World J Surg (2021) 45:2066–2080
https://doi.org/10.1007/s00268-021-06055-x

SCIENTIFIC REVIEW

Are Surgeons Working Smarter or Harder? A Systematic Review Comparing the Physical and Mental Demands of Robotic and Laparoscopic or Open Surgery

Laura Seohyun Park[1] · **Feiyang Pan**[1] · **Daniel Steffens**[2] · **Jane Young**[3] · **Jonathan Hong**[4,5]

Accepted: 26 February 2021 / Published online: 26 March 2021

Abstract

Background Minimally invasive surgical techniques such as robotic surgical platforms have provided favourable outcomes for patients, but the impact on surgeons is not well described. This systematic review aims to synthesize and evaluate the physical and mental impact of robotic surgery on surgeons compared to standard laparoscopic or open surgery.

Methods A search strategy was developed to identify peer-reviewed English articles published from inception to end of December 2019 on the following databases: MEDLINE, PubMed, PsycINFO and Embase. The articles were assessed using a modified Newcastle–Ottawa tool.

Results Of the 6563 papers identified, 30 studies were included in the qualitative synthesis of this review. Most of the included studies presented a high risk of bias. A total of 13 and 21 different physical and mental tools, respectively, were used to examine the impact on surgeons. The most common tool used to measure physical and mental demand were surface electromyography ($N = 9$) and the NASA Task Load Index (NASA-TLX; $N = 8$), respectively. Majority of studies showed mixed results for physical ($N = 10$) and mental impact ($N = 7$). This was followed by eight and six studies favouring RS over other surgical modalities for physical and mental impact, respectively.

Conclusion Most studies showed mixed physical and mental outcomes between the three surgical modalities. There was a high risk of bias and methodological heterogeneity. Future studies need to correlate mental and physical stress with long-term impact on the surgeons.

Introduction

New surgical technology has enabled a shift towards minimally invasive surgery such as laparoscopy and more recently, robotics. These techniques have benefits relating

---

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00268-021-06055-x.

✉ Jonathan Hong
Jonathan.hong@sydney.edu.au

[1] Central Clinical School, The University of Sydney, Sydney, NSW 2050, Australia

[2] Australia & Faculty of Medicine and Health, Surgical Outcomes Research Centre (SOuRCe), Central Clinical School, Royal Prince Alfred Hospital (RPAH), The University of Sydney, Sydney, NSW 2050, Australia

[3] The University of Sydney, Sydney, NSW 2050, Australia

[4] Institute of Academic Surgery at Royal Prince Alfred Hospital (IAS RPA), Missenden Road, Camperdown, NSW 2050, Australia

[5] Australia & Faculty of Medicine and Health, Central Clinical School, The University of Sydney, Sydney, NSW 2050, Australia

to patient outcomes [1, 2] with robotic surgery (RS) demonstrating lower rates of intraoperative and post-operative complications when compared to laparoscopic surgery (LS) or open surgical techniques (OS) [3]. RS techniques have also built a reputation for improved precision and physical comfort for surgeons [4], which has led to its increased use across surgical disciplines. However, some studies have not shown a clear advantage of RS over LS for perioperative [3] or post-operative outcomes [5]. Furthermore, there are a high hospital costs [3] associated with adapting robotic technology.

Surgeon comfort is frequently cited to justify the use of RS over LS and OS. The majority of literature focuses on patient outcomes, but there is an increasing number of studies examining both the cognitive and ergonomic challenges of surgeons when using different surgical modalities. Studies have shown that LS is limited by decreased range of movement, reduced dexterity and two-dimensional views [6], whereas for RS, the three-dimensional optics and comfort of being seated [7] have shown to be associated with reduced muscular workload in the shoulder and neck regions as well as reduced perceived exertion [8]. A recent meta-analysis [9] comparing muscle activation between LS and RS suggests that RS is ergonomically superior with lower muscle activation.

Conceptually, society benefits from surgical methods that provide better patient outcomes while reducing the physical and mental workload for surgeons. While previous reviews and meta-analysis [9–11] have examined the physical ergonomics of RS, this systematic review aims to provide a more comprehensive understanding of the comparative literature on the physical and mental impact of RS compared to LS or OS on surgeons.

Methods

Search strategy and data source

This systematic review was conducted in accordance with the PRISMA-P guidelines [12]. A literature search was conducted using Medline, PubMed, Cochrane database, Embase and PsycINFO. The Medical Subject Headings (MESH) terms and text words from the MEDLINE search strategy were adapted to the other databases and indexing to capture the concept of physical or mental demands on either RS, LS and/or OS on surgeons to identify peer-reviewed articles (Supplementary Tables 1 and 2). In addition, a manual search of the cited references in each article was completed.

Study eligibility criteria

The inclusion criteria for this systematic review were: (1) original studies of comparative study design between RS and LS or OS reporting physical or mental outcomes, (2) published in English language between inception and December 2019, and (3) utilized the da Vinci robotic surgical system. Studies which were not comparative were excluded.

Selection process and data extraction

Two authors (LSP, FYP) independently screened the titles and abstracts for all search results and classified relevant articles based on the eligibility criteria. Studies selected by individual reviewers were then compared and any discrepancies were settled by a third reviewer (JH). Full-text review and data extraction of the final 30 studies were divided between the two reviewers (LSP, FYP), and a summary of the data was recorded in a collective database. Any queries or issues were discussed between the two reviewers (LSP and FYP) with any advice from JH as necessary.

Methodological quality and reporting of results

The quality of all included articles was assessed independently by the two authors using a modified version of the Newcastle–Ottawa risk of bias tool developed by Herzog et al. [13] (Supplementary Fig. 1). The modified Newcastle–Ottawa risk of bias tool is used to evaluate each study in three main categories: (1) selection (maximum score of 5), (2) comparability (maximum score of 2) and (3) outcome (maximum score of 3). Each of the three category scores was added to give a maximum total score of 10. Results of all included studies were synthesized based on their physical or mental impact on surgeons.

