/NS
- Multivariate analysis Y/N

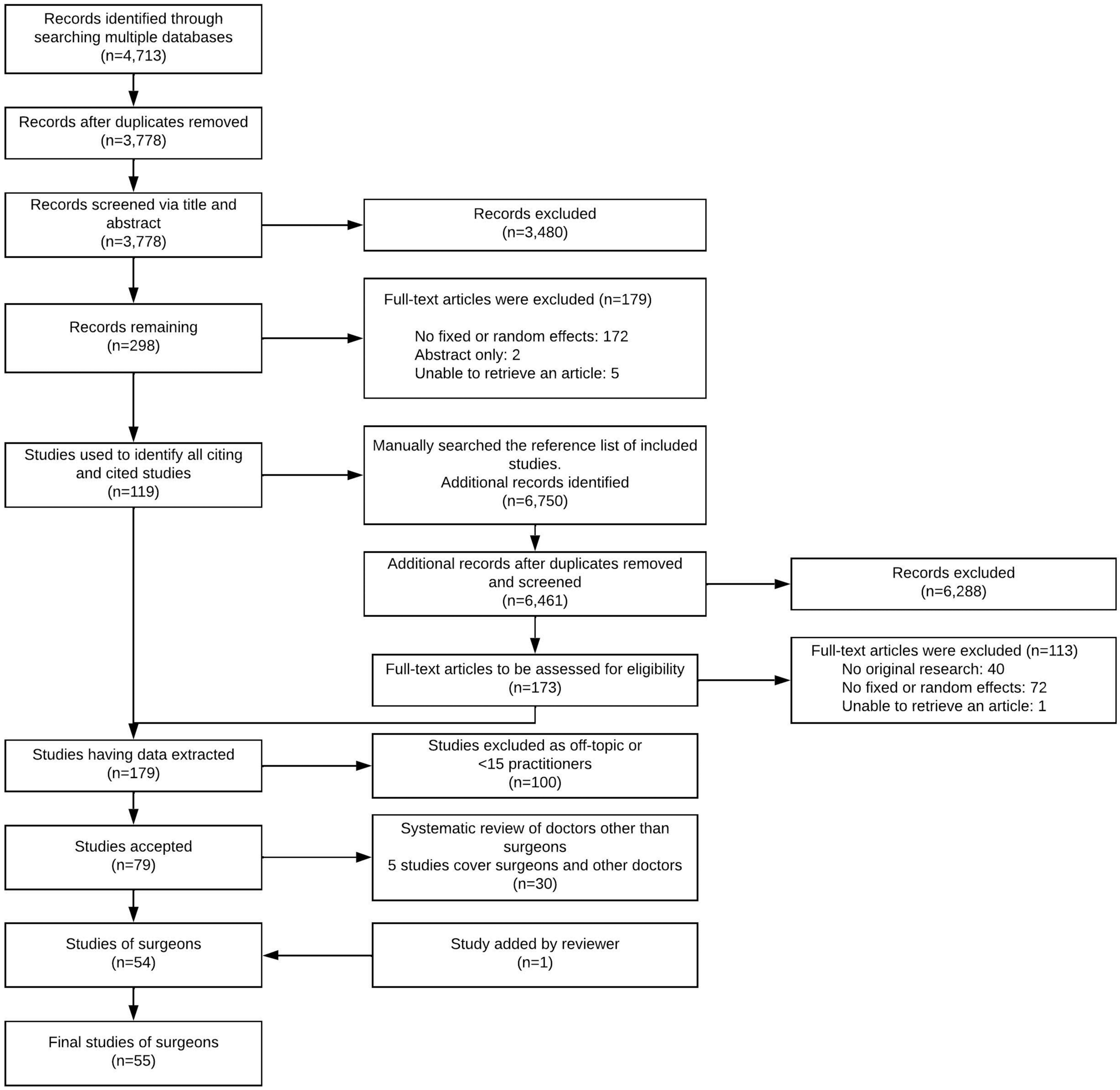

Figure 1 Flow diagram of selection of included documents.

- Number of negative and positive outliers
- Country of origin

Study Risk of Bias Assessment

Two reviewers independently used the Cochrane risk-of-bias assessment tool[2] for the three included randomized controlled studies (Figures 2 and 3), and the Newcastle-Ottawa Scale (NOS) for the included cohort studies.[40,41]

Effect Measures

The metric for the fixed effects is the percentage of positive and negative outliers as defined in the individual study reports. The metric for the random effects is the variance due to the practitioner or the intra-class correlation coefficient, defined as the variation in patient outcome due to the practitioner as a percentage of the total variation.

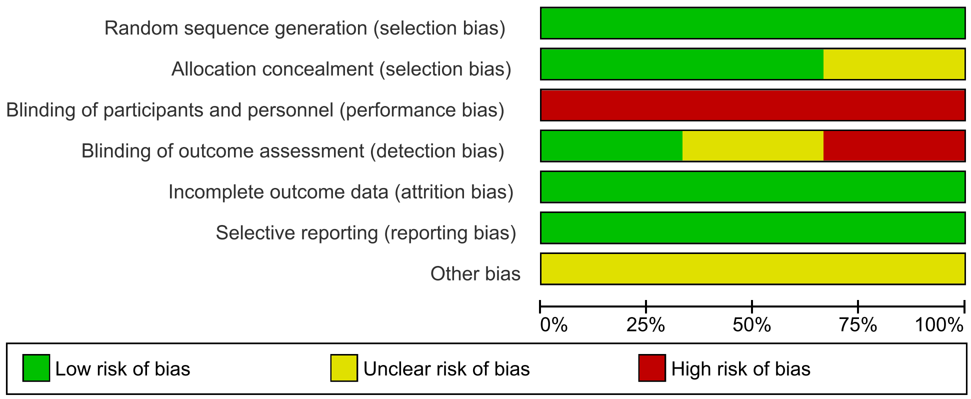

Figure 2 First risk of bias chart for the three randomized controlled trials included.

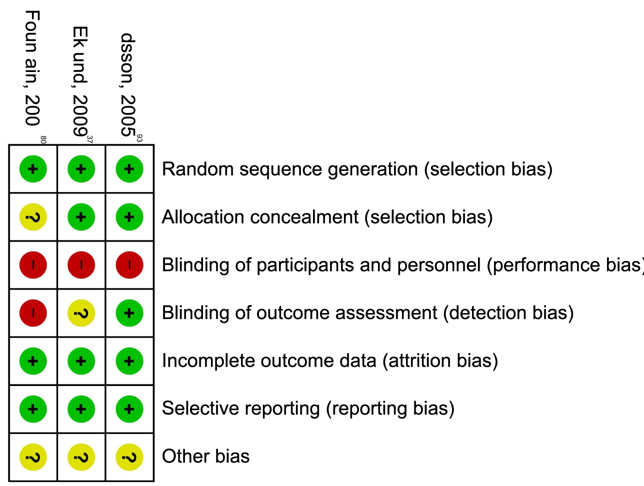

Figure 3 Second risk of bias chart for the three randomized controlled trials included.

Synthesis Methods

As the data are highly heterogenous and there are no established standards on recording doctors' effects or surgeons' effects, no statistical synthesis was used. There were 14 surgical specialties plus two papers covering multiple surgeries, 50 separate interventions and 31 separate outcomes.

The surgeons' effect on patients' physical health is described in two ways, using multilevel mixed effects regression modelling or hierarchical regression to understand both surgeon and system-level variation.[42,43]

Percentage of Variation in Patient Outcome Due to the Surgeon in the Form of the Intra-Class Correlation Coefficient (ICC)

Post-regression estimation gives the ICC, which as a number ranging from 0 to 1, gives the percentage of variation in outcome due to each level in the regression model. For example, in a three-level model of patients clustered per doctor, who in turn were clustered within hospitals, each level has an ICC with the total ICCs adding up to 1. In order to realize this, the studies included random effects for surgeons, and at times hospitals or other aggregators, such as county.

Patient risk scores and other available variables like surgeon demographic data or year of intervention were included as fixed effects in the regression analysis. The quality and depth of the analysis varied greatly between papers. Confidence intervals for the ICC were not reported.[44] A high quality study is Papachristofi et al.[45] There is also substantial other research on the ICC.[46–51]

Grading Surgeons from Best to Worst

In this approach surgeons are ranked by their patient results, usually with a 95% confidence interval and either the raw, unadjusted scores are reported, or patient risk scores and/or surgeon demographic variables and other data, such as year of operation, are included in the model. In the majority of cases the method to calculate the interval is not mentioned, though there are exceptions[52–59] and surgeons whose 95% confidence intervals rank wholly above or below the mean rate of outcomes are considered to be outliers. Reporting is done by listing the count of outliers, or graphically through a caterpillar or a funnel plot,[60] with a caterpillar plot being an outcome-ordered forest plot.

Reporting Bias and Certainty Assessment

Due to there being no synthesis, reporting bias and certainty assessments were not undertaken.

Results

Study Selection

Overall, 4713 records were identified from electronic records, in addition to 6461 from other sources. After removing the 1224 duplicates, 10,239 studies underwent screening for eligibility. Then, full-text versions were retrieved for 471 records. One study was added by a reviewer. Finally, after exclusion of ineligible articles, 55 studies of more than 52,436 surgeons were included in the final synthesis.

