A systematic review of robotic versus open and video assisted thoracoscopic surgery (VATS) approaches for thymectomy

Katie E. O'Sullivan[1], Usha S. Kreaden[2], April E. Hebert[2], Donna Eaton[1], Karen C. Redmond[1]

[1]Department of Thoracic Surgery, Mater Misericordiae University Hospital, Eccles St., Dublin, Ireland; [2]Clinical Affairs, Intuitive Surgical Inc., Sunnyvale, CA, USA

Correspondence to: Karen C. Redmond. Mater Misericordiae University Hospital, Eccles St., Dublin 7, Ireland. Email: karedmond@mater.ie.

Background: Median sternotomy has been the most commonly used approach for thymectomy to date. Recent advances in video-assisted thoracoscopic surgery (VATS) and robotic access with CO_2 insufflation techniques have allowed more minimally invasive approaches. However, prior reviews have not compared robotic to both open and VATS thymectomy.

Methods: A systematic review was conducted in accordance with the PRISMA guidelines using PubMed, Embase and Scopus databases. Original research articles comparing robotic to VATS or to open thymectomy for myasthenia gravis, anterior mediastinal masses, or thymomas were included. Meta-analyses were performed for mortality, operative time, blood loss, transfusions, length of stay, conversion to open, intraoperative and postoperative complication rates, and positive/negative margin rates.

Results: Robotic thymectomy is a valid alternative to the open approach; advantages include: reduced blood loss [weighted mean difference (WMD): –173.03, 95% confidence interval (95% CI): –305.90, –40.17, P=0.01], fewer postoperative complications (odds ratio: 0.37, 95% CI: 0.22, 0.60, P<0.00001), a shorter hospital stay (WMD: –2.78, 95% CI: –3.22, –2.33, P<0.00001), and a lower positive margin rate (relative difference: –0.04, 95% CI: –0.07, –0.01, P=0.01), with comparable operative times (WMD: 6.73, 95% CI: –21.20, 34.66, P=0.64). Robotic thymectomy was comparable with the VATS approach; both have the advantage of avoiding median sternotomy.

Conclusions: While randomized controlled studies are required to make definitive conclusions, current data suggests that robotic thymectomy is superior to open surgery and comparable to a VATS approach. Long-term follow-up is required to further delineate oncological outcomes.

Keywords: Thymectomy; robotic; video-assisted thoracoscopic surgery (VATS); da Vinci

Introduction

Developments in operative technology have enabled rapid advancements in the surgical approach for thymectomy over recent years. Thymectomy is indicated for excision of tumours of the anterior mediastinum, and for the treatment of myasthenia gravis. Median sternotomy is the most widely used surgical approach used to date (1). Additionally, the option to perform a trans cervical approach for thymomas under 4 cm has also been reported in a more limited manner with comparable results (2). However, recent improvements in robotic and video assisted surgical platforms, such as the

addition of CO_2 insufflation and the availability of energy devices, have enabled the use of these approaches for thymectomy.

Initial reports of video-assisted thoracoscopic surgery (VATS) thymectomy began to emerge in 1993 with a number of centers reporting its successful use either alone or in combination with a trans cervical approach (3). Use in the paediatric population was also introduced at this time for treatment of thymic hyperplasia (4). Over subsequent years, the prevalence of this approach has increased, with groups reporting lower intraoperative blood loss, less post-

operative pleural drainage, and a shorter post-operative hospital stay with VATS compared to open operations (5).

The first reports of robotic surgery in the field of thoracic surgery, and particularly its application for thymectomy, emerged in the early 2000's. Ashton *et al.* (6) were first to report the successful use of the da Vinci robotic system in completing a thymectomy for myasthenia gravis in a 28-year-old patient. They used a four-port technique from the right chest, followed by a set of symmetrical incisions on the contralateral side, and used one port to complete the left sided dissection. Since that time, the use of minimally invasive techniques has continued to expand, first for benign thymic pathologies, with a slower adoption for thymic malignancies due to concerns over increased tumour manipulation, capsular disruption, and incomplete resection (7,8). Robotic instruments are reported to offer superior dexterity and are advantageous in the narrow retrosternal anatomical space (9).

In 2011, the International Thymic Malignancy Interest Group published nine principles of minimally invasive thymic resection (10). In this publication, there was concern that dissection of the phrenic nerve and major vessels should not be undertaken using a minimally invasive approach as the possibility of bad outcomes could set the field back. Therefore, it is important to examine the current state of thymectomy.

No study to date has compared both open to robotic and VATS to robotic approaches for thymectomy. We performed a systematic review and meta-analysis comparing open to robotic and VATS to robotic thymectomy.

Methods

Literature search strategy

A systematic review was conducted in accordance with the PRISMA guidelines (11) using a two-pronged approach. Monthly generic robotic searches as well as one-time thymectomy-specific searches were conducted using PubMed, Scopus, and Embase databases to find relevant publications for this clinical evaluation. The monthly robotic generic searches were conducted as listed: (robotic[All Fields] OR robot assist[All Fields] OR robotically assisted[All Fields] OR robot-assist[All Fields] OR da vinci[All Fields] OR "davinci"[All Fields] OR intuitive surgical[All Fields] OR ("robotic"[All Fields] AND "surgery"[all fields]). The one-time thymectomy-specific searches included the search string listed above with the following addition: "AND (thymectomy

OR thymoma OR thymic OR thymus OR "myasthenia gravis")". All citations returned from the above searches were exported into an EndNote library. Duplications were removed and titles and abstracts were reviewed by three authors (KEOS, AEH, USK) for inclusion in the library.

Types of outcomes measured

- ❖ Primary: short term (30-day) mortality.
- ❖ Secondary: operative time, estimated blood loss (EBL), transfusion rate, conversion to open, intra-operative complications and post-operative complications, length of hospital-stay (LOS), and positive/negative margin rate.

Eligibility and exclusion criteria

The literature search was conducted on July 2nd 2018. Inclusion criteria were met if: (I) the English language journal article described robotic-assisted thymectomy in adult humans, (II) publication was a primary source comparative article reporting on robotic versus VATS or open thymectomy. Exclusion criteria were met if: (I) publication was not in English, (II) publication was not a journal article (abstract, book, book chapter), (III) publication was not about da Vinci thymectomy in adult humans, (IV) publication was a health technology assessment that was not published in a peer reviewed journal, (V) study was a review, (VI) publication did not have a comparator, (VII) alternative techniques or approaches were used (i.e., single-port, hand-assist), (VIII) there was no analysis stratified by study arm or by procedure, (IX) the study did not provide quantitative results for at least one of the findings relative to the outcomes of interest, (X) study includes a redundant patient population and similar conclusions. The PRISMA flowchart is outlined in *Figure 1*.

Data extraction and critical appraisal: assessment of risk

Three reviewers independently extracted the data (KEOS, AEH, USK) into a pre-defined excel spreadsheet. We recorded details about trial design, primary and secondary outcomes. Three authors independently evaluated and included studies for the presence of selection, performance, detection, attrition, reporting, or other (learning curve, conflicts of interest) bias using a modified version of the Cochrane Handbook risk of bias tools for non-randomized studies. Summarized criteria for low risk determinations
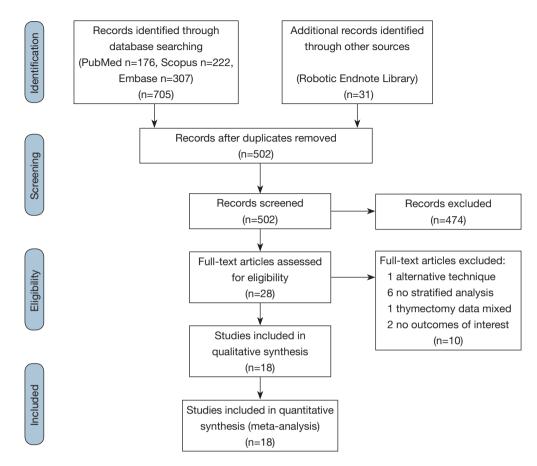
Figure 1 PRISMA flowchart of search.

include: selection bias: the cohorts were contemporary and comparable, matched for patient characteristics, or adjusted for confounding factors; performance bias: the cohorts were matched on operative and/or hospital characteristics, such as surgical technique, care pathways, and length of follow up to make combining them reasonable, or differences were addressed; detection bias: data capture and entry was standardized/performed by trained personnel and precise definitions of outcomes of interest were provided; attrition bias: there was no missing data, or missing data not an issue, no (or few) patients were lost to follow up, and length of follow up comparable and sufficient; reporting bias: all pre-specified outcomes of interest were reported in the pre-specified way regardless of significance and were complete enough for inclusion in a meta-analysis; other bias: there were no funding or industry conflicts of interest that were deemed an issue, the authors accounted for experience/volume of surgeons and/or hospital, no other obvious bias. All disagreements were resolved through discussion.