Results

Search results

A total of 6,563 articles were identified (Fig. 1). Seven additional articles were identified by checking through references of relevant articles. After the removal of duplicated papers, the remaining 5179 abstracts were screened resulting in a list of 71 articles. The 71 full-text articles were further assessed based on the predetermined inclusion/exclusion criteria resulting in a final number of 30 studies that were included in the qualitative synthesis. The 41 studies were excluded due to the use of a different

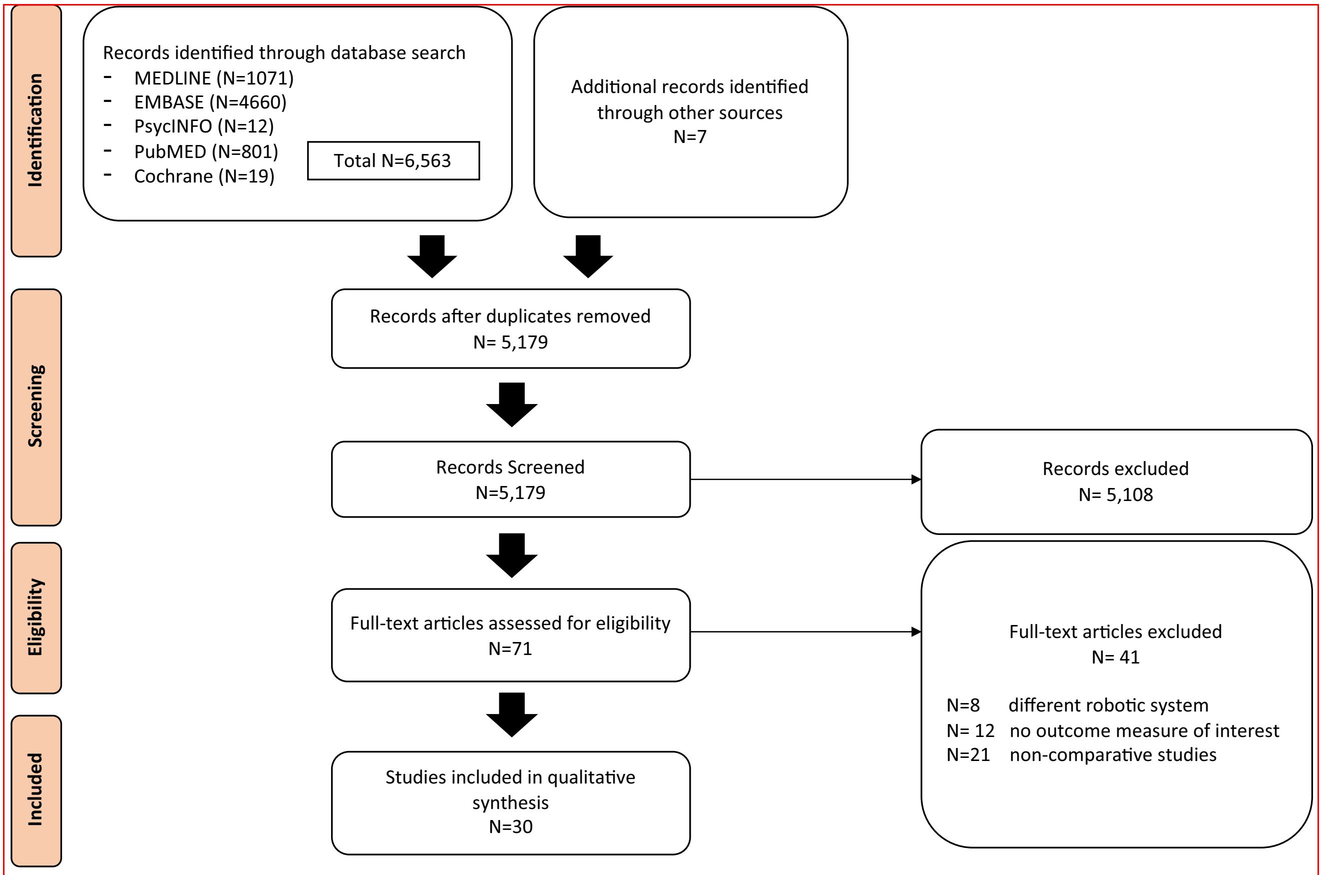

Fig. 1 PRISMA flowchart of literature search and selection of included studies

robotic system ($n = 8$), no outcome measure of interest ($n = 12$), or non-comparative studies ($n = 21$).

Study characteristics

The characteristics of the included studies are summarized in Tables 1 and 2. The number of participants in the final 30 studies ranged from 1 to 117 participants. 25 studies compared RS versus LS, one study compared RS versus OS, and four studies compared all three techniques. The included studies consisted of a wide range of surgery types (Tables 1 and 2) including simulation tasks, general, gynaecological, urological and thyroid surgery. Simulation surgical tasks were the most common type of surgery ($N = 14$) performed.

Physical and mental load assessment tools

Physical and mental impacts were examined in 24 (Table 1) and 19 studies (Table 2), respectively. Of these, 13 studies examined both the physical and mental impact of RS on surgeons. Examples of the tools and the number of studies that used each of these tools to compare the physical and mental impact of RS, LS, and/or OS on surgeons are outlined in Supplementary Tables 3 and 4.

Various types of measures and tools were used to assess the physical (Supplementary Table 3) and mental impact (Supplementary Table 4). 12 studies measured physical impact using quantitative tools, while 15 studies used subjective tools such as self-reported questionnaires or visual analogue scales. In contrast, mental strain was mostly measured using subjective questionnaires ($N = 18$), while quantitative measures such as cortisol levels and cardiovascular responses to stress were used by only five studies. The most commonly used tool was surface electromyography (EMG), an objective measure of physical stress, which was used by 9 studies. The NASA-TLX, a multi-dimensional subjective visual analogue rating scale that measures workload, was the next most common tool used by 8 studies which assesses both mental and physical load.