Study Characteristics

The 55 studies that are included reported 102 outcomes, 33 of which are unique. Of the outcomes, 28 (20 studies)[55–57,59,61–76] graded individual surgeons' performance from best to worst; 38 (12 studies)[34,77–87] recorded an ICC due to surgeons in a multivariate multi-level analysis; 14 (8 studies)[44,45,53,58,88–91] recorded both; 20 (13 studies)[37,52,92–102] provided a non-standard description of fixed effects; and 1 provided an ICC plus a non-standard description of fixed effects.[103] One study[104] graded surgeons from best to worst in one outcome (complications) and used a non-standard fixed effects description for another outcome (mortality).

Of the 55 studies, three were randomized controlled trials,[37,80,93] and 52 were observational cohort studies. The studies included various surgical specialties or aggregates thereof, including 8 or more specialties,[81,91] breast surgery,[59,73,74,101] cardiac surgery,[44,45,52,53,56,65–67,70,89,90,99,100] colorectal surgery,[34,61,68,71,78,85,92,103,104] ENT surgery,[75] gastrointestinal surgery,[34,83] general surgery,[34,37,57,77,79,82,86,93,95,96,102] obstetrics,[80] ophthalmology,[97] orthopedic surgery,[55,76,84,102] rectal surgery,[72,98] spinal surgery,[58] trauma surgery,[87] and urology.[62–64,69,88,94,102] 38 studies were conducted in the USA, 10 in the UK, two in Austria and Sweden, one each in Canada, France, and Germany. The volume of included surgeons ranged from 17 to 14,598. The characteristics of the included studies are summarized in Table 1.

Risk of Bias in Studies

For the cohort studies, of 97 outcomes in 52 studies, (1) scored 7 stars, (21) 8 stars and (75) 9 stars out of a maximum of 9 stars on the Newcastle-Ottawa Scale.[40,41] All studies scored the maximum points on the selection criteria and the outcome criteria. Those with 7 and 8 stars scored either 0 or 1 on comparability while the 9-star studies scored 2 (Table 1). The detailed risk of bias assessment of the three randomized controlled trials, using Cochrane RoB, is described in Figures 2 and 3, and Supplemental File 2.

Results of Individual Studies

Altogether 10 studies published caterpillar plots[59,61,64,67,71,74–76,89,91] and five studies presented funnel plots.[65,66,68–70] The plots showed the performance of surgeons for a particular patient outcome, usually sorted by performance, providing a 95% confidence interval for each surgeon and indicating whether that confidence interval was wholly above or below the average performance. Results ranged from no over- or underperformer[62,66,67,69,70,91] to substantial numbers of both.[59,61,62,64,72,73,75,76]

Table 1 Characteristics of Included Studies

Publication	Specialty	Detailed Intervention	Surgeons	Patients/ Procedures	Institutions	Outcome	NOS **
Anderson, 2016[52]	Cardiac surgery	Norwood operation	NS	2880	35	Mortality	9
Aquina, 2015a[92] pg e163	Colorectal surgery	Colorectal resection	NS	158,596	NS	C. difficile infection	9
Aquina, 2015b[61]	Colorectal surgery	Upper GI cancer resection	223	14,875	99	Blood transfusion, wound infection, pneumonia, sepsis	9
Aquina, 2016[103]	Colorectal surgery	Colorectal resection	3481	125,160	210	Blood transfusion, wound infection, pneumonia, sepsis, intra-abdominal abscess	9
Aquina, 2017[77]	General surgery	Inguinal hernia operation	1572	124,416	260	Reoperation	9
		Ventral hernia operation	2012	78,267	256	Reoperation	9
Arvidsson, 2005[93]	General surgery	Hernia operation	25	1068	7	Recurrence	RCT
Becerra, 2017[78]	Colorectal surgery	Lymph node examination in colectomy	1503	12,332	187	Suboptimal care	9
Begg, 2002[62]	Urology	Radical prostatectomy	159	10,737	72	Postoperative complications	9
						Incontinence	9
						Late urinary complications	9
						Mortality	9
Bianco, 2005[63]	Urology	Radical prostatectomy	159	5238	NS	Complications	9
						Incontinence	9
						Late urinary complications	9
Bianco, 2010[64]	Urology	Radical prostatectomy	54	7725	4	Cancer recurrence	9
Bolling, 2010[65]	Cardiac surgery	Mitral valve repair/ replacement	1088	28,507	639	Mitral valve repair rates	9
Bridgewater, 2003[67]	Cardiac surgery	Coronary artery surgery	23	8572	4	Mortality	9
Bridgewater, 2005[66]	Cardiac surgery	Aortic valve surgery	25	1097	4	Mortality	8
		Coronary artery surgery	25	9066	4	Mortality	8
Burns, 2011[68]	Colorectal surgery	Colorectal surgery	1557	246,469	156	Reoperation	9
Cromwell, 2013[69]	Urology and Gynecology	Urinary-genital tract fistula	490	1194	129	Reoperation	8
Dagenais, 2019[88]	Urology	Partial nephrectomy	19	1461	1	Estimated blood loss	9
Duclos, 2012[79]	General surgery	Thyroid surgery	28	3574	5	Hypoparathyroidism	9
						Recurrent laryngeal nerve palsy	9
Eastham, 2003[94]	Urology	Radical prostatectomy	44	4629	2	Positive surgical margins	9

(*Continued*)

Table 1 (Continued).

Publication	Specialty	Detailed Intervention	Surgeons	Patients/ Procedures	Institutions	Outcome	NOS **
Eklund, 2009[37]	General surgery	Inguinal hernia repair surgery	48	1275	NS	Recurrence	RCT
Faschinger, 2011[95]	General surgery	Cataract surgery	17	36,329	1	Capsule rupture	9
Fountain, 2004[80]	Obstetrics	Hysterectomy, Abdominal	43*	876	28*	Complications	RCT
		Hysterectomy, Vaginal	43*	504	28*	Complications	RCT
Gani, 2015[81]	[See on right]	8 (cardiac, GI* surgery, trauma, HPB*, BME*, thoracic, transplant, vascular)	56	22,559	1	Readmission	9
Glance, 2006[53]	Cardiac surgery	Cardiac surgery	138	51,750	33	Mortality	9*
Glance, 2016[89]	Cardiac surgery	CABG*	241	55,436	40	Major complications or mortality	9
Grant, 2008[70]	Cardiac surgery	Cardiac surgery	31	14,637	4	Mortality	9
Healy, 2017[71]	Colorectal surgery	Minimally invasive colectomy	97	3118	46	Complications	8
		Open colectomy	97	2078	46	Complications	8
Hermanek, 1999[72]	Rectal surgery	Rectal carcinoma resection	43	1121	7	Mortality	9
Hermann, 2002[96]	General surgery	Primary surgery for benign thyroid disease	20	16,443	1	Recurrent laryngeal nerve injury (RLNI)	8
Hoffman, 2017[82]	General surgery	General surgery	1128	183,283	601	Complications	9
Huesch, 2009[56]	Cardiac surgery	CABG*	398	221,327	75	Mortality	8
Hyder, 2013[83]	Gastrointestinal surgery	Pancreatoduodenectomy	575	1488	298	Readmission	9
Johnston, 2010[97]	Ophthalmologist	Cataract surgery	404	55,515	12	Posterior capsule rupture (PCR)	8
Justiniano, 2019[98]	Rectal surgery	Rectal carcinoma resection	345	1251	118	Mortality	9
Kaczmarski, 2019[73]	Breast surgery	Breast-conserving surgery	5337	291,065	NS	Reoperation	9
Kissenberth, 2018[84]	Orthopedic surgery	Rotator cuff repair	57	1703	NS	Single Assessment Numeric Evaluation (SANE) score	8
Landercasper, 2019[74]	Breast surgery	Breast-conserving surgery	71	3954	NS	Reoperation	9
LaPar, 2014[99]	Cardiac surgery	Mitral valve repair/ replacement	93	4194	17	Lack of repair	8
Likosky, 2012[100]	Cardiac surgery	CABG*	32	11,838	8	Postoperative low-output failure	9
Luan, 2019[57]	General surgery	Bariatric surgery	38	1277	21	Complications	9
Martin, 2013[58]	Spinal surgery	Lumbar fusion	298	6091	43	Complications	9
Martin, 2013[58]	Spinal surgery	Lumbar fusion	298	6091	43	Reoperation	9
McCahill, 2012[59]	Breast surgery	Breast-conserving surgery	54	2206	4	Reoperation	9

(*Continued*)

Table 1 (Continued).

Publication	Specialty	Detailed Intervention	Surgeons	Patients/ Procedures	Institutions	Outcome	NOS **
Papachristofi, 2014[90]	Cardiac surgery	Cardiac surgery	18	18,426	1	Mortality	9
Papachristofi, 2016[44]	Cardiac surgery	Cardiac surgery	127	110,769	10	Mortality	9
Papachristofi, 2017[45]	Cardiac surgery	Cardiac surgery	127	107,038	10	Length of stay	9
Quinn, 2018[91]	All surgeries	All surgeries	2724	123,141	51	Any morbidity	9
						Death or serious morbidity	9
						Mortality	9
						Readmission	9
						Reoperation	9
						Surgical site infection	9
Rudmik, 2017[75]	ENT surgery	Endoscopic sinus surgery	43	2168	NS	ESS* revision rate	9
Schumacher, 2017[101]	Breast surgery	Breast-conserving surgery	93	3470	56	Reoperation	7
Shih, 2015[85]	Colorectal surgery	Colectomy	345	5033	24	Complications	9
Singh, 2018[76]	Orthopedic surgery/ Neurosurgery	Spine surgery	3987	39,884	NS	Length of stay	8
						Readmission	8
Thigpen, 2018[55]	Orthopedic surgery	Rotator cuff repair	34	995	1	ASES* performance score	8
Udyavar, 2018a[34]	Colorectal surgery	Colectomy	2149*	569,767*	225*	Complications	9
						Mortality	9
						Readmission	9
	Gastrointestinal surgery	Peptic ulcer disease	2149*	569,767*	225*	Complications	9
						Mortality	9
						Readmission	9
		Small bowel resection	2149*	569,767*	225*	Complications	9
						Mortality	9
						Readmission	9
	General surgery	Appendectomy	2149*	569,767*	225*	Complications	9
						Mortality	9
						Readmission	9
		Cholecystectomy	2149*	569,767*	225*	Complications	9
						Mortality	9
						Readmission	9
		Laparotomy	2149*	569,767*	225*	Complications	9
						Mortality	9
						Readmission	9
		Lysis of adhesions	2149*	569,767*	225*	Complications	9
						Mortality	9
						Readmission	9

(*Continued*)

Table 1 (Continued).