Statistical analysis

Pooled analysis was performed whenever at least two papers reported an outcome of interest in sufficient detail (data were reported for both cohorts, point estimate with variance for continuous variables, total n and event n or % for dichotomous variables). For continuous variables, a weighted mean difference and 95% confidence interval (WMD, 95% CI) were calculated using the inverse variance method. For dichotomous variables, an odds ratio with 95% confidence interval (OR, 95% CI) was calculated using the Mantel-Haenszel (M-H) method, except when at least two papers reported zero events for both cohorts. In that case, a risk difference (RD, 95% CI) was calculated. A random effects model was used when heterogeneity was statistically significant (Chi^2 P<0.05, I^2 >50%) and a fixed effects model was used when heterogeneity was not statistically significant. In all cases, a P value of less than 0.05 was considered statistically significant. All analyses, forest plots, and funnel plots were done using Review Manager

(Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014). Continuous data not reported as mean with standard deviation were converted using Review Manager calculators when possible.

Results

Quantity of evidence

A total of 736 results were returned following application of the search terms to PubMed (n=176), Scopus (n=222), Embase (n=307), and from the generic robotic searches (n=31) (*Figure 1*). After removal of duplicates, 502 studies remained. Preliminary screening resulted in the removal of 474 further studies and 28 were then assessed for eligibility. Of these, one study by Suda *et al.* (12) that compared single-port thymectomy using conventional laparoscopic instruments with multi-port robotic thymectomy was removed because the use of an alternative technique in the laparoscopic arm was a confounding factor in comparing laparoscopy to robotics. However, the authors reported that there were no differences between groups for EBL, LOS, oral analgesic use, or complications. The only difference found was a shorter length of operative time in the single-port VATS group. The authors concluded that the multiport robotic approach was as "equally minimally invasive as" the single-port VATS approach. An additional six studies (13-18) were excluded because none of the results were stratified according to surgical approach, one study (19) was excluded because the thymectomy data was mixed with data for other procedures in the analysis, and two studies (20,21) were excluded because they did not report on any of our outcomes of interest. A total of 18 studies (22-39) remained for further analysis, reporting on 776 robotic, 566 VATS, and 2,872 open cases. Study characteristics are outlined in *Table 1*. For analysis, the studies were separated into papers comparing robotic *vs.* open and papers comparing robotic *vs.* VATS thymectomy.

Quality of evidence: risk of bias

The risk of bias assessment is summarized in *Figure 2*. All of the studies were non-randomized and most were retrospective, so they were at higher risk of bias than a randomized-controlled trial (RCT). Several of the studies had a high risk of selection bias due to historical controls (24,32,34,36,37), differences in criteria for patient selection (22,33), or patient characteristic differences between groups (23). Studies that performed propensity score matching for outcomes of interest were considered at low risk for selection bias (25,27-30,35), as were the papers by Qian *et al.* (31) and Ye *et al.* (39), due to contemporary controls, the same selection criteria, and no differences in patient characteristics. Most of the studies were rated as unknown or high risk for performance bias due to a lack of information provided on surgical techniques and care pathways. Studies that described similar surgical procedures and postoperative care pathways (22,24,35), utilized the same surgical team (32), or specifically mentioned that surgical decisions were standardized regardless of surgical approach (25), were considered at low risk for performance bias. The paper by Weksler *et al.* (36) was assessed as having a high risk of detection bias because the authors specifically mentioned not being able to obtain accurate operative times for the transsternal group and did not report all of the same outcomes for each group. All of the other studies were rated as unclear risk due to a lack of information (22,24,26,28,35,39), or as low risk of detection bias due to prospectively collected data (27,33,34,37), data entry into a registry (23), precise definitions of outcomes provided (25,29,30,38), or because any missing definitions were unlikely to affect outcomes (31,32). The study by Weksler *et al.* (36) was also assessed as high risk for attrition bias because of incomplete data that was likely due to intervention type. All of the other studies were assessed as unclear or low risk for attrition bias for perioperative outcomes due to a sufficient follow up and infrequent or equivalent loss of patients/missing data between groups. Five studies were assessed as having a high risk of reporting bias. The study by Burt *et al.* (23) did not report data separately for the robotic group, except for an unmatched margin rate. In many studies (24,26,32,36), outcomes of interest were reported incompletely (e.g., missing a measure of variance), such that they could not be entered into the meta-analysis. Studies that included the robotic learning curve (22,24,32,33) were assessed at high risk of "other" bias. Publication bias was assessed using funnel plots (*Figure S1*) for all analyses that included at least 10 studies, all four showed symmetry and two showed no publication bias.

Assessment of primary and secondary endpoints

Robotic versus open thymectomy

A total of 14 publications (22-25,27-32,35-38) including 3,487 patients (n=615 robotic and n=2,872 open) compared robotic and open approaches. The results are detailed in

Table 1 Study characteristics

Author	Study design	Time period	Study arms	Patient number	Operative approach	Indication	Thymoma pathology (n)/or MGFA classification (n)	Age, mean ± SD or median [range]	Gender (# male)	Length of follow up
Balduyck (22)	PC	2004–2008	R	14	Rt or L multiport	AMM	A: 1, B1: 2, B2: 1, AB: 1	49 [18–63]	4	34 mo
			O	22	Median sternotomy		A: 1, B1: 2, B2: 5, B3: 1, AB: 3	56 [23–84]	12	50 mo
Burt (23)	ITMIG-DB	1997–2012	R	146	NR	Thymoma	MI: I: 199, II: 186, III: 27, IV: 12	MI: 52 [15–85]	MI: 207	NR
			VATS	315	NR		—	—	—	
			O	2,053	Sternotomy/thoracotomy		I: 669, II: 654, III: 344, IV: 130	54 [8–88]	1,032	
Cakar (24)	RC	1996–2006	R	9	4-port	MG (some w/thymoma)	A: 2, B2: 1, AB: 1	NR	NR	R: 13±10 mo
			O	10	Median sternotomy		A: 2, B2: 1			O: 74±23 mo
Casiraghi (25)	RC-PM	1998–2017	R	24*	3-port	Thymoma	I: 8, IIA: 10, IIB: 6*	62±11*	10*	1.3 yr
			O	24*	Sternotomy/thoracotomy		I: 9, IIA: 10, IIB: 5*	59±11*	7*	6.1 yr
Jun (26)	RC	2010–2012	R	55	Rt or L 3-port	Thymoma, MG, other	R: B1: 4, B2: 6, B3: 4, AB: 3, B1/2: 1, B2/3: 1, B1/2/3: 1, SCC: 1	41	25	NR
			VATS	60	NR			43	NR	
Kamel (27)	RPC-PM	2012–2016	R	70	Rt or L, 3- or 4-port	AMM, MG	I/II: 17, III/IVA: 5*	52 [34–61]	23	30 days
			R	22*	—	—	—	58 [50–67]*	8*	
			VATS	7	NR	—	—	55 [28–64]	5	
			O	12/22*	Transsternal		I/II: 15, III/IVA: 7*	59 [51–72]*	9*	
Kang (28)	RC-PM	2006–2015	R	100*	Rt or L or Bi, 3-port	AMM	I: 19, II: 21, III: 4, IV: 2	52±14*	48*	36 mo
			O	100*	Transsternal		I: 22, II: 28, III: 3, IV: 3	52±13*	51*	
Kneurtz (29)	RC-PM	2004–2016	R	20	3-port + assist port	Thymoma	I: 5, II: 9, III: 5, IVA: 1	59 [47–65]**	5	24 mo
			O	34	Transsternal		I: 11, II: 11, III: 9, IVA: 3	61 [47–73]**	14	
Marulli (30)	RC-PM	1982–2017	R	41*	Rt or L multiport	Thymoma	I: 8, II: 33*	58±11*	18*	28 mo [18–61]*, **
			O	41*	Median sternotomy		I: 9, II: 32*	58±10*	19*	88 mo [62–116]*, **
Qian (31)	RC	2009–2014	R	51	Rt or L, 3-port	Thymoma	I: 19, IIA: 21, IIB: 11	49±13	21	421±469 d
			VATS	35	Rt or L, 3-port		I: 10, IIA: 14, IIB: 11	50±13	19	701±382 d
			O	37	Median sternotomy		I: 10, IIA: 12, IIB: 15	47±14	15	818±592 d