Table 1 Summary of characteristics and findings of studies examining physical demand

Author, year	Subjects	Type of surgery	Tool	p value	Summary of outcome	Favours
Robotic (intervention) versus laparoscopic surgery (control): N = 19						
Stefanidis 2010 [4]	34 medical students	Porcine Model: *Nissen fundoplication*	NASA-TLX summative score	$p < 0.001$	Less workload in RS	RS
van der Schatte, 2009 [14]	16 medical students	Simulation Tasks: *Rope passing, Needle capping, Bead dropping*	LED, # of physical complaints	$p < 0.01$	RS had lower LED Score and median score of physical complaints demonstrating less physical workload	RS
Hubert 2013 [15]	11 senior surgeons (7 general surgeons, 4 urology surgeons)	Porcine Model: *varied type of procedure depending on surgeon's specialty*	NASA-TLX: Physical Demand	$p < 0.05$	Less physical demand in RS	RS
			BORG CR-10: shoulders, neck, back		Less activity in RS for all regions	
			CV Measurement: Heart Rate		Slower HR and lower mean heart rate cost in RS	
			EMG: bilateral erector spinae, trapezius, flexor digitorum (RMS)		Less muscle activation in RS	
Zihni 2014 [16]	1 general surgeon	18 General Surgeries (13 LS, 5 RS)	EMG: bilateral biceps, triceps, deltoids, trapezius (% MVC)	$p < 0.05$	Less muscle activation in RS	RS
Zihni 2014 [23]	6 surgeons (2 attending surgeons, 2 mid-level surgical residents, 2 novices)	Simulation Tasks: *FLS*	EMG: Right trapezius, biceps, deltoid (% MVC)	$p < 0.05$	Mean activation of right bicep deltoid more elevated in LS Mean activation of right trapezius more elevated in RS	MIXED
Lawson 2007 [24]	1 surgeon	Roux-en-Y gastric bypass surgery (4 LS, 4 RS)	BPD: neck, upper back	$p < 0.05$	Neck: less discomfort in LS Upper back: less discomfort in RS	MIXED
			BPD: lower back, bilateral shoulders, buttocks, wrists, hands	$p > 0.05$	No difference	
			RULA: upper & lower arm, wrist, trunk	$p < 0.05$	Upper & lower arm, wrist: more ergonomical in RS Trunk: more ergonomical in LS	

Table 1 continued

Author, year	Subjects	Type of surgery	Tool	p value	Summary of outcome	Favours
Szeto 2013 [25]	2 surgeons	Colorectal Surgery (low anterior resection): each surgeon performed 1 RS and 1 LS	EMG: erector spinae, upper trapezius, anterior deltoid	N/A	Erector Spinae: no difference in surgeon 1, less activation in RS for surgeon 2 Upper Trapezius: Less activity in LS for surgeon 1, less activation in RS for surgeon 2 Anterior Deltoid: No difference in surgeon 1, less activation in RS for surgeon 2	MIXED
Sanchez 2018 [17]	14 surgeons	Simulation Task: *surgical repair of incisional hernia in an inanimate model*	LED: upper limb	$p < 0.01$	Less upper limb discomfort in RS	RS
Zarate Rodriguez 2019 [26]	14 novices	Simulation Tasks: *FLS*	EMG	$p < 0.05$	Higher %MVC in LS for most muscle groups except for right trapezius muscle Higher %MVC in RS for right trapezius muscle for most tasks except for the intracorporeal suturing task	MIXED
	12 LS surgeons			$p > 0.05$ except for right trapezius	Higher %MVC in RS for peg transfer and pattern cutting tasks but not for intracorporeal suturing task. Not statistically significant. Only stat significance in right trapezius ($p < 0.05$)	
	5 RS surgeons			$p > 0.05$ except for right trapezius	Muscle activation greater on LS but not statistically significant for most muscle groups Only statistically significant difference was for right trapezius muscle for peg transfer and pattern cutting ($p < 0.05$)	
	14 novices		NASA-TLX: physical demand	$p < 0.01$	Less workload in RS	
	12 LS surgeons			$p > 0.05$	No difference	
	5 RS surgeons					
Stefanidis 2011 [18]	117 attendees at the SAGES 2006 Learning Centre	Simulation Tasks: *FLS*	NASA-TLX: physical demand	$p < 0.001$	Less physical workload in RS	RS

Table 1 continued

Author, year	Subjects	Type of surgery	Tool	*p* value	Summary of outcome	Favours
Moss 2019 [19]	4 consultant gynaecologic oncology surgeons	Simulation Tasks: *Beads, Hoops, Wire Chase (normal & high BMI model)*	EMG (normalized peak)	Beads (normal BMI model): N.S. All others p < 0.01	Beads (normal BMI model): no difference. All other tasks: significantly higher muscle usage in LS	RS with NO DIFFERENCE on just one task/model
Mendes 2019 [27]	22 experienced & young surgeons (8 urologists, 11 gynaecologists, 3 paediatric surgeons)	Any surgery lasting greater than 60 min (88 RS, 82 LS)	BORG Scale	$p < 0.05$	Higher scores during LS in experienced surgeons No significant difference in young surgeons	MIXED
			NASA-TLX: Physical Demand		Less demand in RS in experienced surgeons Less demand in LS in young surgeons	
Butler 2013 [32]	6 gynaecologic surgeons (4 attending, 2 fellows)	Gynaecologic Surgeries	Quantitative Grip Dynamometer, Single-leg stance, Subjective visual analogue scale	$p > 0.05$	No difference between groups Single-leg stance shows trend towards decreased postural stability following laparoscopy (balance error scores 33%)	NO DIFFERENCE
Armijo 2019 [28]	16 various surgeons	28 various surgeries (18 LS, 10 RS)	EMG: upper trapezium, anterior deltoid, flexor carpi radialis (% MVC)	$p < 0.05$	Higher activations in RS	MIXED
			EMG: extensor digitorum (% MVC)	$p > 0.05$	No difference	
			EMG: upper trapezium, anterior deltoid, flexor carpi radialis (median frequency)	$p > 0.05$	No difference	
			EMG: extensor digitorum (median frequency)	$p < 0.001$	Lower median frequency in LS which correlates with increased muscle fatigue	
			Piper Fatigue Scale-12 Overall Score	$p = 0.869$	No difference	

Table 1 continued

Author, year	Subjects	Type of surgery	Tool	p value	Summary of outcome	Favours
Van Koughnett 2009 [29]	1 hepatobiliary surgeon	Porcine Model: 20 *choledochojejunal anastomoses*	SAS: Image quality, depth perception, comfort, eye fatigue, dexterity, precision of motion, speed of motion, range of motion	$p < 0.001$	Higher difficulty in LS	MIXED
			SAS: tactile awareness	$p < 0.001$	Higher difficulty in RS	
			SAS: fluidity of motion	$p = 0.07$	No difference; Non-significant trends favouring RS	
			Visual Analogue Scale: Degree of Difficulty	$p < 0.001$	Higher difficulty in LS	
Gonzalez-Sanchez 2017 [35]	1 general surgeon (chief role and assistant role)	Digestive surgery (1 RS/chief, 1 RS/assistant, 1 LS/chief, 1 LS/assistant)	PROMS: POMS Index, VAS, QPFS 9	$p < 0.05$	Chief role: significantly higher changes in RS for all except POMS confusion and QPFS Assistant role: higher changes in LS for POMS index, fatigue, vigour, friendliness, anxiety and QPFS; Higher changes in RS for POMS confusion and VAS; No difference in POMS anger and depression	INCONCLUSIVE
			OCOM: handgrip test & SLBT	$p < 0.05$	Chief role: high functional fatigue during RS Assistant role: no significant changes in grip strength for both RS and LS; higher SLBT runtime for dominant leg in LS but no difference in non-dominant leg	
			EMG (%MVC)	N/A	Higher activation in RS but statistic test not completed	