Publication	Specialty	Detailed Intervention	Surgeons	Patients/ Procedures	Institutions	Outcome	NOS **
		Overall	2149	569,767	225	Complications	9
						Mortality	9
						Readmission	9
Udyavar, 2018b[87]	Trauma surgery	Trauma surgery	175	65,706	31	Mortality	9
Udyavar, 2019[86]	General surgery	Emergency surgery	5816	215,745	198	Complications	9
Xu, 2016[104]	Colorectal surgery	Colectomy	276	2525	44	Complications	9
						Mortality	9
Xu, 2019[102]	General surgery	Laparoscopic cholecystectomy	2476	1,884,842*	NS	Complications	8
	Urology	Other transurethral prostatectomy	1663*	1,884,842*	NS	Complications	8
		Radical prostatectomy	1663*	1,884,842*	NS	Complications	8
	Orthopedic surgery	Cervical spinal fusion	10,459*	1,884,842*	NS	Complications	8
		Lumbar spinal fusion, anterior column	10,459*	1,884,842*	NS	Complications	8
		Lumbar spinal fusion, posterior column	10,459*	1,884,842*	NS	Complications	8
		Total hip arthroplasty	10,459*	1,884,842*	NS	Complications	8
		Total knee replacement	10,459*	1,884,842*	NS	Complications	8

Notes: *The values are for the whole study population. Values for each subgroup were not reported.
Abbreviations: **ASES, American Shoulder and Elbow Surgeons; BME, breast, melanoma, and endocrine surgery; CABG, coronary artery bypass grafting; ESS, endoscopic sinus surgery; EQ-5D, quality of life via the Euro-Qol; GI, gastro-intestinal; HPB, hepatopancreatic biliary; NOS, Newcastle-Ottawa Scale for assessing risk of bias of cohort studies; NS, not stated, the number is not given and most likely greater than one.

Of the papers that reported fixed effects, 15 recorded exceptional performers after taking account of all known variables, including demographic variables of the practitioners, such as experience, volume of patients/procedures, and hospital effects (which themselves can be substantial).[55,59,61–66,68,71,73–76,91] Other studies (n=22) published a random effect, worded many different ways, that showed the Intra-Class Correlation (ICC) effect.[34,44,45,53,57,58,77–91,103] The random effects reported ranged from zero (ICC of 0.0%) to substantial (ICC of 10% or higher). (Tables 2 and 3, and Figures 4 and 5).

Only for a. complications after colectomy and b. mortality after cardiac surgery was there more than one study included that reported an ICC. As these are the only outcomes with multiple ICCs, a more detailed analysis follows:

For colectomy, Shih et al[85] reported an ICC of 14.0% and Udyavar et al[34] an ICC of 2.3%. Udyavar defined complications as any of "*pulmonary embolism, sepsis, myocardial infarction, acute renal failure, and cardiac arrest*" while Shih defined a much longer list of items as complications, including surgical site infection; wound disruption; multiple types of infection; unplanned intubation; transfusion; multiple stroke or clotting diagnoses; multiple heart issues; renal complications or failure; extended coma or mechanical ventilation; nerve damage; failure of the graft or prosthesis; bowel obstruction; and anastomotic leak. For mortality after cardiac surgery three studies[44,53,90] reported an ICC of 2.8% to 5.9% (Table 2).

Results of Syntheses, Reporting Biases and Certainty of Evidence

Not applicable as there was no synthesis.

Table 2 Publications by Outcome and Numerical Results

	Outcome	**Specialty**	**Detailed Intervention**	**Publication**	**ICC^**	**Outliers %**	
						Negative	**Positive**
Complications	Any morbidity	All surgeries	All surgeries	Quinn, 2018[91]	2.2%	0.18%	0.22%
	Blood transfusion, wound infection, pneumonia, sepsis	Colorectal surgery	Upper GI cancer resection	Aquina, 2015b[61]		13.0%	28.0%
	Blood transfusion, wound infection, pneumonia, sepsis, intra-abdominal abscess	Colorectal surgery	Colorectal resection	Aquina, 2016[103]	24.3%	Other	Other
	C. difficile infection	Colorectal surgery	Colorectal resection	Aquina, 2015a[92] pg e163		Other	Other
	Capsule rupture	General surgery	Cataract surgery	Faschinger, 2011[95]		Other	Other
	Complications (postoperative)	Colorectal surgery	Colectomy	Shih, 2015[85]	14.0%		
				Udyavar, 2018a[34]	2.3%		
				Xu, 2016[104]		3.3%	NS
			Minimally invasive colectomy	Healy, 2017[71]		10.3%	7.2%
			Open colectomy	Healy, 2017[71]		9.3%	5.2%
		Gastrointestinal surgery	Peptic ulcer disease	Udyavar, 2018a[34]	0.03%		
			Small bowel resection	Udyavar, 2018a[34]	0.02%		
		General surgery	Appendectomy	Udyavar, 2018a[34]	0.2%		
			Bariatric surgery	Luan, 2019[57]		2.6%	15.8%
			Cholecystectomy	Udyavar, 2018a[34]	0.1%		
			Emergency surgery	Udyavar, 2019[86]	27.3%		
			General surgery	Hoffman, 2017[82]	6.2%		
			Laparoscopic cholecystectomy	Xu, 2019[102]		Other	Other
			Laparotomy	Udyavar, 2018a[34]	0.1%		
			Lysis of adhesions	Udyavar, 2018a[34]	0.0%		

(*Continued*)

https://doi.org/10.2147/TCRM.S357934

https://doi.org/10.2147/TCRM.S357934

Table 2 (Continued).

	Outcome	Specialty	Detailed Intervention	Publication	ICC^	Outliers %	
						Negative	Positive
			Overall (Emergency general surgeries)	Udyavar, 2018a[34]	0.1%		
		Obstetrics	Hysterectomy, Abdominal	Fountain, 2004[80]	7.4%		
			Hysterectomy, Vaginal	Fountain, 2004[80]	0.5%		
		Spinal surgery	Lumbar fusion	Martin, 2013[58]	2.6%	3.7%	0.0%
		Urology	Other transurethral prostatectomy	Xu, 2019[102]		Other	Other
			Radical prostatectomy	Begg, 2002[62]		8.0%	3.0%
				Bianco, 2005[63] *		7.5%	2.5%
				Xu, 2019[102]		Other	Other
		Orthopedic/Neuro-surgery	Cervical spinal fusion	Xu, 2019[102]		Other	Other
			Lumbar spinal fusion, anterior column	Xu, 2019[102]		Other	Other
			Lumbar spinal fusion, posterior column	Xu, 2019[102]		Other	Other
			Total hip arthroplasty	Xu, 2019[102]		Other	Other
			Total knee replacement	Xu, 2019[102]		Other	Other
	Death or serious morbidity	All surgeries	All surgeries	Quinn, 2018[91]	2.0%	0.15%	0.15%
	Estimated blood loss	Urology	Partial nephrectomy	Dagenais, 2019[88] ##	14.4%	10.5%	10.5%
	Hypoparathyroidism	General surgery	Thyroid surgery	Duclos, 2012[79]	32.0%		
	Incontinence	Urology	Radical prostatectomy	Begg, 2002[62]		9.0%	3.0%
				Bianco, 2005[63] *		9.4%	2.5%
	Late urinary complications	Urology	Radical prostatectomy	Begg, 2002[62]		13.0%	14.0%
				Bianco, 2005[63] *		13.2%	14.5%
	Major complications or mortality	Cardiac surgery	Cardiac surgery	Glance, 2016[89]	1.76%	3.3%	1.7%
	Posterior capsule rupture (PCR)	Ophthalmologist	Cataract surgery	Johnston, 2010[97]		Other	Other

	Recurrent laryngeal nerve injury (RLNI)	General surgery	Primary surgery for benign thyroid disease	Hermann, 2002[96]		Other	Other
	Recurrent laryngeal nerve palsy	General surgery	Thyroid surgery	Duclos, 2012[79]	10.0%		
	Surgical site infection	All surgeries	All surgeries	Quinn, 2018[91]	4.5%	0.29%	0.07%
Length of stay		Cardiac surgery	Cardiac surgery	Papachristofi, 2017[45]	2.79%	11.8%	14.2%
		Orthopedic/Neuro-surgery	Spine surgery	Singh, 2018[76]		10.0%	7.2%
Mortality		All surgeries	All surgeries	Quinn, 2018[91]	1.4%	0.0%	0.0%
		Cardiac surgery	Aortic valve surgery	Bridgewater, 2005[66] *		0.0%	0.0%
			CABG	Huesch, 2009[56] ***		1.2%	Other
				Glance, 2006[53]	5.9%	3.3%	8.7%
				Grant, 2008[70] *		0.0%	0.0%
			Cardiac surgery	Papachristofi, 2014[90]	2.79%	16.7%	0.0%
				Papachristofi, 2016[44]	4.0%	15.0%	6.3%
			Coronary artery surgery	Bridgewater, 2003[67]		0.0%	0.0%
				Bridgewater, 2005[66] *		0.0%	16.0%
			Norwood operation	Anderson, 2016[52]		Other	Other
		Colorectal surgery	Colectomy	Udyavar, 2018a[34]	22.9%		
				Xu, 2016[104]		Other	Other
		Gastrointestinal surgery	Peptic ulcer disease	Udyavar, 2018a[34]	47.3%		
			Small bowel resection	Udyavar, 2018a[34]	23.1%		
		General surgery	Appendectomy	Udyavar, 2018a[34]	6.9%		

(*Continued*)

DovePress https://doi.org/10.2147/TCRM.S357934

Table 2 (Continued).