Table 1 (*continued*)

Table 1 (*continued*)

Author	Study design	Time period	Study arms	Patient number	Operative approach	Indication	Thymoma pathology (n)/or MGFA classification (n)	Age, mean ± SD or median [range]	Gender (# male)	Length of follow up
Renaud (32)	RC	1998–2010	R	6	L, 3-port	MG	MGFA: I: 1, II: 1, III: 4	40 [27–57]	1	12 mo
			O	15	Median sternotomy		I: 0, II: 6, III: 13, IV: 2	28 [6–46]	6	
Rowse (33)	RPC	1995–2015	R	11	Rt or L or Bi, 3-port	Thymoma, MG, other	R/VATS: I: 18, IIA: 9	52 [23–74]	6	18 mo [1–79]
			VATS	45	Rt or L or Bi, 3-port			50 [23–87]	19	
Ruckert (34)	RPC	1994–2006	R	74	Rt or L, 3-port	MG	MGFA: I: 4, II: 25, III: 35, IV: 10	39 [7–75]	32	42 mo
			VATS	79	Rt or L, 3-port		I: 3, II: 26, III: 34, IV: 15, V: 1	37 [11–74]	23	
Seong (35)	RC-PM	2008–2012	R	34*	Rt or L or Bi, 3-port	AMM	NR	54±2*	15*	1.1±0.2 yr
			O	34*	Median sternotomy			52±2*	18*	1.8±0.2 yr
Weksler (36)	RC	2001–2010	R	15	Rt or L, 3-port	AMM	I: 3, II: 7, III: 0	57±16	7	30 days
			O	35	Transsternal		I: 3, II: 6, III: 2	51±18	18	
Wilshire (37)	RPC	2004–2014	R	23	Rt or L, 3-port	Thymoma	I: 9, II: 13, III: 1	58 [50–67]**	11	12 mo [8–34]**
			O	17	Sternotomy		I: 7, II: 8, III: 2	59 [52–69]**	12	83 mo [22–94]**
Ye 2013 (39)	RC	2009–2012	R	21	Rt or L, 3-port	Thymoma	I: 21	53±8	9	17 mo [6–48]
			VATS	25	Rt or L, 3-port		I: 25	53±5	13	25 mo [6–48]
Ye 2014 (38)	RC	2009–2012	R	23	Rt or L, 3-port	Thymoma	I: 21, II: 2	52±7	11	17 mo [1–48]
			O	51	Transsternal		I: 43, II: 8	50±13	31	18 mo [1–48]

PC, prospective cohort study; ITMIG-DB, international thymic malignancy interest group database study; RC, retrospective cohort study; RPC, retrospective review of a prospectively collected institutional database cohort study; PM, propensity matched; R, robotic; V or VATS, video-assisted thoracic surgery; O, open; Rt, right; L, left; Bi, bilateral; U, unmatched; M, matched; MI; minimally invasive; AMM, anterior mediastinal mass; MG, myasthenia gravis; NR, not recorded; NA, not applicable; d, days; mo, months; yr, years; SCC, squamous-cell carcinoma. WHO classification of thymomas is reported as type A, B1, B2, B3, AB, stage was preferentially reported. Stage and Myasthenia Gravis Foundation of America (MGFA) classification are reported with roman numerals (i.e., I, II, III, IV). Matched data were preferentially included (marked with *). Casiraghi matched: tumor size and stage; Kamel matched: age (q52 versus <52 years "median age of the whole cohort"), gender, pathologic size of tumor (p5 versus >5 cm), Masaoka-Koga stage (I/II versus III/IV), and World Health Organization (WHO)–Histologic subtypes (B3/C versus others); Kang matched: age, sex, size of tumor, clinical stage, Charlson-Deyo comorbidity index, cell type, and combined operation; Kneurtz matched: age, sex, tumor size, and Masaoka stages (grouped as I, II, III/IV); Marulli matched: age at surgery, Masaoka stage, World Health Organization (WHO) histological classification, tumour dimension and MG; Seong matched: age, size of the mass, gender, presence of MG, resection of additional organ/tissue and pathology. **, interquartile range.

Figure 2 Risk of bias assessment. Graph lists each paper by first author and year of publication and shows high (red circle with a minus sign), low (green circle with a plus sign), and unclear (yellow circle with a question mark) risk of: selection (systematic differences between groups in baseline characteristics, comparability of groups), performance (systematic differences between groups in the care that is provided), detection (systematic differences between groups in how outcomes are determined), attrition (systematic differences between groups in withdrawals from a study, completeness of sample, follow-up, or data), reporting (systematic differences between reported and unreported findings, selective reporting of results), and other (learning curve, conflicts of interest, funding) bias. Summarized criteria for low risk determinations were: selection bias: the cohorts were contemporary and comparable, matched for patient characteristics, or adjusted for confounding factors. Performance bias: the cohorts were matched on operative and/or hospital characteristics, such as surgical technique, care pathways, and length of follow up to make combining them reasonable, or differences were addressed. Detection bias: data capture and entry was standardized/performed by trained personnel and precise definitions of outcomes of interest were provided. Attrition bias: there was no missing data, or missing data not an issue, no (or few) patients were lost to follow up, and length of follow up comparable and sufficient. Reporting bias: all pre-specified outcomes of interest (and meta-analyses) were reported in the pre-specified way regardless of significance and complete enough for inclusion in a meta-analysis. Other bias: there were no funding or industry conflicts of interest that were deemed an issue, they authors accounted for experience/volume of surgeons and/or hospital. No other obvious bias.

Table 2 and forest plots are presented in *Figure 3*. Operative time (*Figure 3A*) was reported in all studies except one (23). Of those, 12 reported data in sufficient detail for meta-analysis (22,24,25,27-32,35,37,38), which showed equivalent operative time for robotic thymectomy *vs.* open surgery (WMD: 6.73; 95% CI: –21.20, 34.66, P=0.64) with significant heterogeneity (P<0.00001, I^2=97%), perhaps due to differences in definitions of operative time between studies. Eight studies reported on intraoperative blood loss (EBL) (27-29,31,32,36,37,39), of these, seven reported EBL in enough detail for pooled analysis (*Figure 3B*) (27-29,31,36,37,39), which showed significantly lower blood loss in the robotic group (WMD: –173.03; 95% CI: –305.90, –40.17, P=0.01), but with significant heterogeneity (P<0.00001, I^2=99%) due to large variance in the magnitude of the robotic benefit; all seven papers reported significantly lower EBL for the robotic group. Only one study reported on blood transfusions (R 0% *vs.* O 0%, ns) (38). Length of hospital stay (LOS) was reported in 13 publications (22,24,25,27-32,35-38). Pooled analysis showed a significantly shorter hospital stay (*Figure 3C*) in the robotic group (WMD: –2.78; 95% CI: –3.22, –2.33, P<0.00001), with significant heterogeneity (P=0.0002, I^2=67%), which appears to be due to variability in the degrees of benefit in the robotic group as all studies reported a shorter LOS in the robotic group. Intraoperative complications were reported by six studies (22,24,27,30-32) and the pooled analysis (*Figure 3D*) was not statistically different between groups (RD: –0.00; 95% CI: –0.05, 0.04, P=0.84) with no heterogeneity (P=1.0, I^2=0%). Post-operative complications were reported by 13 studies (22,24,25,27-32,35-38), with meta-analysis (*Figure 3E*) demonstrating a significantly lower postoperative complication rate in the robotic group (OR: 0.37; 95% CI: 0.22, 0.60, P<0.0001) with low heterogeneity (P=0.17, I^2=27%). Mortality was reported in ten studies (22,28-32,35-38) with a pooled analysis (*Figure 3F*) showing equivalent results between groups (RD: –0.00; 95% CI: –0.02, 0.02, P=0.86) and no heterogeneity (P=1.0, I^2=0%). Only one study reported a death and this was in their open cohort (36). There were no deaths in any of the robotic patients included. There were nine studies (23-25,28-30,36-38) that reported on margin status, with the robotic group showing a significantly decreased positive margin rate (*Figure 3G*) (RD –0.04; 95% CI: –0.07, –0.01, P=0.01) with low heterogeneity (P=0.6, I^2=0%). However, sensitivity analysis removing Burt *et al.* (23), which is a large international registry paper and was the only paper reporting a significantly lower rate for robotic surgery prior