Table 1 continued

Author, year	Subjects	Type of surgery	Tool	p value	Summary of outcome	Favours
Lee 2014 [7]	13 MIS surgeons (6 LS experts, 4 surgical residents, 3 RS experts)	Simulation Task: *6 surgical training tasks*	EMG: biceps, flexor carpi ulnaris (cumulative workload)	$p < 0.05$	Higher workload in LS	MIXED
			EMG: trapezius (cumulative workload)	$p < 0.05$	Higher workload in RS	
			EMG: triceps, deltoid, extensor digitorum, thenar compartment, erector spinae (cumulative workload)	$p > 0.05$	No difference	
			NASA-TLX: physical demand	$p < 0.05$	Higher physical workload in LS	
Tarr 2015 [30]	16 surgeons	86 Sacrocolpopexy (33 RS, 53 LS)	NASA-TLX: physical demand, BPD (median)	$p > 0.05$	No difference	MIXED
			BPD (final model for change in BPD scale)	No p value reported but significant t-value scores of -2.49 and -2.38	Lower neck/shoulder, back discomfort scores in RS	
Moore 2015 [33]	32 surgeons	Simulation Tasks: *ball pick-and-drop task, rope-threading task*	SURG-TLX: physical demand	$p > 0.05$	No difference	NO DIFFERENCE
Robotic (intervention) versus open surgery (control): N = 1						
Collins 2012 [34]	8 gynaecological surgeons	Gynaecological Surgeries (8 OS, 8 RS)	Accelerometer	$p > 0.05$	No difference	NO DIFFERENCE
Robotic (intervention) versus laparoscopic surgery (control I) versus open surgery (control II): N = 4						
Lee 2011 [21]	7 surgeons	Thyroid Surgeries	MSK Questionnaire (neck/back pain)	Statistical analysis not performed	71.4% found LS and 28.6% found OS associated with the most MSK pain	RS (trend)
Law 2018 [20]	7 colorectal surgeons	Colorectal Surgeries (87 OS, 70 LS, 28 RS)	NASA-TLX: physical demand	$p < 0.005$	Less physical demand in RS	RS
Marcon 2019 [31]	multi-centre surgeons	Nephrectomies (65 OS, 65 SL, 65 HAL, 69 RS)	NASA-TLX: physical demand	$p < 0.005$	The least physical demand in RS	MIXED
			BORG Scale: bilateral shoulder/arm/forearm/hand, lower back, leg exertion	$p < 0.05$	Less left shoulder and arm exertion, left forearm and hand exertion but greater lower back exertion in RS	

Table 1 continued

Author, year	Subjects	Type of surgery	Tool	*p* value	Summary of outcome	Favours
Elhage 2015 [22]	6 urological surgeons	Simulation Task: In vitro vesico-urethral anastomosis	BORG Scale	OS versus LS, LS versus RS: $p < 0.005$ OS versus RS: $p > 0.005$	The greatest discomfort in LS. No difference in RS and OS	RS/OS

% MVC Percent Maximum Voluntary Contraction, *BMI* Body Mass Index, *BPD* Body Part Discomfort Scale, *CV* Cardiovascular, *EMG* Electromyography, *FLS* Fundamentals of Laparoscopic Surgery, *HAL* Hand-assisted Laparoscopy, *LS* Laparoscopic Surgery, *LED* Local Experienced Discomfort Scale, *MIS* Minimally Invasive Surgery, *MSK* Musculoskeletal, *N* Number of Studies, *N/A* Not Available, *NASA-TLX* National Aeronautics and Space Administration Task Load Index, *OCOM* Objective Clinical Outcome Measure, *OS* Open Surgery, *POMS* Profile of Mood States, *QPFS 9* Quick Questionnaire Piper Fatigue Scale, *RMS* Root Mean Square, *RS* Robotic Surgery, *RULA* Rapid Upper-Limb Assessment Tool, *SAS* Subjective Assessment Scale, *SL* Standard Laparoscopy, *SLBT* Single Leg Balance Test, *SURG-TLX* Surgery Task Load Index, *VAS* Visual Analogue Scale

Risk of bias

The modified Newcastle–Ottawa scores of the 30 studies ranged from 5 to 9 with a mean score of 7.5, showing moderate risk of bias (Table 3). As a modified version of the Newcastle–Ottawa Scale for cross-sectional studies [13], there were no predetermined threshold scores to determine a “good” quality study. None of the 30 studies justified their sample size. 5 studies did not choose a representative sample for the target population. 15 studies lost points as the ascertainment of exposure was based on self-report. All studies used appropriate statistical tests to analyse the data and were clearly described.

Study findings

A summary of the study findings is shown separately for the physical and mental demand in Tables 2 and 3, respectively.

Physical Demand. A total of 19 studies compared RS versus LS, one study compared RS versus OS, and four studies compared all three surgical modalities for physical impact on surgeons. Among these studies, eight studies favoured RS [4, 14–20], one study showed a trend towards favouring RS [21], and one study favoured RS over LS but showed no difference to OS [22]. Most studies ($N = 10$) showed mixed results [7, 23–31] with only three studies showing no difference between the surgical modalities [32–34], and one study with inconclusive results [35].

EMG was the most common tool used to measure physical demand in studies, which either favoured RS over LS [15, 16, 19, 23, 25, 26, 28], or produced mixed results [7, 23, 25, 28] as physical demand highly depended on which muscle was being measured. Less muscle activation in trapezius muscle [7, 23] but higher activation of arm muscles was seen for LS compared to RS [7, 23, 25]. On the other hand, one study [28] showed higher activation of the trapezium, anterior deltoid and flexor carpi radialis in RS, while there were no significant differences between the two surgical techniques in activation of other measure muscle groups.