	Outcome	Specialty	Detailed Intervention	Publication	ICC^	Outliers %	
						Negative	**Positive**
			Cholecystectomy	Udyavar, 2018a[34]	3.5%		
			Laparotomy	Udyavar, 2018a[34]	33.2%		
			Lysis of adhesions	Udyavar, 2018a[34]	35.5%		
			Overall (Emergency general surgeries)	Udyavar, 2018a[34]	32.7%		
		Rectal surgery	Rectal carcinoma resection	Hermanek, 1999[72]		9.3%	16.3%
				Justiniano, 2019[98]		Other	Other
		Trauma surgery	Trauma surgery	Udyavar, 2018b[87]	8.7%		
		Urology	Radical prostatectomy	Begg, 2002[62]		0.0%	0.0%
Readmission		8 (Cardiac, GIS, Trauma, HPB, BME, Thoracic, Transplant, Vascular)	8 (Cardiac, GIS, Trauma, HPB, BME, Thoracic, Transplant, Vascular)	Gani, 2015[81]	2.8%		
		All surgeries	All surgeries	Quinn, 2018[91]	0.7%	0.0%	0.0%
		Colorectal surgery	Colectomy	Udyavar, 2018a[34]	3.1%		
		Gastrointestinal surgery	Pancreatoduodenectomy	Hyder, 2013[83]	0.3%		
			Peptic ulcer disease	Udyavar, 2018a[34]	6.8%		
			Small bowel resection	Udyavar, 2018a[34]	2.9%		
		General surgery	Appendectomy	Udyavar, 2018a[34]	3.5%		
			Cholecystectomy	Udyavar, 2018a[34]	3.0%		
			Laparotomy	Udyavar, 2018a[34]	6.0%		
			Lysis of adhesions	Udyavar, 2018a[34]	4.9%		

			Overall (Emergency general surgeries)	Udyavar, 2018a[34]	2.3%		
Reoperation	Reoperation	All surgeries	All surgeries	Quinn, 2018[91]	3.8%	0.04%	0.0%
		Breast surgery	Breast-conserving surgery	Kaczmarski, 2019[73]		17.5%	3.7%
				Landercasper, 2019[74]		5.7%	4.3%
				McCahill, 2012[59]		13.0%	31.5%
				Schumacher, 2017[101]		Other	Other
		Colorectal surgery	Colorectal surgery	Burns, 2011[68] **		0.7%	4.5%
		General surgery	Hernia operation	Arvidsson, 2005[93]		Other	Other
			Inguinal hernia operation	Aquina, 2017[77]	40.5%		
			Ventral hernia operation	Aquina, 2017[77]	14.0%		
		Spinal surgery	Lumbar fusion	Martin, 2013[58]	9.0%	#	#
		Urology	Urinary-genital tract fistula	Cromwell, 2013[69] **		0.0%	0.0%
Suboptimal care		Colorectal surgery	Lymph node examination in colectomy	Becerra, 2017[78]	7.9%		
Success or failure	ASES score	Orthopedic surgery	Rotator cuff repair	Thigpen, 2018[55]		5.9%	8.8%
	Cancer recurrence	Urology	Radical prostatectomy	Bianco, 2010[64]		8.3%	36.1%
	ESS revision rate	ENT surgery	Endoscopic sinus surgery	Rudmik, 2017[75]		16.3%	4.7%
	Mitral valve repair rates	Cardiac surgery	Mitral valve repair/replacement	Bolling, 2010[65]		6.6%	7.4%
				LaPar, 2014[99]		Other	Other
	Positive surgical margins	Urology	Radical prostatectomy	Eastham, 2003[94]		Other	Other

(*Continued*)

https://doi.org/10.2147/TCRM.S357934

Table 2 (Continued).

	Outcome	Specialty	Detailed Intervention	Publication	ICC^	Outliers %	
						Negative	Positive
Score	Postoperative low-output failure	Cardiac surgery	CABG	Likosky, 2012[100]		Other	Other
	Readmission	Orthopedic/Neuro-surgery	Spine surgery	Singh, 2018[76]		0.1%	0.03%
	Recurrence	General surgery	Inguinal hernia repair surgery	Eklund, 2009[37]		2.1%	
	ASES^ score	Orthopedic surgery	Rotator cuff repair	Thigpen, 2018[55]		5.9%	8.8%
	Single Assessment Numeric Evaluation (SANE) score	Orthopedic surgery	Rotator cuff repair	Kissenberth, 2018[84]	44.0%		

Notes: ^ASES score is American Shoulder and Elbow Surgeons (ASES) performance score; ICC is intra-class correlation coefficient and shows percentage of variance due to practitioner as percentage of total variance after accounting for all known variables. Outliers are listed for papers where the surgeons were ordered in their effect on patients' physical health from best to worst or vice versa. The percentages listed are those practitioners whose 95% confidence interval is wholly below or above the mean. Outliers listed as "Other" sorted their surgeons by physical patient effect but used a different way to present their data. Common examples are a caterpillar plot without confidence intervals or a bar chart. *99% confidence interval to define outliers used. **99.8% confidence interval to define outliers used. ***90% confidence interval to define outliers used. #Graph too small to calculate positive or negative outliers. ##Dagenais et al[88] also shows precisely 0.00 between-surgeon variance for length of stay, glomerular filtration rate (GFR) preservation, positive margins, chronic kidney disease (CKD) upstaging, Clavien grade ≥ 1 complications, and 30-day readmission. Operative time had an ICC of 33.4%.

https://doi.org/10.2147/TCRM.S357934
DovePress

Table 3 ICC Summary Statistics

ICC n=53 Outcomes	
Minimum	0.001%
Maximum	47.3%
Average	10.2%
Median	4.0%
IQR	2.2–14.0%
Standard deviation	0.13

Discussion

In this review, the objective was to determine whether there is a surgeons' effect on patients' physical health that is apparent even after accounting for all known variables, such as level of experience. Included studies graded surgeons in order of performance or listed the proportion of variation that is due to practitioners after taking account of all known variables. All but three studies were cohort studies. The other three studies were randomized controlled trials. Findings showed substantial heterogeneity that may be related to type of surgery and type of outcome. After accounting for surgeons' experience, patients' risk, and all other known variables, there remained at times substantial differences in patients' physical health outcomes between surgeons. More than a quarter of all studies (15 out of 55) showed high-volume outliers whose performance is well above the average. In contrast, there were types of surgery/intervention/outcome combinations that showed little evidence of a surgeons' effect on patients' physical health. These findings are somewhat consistent with the substantial body of research on a therapist effect in psychotherapy showing a wide variation in patient outcomes.

With two exceptions the authors only found one study per combination of surgical specialty, intervention and patient outcome. The first exception was two studies covering complications after colectomy and they had very different ICCs of 14.0%[85] and 2.3%.[34] It seems the much wider definition of "complication" in Shih led to a bigger influence of surgeons

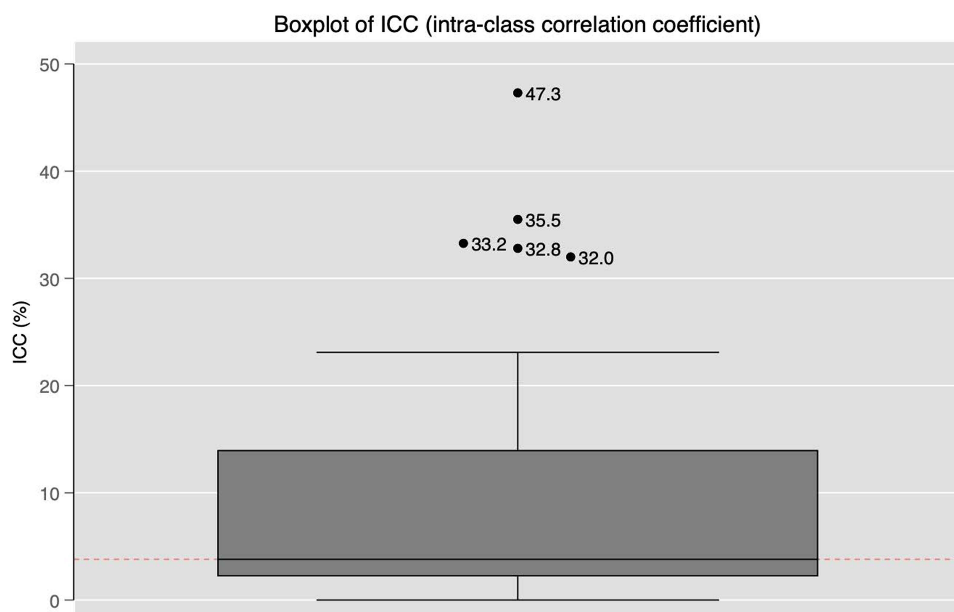

Figure 4 Boxplot of ICC (intra-class correlation coefficient).