to pooling, eliminates the significant difference (RD: –0.02; –0.05, 0.02, P=0.33, Chi^2 P=0.99, I^2=0%).

Robotic versus VATS thymectomy

A total of seven publications (23,26,27,31,33,34,39) inclusive of 994 patients (n=428 robotic and n=566 VATS) are included and results are detailed in *Table 3* and forest plots are presented in *Figure 4*. Operative time was reported in six studies (26,27,31,33,34,39), with pooled analysis (*Figure 4A*) showing no significant difference between groups (WMD: 8.99; 95% CI: –10.53, 28.51, P=0.37), but with high heterogeneity (P<0.0001, I^2=83%). Blood loss was reported by four studies (27,31,33,39), with meta-analysis (*Figure 4B*) showing no difference between groups (WMD: –9.35; 95% CI: –48.20, 29.51, P=0.64), but with high heterogeneity (P=0.005, I^2=77%). Blood transfusions were only reported in two studies (33,39), with no differences between groups (*Figure 4C*) (RD: –0.02; 95% CI: –0.11, 0.06, P=0.60) and low heterogeneity (P=0.62, I^2=0%). Length of hospital stay was reported by five studies (26,27,31,33,39), with a pooled analysis (*Figure 4D*) showing no difference between groups (WMD –0.81; 95% CI: –2.22, 0.59, P=0.26), but with high heterogeneity (P<0.00001, I^2=93%). Conversion to open was reported in both cohorts in four studies (31,33,34,39). Meta-analysis showed no difference between robotic and VATS (RD: –0.01; 95% CI: –0.04, 0.03, P=0.73), with low heterogeneity (P=0.91, I^2=0%). There was no significant difference in the pooled analysis (*Figure 4F*) of three studies (27,31,33) reporting intraoperative complication rates for both cohorts (OR: 0.74; 95% CI: 0.19, 2.85, P=0.66) or in the pooled analysis (*Figure 4G*) of the five studies (26,31,33,34,39) reporting postoperative complication rates (OR: 1.18; 95% CI: 0.48, 2.91, P=0.71). None of six studies (26,27,31,33,34,39) reported any mortalities in either robotic or VATS groups. Three of the studies reported mortality rates for both cohorts (31,33,34) and were included in the pooled analysis (*Figure 4H*). Margin rates were reported for both cohorts in three studies (23,33,39), with pooled analysis (*Figure 4I*) demonstrating no significant difference between the robotic and VATS cohorts (RD: 0.02; 95% CI: –0.02, 0.07, P=0.3) with low heterogeneity (P=0.78, I^2=0%).

Discussion

Statement of principal findings

We performed a systematic review and met-analysis of data

Table 2 Comparison of robotic *vs.* open thymectomy

Author	Study size	Operative time (min)	Blood loss (mL)	BTx (%)	Length of stay (days)	Conversion to open (%)	Intraoperative complication rate (%)	Postoperative complication rate (%)	Mortality (in-hospital or 30-day) (%)	R0 resection (%)
Balduyck (22)	Robotic 14	224.2±66.5	N/A	N/A	9.6±3.9	7.1	0	21.4	0	N/A
	Open 22	243.8±55.5	N/A	N/A	11.8±5.7	N/A	0	13.5	0	N/A
	P value	NS	N/A	N/A	NS	N/A	NS	NS	NS	N/A
Burt (23)	Robotic 146	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	92
	Open 2,028	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	86
	P value	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.2
Cakar (24)	Robotic 9	154 [94–312]	<50	N/A	5 [4–15]	0	0	11	N/A	100
	Open 10	110 [42–152]	<50	N/A	10 [10–23]	N/A	0	30	N/A	100
	P value	<0.05	ns	N/A	<0.05	N/A	NS	N/A	N/A	N/A
Casiraghi (25)	Robotic 24	117±40	N/A	–	4.0±1.9	0	N/A	16.7	N/A	100
	Open 24	141±46	N/A	N/A	5.9±1.7	N/A	N/A	12.5	N/A	100
	P value	0.06	N/A	N/A	0.00009	N/A	N/A	NS	N/A	N/A
Kamel (27) Unmatched	Robotic 70	102 [77–125]	20 [20–50]	N/A	3 [2–4]	7.1	4.2	11	0	N/A
	Open 12	105 [79–263]	125 [38–550]	N/A	5 [3–8]	N/A	17	0	N/A	N/A
	P value	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Kamel (27) Matched	Robotic 22	104 [71–121]	50 [20–70]	N/A	2 [2–4]	0	4.5	4.5	0	N/A
	Open 22	85 [71–137]	150 [100–300]	N/A	4 [3–6]	N/A	4.5	0	N/A	N/A
	P value	0.79	<0.001	N/A	0.01	NS	NS	NS	N/A	N/A
Kang (28) Unmatched	Robotic 117	164.3±96.0	103.6±107.4	N/A	2 [2–3]	0	N/A	1	0	100
	Open 312	191±102.1	473.5±684.3	N/A	5 [4–7]	N/A	N/A	14	0	95
	P value	0.01	<0.001	N/A	<0.001	N/A	N/A	<0.001	NS	0.03
Kang (28) Matched	Robotic 100	150.1±65.6	100.9±105.4	N/A	2.5±1.2	0	N/A	1	0	100
	Open 100	160.8±59.5	354.5±412.4	N/A	6.4±6.6	N/A	N/A	12	0	98
	P value	0.23	<0.001	N/A	<0.001	N/A	N/A	0.002	NS	0.16
Kneurtz (29)	Robotic 20	119 [92–142]	25 [10–94]	N/A	3 [2–5]	5	N/A	15	0	90
	Open 34	90 [75–124]	150 [75–275]	N/A	4 [3–6]	N/A	N/A	24	0	85
	P value	0.13	0.001	N/A	0.034	N/A	N/A	0.45	NS	NS

Table 2 (*continued*)

Table 2 (*continued*)