Another common tool was the NASA-TLX, a subjective measure for physical workload, which showed that laparoscopic surgery was either more physically demanding [4, 7, 15, 18, 27] than robotic surgery or there were no significant differences between the two types of surgery [26, 30]. Interestingly, two studies [26, 27] compared the physical demand of robotic and laparoscopic surgery between novices and expert surgeons. While the study by Mendes et al. [27] showed that both novice and expert surgeons both showed less physical demand in robotic surgery measured by the NASA-TLX, the study by Zarate Rodrigues et al. [26] showed that novices found robotic surgery less physically demanding than laparoscopic surgery, while experts found no significant differences in physical demand between the two surgical techniques.

One study [34] comparing RS to OS showed no significant difference in physical activity levels measured by accelerometers between the two surgical techniques. The four studies [20–22, 31] comparing all three types of surgery showed the least physical discomfort or a trend towards the least physical discomfort in RS [20–22, 31] although one study recorded the greatest pain in the lower back in RS compared to OS or LS [31].

Mental demand. A total of 17 studies compared RS versus LS and two studies compared all three surgical

Table 2 Summary of characteristics and findings of studies examining mental demand

Author, year	Subjects	Type of surgery	Tool	p value	Summary of outcome	Favours
Robotic (intervention) versus laparoscopic surgery (control): N = 17						
Passerotti 2015 [39]	31 medical students and 12 surgeons	Simulation Tasks: *Peg transfer, precision cutting, simple suturing with intracorporeal knot tying*	ISAT: Frustration (during & after session) VAS: Mood (after session)	$p < 0.01$	More frustration in LS Better mood in RS	RS
Sanchez 2018 [17]	14 surgeons	Simulation Task: *surgical repair of incisional hernia in an inanimate model*	SMEQ	$p < 0.01$	Less mental effort in RS	RS
Gonzalez-Sanchez 2017 [35]	1 general surgeon (chief role and assistant role)	Digestive surgery (1 RS/chief, 1 RS/ assistant, 1 LS/chief, 1 LS/assistant)	QPFS VAS & POMS	$p > 0.05$ $p < 0.001$	No difference Less fatigue in LS	MIXED
Armijo 2019 [28]	16 various surgeons	28 various surgeries (18 LS, 10 RS)	QPFS	$p > 0.05$	No difference	NO DIFFERENCE
Klein 2012 [36]	15 medical students	Simulation Task: *peg transfer*	MRQ DSSQ	$p > 0.05$ $P < 0.05$	No difference Less stress in RS	MIXED
Klein 2014 [40]	10 residents, 6 expert surgeons	Simulation Task: *peg transfer*	DSSQ	$p < 0.05$	Less stress in RS	RS
Stefanidis, 2010 [4]	34 medical students	Porcine Model: *Nissen fundoplication*	NASA-TLX summative score	$p < 0.001$	Less workload in RS	RS
Hubert 2013 [15]	11 senior surgeons (7 general surgeons, 4 urology surgeons)	Porcine Model: *varied type of procedure depending on surgeon's specialty*	NASA-TLX: mental demand	$p > 0.05$	No difference	NO DIFFERENCE
Lee 2014 [7]	13 MIS surgeons (6 LS experts, 4 surgical residents, 3 RS experts)	Simulation Task: *6 surgical training tasks*	NASA-TLX: mental demand	$p > 0.05$	No difference	NO DIFFERENCE
Stefanidis 2011 [18]	117 attendees at the SAGES 2006 Learning Centre	Simulation Tasks: *FLS*	NASA-TLX: mental demand	$p > 0.05$	No difference	NO DIFFERENCE
Tarr 2015 [30]	16 surgeons	86 Sacrocolpopexy (33 RS, 53 LS)	NASA-TLX: mental demand	$p > 0.05$	No difference	NO DIFFERENCE
Mendes 2019 [27]	22 experienced & young surgeons (8 urologists, 11 gynaecologists, 3 paediatric surgeons)	Any surgery lasting greater than 60 min (88 RS, 82 LS)	NASA-TLX: mental demand	Experienced surgeons: $p > 0.05$ Young surgeons: $p < 0.05$	Experienced surgeons: no difference Young surgeons: less mental effort in LS	MIXED
Moore 2015 [37]	32 surgeons	Simulation Tasks: *ball pick-and-drop, rope-threading*	SURG-TLX: mental demand RSME	$p > 0.051$ $p < 0.001$	No difference Less demand in RS	MIXED
Moore 2015 [33]	32 surgeons	Simulation Task: *ball pick-and-drop*	STAI CV Measurements: challenge threat index	$p > 0.051$ $p < 0.05$	No difference More adaptive CV response in RS	MIXED

Table 2 continued

Author, year	Subjects	Type of surgery	Tool	p value	Summary of outcome	Favours
van der Schatte, 2009 [14]	16 medical students	Simulation Tasks: *Rope passing, Needle capping, Bead dropping*	CV Measurements: MSSD, PEP, HRA (mean)	$p < 0.01$	Lower CV stress in RS	RS
			SMEQ (median)		Less mental effort in RS	
Hurley 2015 [38]	16 medical students	Simulation Tasks: *The 3-Dmed® task kit*	Skin conductance, Heart rate (avg), Heart rate variability (SDNN, RMSSD)	$p < 0.01$	Less demand in RS	MIXED
			Mean arterial pressure, Cortisol	$p > 0.05$	No difference	
Heemskerk 2014 [41]	2 surgeons	Elective Cholecystectomies (11 LS and 11 RS)	HR (mean)	$p < 0.05$	Less demand in RS	RS
			LF/HF ratio	$p < 0.05$	Less demand in RS	
Robotic (Intervention) versus Laparoscopic surgery (Control I) versus Open surgery (Control II): N = 2						
Law 2018 [20]	7 colorectal surgeons	Colorectal Surgeries (87 OS, 70 LS, 28 RS)	NASA-TLX: mental demand	RS vs. LS: $p < 0.05$ RS vs. OS: $p > 0.05$	Less mental demand in RS than LS but no difference between RS and OS	MIXED
Marcon 2019 [31]	multi-centre surgeons	Nephrectomies (65 OS, 65 SL, 65 HAL, 69 RS)	NASA-TLX: mental demand	$p < 0.01$	The most mental demand in OS	INCONCLUSIVE