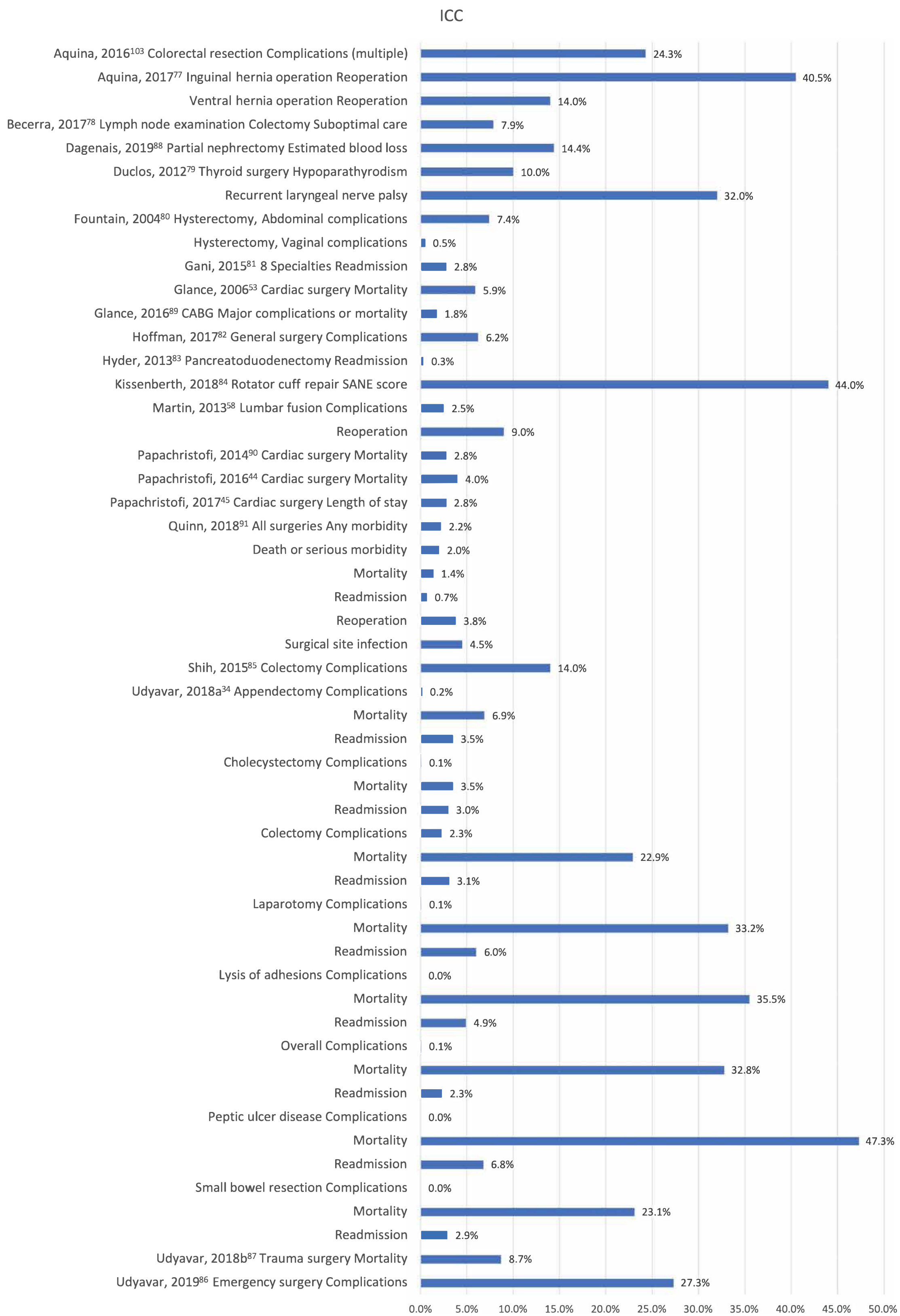

Figure 5 Intra-class correlation coefficients (ICCs) by paper, intervention, and patient outcome.

on the outcome and therefore a higher ICC. The second exception was for mortality after cardiac surgery with three studies[44,53,90] reporting an ICC of 2.8% to 5.9%. Clearly, standardized definitions of physical patient outcomes would assist comparisons across studies.

A major limitation of the evidence identified in this review is that there is currently no standard way to report surgeons' performance on patient's physical health. What does get reported can be divided into either grading individual surgeons by performance or calculating the percentage of variation in patients' outcome that is due to the surgeon after all known variables have been taken into account. Both types of reporting are worded in many different ways, making discovery of such research difficult as can be seen in that more than 10,000 publications had to be reviewed.

A further limitation is that only for very few papers the primary purpose was to report surgeons' performance after taking account of all known variables. Much of the time the reviewed publications' authors emphasized other aspects of healthcare.

Summary

In terms of this systematic review, it was revealed that surgeons' performance data on physical patient health is available to the authors of many published research studies. However, this data is in most cases either not at all reported or only in a limited way. This data could easily be included in an article prepared for publication as the data is already available and often requires minimal or no extra analysis to provide it in the format recommended in the methodological review that is reviewed for publishing. Publishing this data will also allow these studies to be part of future meta-analyses, gaining further dissemination of the work.

It seems that the possibility that surgeons are an intervention in their own right, an intervention that can be more or less effective and an intervention whose effect can be measured, is an area where there has been little systematic research. This is despite the fact that in psychotherapy it is well established that doctors (therapists) constitute an intervention in their own right, independent of the actual intervention they use.[24,25]

Furthermore, if the intervention is held constant, then surgeons are an effect modifier whose strength varies substantially depending on the intervention and the patients' physical health outcome measured.

If it can be established when and how much surgeons constitute an intervention or a substantial effect modifier in their own right, independent of the intervention they use, then this opens up the possibility that this intervention (surgeons) can be systematically managed and improved to the benefit of patients, the surgeons themselves, and the entire health system.

None of the studies that identified outstanding surgeons[61,64–66,68,71,74–76,89,91] made any recommendations on how to use this potential quality improvement resource. So far, we see little or no evidence in the literature that even when exceptional performers have been quantitatively identified, these exceptional performers are used as role models or as research subjects for qualitative research in order to find out what makes them exceptional.

A key point of this systematic review is that the authors specifically looked for studies that showed a surgeons' effect for which there was no explanation, ie a residual effect after all known information had been included in the statistical analysis. Therefore, the cause of the surgeons' effect measured is, by the definition of the research question for this systematic review, not known. This leaves open the question whether the cause is unknowable, or if there are one or more causes that could be identified in future research.

If we want to know what makes a good surgeon beyond the well-founded opinions of surgeons[12] or those who work with surgeons – and how to train surgeons to be good surgeons – then the first step beyond all the current measures taken to train surgeons could be to reliably identify outstanding surgeons. Consequently, we can find out if their ability can be passed on to others and, if yes, to lift the overall standard of healthcare by transferring their exceptional ability to other surgeons. This is especially so as identifying data is already available in the many datasets consisting of medical records, some of which were accessed in the cohort studies covered here.

Exceptional performances may be due to personality characteristics that may be hard or impossible to emulate, or we may find out that the surgeons employ easy to emulate techniques like connecting with patients, or simply have higher expectations of patient outcomes,[26] or we may find that they live stress-resistant lives, or that they are rarely exhausted, or any other of a myriad of possibilities. If research that investigates exceptional performers identifies simple techniques or choices made at work, or out of work, that could be emulated relatively easily by many other surgeons, then this could lead to fewer complications and more successful surgeries, and there could be large beneficial effects on healthcare costs and patient health.

However, the misuse of identifying supposed underperformers, for example by disciplining or evicting practitioners whose performance appears substandard but who are not statistical outliers or whose performance appears substandard due to a small number of high-risk patients, or due to other confounders like incomplete case-mix or risk score data, is a danger that can cause substantial harm to the surgeons. Further, an old saying is that what gets measured gets managed.[105] If more data is available for each surgeon, then this data can be misused to disempower practitioners by adding more and more rules and regulations, and by giving practitioners less opportunity to use their experience and ability. Such data can also be misused in being available online, especially with insufficient explanations of proper usage; or being very much out of date, as is the case for two publicly available databases of surgeons whose data in 2021 only went until 2013[106] and 2014.[6] Moreover, giving surgeons key performance indicators of patient outcomes could be an unwarranted intrusion into the doctor/patient relationship and lead to surgeons avoiding high-risk patients, as even a few such patients can skew an individual surgeon's patient outcome statistics, confirmed anecdotally here.[66] However, this fact is denied if patients' risk was accounted for.[107] Hence crude performance data should not be published.[67]

Strengths and Limitations of This Study

The strength of this work lies in the broad search of the literature, the condensed and clear reporting of effect size, and the importance behind the finding that the surgeons' effect at times has a significant effect size, as big as many non-surgical interventions themselves. The search term strategy used to identify studies was a complex and complete combination of terms that should have identified most of the relevant published studies. Furthermore, the references list of relevant articles and studies citing these articles were screened. This review was not limited by language or by timeframe.