Author	Study size	Operative time (min)	Blood loss (mL)	BTx (%)	Length of stay (days)	Conversion to open (%)	Intraoperative complication rate (%)	Postoperative complication rate (%)	Mortality (in-hospital or 30-day) (%)	R0 resection (%)
Marulli (30) Unmatched	Robotic 56	132.5 [115–170]	≥50 mL in 5% ≥100 mL in 3% ≥200 mL in 3%	N/A	3 [3–4]	3.5	0	4	0	100
	Open 108	115 [90–137]	≥50 mL in 19% ≥100 mL in 16% ≥200 mL in 10%	N/A	6 [5–7]	N/A	0	15	0	100
	P value	<0.001	0.01, 0.02, 0.13	N/A	<0.001	N/A	ns	0.03	ns	N/A
Marulli (30) Matched	Robotic 41	125 [115–160]	≥50 mL in 7%	N/A	3 [3–4]	2.4	0	5	0	100
	Open 41	120 [95–135]	≥50 mL in 22%	N/A	6 [5–6]	N/A	0	7	0	100
	P value	ns	ns	N/A	0.04	N/A	ns	ns	ns	N/A
Qian (31)	Robotic 51	71.2±39.8	77.5±69.5	N/A	4.3±1.1	0	3.9	0	0	N/A
	Open 37	88.5±37.6	246±316.5	N/A	6.6±1.4	N/A	5.4	8.1	0	N/A
	P value	ns	N/A	N/A	<0.001	N/A	N/A	N/A	Ns	N/A
Renaud (32)	Robotic 6	189 [135–240]	<10	N/A	5 [4–6]	0	0	0	0	N/A
	Open 15	55 [45–80]	24.6 [10–200]	N/A	8.7 [3–16]	N/A	0	20	0	N/A
	P value	<0.001	ns	N/A	0.02	N/A	ns	ns	Ns	N/A
Seong (35)	Robotic 34	157.2±12.6	N/A	N/A	2.6±0.2	0	N/A	0	0	N/A
	Open 34	139.3±8.86	N/A	N/A	5.5±0.7	N/A	N/A	14.7	0	N/A
	P value	Ns	N/A	N/A	0.001	ns	N/A	Ns	Ns	N/A
Weksler (36)	Robotic 15	130±23	41.7±29.7	N/A	1 [1–7]	0	0	6.6	0	100
	Open 35	N/A	151.4±107.2	N/A	4 [2–27]	N/A	0	57	2.8	100
	P value	N/A	0.001	N/A	0.002	N/A	ns	0.001	N/A	N/A
Wilshire (37)	Robotic 23	181 [142–232]	50 [45–100]	N/A	2 [1–3]	4	N/A	26	0	100
	Open 17	157 [136–171]	100 [100–150]	N/A	5 [4–8]	N/A	N/A	41	0	100
	P value	ns	0.001	N/A	<0.001	N/A	N/A	Ns	Ns	N/A
Ye (38)	Robotic 23	97±38	61.3±21.8	0	3.7±1.1	0	N/A	4.3	0	100
	Open 51	214.5±35.4	466.1±91.4	0	11.6±10.4	N/A	N/A	3.9	0	100
	P value	<0.01	<0.01	ns	<0.01	N/A	N/A	ns	ns	N/A

BTx, blood transfusions; NS, not significant; N/A, not applicable/not reported. Continuous data are reported as mean ± SD, median (range) or (interquartile range).

184

O'Sullivan et al. Comparing robotic to VATS and open approaches for thymectomy

A

Study or Subgroup	Robot (OR Time) Mean [min]	SD [min]	Total	Open (OR Time) Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Balduyck 2011 (1)	224.2	66.5	14	243.8	55.5	22	7.4%	-19.60 [-61.45, 22.25]
Cakar 2007	154	54.5	9	110	27.5	10	7.6%	44.00 [4.52, 83.48]
Casiraghi 2018 (2)	117	40	24	141	46	24	8.3%	-24.00 [-48.39, 0.39]
Kamel 2017 (3)	104	37	22	85	48.9	22	8.3%	19.00 [-6.62, 44.62]
Kang 2016 (4)	150.1	65.6	100	160.8	59.5	100	8.6%	-10.70 [-28.06, 6.66]
Kneuertz 2017 (5)	119	37	20	90	36.3	34	8.5%	29.00 [8.71, 49.29]
Marulli 2018 (6)	125	33.3	41	120	29.6	41	8.7%	5.00 [-8.64, 18.64]
Qian 2017 (7)	71.2	39.8	51	88.5	37.6	37	8.6%	-17.30 [-33.61, -0.99]
Renaud 2013	189.2	26.3	6	55	8.75	15	8.5%	134.20 [112.70, 155.70]
Seong 2014 (8)	157.2	12.6	34	139.3	8.86	34	8.9%	17.90 [12.72, 23.08]
Wilshire 2016	181	66.7	23	157	25.9	17	8.1%	24.00 [-5.91, 53.91]
Ye 2014	97	38	23	214.5	35.4	51	8.6%	-117.50 [-135.82, -99.18]
Total (95% CI)			367			407	100.0%	6.73 [-21.20, 34.66]

Heterogeneity: Tau² = 2282.38; Chi² = 356.83, df = 11 (P < 0.00001); I² = 97%
Test for overall effect: Z = 0.47 (P = 0.64)

Mean Difference IV, Random, 95% CI [min]
-200 -100 0 100 200
Favours Robotic Favours Open

B

Study or Subgroup	Robot (EBL) Mean [mL]	SD [mL]	Total	Open (EBL) Mean [mL]	SD [mL]	Total	Weight	Mean Difference IV, Random, 95% CI [mL]
Kamel 2017 (1)	50	37	22	150	148.1	22	14.3%	-100.00 [-163.79, -36.21]
Kang 2016 (2)	100.9	105.4	100	354.5	412.4	100	13.9%	-253.60 [-337.03, -170.17]
Kneuertz 2017 (3)	25	62.2	20	150	148.1	34	14.4%	-125.00 [-181.76, -68.24]
Qian 2017	77.5	69.5	51	246	316.5	37	13.5%	-168.50 [-272.25, -64.75]
Weksler 2012	41.7	29.7	15	151.4	107.2	35	14.6%	-109.70 [-148.26, -71.14]
Wilshire 2016	50	40.7	23	100	37.04	17	14.7%	-50.00 [-74.22, -25.78]
Ye 2014	61.3	21.8	23	466.1	91.4	51	14.7%	-404.80 [-431.42, -378.18]
Total (95% CI)			254			296	100.0%	-173.03 [-305.90, -40.17]

Heterogeneity: Tau² = 31158.96; Chi² = 409.69, df = 6 (P < 0.00001); I² = 99%
Test for overall effect: Z = 2.55 (P = 0.01)

Mean Difference IV, Random, 95% CI [mL]
-500 -250 0 250 500
Favours Robotic Favours Open

C

Study or Subgroup	Robot (LOS) Mean [days]	SD [days]	Total	Open (LOS) Mean [days]	SD [days]	Total	Weight	Mean Difference IV, Random, 95% CI [days]
Balduyck 2011	9.6	3.9	14	11.8	5.7	22	1.8%	-2.20 [-5.34, 0.94]
Cakar 2007 (1)	5	2.75	9	10	3.25	10	2.4%	-5.00 [-7.70, -2.30]
Casiraghi 2018 (2)	4	1.9	24	5.9	1.7	24	9.3%	-1.90 [-2.92, -0.88]
Kamel 2017 (3)	2	1.48	22	4	2.22	22	8.5%	-2.00 [-3.11, -0.89]
Kang 2016 (4)	2.5	1.2	100	6.4	6.6	100	7.1%	-3.90 [-5.21, -2.59]
Kneuertz 2017 (5)	3	2.22	20	4	2.22	34	7.7%	-1.00 [-2.23, 0.23]
Marulli 2018 (6)	3	0.74	41	6	0.74	41	16.4%	-3.00 [-3.32, -2.68]
Qian 2017	4.3	1.1	51	6.6	1.4	37	14.2%	-2.30 [-2.84, -1.76]
Renaud 2013	5	0.5	6	8.7	3.25	15	5.0%	-3.70 [-5.39, -2.01]
Seong 2014 (7)	2.65	0.18	34	5.53	0.75	34	16.9%	-2.88 [-3.14, -2.62]
Weksler 2012	1	1.5	15	4	6.25	35	3.4%	-3.00 [-5.21, -0.79]
Wilshire 2016	2	1.48	23	5	3.2	17	5.3%	-3.00 [-4.64, -1.36]
Ye 2014	3.7	1.1	23	11.6	10.4	51	2.1%	-7.90 [-10.79, -5.01]
Total (95% CI)			382			442	100.0%	-2.78 [-3.22, -2.33]

Heterogeneity: Tau² = 0.29; Chi² = 36.84, df = 12 (P = 0.0002); I² = 67%
Test for overall effect: Z = 12.13 (P < 0.00001)

Mean Difference IV, Random, 95% CI [days]
-20 -10 0 10 20
Favours Robotic Favours Open