CV Cardiovascular, *DSSQ* Dundee Stress State Questionnaire, *FLS* Fundamentals of Laparoscopic Surgery, *HAL* Hand-assisted Laparoscopy, *HR* Heart Rate, *HRA* Heart Rate Average, *ISAT* Imperial Stress Assessment Tool, *LS* Laparoscopic Surgery, *LF/HF* Low frequency/High frequency, *MIS* Minimally Invasive Surgery, *MRQ* Multiple Resources Questionnaire, *MSSD* Mean Square of Successive Differences between Consecutive Heartbeats, *NASA-TLX* National Aeronautics and Space Administration Task Load Index, *OS* Open Surgery, *PEP* Pre-Ejection Period, *POMS* Profiles of Mood States, *QPFS* Quick Questionnaire Piper Fatigue Scale, *RMSSD* Root Mean Square of Successive Differences between Consecutive Heartbeats, *RS* Robotic Surgery, *RSME* Rating Scale of Mental Effort, *SDNN* Standard Deviation of the NN (R-R) Intervals, *SL* Standard Laparoscopy, *SMEQ* Subjective Mental Effort Questionnaire, *STAI* State-Trait Anxiety Inventory, *SURG-TLX* Surgery Task Load Questionnaire, *VAS* Visual Analogue Scale

modalities for mental impact on surgeons. Of these, seven studies [20, 27, 33, 35–38] showed mixed results, six studies [4, 14, 17, 39–41] favoured better mental outcomes in RS, and five studies [7, 15, 18, 28, 30] showed no difference. One study [31] comparing all three modalities showed the most mental demand in OS but did not report any statistical differences between RS and LS.

Mental demand was mostly measured by subjective measures such as self-report questionnaires, most commonly the NASA-TLX [7, 15, 18, 20, 27, 30, 31]. These studies all showed no differences in mental demand between the surgical techniques except for in young surgeons [27] and one study showing less mental demand in RS than LS [20]. Studies using physiological measures of mental stress [14, 33, 38] all favoured robotic surgery with the exception of mean arterial pressure and cortisol levels in one study [38].

Discussion

This systematic review aimed to synthesize the current literature on the physical and mental demand of robotic surgery on surgeons compared to laparoscopic and/or open surgery. Although systematic reviews examining the impact of robotic surgery to the patient outcomes have

Table 3 Risk of Bias scored by the modified Newcastle–Ottawa Tool

Author	Selection	Comparability	Exposure	Total
Hubert 2013 [15]	4	2	3	9
Heemskerk 2014 [41]	4	2	3	9
Elhage 2015 [22]	4	2	3	9
Gonzalez-Sanchez 2017 [35]	4	2	3	9
Armijo 2019 [28]	4	2	3	9
Moss 2019 [19]	4	2	3	9
van der Schatte 2009 [14]	4	2	2	8
Stefanidis 2010 [4]	4	2	2	8
Stefanidis 2011[18]	4	2	2	8
Collins 2012 [34]	3	2	3	8
Butler 2013 [32]	4	1	3	8
Zihni 2014 [16]	3	2	3	8
Hurley 2015 [38]	3	2	3	8
Moore 2015 [37]	4	1	3	8
Zarate Rodriguez 2019 [26]	4	1	3	8
Mendes 2019 [27]	4	2	2	8
Lawson 2007 [24]	3	2	2	7
Klein 2012 [36]	3	2	2	7
Zihni 2014 [23]	2	2	3	7
Klein 2014 [40]	3	1	3	7
Lee 2014 [7]	3	2	2	7
Tarr 2015 [30]	4	1	2	7
Law 2018 [20]	3	2	2	7
Van Koughnett 2009 [29]	2	2	2	6
Lee 2011 [21]	3	1	2	6
Passerotti 2015 [39]	3	1	2	6
Moore 2015 [33]	3	1	2	6
Sanchez 2018 [17]	3	1	2	6
Marcon 2019 [31]	4	0	2	6
Szeto 2013 [25]	1	1	3	5

been published [15], the benefits for surgeons have yet to be critically assessed. One recent systematic review [10] appraised musculoskeletal pain in surgeons performing robotic surgery and a more recent meta-analysis [9] compared muscle activation between robotic and laparoscopic surgery using EMG; however, no studies to date have evaluated both the physical and mental impact of robotic surgery on surgeons.

Majority of the included studies in the current systematic review produced mixed results. Although many of the studies showed a general trend towards favouring robotic surgery, it is evident from this review that there is high heterogeneity in study size and methodology, surgical specialties, procedures, techniques, as well as measures to evaluate physical and mental demand. Therefore, the results of the current systematic review need to be interpreted with caution. The inconsistency in the study findings as well as the variability of outcome measures makes interpretation of the study results and generalizability challenging.

One of the major contributors to the heterogeneity were the tools used to measure physical and mental impact. Seven different subjective scales and six different objective tools were used to measure physical workload alone, while there were even greater variety in tools used to measure mental workload. Such variety in tools prevented performing a meta-analysis.

The NASA-TLX is a commonly used subjective measure of workload in human factors that is being increasingly used in surgical research [42]. The NASA-TLX consists of general questions on the subjective experience of physical or mental workload without specifying locations of the body or aspects of cognitive demand, respectively. It measures a general impression of physical and mental demand rather than specific areas of discomfort that can be compared between surgical technique types. Lawson et al. [24] used a different questionnaire, the Body Part Discomfort Questionnaire, which is a more targeted survey to measure physical discomfort. The authors reported less physical discomfort in the upper back and extremities but more discomfort in the neck and trunk region for robotic surgery compared to laparoscopic surgery. [24] The physical comfort surgeons experience may differ depending on the body part, which the NASA-TLX fails to capture. Despite this limitation, the NASA-TLX was one of the most commonly used tools. The Surgery Task Load Index is more specific to surgery but was used in two studies by the same group of researchers. [33, 37]

Surface EMG, an objective measure of muscle activity in a specific body part, was the most common tool used in this systematic review to assess physical workload. Some studies using surface EMG [15, 16, 19] were favourable for robotic surgery compared to laparoscopic or open surgery. However, majority of the studies using EMG [7, 23, 25, 26, 28] showed mixed results where the reduced muscle activation depended on which body part was being measured. This is consistent with the recent meta-analysis by Hislop et al. [9] which showed that the biceps were the only muscle group that consistently demonstrated lower muscle activation for robotic surgery.