On the other hand, there are at least three broad limitations. First, the Newcastle-Ottawa Scale (NOS) was used for quality assessment with the majority of studies scoring between 8–9 (9 being the maximum total); however, the NOS has been critiqued for being "difficult to use and [having] vague decision rules"[1] which derived from poor or fair inter-rater reliability between reviewers. However, it is important to note that associations between individual quality domains or overall quality score and effect estimates were not found. Moreover, the NOS has been endorsed by The Cochrane Collaboration[2] for its implementation in systematic reviews of non-randomized studies.

Second, as all of the review's studies were conducted in North America and Europe, it is unclear whether the findings can be generalized to other regions, particularly in developing nations.

Finally, while the outcome data was heterogeneous and did not enable a meta-analysis, there was also heterogeneity regarding surgical specialty, type of intervention, and type of outcome. Thereby, it is difficult to draw conclusions and synthesize studies with inconsistent outcome measures, and these characteristics have often been found attributable to a lack of a high level of evidence on the specific research subject.

Conclusions and Implications

Even after accounting for surgeons' experience, patients' risk and all other known variables there remain sometimes substantial differences in patients' physical health outcomes between surgeons. Therefore it can matter which surgeon is chosen. At times it is possible to identify high-volume outliers whose performance is well above the average, and it could be worthwhile to study these surgeons to see whether their excellence can be passed on to their peers. It is evident that there are currently no well-established standards on how to assess surgeons as an intervention in their own right, thus systematic approaches to establishing standardized measures are needed, and researching the surgeons' effect on patients' physical health is still in its early stages.

Support

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Lead Author Statement

The lead author affirms that the manuscript is an honest, accurate and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as originally planned (and if relevant, registered) have been explained.

Data Sharing Statement

No additional data available.

Ethical Approval

As this is a systematic review of published studies, no ethical approval was required.

Acknowledgments

The authors thank Dr Aya Ashraf Ali and Tulia Gonzalez Flores for their excellent editorial contributions.

The authors thank Dr Jeremy Howick for his support and advice during the conception of this research and the many helpful additions afterwards. Without him this paper would not have happened.

Funding

This review has been funded by the first author as part of his PhD studies. No external funding was received.

Disclosure

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf, and declare: no support from any organization for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

References

1. Hartling L, Milne A, Hamm MP, et al. Testing the Newcastle Ottawa Scale showed low reliability between individual reviewers. *J Clin Epidemiol*. 2013;66(9):982–993. doi:10.1016/j.jclinepi.2013.03.003
2. Higgins JPT, Chandler J, Cumpston M, Li T, Page MJ, Welch VA. Cochrane Handbook for Systematic Reviews of Interventions version 6.1 (updated September 2020). Cochrane; 2020 [updated September, 2020]. Available from: https://training.cochrane.org/cochrane-handbook-systematic-reviews-interventions https://training.cochrane.org/cochrane-handbook-systematic-reviews-interventionshttps://training.cochrane.org/handbook. Accessed April 20, 2022.
3. Gordon M Do you need that surgery? How to decide, and how to pick a surgeon if you do. NPR; 2019. Available from: https://www.npr.org/sections/health-shots/2019/07/19/743248074/do-you-need-that-surgery-how-to-decide-and-how-to-pick-a-surgeon-if-you-do. Accessed January 23, 2021.
4. Whitlock J Tips for choosing a surgeon. Verywellhealth; 2020. Available from: https://www.verywellhealth.com/how-to-find-a-great-surgeon-3156937. Accessed January 23, 2021.
5. Marsa L How to pick the right surgeon. AARP; 2017. Available from: https://www.aarp.org/health/conditions-treatments/info-2017/choose-a-surgeon-doctor-surgeries.html. Accessed January 23, 2021.
6. Search for the Best (Surgeon, Doctor, Hospital, Dentist). Consumers' Checkbook; 2021. Available from: https://www.checkbook.org/surgeonratings/. Accessed April 14, 2022.
7. Simpson L. What makes a good surgeon? *J Natl Med Assoc*. 2008;100(2):261–264. doi:10.1016/s0027-9684(15)31216-5
8. Crile GW. The most important factor in the treatment of war wounds and the most important factor in civilian surgery-The Good Surgeon. *Ann Surg*. 1919;70(4):385–387. doi:10.1097/00000658-191910000-00001
9. Darzi A, Smith S, Taffinder N. Assessing operative skill. *Needs to become more objective*. *BMJ*. 1999;318(7188):887–888.
10. Gandhi J. Making of a surgeon. *Al Ameen J Med Sci*. 2019;12(2):54–55.
11. Arora S, Sevdalis N, Suliman I, Athanasiou T, Kneebone R, Darzi A. What makes a competent surgeon?: Experts' and trainees' perceptions of the roles of a surgeon. *Am J Surg*. 2009;198(5):726–732. doi:10.1016/j.amjsurg.2009.01.015
12. Jackson B. What makes an excellent surgeon? *Obes Surg*. 2019;29(4):1087–1089. doi:10.1007/s11695-019-03778-8
13. Wilt TJ, Shamliyan TA, Taylor BC, MacDonald R, Kane RL. Association between hospital and surgeon radical prostatectomy volume and patient outcomes: a systematic review. *J Urol*. 2008;180(3):820–829. doi:10.1016/j.juro.2008.05.010
14. Maruthappu M, Gilbert BJ, El-Harasis MA, et al. The influence of volume and experience on individual surgical performance: a systematic review. *Ann Surg*. 2015;261(4):642–647. doi:10.1097/SLA.0000000000000852
15. Churchill LR, Schenck D. Healing skills for medical practice. *Ann Intern Med*. 2008;149(10):720–724. doi:10.7326/0003-4819-149-10-200811180-00006
16. Schenck D, Churchill L. *Healers: Extraordinary Clinicians at Work*. Oxford University Press; 2011.

17. Hanyok LA, Hellmann DB, Rand C, Ziegelstein RC. Practicing patient-centered care: the questions clinically excellent physicians use to get to know their patients as individuals. *Patient*. 2012;5(3):141–145. doi:10.1007/BF03262487
18. De Vries EN, Ramrattan MA, Smorenburg SM, Gouma DJ, Boermeester MA. The incidence and nature of in-hospital adverse events: a systematic review. *Qual Saf Health Care*. 2008;17(3):216–223. doi:10.1136/qshc.2007.023622
19. Tam VC, Knowles SR, Cornish PL, Fine N, Marchesano R, Etchells EE. Frequency, type and clinical importance of medication history errors at admission to hospital: a systematic review. *CMAJ*. 2005;173(5):510–515. doi:10.1503/cmaj.045311
20. Van Walraven C, Bennett C, Jennings A, Austin PC, Forster AJ. Proportion of hospital readmissions deemed avoidable: a systematic review. *CMAJ*. 2011;183(7):E391–E402. doi:10.1503/cmaj.101860
21. Leppin AL, Gionfriddo MR, Kessler M, et al. Preventing 30-day hospital readmissions: a systematic review and meta-analysis of randomized trials. *JAMA Intern Med*. 2014;174(7):1095–1107. doi:10.1001/jamainternmed.2014.1608
22. Tjarda Van Heek N, Kuhlmann KFD, Scholten RJ, et al. Hospital volume and mortality after pancreatic resection: a systematic review and an evaluation of intervention in The Netherlands. *Ann Surg*. 2005;242(6):781–790. doi:10.1097/01.sla.0000188462.00249.36
23. Nilsen SM, Bjørngaard JH, Carlsen F, et al. Hospitals´ discharge tendency and risk of death-an analysis of 60,000 Norwegian hip fracture patients. *Clin Epidemiol*. 2020;12:173–182. doi:10.2147/CLEP.S237060
24. Wampold BE, Imel ZE. *The Great Psychotherapy Debate: The Evidence for What Makes Psychotherapy Work: Second Edition*. Taylor and Francis Inc.; 2015.
25. Baldwin SA, Imel Z. Therapist effects: findings and methods. In: *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change*. Vol. 6; 2013:258–297.
26. Chen P-HA, Cheong JH, Jolly E, Elhence H, Wager TD, Chang LJ. Socially transmitted placebo effects. *Nat Hum Behav*. 2019;3 (12):1295–1305. doi:10.1038/s41562-019-0749-5
27. Howick J, Friedemann C, Tsakok M, et al. Are treatments more effective than placebos? A systematic review and meta-analysis. *PLoS One*. 2013;8(5):e62599. doi:10.1371/journal.pone.0062599
28. Wartolowska K, Judge A, Hopewell S, et al. Use of placebo controls in the evaluation of surgery: systematic review. *BMJ*. 2014;348:g3253. doi:10.1136/bmj.g3253
29. Probst P, Grummich K, Harnoss JC, et al. Placebo-controlled trials in surgery: a systematic review and meta-analysis. *Medicine (Baltimore)*. 2016;95(17):e3516. doi:10.1097/MD.0000000000003516
30. Wartolowska K, Collins GS, Hopewell S, et al. Feasibility of surgical randomised controlled trials with a placebo arm: a systematic review. *BMJ Open*. 2016;6(3):e010194. doi:10.1136/bmjopen-2015-010194
31. Jonas WB, Crawford C, Colloca L, et al. To what extent are surgery and invasive procedures effective beyond a placebo response? A systematic review with meta-analysis of randomised, sham controlled trials. *BMJ Open*. 2015;5(12):e009655. doi:10.1136/bmjopen-2015-009655
32. Sochacki KR, Mather RC, Nwachukwu BU, et al. Sham surgery studies in orthopaedic surgery may just be a Sham: a systematic review of randomized placebo-controlled trials. *Arthroscopy*. 2020;36(10):2750–2762.e2752. doi:10.1016/j.arthro.2020.05.001
33. Cook JA, Bruckner T, MacLennan GS, Seiler CM. Clustering in surgical trials - database of intracluster correlations. *Trials*. 2012;13. doi:10.1186/1745-6215-13-2
34. Udyavar R, Cornwell EE, Havens JM, et al. Surgeon-driven variability in emergency general surgery outcomes: does it matter who is on call? *Surgery*. 2018;164(5):1109–1116. doi:10.1016/j.surg.2018.07.008
35. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ*. 2020;368:l6890. doi:10.1136/bmj.l6890
36. Gill L, White L. A critical review of patient satisfaction. *Leadersh Health Serv*. 2009;22(1):8–19. doi:10.1108/17511870910927994
37. Eklund AS, Montgomery AK, Rasmussen IC, Sandbue RP, Bergkvist LÅ, Rudberg CR. Low recurrence rate after laparoscopic (TEP) and open (Lichtenstein) Inguinal Hernia repair: a randomized, multicenter trial with 5-year follow-up. *Ann Surg*. 2009;249(1):33–38. doi:10.1097/SLA.0b013e31819255d0
38. Murphy WS, Cheng T, Lin B, Terry D, Murphy SB. Higher volume surgeons have lower medicare payments, readmissions, and mortality after THA [Hip]. *Clin Orthop Relat Res*. 2019;477(2):334–341. doi:10.1097/CORR.0000000000000370
39. Sivaganesan A, Asher AL, Bydon M, et al. A strategy for risk-adjusted ranking of surgeons and practices based on patient-reported outcomes after elective lumbar surgery. *Spine*. 2019;44(9):670–677. doi:10.1097/BRS.0000000000002894
40. Wells GA, Shea B, O'Connell D, et al. *The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-Analyses*. University of Eastern Piedmont . http://www3.med.unipmn.it/dispense_ebm/2009-2010/Corso%20Perfezionamento%20EBM_Faggiano/NOS_oxford.pdfAccessed April 20, 2000.
41. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol*. 2010;25(9):603–605. doi:10.1007/s10654-010-9491-z
42. Rasbash J, Goldstein H. Efficient analysis of mixed hierarchical and cross-classified random structures using a multilevel model. *J Educ Behav Stat*. 1994;19(4):337–350. doi:10.3102/10769986019004337
43. Goldstein H, McDonald RP. A general model for the analysis of multilevel data. *Psychometrika*. 1988;53(4):455–467. doi:10.1007/BF02294400
44. Papachristofi O, Sharples LD, Mackay JH, Nashef SAM, Fletcher SN, Klein AA. The contribution of the anaesthetist to risk-adjusted mortality after cardiac surgery. *Anaesthesia*. 2016;71(2):138–146. doi:10.1111/anae.13291
45. Papachristofi O, Klein AA, Mackay J, Nashef S, Fletcher N, Sharples LD. Effect of individual patient risk, centre, surgeon and anaesthetist on length of stay in hospital after cardiac surgery: Association of Cardiothoracic Anaesthesia and Critical Care (ACTACC) consecutive cases series study of 10 UK specialist centres. *BMJ Open*. 2017;7(9):e016947. doi:10.1136/bmjopen-2017-016947
46. Gulliford MC, Adams G, Ukoumunne OC, Latinovic R, Chinn S, Campbell MJ. Intraclass correlation coefficient and outcome prevalence are associated in clustered binary data. *J Clin Epidemiol*. 2005;58(3):246–251. doi:10.1016/j.jclinepi.2004.08.012
47. Chakraborty H, Hossain A. R package to estimate intracluster correlation coefficient with confidence interval for binary data. *Comput Methods Programs Biomed*. 2018;155:85–92. doi:10.1016/j.cmpb.2017.10.023
48. Adams G, Gulliford MC, Ukoumunne OC, Eldridge S, Chinn S, Campbell MJ. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *J Clin Epidemiol*. 2004;57(8):785–794. doi:10.1016/j.jclinepi.2003.12.013