D

Study or Subgroup	Robot (Intraop Comps) Events	Total	Open (Intraop Comps) Events	Total	Weight	Risk Difference M-H, Fixed, 95% CI
Balduyck 2011	0	14	0	22	12.1%	0.00 [-0.11, 0.11]
Cakar 2007	0	9	0	10	6.7%	0.00 [-0.18, 0.18]
Kamel 2017 (1)	1	22	1	22	15.6%	0.00 [-0.12, 0.12]
Marulli 2018 (2)	0	41	0	41	29.1%	0.00 [-0.05, 0.05]
Qian 2017	2	51	2	37	30.4%	-0.01 [-0.11, 0.08]
Renaud 2013	0	6	0	15	6.1%	0.00 [-0.21, 0.21]
Total (95% CI)		143		147	100.0%	-0.00 [-0.05, 0.04]
Total events	3		3			

Heterogeneity: Chi² = 0.10, df = 5 (P = 1.00); I² = 0%
Test for overall effect: Z = 0.20 (P = 0.84)

Risk Difference M-H, Fixed, 95% CI
-0.5 -0.25 0 0.25 0.5
Favours Robotic Favours Open

E

Study or Subgroup	Robot (Postop Comps) Events	Total	Open (Postop comps) Events	Total	Weight	Odds Ratio M-H, Fixed, 95% CI
Balduyck 2011	3	14	3	22	3.2%	1.73 [0.30, 10.08]
Cakar 2007	1	9	3	10	4.4%	0.29 [0.02, 3.48]
Casiraghi 2018 (1)	4	24	3	24	4.4%	1.40 [0.28, 7.06]
Kamel 2017 (2)	1	22	0	22	0.8%	3.14 [0.12, 81.35]
Kang 2016 (3)	1	100	12	100	20.9%	0.07 [0.01, 0.58]
Kneuertz 2017 (4)	3	20	8	34	8.9%	0.57 [0.13, 2.47]
Marulli 2018 (5)	2	41	3	41	5.0%	0.65 [0.10, 4.11]
Qian 2017	0	51	3	37	7.0%	0.10 [0.00, 1.91]
Renaud 2013	0	6	3	15	3.5%	0.27 [0.01, 6.17]
Seong 2014 (6)	0	34	5	34	9.5%	0.08 [0.00, 1.47]
Weksler 2012 (7)	1	15	20	35	19.7%	0.05 [0.01, 0.45]
Wilshire 2016 (8)	6	23	7	17	10.5%	0.50 [0.13, 1.93]
Ye 2014	1	23	2	51	2.1%	1.11 [0.10, 12.94]
Total (95% CI)		**382**		**442**	**100.0%**	**0.37 [0.22, 0.60]**
Total events	23		72			

Heterogeneity: Chi² = 16.37, df = 12 (P = 0.17); I² = 27%
Test for overall effect: Z = 3.98 (P < 0.0001)

F

Study or Subgroup	Robot (Mortality) Events	Total	Open (Mortality) Events	Total	Weight	Risk Difference M-H, Fixed, 95% CI
Balduyck 2011 (1)	0	14	0	22	4.6%	0.00 [-0.11, 0.11]
Kang 2016 (2)	0	100	0	100	26.8%	0.00 [-0.02, 0.02]
Kneuertz 2017 (3)	0	20	0	34	6.7%	0.00 [-0.08, 0.08]
Marulli 2018 (4)	0	56	0	108	19.7%	0.00 [-0.03, 0.03]
Qian 2017	0	51	0	37	11.5%	0.00 [-0.04, 0.04]
Renaud 2013	0	6	0	15	2.3%	0.00 [-0.21, 0.21]
Seong 2014 (5)	0	34	0	34	9.1%	0.00 [-0.06, 0.06]
Weksler 2012	0	15	1	35	5.6%	-0.03 [-0.14, 0.08]
Wilshire 2016	0	23	0	17	5.2%	0.00 [-0.10, 0.10]
Ye 2014	0	23	0	51	8.5%	0.00 [-0.06, 0.06]
Total (95% CI)		**342**		**453**	**100.0%**	**-0.00 [-0.02, 0.02]**
Total events	0		1			

Heterogeneity: Chi² = 0.30, df = 9 (P = 1.00); I² = 0%
Test for overall effect: Z = 0.18 (P = 0.86)

G

Study or Subgroup	Robot (PSM) Events	Total	Open (PSM) Events	Total	Weight	Risk Difference M-H, Fixed, 95% CI
Burt 2017	12	146	280	2028	51.6%	-0.06 [-0.10, -0.01]
Cakar 2007	0	4	0	3	0.6%	0.00 [-0.42, 0.42]
Casiraghi 2018 (1)	0	24	0	24	4.5%	0.00 [-0.08, 0.08]
Kang 2016 (2)	0	100	2	100	19.0%	-0.02 [-0.05, 0.01]
Kneuertz 2017 (3)	2	20	5	34	4.8%	-0.05 [-0.22, 0.13]
Marulli 2018 (4)	0	41	0	41	7.8%	0.00 [-0.05, 0.05]
Weksler 2012	0	10	0	14	2.2%	0.00 [-0.15, 0.15]
Wilshire 2016	2	23	2	15	3.4%	-0.05 [-0.25, 0.16]
Ye 2014	0	23	0	51	6.0%	0.00 [-0.06, 0.06]
Total (95% CI)		**391**		**2310**	**100.0%**	**-0.04 [-0.07, -0.01]**
Total events	16		289			

Heterogeneity: Chi² = 6.38, df = 8 (P = 0.60); I² = 0%
Test for overall effect: Z = 2.50 (P = 0.01)

Figure 3 Robotic *vs.* open thymectomy forest plots. Forest plots showing comparisons between robotic and open cohorts for outcomes of interest. For operative time (A), [1] is operative room time, [2-6,8] are matched, and [7] does not include robotic set up time. For estimated blood loss (EBL) (B), [1-3] are matched. For length of hospital stay (LOS) (C), [1] LOS in Austria is prolonged due to less pressure from insurance companies, [2-7] are matched. For intraoperative complications (D), [1,2] are matched. For postoperative complications (E), [1-6] are matched, [7] does not specify if postoperative, [8] is the perioperative complication rate. For mortality (F), [1] is the intraoperative rate, [2-5] are matched. For positive margin rate (G), [1-4] are matched, n was based on the number of thymoma cases.

Table 3 Comparison of robotic *vs.* VATS thymectomy

Author	Study size	Operative time (min)	Blood loss (mL)	BTx (%)	Length of stay (days)	Conversion to open (%)	Intraoperative complication rate (%)	Postoperative complication rate (%)	Mortality (in-hospital or 30-day) (%)	R0 resection (%)
Burt (23)	Robotic 146	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	92
	VATS 315	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	95
	P value	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.2
Jun (26)	Robotic 55	139.8±45	N/A	N/A	7.18 [4–14]	0	0	10.9	0	N/A
	VATS 60	121±20.5	N/A	N/A	7.23	N/A	N/A	6.7	N/A	N/A
	P value	<0.05	N/A	N/A	ns	N/A	N/A	ns	ns	N/A
Kamel (27) Unmatched	Robotic 70	102 [77–125]	20 [20–50]	N/A	3 [2–4]	7.1	4	N/A	0	3
	VATS 7	101 [59–112]	10 [10–10]	N/A	3 [2–4]	N/A	0	N/A	N/A	N/A
	P value	N/A	N/A	N/A	N/A	N/A	N/A	N/A	ns	N/A
Qian (31)	Robotic 51	71.2±39.8	77.5±69.5	N/A	4.3±1.1	0	3.9	0	0	N/A
	VATS 35	79.1±41.0	127±115.2	N/A	5.5±1.2	0	5.7	0	0	N/A
	P value	ns	ns	N/A	<0.001	ns	ns	ns	ns	N/A
Rowse (33)	Robotic 11	178±53	160±205	0	2.1 [1–5]	0	9	9	0	100
	VATS 45	102±39	65±41	0	1.5 [1–4]	0	11	15.6	0	100
	P value	0.001	0.042	ns	0.05	ns	N/A	N/A	ns	N/A
Ruckert (34)	Robotic 74	187±48	N/A	N/A	N/A	1.4	N/A	2.7	0	N/A
	VATS 79	198±48	N/A	N/A	N/A	1.3	N/A	2.5	0	N/A
	P value	N/A	N/A	N/A	N/A	N/A	N/A	ns	ns	N/A
Ye (39)	Robotic 21	96.2±39.8	58.6±20.6	0	3.7±1.1	0	N/A	4.7	0	100
	VATS 25	103.6±36	86.8±97.1	4	6.7±1.4	4	N/A	4	N/A	100
	P value	ns	ns	N/A	<0.01	N/A	ns	N/A	ns	N/A

BTx, blood transfusions; ns, not significant; N/A, not applicable/not reported. Continuous data are reported as mean ± SD, median (range) or (interquartile range).