Other studies have used tools such as the RULA, accelerometers, quantitative grip dynamometers, single-leg stance and cardiovascular measurements to examine the physical stress of surgical techniques on surgeons (Supplementary Table 3). Since there is no “gold standard” tool for measuring physical stress in surgeons, several different tools have been utilized, which make comparison of results across studies difficult. Moreover, increased activities in these objective measures may indicate more movement or muscle strain; however, they may not necessarily mean that the surgeon subjectively experiences greater physical strain. Only nine studies [7, 14, 15, 24, 26–28, 31, 32] included in this review used both subjective and objective measures to correlate the physical workload findings. It would be informative for future studies to examine correlations between these objective measures of physical stress and subjective measurement tools.

19 studies examined the mental impact of robotic surgery on surgeons most commonly using self-reported rating scales, while only five studies [14, 33, 37, 38, 41] utilized objective measures of mental stress. Various validated self-reported scales (Supplementary Table 4) including visual analogue scales [29, 32] were used to evaluate mental stress, fatigue, frustration or mental effort in surgeons, which may be subject to personal bias and preference in surgical technique. Furthermore, the variety of tools used make it difficult to interpret the results across the studies.

Studies using physiological measures of stress such as cardiovascular measures, cortisol, and skin conductance have been used to objectively measure stress. Although a few studies favoured robotic surgery, which showed lower mental effort and more adaptive cardiovascular responses to stress conditions, each study used a different calculation making interpretation across the studies challenging. It is important to note that these studies used heart rate variability as a measure of mental strain under the assumption that its increase is closely associated with increased sympathetic activation. However, cardiovascular measures may be influenced by various factors and the validity of such physiological measures of stress remains unclear. As such more studies using the same type of physiological tool measure in the same way in a larger number of subjects are necessary for better interpretation of the data.

There is no doubt that comparative surgical studies are challenging due to various patient, surgeon, environmental and skill factors. This is reflected in this systematic review which has demonstrated high heterogeneity in study size, methodology, surgical specialties, surgical expertise (from medical students to experienced surgeons) and tools used. In addition to the methodological weakness of the included studies, the majority of the studies used simulations which may underestimate both the physical and mental stress experienced by the surgeons compared to real surgeries. Even within the simulation studies, there were a wide range of outcome measures, expertise level and type of simulation used which resulted in significant heterogeneity and prevented meta-analysis and consistent interpretation of the results. It is possible that certain surgical specialties may benefit greater from robotic surgical techniques than others. For example, laparoscopic prostatectomy was performed by a small number of surgeons because of the technical difficulty but many perform robotic prostatectomy. This review is specific to the da Vinci robotic system and cannot be generalized to other platforms.

Further studies are necessary to better understand how different surgical approaches impact the surgeon’s physical and mental load during surgery. In the surgeons’ lifetime, physical pain and fatigue may increase the risk of complications and mistakes during surgery. [7] This level of physical pain and fatigue may vary depending on the level of surgeons’ expertise and setting of the surgery, which both need to be further ascertained in future studies. Additionally, mental stress and mental wellbeing can affect the efficiency, productivity and longevity of the surgeons’ career, [40] which may have economic benefits as training a surgeon is costly.

Conclusion

This systematic review identified 30 studies that examined the physical and/or mental impact of the da Vinci robotic surgical system on surgeons compared to laparoscopic and/or open surgical techniques. Most studies showed mixed physical and mental outcomes between the three surgical modalities. This is most likely due to the high heterogeneity in methodology and measurement tools used in the included studies, which makes comparison of results between studies challenging. Overall, the available evidence regarding the physical and mental demand for the different surgical approaches is of relatively low quality and it is not possible to definitely state that robotic surgery has less physical or mental fatigue based on the current evidence. Studies on long-term outcomes are needed to better understand the differences in cognitive or physical demand of surgeons between the three surgical modalities and their impact over time.

Acknowledgements We would like to thank the librarians at the University of Sydney for their time and assistance in the systematic literature search.

Funding There are no sources of funding to declare for this current study.