49. Turner RM, Omar RZ, Thompson SG. Constructing intervals for the intracluster correlation coefficient using Bayesian modelling, and application in cluster randomized trials. *Stat Med*. 2006;25(9):1443–1456. doi:10.1002/sim.2304
50. Donner A. A review of inference procedures for the intraclass correlation coefficient in the one-way random effects model. *Int Stat Rev*. 1986;54 (1):67–82.
51. Eldridge SM, Ukoumunne OC, Carlin JB. The intra-cluster correlation coefficient in cluster randomized trials: a review of definitions. *Int Stat Rev*. 2009;77(3):378–394. doi:10.1111/j.1751-5823.2009.00092.x
52. Anderson BR, Ciarleglio AJ, Cohen DJ, et al. The Norwood operation: relative effects of surgeon and institutional volumes on outcomes and resource utilization. *Cardiol Young*. 2016;26(4):683–692. doi:10.1017/S1047951115001031
53. Glance LG, Dick A, Osler TM, Li Y, Mukamel DB. Impact of changing the statistical methodology on hospital and surgeon ranking: the case of the New York State cardiac surgery report card. *Med Care*. 2006;44(4):311–319. doi:10.1097/01.mlr.0000204106.64619.2a
54. Singh S, Goodwin JS, Zhou J, Kuo YF, Nattinger AB. Variation among primary care physicians in 30-day readmissions. *Ann Intern Med*. 2019;170(11):749–755. doi:10.7326/M18-2526
55. Thigpen CA, Floyd SB, Chapman C, et al. Comparison of surgeon performance of rotator cuff repair: risk adjustment toward a more accurate performance measure. *J Bone Joint Surg Am*. 2018;100(24):2110–2117. doi:10.2106/JBJS.18.00211
56. Huesch MD. Can managed care plans reliably infer the quality of cardiac surgeons' outcomes? *Am J Manag Care*. 2009;15(12):890–896.
57. Luan WP, Leroux TC, Olsen C, Robb D, Skinner JS, Richard P. Variation in bariatric surgery costs and complication rates in the military health system. *Mil Med*. 2019;185(7–8):e1057–e1064.
58. Martin BI, Mirza SK, Franklin GM, Lurie JD, MacKenzie TA, Deyo RA. Hospital and surgeon variation in complications and repeat surgery following incident lumbar fusion for common degenerative diagnoses. *Health Serv Res*. 2013;48(1):1–25. doi:10.1111/j.1475-6773.2012.01434.x
59. McCahill LE, Single RM, Aiello Bowles EJ, et al. Variability in reexcision following breast conservation surgery. *JAMA*. 2012;307(5):467–475. doi:10.1001/jama.2012.43
60. Fernández-Castilla B, Declercq L, Jamshidi L, Beretvas N, Onghena P, Van den Noortgate W. Visual representations of meta-analyses of multiple outcomes: extensions to forest plots, funnel plots, and caterpillar plots. *Methodology*. 2020;16(4):299–315. doi:10.5964/meth.4013
61. Aquina CT, Blumberg N, Probst CP, et al. Significant variation in blood transfusion practice persists following upper GI cancer resection. *J Gastrointest Surg*. 2015;19(11):1927–1937. doi:10.1007/s11605-015-2903-3
62. Begg CB, Riedel ER, Bach PB, et al. Variations in morbidity after radical prostatectomy. *New Engl J Med*. 2002;346(15):1138–1144. doi:10.1056/NEJMsa011788
63. Bianco FJ Jr, Riedel ER, Begg CB, Kattan MW, Scardino PT. Variations among high volume surgeons in the rate of complications after radical prostatectomy: further evidence that technique matters. *J Urol*. 2005;173(6):2099–2103. doi:10.1097/01.ju.0000158163.21079.66
64. Bianco JFJ, Vickers AJ, Cronin AM, et al. Variations among experienced surgeons in cancer control after open radical prostatectomy. *J Urol*. 2010;183(3):977–983. doi:10.1016/j.juro.2009.11.015
65. Bolling SF, Li S, O'Brien SM, Brennan JM, Prager RL, Gammie JS. Predictors of mitral valve repair: clinical and surgeon factors. *Ann Thorac Surg*. 2010;90(6):1904–1911. doi:10.1016/j.athoracsur.2010.07.062
66. Bridgewater B. Mortality data in adult cardiac surgery for named surgeons: retrospective examination of prospectively collected data on coronary artery surgery and aortic valve replacement. *Br Med J*. 2005;330(7490):506–510. doi:10.1136/bmj.330.7490.506
67. Bridgewater B, Grayson AD, Jackson M, et al. Surgeon specific mortality in adult cardiac surgery: comparison between crude and risk stratified data. *BMJ*. 2003;327(7405):13–17. doi:10.1136/bmj.327.7405.13
68. Burns EM, Bottle A, Aylin P, Darzi A, John Nicholls R, Faiz O. Variation in reoperation after colorectal surgery in England as an indicator of surgical performance: retrospective analysis of hospital episode statistics. *BMJ (Online)*. 2011;343(7820):d4836.
69. Cromwell D, Hilton P. Retrospective cohort study on patterns of care and outcomes of surgical treatment for lower urinary-genital tract fistula among English National Health Service hospitals between 2000 and 2009. *BJU Int*. 2013;111(4 B):E257–E262. doi:10.1111/j.1464-410X.2012.11483.x
70. Grant SW, Grayson AD, Jackson M, et al. Does the choice of risk-adjustment model influence the outcome of surgeon-specific mortality analysis? A retrospective analysis of 14 637 patients under 31 surgeons. *Heart*. 2008;94(8):1044–1049. doi:10.1136/hrt.2006.110478
71. Healy MA, Regenbogen SE, Kanters AE, et al. Surgeon variation in complications with minimally invasive and open colectomy: results from the Michigan Surgical Quality Collaborative. *JAMA Surg*. 2017;152(9):860–867. doi:10.1001/jamasurg.2017.1527
72. Hermanek P. Impact of surgeon's technique on outcome after treatment of rectal carcinoma. *Dis Colon Rectum*. 1999;42(5):559–562. doi:10.1007/BF02234128
73. Kaczmarski K, Wang P, Gilmore R, et al. Surgeon re-excision rates after breast-conserving surgery: a measure of low-value care. *J Am Coll Surg*. 2019;228(4):504–512.e502. doi:10.1016/j.jamcollsurg.2018.12.043
74. Landercasper J, Borgert AJ, Fayanju OM, et al. Factors associated with reoperation in breast-conserving surgery for cancer: a prospective study of American Society of Breast Surgeon Members. *Ann Surg Oncol*. 2019;26(10):3321–3336. doi:10.1245/s10434-019-07547-w
75. Rudmik L, Xu Y, Alt JA, et al. Evaluating surgeon-specific performance for endoscopic sinus surgery. *JAMA Otolaryngol Head Neck Surg*. 2017;143(9):891–898. doi:10.1001/jamaoto.2017.0752
76. Singh S, Sparapani R, Wang MC. Variations in 30-day readmissions and length of stay among spine surgeons: a national study of elective spine surgery among US Medicare beneficiaries. *J Neurosurg Spine*. 2018;29(3):286–291. doi:10.3171/2018.1.SPINE171064
77. Aquina CT, Fleming FJ, Becerra AZ, et al. Explaining variation in ventral and inguinal hernia repair outcomes: a population-based analysis. *Surgery*. 2017;162(3):628–639. doi:10.1016/j.surg.2017.03.013
78. Becerra AZ, Aquina CT, Berho M, et al. Surgeon-, pathologist-, and hospital-level variation in suboptimal lymph node examination after colectomy: compartmentalizing quality improvement strategies. *Surgery (United States)*. 2017;161(5):1299–1306.
79. Duclos A, Peix JL, Colin C, et al. Influence of experience on performance of individual surgeons in thyroid surgery: prospective cross sectional multicentre study. *BMJ (Online)*. 2012;344(7843):d8041.
80. Fountain J, Gallagher J, Brown J. A practical approach to a multi-level analysis with a sparse binary outcome within a large surgical trial. *J Eval Clin Pract*. 2004;10(2):323–327. doi:10.1111/j.1365-2753.2003.00462.x
81. Gani F, Lucas DJ, Kim Y, Schneider EB, Pawlik TM. Understanding variation in 30-day surgical readmission in the era of accountable care: effect of the patient, surgeon, and surgical subspecialties. *JAMA Surg*. 2015;150(11):1042–1049. doi:10.1001/jamasurg.2015.2215