A

Study or Subgroup	Robot (OR Time) Mean [min]	SD [min]	Total	VATS (OR Time) Mean [min]	SD [min]	Total	Weight	Mean Difference IV, Random, 95% CI [min]
Jun 2014 (1)	139.8	45.3	55	121.07	20.6	60	19.3%	18.73 [5.67, 31.79]
Kamel 2017	102	35.6	70	101	39.3	7	14.0%	1.00 [-29.28, 31.28]
Qian 2017 (2)	71.2	39.8	51	79.1	41	35	18.1%	-7.90 [-25.33, 9.53]
Rowse 2015 (3)	178	53	11	102	39	45	13.1%	76.00 [42.67, 109.33]
Ruckert 2011	187	48	74	198	48	79	18.8%	-11.00 [-26.22, 4.22]
Ye 2013	96.2	39.8	21	103.6	36	25	16.7%	-7.40 [-29.51, 14.71]
Total (95% CI)			**282**			**251**	**100.0%**	**8.99 [-10.53, 28.51]**

Heterogeneity: Tau² = 468.32; Chi² = 29.30, df = 5 (P < 0.0001); I² = 83%
Test for overall effect: Z = 0.90 (P = 0.37)

Favours Robotic Favours VATS

B

Study or Subgroup	Robot (EBL) Mean [mL]	SD [mL]	Total	VATS (EBL) Mean [mL]	SD [mL]	Total	Weight	Mean Difference IV, Random, 95% CI [mL]
Kamel 2017 (1)	20	22.2	70	10	0.001	7	38.1%	10.00 [4.80, 15.20]
Qian 2017	77.5	69.5	51	127.1	115.2	35	26.2%	-49.60 [-92.27, -6.93]
Rowse 2015	160	205	11	65	41	45	8.0%	95.00 [-26.74, 216.74]
Ye 2013	58.6	20.6	21	86.8	97.1	25	27.6%	-28.20 [-67.27, 10.87]
Total (95% CI)			**153**			**112**	**100.0%**	**-9.35 [-48.20, 29.51]**

Heterogeneity: Tau² = 1024.47; Chi² = 12.77, df = 3 (P = 0.005); I² = 77%
Test for overall effect: Z = 0.47 (P = 0.64)

Favours Robotic Favours VATS

C

Study or Subgroup	Robot (BTx) Events	Total	VATS (BTx) Events	Total	Weight	Risk Difference M-H, Fixed, 95% CI
Rowse 2015	0	11	0	45	43.6%	0.00 [-0.12, 0.12]
Ye 2013	0	21	1	25	56.4%	-0.04 [-0.15, 0.07]
Total (95% CI)		**32**		**70**	**100.0%**	**-0.02 [-0.11, 0.06]**
Total events	0		1			

Heterogeneity: Chi² = 0.24, df = 1 (P = 0.62); I² = 0%
Test for overall effect: Z = 0.52 (P = 0.60)

Favours Robotic Favours VATS

D

Study or Subgroup	Robot (LOS) Mean [days]	SD [days]	Total	VATS (LOS) Mean [days]	SD [days]	Total	Weight	Mean Difference IV, Random, 95% CI [days]
Jun 2014 (1)	7.18	2.5	55	7.23	11	60	11.9%	-0.05 [-2.91, 2.81]
Kamel 2017	3	1.48	70	3	1.48	7	20.4%	0.00 [-1.15, 1.15]
Qian 2017	4.3	1.1	51	5.5	1.2	35	22.9%	-1.20 [-1.70, -0.70]
Rowse 2015	2.1	1	11	1.5	0.75	45	22.5%	0.60 [-0.03, 1.23]
Ye 2013	3.7	1.1	21	6.7	1.4	25	22.2%	-3.00 [-3.72, -2.28]
Total (95% CI)			**208**			**172**	**100.0%**	**-0.81 [-2.22, 0.59]**

Heterogeneity: Tau² = 2.18; Chi² = 58.08, df = 4 (P < 0.00001); I² = 93%
Test for overall effect: Z = 1.13 (P = 0.26)

Favours Robotic Favours VATS

E

Study or Subgroup	Robot (Convert) Events	Total	VATS (Convert) Events	Total	Weight	Risk Difference M-H, Fixed, 95% CI
Qian 2017	0	51	0	35	26.2%	0.00 [-0.05, 0.05]
Rowse 2015	0	11	0	45	11.2%	0.00 [-0.12, 0.12]
Ruckert 2011	1	74	1	79	48.2%	0.00 [-0.04, 0.04]
Ye 2013	0	21	1	25	14.4%	-0.04 [-0.15, 0.07]
Total (95% CI)		**157**		**184**	**100.0%**	**-0.01 [-0.04, 0.03]**
Total events	1		2			

Heterogeneity: Chi² = 0.56, df = 3 (P = 0.91); I² = 0%
Test for overall effect: Z = 0.34 (P = 0.73)

Favours Robotic Favours VATS

Figure 4 Robotic *vs.* VATS thymectomy forest plots. Forest plots showing comparisons between robotic and VATS cohorts for outcomes of interest. For operative time (A), [1] included robotic set up time, [2] did not include robotic set up time, [3] authors only completed 11 robotic cases in 20 years. For estimated blood loss (B), [1] was reported as median (interquartile range) and standard deviation was calculated as zero and estimated at 0.001 to allow for calculation of a mean difference and 95% confidence interval. Blood transfusions were reported in two studies (C). For length of hospital stay (D), [1] standard deviation was extrapolated from Figure 3 in paper. For conversions to open (E), only papers reporting both a robotic and a VATS rate were included in the analysis. Intraoperative complications (F), postoperative complications (G), mortality (H), and positive surgical margin rate (I) were also reported.

from 18 papers comparing robotic thymectomy to VATS or to open thymectomy and found that robotic surgery resulted in significantly reduced blood loss, postoperative complication rate, length of hospital-stay, and positive margin rate as compared to open thymectomy. Robotic surgery was comparable to VATS for all outcomes.

Comparison to prior systematic reviews

To our knowledge, there have been five prior systematic reviews comparing robotic thymectomy with VATS (40-42) or open (7,42,43) thymectomy. None compare robotics to both VATS and open, and none include all of the papers included in our analysis. The paper by Giuoutsos et al. (42) is a descriptive narrative of single cohort and comparison thoracic papers. The authors concluded that robotic surgery has been shown to be feasible and safe and is less expensive than open surgery due to a shorter length of hospital stay, but is more expensive than VATS, due to the purchase cost of the robotic system. The review by Friedant et al. (7) combined robotic and VATS cohorts in their comparison to open surgery and found decreased blood loss and length of hospital stay in the minimally invasive group. The paper by Fok et al. (41) compared robotic to VATS and found equivalent conversions, mortality, length of hospital-stay, and pneumonia rates. They also reported a shorter operative time for VATS, but that appears to be because they accidentally switched VATS and robotic operative times for the paper by Ruckert et al. 2011 (34) in their analysis; Ruckert et al. reported a longer operative time in the VATS group (187±48 robotic vs. 198±48 min VATS). Correcting this error would result in no difference in operative time between groups. There are two systematic reviews by Buentzel et al., one comparing robotics to VATS (40) and one comparing robotics to open (43). The authors reported no differences in operative time, blood loss, or postoperative complications in their comparisons to either VATS or open but reported a significantly shorter hospital stay for robotics vs. open thymectomy.