References

1. Darzi A, Munz Y (2004) The impact of minimally invasive surgical techniques. Annu Rev Med 55:223–237
2. Wagenaar S, Nederhoed JH, Hoksbergen AWJ et al (2017) Minimally invasive, laparoscopic, and robotic-assisted techniques versus open techniques for kidney transplant recipients: a systematic review. Eur Urol 72:205–217
3. Gershman B, Bukavina L, Chen Z et al (2018) The association of robot-assisted versus pure laparoscopic radical nephrectomy with perioperative outcomes and hospital costs. Eur Urol Focus 6:305–312
4. Stefanidis D, Wang F, Korndorffer JR Jr et al (2010) Robotic assistance improves intracorporeal suturing performance and safety in the operating room while decreasing operator workload. Surg Endosc 24:377–382
5. Coughlin GD, Yaxley JW, Chambers SK et al (2018) Robot-assisted laparoscopic prostatectomy versus open radical retropubic prostatectomy: 24-month outcomes from a randomised controlled study. Lancet Oncol 19:1051–1060
6. Berguer R, Forkey DL, Smith WD (1999) Ergonomic problems associated with laparoscopic surgery. Surg Endosc 13:466–468
7. Lee GI, Lee MR, Clanton T et al (2014) Comparative assessment of physical and cognitive ergonomics associated with robotic and traditional laparoscopic surgeries. Surg Endosc 28:456–465
8. Dalsgaard T, Jensen MD, Hartwell D et al (2020) Robotic surgery is less physically demanding than laparoscopic surgery: paired cross sectional study. Ann Surg 271:106–113
9. Hislop J, Tirosh O, McCormick J et al (2019) Muscle activation during traditional laparoscopic surgery compared with robot-assisted laparoscopic surgery: a meta-analysis. Surg Endosc 03:03
10. Dalager T, Sogaard K, Bech KT et al (2017) Musculoskeletal pain among surgeons performing minimally invasive surgery: a systematic review. Surg Endosc 31:516–526
11. Abdelrahman AM, Lowndes B, Rand C et al (2017) Impact of robotic surgery versus laparoscopic surgery on surgeon musculoskeletal symptoms and workload: a systematic review and meta-analysis. Surg Endosc Other Interv Tech 31(1):S324
12. Shamseer L, Moher D, Clarke M et al (2015) Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 350:g7647
13. Herzog R, Álvarez-Pasquin MJ, Díaz C et al (2013) Are healthcare workers' intentions to vaccinate related to their knowledge, beliefs and attitudes? A systematic review. BMC Public Health 13:154
14. van der Schatte Olivier RH, Van't Hullenaar CD, Ruurda JP et al (2009) Ergonomics, user comfort, and performance in standard and robot-assisted laparoscopic surgery. Surg Endosc 23:1365–1371
15. Hubert N, Gilles M, Desbrosses K et al (2013) Ergonomic assessment of the surgeon's physical workload during standard and robotic assisted laparoscopic procedures. Int J Med Robot 9:142–147
16. Zihni AM, Ohu I, Cavallo JA et al (2014) Ergonomic analysis of robot-assisted and traditional laparoscopic procedures. Surg Endosc 28:3379–3384
17. Sanchez A, Rodriguez O, Jara G et al (2018) Robot-assisted surgery and incisional hernia: a comparative study of ergonomics in a training model. J Robot Surg 12:523–527
18. Stefanidis D, Hope WW, Scott DJ (2011) Robotic suturing on the FLS model possesses construct validity, is less physically demanding, and is favored by more surgeons compared with laparoscopy. Surg Endosc 25:2141–2146
19. Moss EL, Sarhanis P, Ind T et al (2019) Impact of obesity on surgeon ergonomics in robotic and straight-stick laparoscopic surgery. J Minim Invasive Gynecol. https://doi.org/10.1016/j.jmig.2019.07.009
20. Law KE, Lowndes BR, Kelley SR et al (2018) NASA-task load index differentiates surgical approach opportunities for improvement in colon and rectal surgery. Ann Surg. https://doi.org/10.1097/SLA.0000000000003173
21. Lee J, Kang SW, Jung JJ et al (2011) Multicenter study of robotic thyroidectomy: Short-term postoperative outcomes and surgeon ergonomic considerations. Ann Surg Oncol 18:2538–2547
22. Elhage O, Challacombe B, Shortland A et al (2015) An assessment of the physical impact of complex surgical tasks on surgeon errors and discomfort: a comparison between robot-assisted, laparoscopic and open approaches. BJU Int 115:274–281
23. Zihni AM, Ohu I, Cavallo JA et al (2014) FLS tasks can be used as an ergonomic discriminator between laparoscopic and robotic surgery. Surg Endosc 28:2459–2465
24. Lawson EH, Curet MJ, Sanchez BR et al (2007) Postural ergonomics during robotic and laparoscopic gastric bypass surgery: a pilot project. J Robot Surg 1:61–67
25. Szeto GP, Poon JT, Law WL (2013) A comparison of surgeon's postural muscle activity during robotic-assisted and laparoscopic rectal surgery. J Robot Surg 7:305–308
26. Zarate Rodriguez JG, Zihni AM, Ohu I et al (2019) Ergonomic analysis of laparoscopic and robotic surgical task performance at various experience levels. Surg Endosc 33:1938–1943
27. Mendes V, Bruyere F, Escoffre JM et al (2019) Experience implication in subjective surgical ergonomics comparison between laparoscopic and robot-assisted surgeries. J Robot Surg. https://doi.org/10.1007/s11701-019-00933-2
28. Armijo PR, Huang CK, High R et al (2019) Ergonomics of minimally invasive surgery: an analysis of muscle effort and

fatigue in the operating room between laparoscopic and robotic surgery. Surg Endosc 33:2323–2331

29. Van Koughnett JA, Jayaraman S, Eagleson R et al (2009) Are there advantages to robotic-assisted surgery over laparoscopy from the surgeon's perspective? J Robot Surg 3:79–82
30. Tarr ME, Brancato SJ, Cunkelman JA et al (2015) Comparison of postural ergonomics between laparoscopic and robotic sacrocolpopexy: a pilot study. J Minim Invasive Gynecol 22:234–238
31. Marcon B, Sime WN, Guillemin F et al (2019) An ergonomic assessment of four different donor nephrectomy approaches for the surgeons and their assistants. Res Rep Urol 11:261–268
32. Butler KA, Kapetanakis VE, Smith BE et al (2013) Surgeon fatigue and postural stability: is robotic better than laparoscopic surgery? J Laparoendosc Adv Surg Tech A 23:343–346
33. Moore LJ, Wilson MR, Waine E et al (2015) Robotically assisted laparoscopy benefits surgical performance under stress. J Robot Surg 9:277–284
34. Collins SA, O'Sullivan DM, Tulikangas PK (2012) Surgeon activity in robotic versus abdominal gynecologic surgery. J Robot Surg 6:333–336
35. Gonzalez-Sanchez M, Gonzalez-Poveda I, Mera-Velasco S et al (2017) Comparison of fatigue accumulated during and after prolonged robotic and laparoscopic surgical methods: a cross-sectional study. Surg Endosc 31:1119–1135
36. Klein MI, Warm JS, Riley MA et al (2012) Mental workload and stress perceived by novice operators in the laparoscopic and robotic minimally invasive surgical interfaces. J Endourol 26:1089–1094
37. Moore LJ, Wilson MR, McGrath JS et al (2015) Surgeons' display reduced mental effort and workload while performing robotically assisted surgical tasks, when compared to conventional laparoscopy. Surg Endosc 29:2553–2560
38. Hurley AM, Kennedy PJ, O'Connor L et al (2015) SOS save our surgeons: stress levels reduced by robotic surgery. Gynecol Surg 12:197–206
39. Passerotti CC, Franco F, Bissoli JC et al (2015) Comparison of the learning curves and frustration level in performing laparoscopic and robotic training skills by experts and novices. Int Urol Nephrol 47:1075–1084
40. Klein MI, Mouraviev V, Craig C et al (2014) Mental stress experienced by first-year residents and expert surgeons with robotic and laparoscopic surgery interfaces. J Robot Surg 8:149–155
41. Heemskerk J, Zandbergen HR, Keet SW et al (2014) Relax, it's just laparoscopy! a prospective randomized trial on heart rate variability of the surgeon in robot-assisted versus conventional laparoscopic cholecystectomy. Dig Surg 31:225–232
42. Lowndes BR, Forsyth KL, Blocker RC et al (2020) NASA-TLX Assessment of surgeon workload variation across specialties. Ann Surg 271:686–692