82. Hoffman RL, Kelz RR, Wirtalla CJ, et al. Variations in surgical outcomes: is it the residency program, the surgeon or the practice venue? *J Am Coll Surg*. 2017;225(4):S185. doi:10.1016/j.jamcollsurg.2017.07.421
83. Hyder O, Dodson RM, Nathan H, et al. Influence of patient, physician, and hospital factors on 30-day readmission following pancreatoduodenectomy in the United States. *JAMA Surg*. 2013;148(12):1095–1102. doi:10.1001/jamasurg.2013.2509
84. Kissenberth M, Thigpen C, Brooks J, Floyd S, Hawkins RJ, Tokish JM. Comparing surgeon performance of rotator cuff repair: risk adjustment toward a fair performance measure. *Arthroscopy*. 2018;34(12):e3. doi:10.1016/j.arthro.2018.10.022
85. Shih T, Cole AI, Al-Attar PM, et al. Reliability of surgeon-specific reporting of complications after colectomy. *Ann Surg*. 2015;261(5):920–925. doi:10.1097/SLA.0000000000001032
86. Udyavar NR, Salim A, Cornwell EE, et al. Racial differences in complication risk following emergency general surgery: who your surgeon is may matter. *J Surg Res*. 2019;235:424–431. doi:10.1016/j.jss.2018.05.086
87. Udyavar NR, Salim A, Havens JM, et al. The impact of individual physicians on outcomes after trauma: is it the system or the surgeon? *J Surg Res*. 2018;229:51–57. doi:10.1016/j.jss.2018.02.051
88. Dagenais J, Bertolo R, Garisto J, et al. Variability in partial nephrectomy outcomes: does your surgeon matter? *Eur Urol*. 2019;75(4):628–634. doi:10.1016/j.eururo.2018.10.046
89. Glance LG, Hannan EL, Fleisher LA, et al. Feasibility of report cards for measuring anesthesiologist quality for cardiac surgery. *Anesth Analg*. 2016;122(5):1603–1613. doi:10.1213/ANE.0000000000001252
90. Papachristofi O, Mackay JH, Powell SJ, Nashef SAM, Sharples L. Impact of the anesthesiologist and surgeon on cardiac surgical outcomes. *J Cardiothorac Vasc Anesth*. 2014;28(1):103–109. doi:10.1053/j.jvca.2013.07.004
91. Quinn CM, Bilimoria KY, Chung JW, Ko CY, Cohen ME, Stulberg JJ. Creating individual surgeon performance assessments in a statewide hospital surgical quality improvement collaborative. *J Am Coll Surg*. 2018;227(3):303–312.e303. doi:10.1016/j.jamcollsurg.2018.06.002
92. Aquina C, Probst C, Hensley B, et al. High variability in nosocomial clostridium difficile infection rates among both surgeons and hospitals following colorectal resection. *Dis Colon Rectum*. 2015;58(5):e163.
93. Arvidsson D, Berndsen FH, Larsson LG, et al. Randomized clinical trial comparing 5-year recurrence rate after laparoscopic versus Shouldice repair of primary inguinal hernia. *Br J Surg*. 2005;92(9):1085–1091. doi:10.1002/bjs.5137
94. Eastham JA, Kattan MW, Riedel E, et al. Variations among individual surgeons in the rate of positive surgical margins in radical prostatectomy specimens. *J Urol*. 2003;170(6 I):2292–2295. doi:10.1097/01.ju.0000091100.83725.51
95. Faschinger C. Quality assessment of cataract surgery of the Department of Ophthalmology, Medical University of Graz. *Spektrum Augenheilkd*. 2011;25(3):215–219. doi:10.1007/s00717-011-0013-5
96. Hermann M, Alk G, Roka R, Glaser K, Freissmuth M. Laryngeal recurrent nerve injury in surgery for benign thyroid diseases: effect of nerve dissection and impact of individual surgeon in more than 27,000 nerves at risk. *Ann Surg*. 2002;235(2):261–268. doi:10.1097/00000658-200202000-00015
97. Johnston RL, Taylor H, Smith R, Sparrow JM. The Cataract National Dataset electronic multi-centre audit of 55,567 operations: variation in posterior capsule rupture rates between surgeons. *Eye (Lond)*. 2010;24(5):888–893. doi:10.1038/eye.2009.195
98. Justiniano CF, Aquina CT, Fleming FJ, et al. Hospital and surgeon variation in positive circumferential resection margin among rectal cancer patients. *Am J Surg*. 2019;218(5):881–886. doi:10.1016/j.amjsurg.2019.02.029
99. LaPar DJ, Ailawadi G, Isbell JM, et al. Mitral valve repair rates correlate with surgeon and institutional experience. *J Thorac Cardiovasc Surg*. 2014;148(3):995–1003; discussion 1003–1004. doi:10.1016/j.jtcvs.2014.06.039
100. Likosky DS, Goldberg JB, DiScipio AW, et al. Variability in surgeons' perioperative practices may influence the incidence of low-output failure after coronary artery bypass grafting surgery. *Circ Cardiovasc Qual Outcomes*. 2012;5(5):638–644. doi:10.1161/CIRCOUTCOMES.112.967091
101. Schumacher J, Neuman H, Landercasper J, Wilke L, Steiman J, Greenberg C. Assessing variation in provider and institution-level re-excision rates: opportunity for a statewide surgical collaborative to improve breast cancer care. *Ann Surg Oncol*. 2017;24(2):224–225.
102. Xu T, Mehta A, Park A, Makary MA, Price DW. Association between board certification, maintenance of certification, and surgical complications in the United States. *Am J Med Qual*. 2019;34(6):545–552. doi:10.1177/1062860618822752
103. Aquina CT, Blumberg N, Probst CP, et al. Large variation in blood transfusion use after colorectal resection: a call to action. *Dis Colon Rectum*. 2016;59(5):411–418. doi:10.1097/DCR.0000000000000588
104. Xu T, Makary MA, Al Kazzi E, Zhou M, Pawlik TM, Hutfless SM. Surgeon-level variation in postoperative complications. *J Gastrointest Surg*. 2016;20(7):1393–1399. doi:10.1007/s11605-016-3139-6
105. Catasús B, Ersson S, Gröjer JE, Yang Wallentin F. What gets measured gets … on indicating, mobilizing and acting. *Account Audit Account J*. 2007;20(4):505–521. doi:10.1108/09513570710762566
106. Surgeon Scorecard. ProPublica; 2021. Available from: https://projects.propublica.org/surgeons. Accessed April 14, 2022.
107. Hannan EL, Siu AL, Kumar D, Racz M, Pryor DB, Chassin MR. Assessment of coronary artery bypass graft surgery performance in New York: is there a bias against taking HIGH-RISK PAtients? *Med Care*. 1997;35(1):49–56. doi:10.1097/00005650-199701000-00004