Robotic versus open thymectomy
Operative times reported in the literature to date tend to be longer in the robotic cohorts. This trend is mirrored when examining robotics across other specialties (44). This is possibly due to the more complex operative set up required for robotic surgery, but has also been demonstrated to be as a result of the inevitable learning curve that comes with the use of a new technique. We found no statistically significant difference

in operative time between the robotic and open groups, but there was high heterogeneity. There is inconsistency in how operative duration is defined across studies, with some authors choosing to include robotic set up and anaesthetic time, and some opting to report skin-to-skin durations.

Kamel et al. (27) have reported that approximately 15–20 cases are required to overcome the robotic learning curve and that operative times reduce significantly comparing early versus late operative experience in robotics. This could also influence the heterogeneity we observed. Cerfolio et al. reviewed their experience starting a thoracic robotics program and described that it requires training for the entire operating room staff and steep learning with respect to port placement, use of proper instrumentation and use of the correct robotic arms to overcome this initial period (45). Further supporting this, Ro et al. also describe significant reductions in operating room and operative time with improved experience comparing early versus late operative experience in robotic thymectomy (46). It seems therefore likely that overall operative times will continue to reduce as experience with robotic surgery increases. There should, however, be agreement on the appropriate standardized reporting of operative durations to allow comparability.

Although blood loss was lower in the robotic cohorts, the clinical significance of a blood loss difference of 173 mL is unclear. Length of hospital stay was significantly shorter for patients undergoing robotic thymectomy. This offers a clear advantage over the open approach and is a proxy marker of reduced post-operative pain and patient recovery burden with the robotic approach and avoidance of sternotomy (47). Sternotomy itself is associated with a number of complications including sternal wound infection, instability, dehiscence, mediastinitis and osteomyelitis (48). Its avoidance with the robotic approach is a clear advantage.

Conversion to open was reported by a number of authors. Baldyuck et al. (22) and Kneurtz et al. (29) both reported one case for invasion into the innominate vein. Kamel et al. (27) report five conversions; for invasion into the innominate or pleural adhesions. Similarly, Marulli et al. (30) report conversion in two patients because of large dimensions of the specimen and suspicious pericardial infiltration and Wilshire et al. (37) one because of difficulty delineating tumour margins. Examining the principles of thymic resection previously cited, in all cases the conversions reported were justified and legitimate. No cases of conversion due to uncontrollable bleeding were described in the papers included in this study.

It is not surprising that we reported a significantly lower

postoperative complication rate in the robotic group; minimally invasive approaches tend to have lower morbidity than an open approach. However, variability in the types of open approaches, the method of recording, and of reporting complications, make it difficult to form further conclusions without further study. Mortality rates between groups were comparable. We saw a significantly lower positive margin rate in the robotic group; however, this result was driven by a single study (23), and appears to be the result of selection bias, where patients in that study with earlier stage, smaller tumors were chosen preferentially for a minimally invasive approach. Histological type, stage, and tumor size were all different between minimally invasive thymectomy (MIT) (robotic and VATS combined) and open thymectomy. After propensity score matching the MIT and open groups, the authors reported that the patient characteristic and margin rate data were both no longer significantly different. Matching was not available for the separately reported robotic data. Our results do suggest that robotic thymectomy is not inferior to open thymectomy for short-term oncological outcomes.

Robotic versus VATS thymectomy

Whilst there are some clear advantages to robotic versus open approaches, there were no significant differences between robotic and VATS from studies examined. This review identified a low number of studies comparing these approaches and none were randomized or reported propensity-matched analyses. For operative time, the one outlier paper by Rowse et al. (33) reported a significantly shorter operative time for the VATS group; however, this study included their robotic learning curve. The surgical approach was chosen based on surgeon preference and technical ability, and only 11 robotic cases were performed over 20 years, with 4 times more VATS performed over the same time-period. This paper was also an outlier for estimated blood loss and length of hospital stay. Those three outcomes (operative time, EBL, LOS) are the only pooled analyses with significant heterogeneity, suggesting that currently robotic and VATS approaches are comparable. It is possible that this will change as surgeons gain experience with the robotic approach, as two out of the seven studies comparing robotic and VATS were at a high risk of bias due to learning curve. Also, because these two techniques are both minimally invasive, it is also possible that the studies included in this review were underpowered to detect small differences between groups. Variance was high in many of the studies for many of the outcomes

and it is possible that large sample size studies would be needed to detect differences between groups. However, the clinical significance of small differences may be negligible, especially when the additional cost of the robotic system is taken into account (39). The development of alternative incision locations and fewer incisions could also impact this comparison. None of the studies included in this review utilized a subxiphoid or single-port technique. The one study that did utilize these techniques by Suda et al. (12) was excluded because the single-port technique was not utilized for both the robotic and VATS cohorts. The robotic approach does offer a number of advantages for the surgeon over the VATS approach such as increased dexterity, maneuverability and articulation with vastly superior freedom to navigate anatomical structures (26). It has also been reported that robotics is associated with a shorter learning curve than VATS, although this is possibly due to the benefit of most surgeons having VATS experience as they embark upon their robotic learning curve (26).

Study limitations

No study included in this review was randomized to surgical approach. With a lack of randomization in surgical approach there is an inherent risk of bias. Many papers were deemed at high risk for selection bias, reporting bias, and bias due to unequal surgeon experience levels between groups; however, publication bias did not seem to be an issue. There was limited data regarding the robotic vs. VATS comparison and not all incision locations or techniques were represented. The majority of studies included in this review did not perform propensity matching and also included a heterogeneous group of patients in terms of indication. The long-term oncological outcomes associated with thymectomy for thymoma and anterior mediastinal tumours were not examined here due to limitations in the studies identified. This would be of benefit when such an analysis is possible. A full assessment of patient pain scores and quality of life is also required to determine the optimal surgical approach for this procedure, as this was not reported by any of the studies examined. The details of complications, such as individual complication rates were not assessed due to limited data and heterogeneity in reporting. Full cost benefit analysis of the robotic approach is also required. It must be highlighted that of the 17 studies reporting operative time, only eight reported their definition of operative time. The majority (7/8) reported it as surgical operative time [four skin-to-skin (25,27,30,36), two robotic docking and procedure

time (26,35), one procedure time not including robotic docking (31)], and Balduyck (22) reported it as total operating room occupation. This could have contributed to why the meta-analysis of operative time showed significant heterogeneity in both the robotic *vs.* VATS (P<0.0001, I^2=82%) and the robotic *vs.* open (P<0.00001, I^2=97%) analyses. For the most part, the operative details were similar between papers that is most robotic and VATS papers used a 3-port technique and most open papers utilized a median sternotomy. None of the included papers reported on the use of a robotic or VATS subxiphoid approach, which provides a surgical field comparable to that in a median sternotomy. While this decreases the heterogeneity between groups, it also points to lack of complete representation of the available approaches to thymectomy.

Conclusions

Examining available evidence; robotic thymectomy offers several advantages over the open approach, including reduced hospital stay and blood loss. Although complication and mortality rates are comparable when compared to VATS, robotic thymectomy is likely to offer the surgeon technical advantages such as autonomous control of a tridimensional camera, an in-built tremor filter and Endowrist instrumentation. Capital investment with annual maintenance costs remain a principal drawback with further economic analysis needed to determine long-term cost implications of RATS *vs.* VATS. A randomized controlled study has not yet been undertaken to compare approaches before any definitive conclusions can be drawn.

Acknowledgements

None.

Footnote

References

1. Maurizi G, D'Andrilli A, Sommella L, et al. Transsternal thymectomy. Thorac Cardiovasc Surg 2015;63:178-86.
2. Shrager JB, Nathan D, Brinster CJ, et al. Outcomes after 151 extended transcervical thymectomies for myasthenia gravis. Ann Thorac Surg 2006;82:1863-9.
3. Coosemans W, Lerut TE, Van Raemdonck DE. Thoracoscopic surgery: the Belgian experience. Ann Thorac Surg 1993;56:721-30.
4. Yim AP, Low JM, Ng SK, et al. Video-assisted thoracoscopic surgery in the paediatric population. J Paediatr Child Health 1995;31:192-6.
5. Ye B, Tantai JC, Ge XX, et al. Surgical techniques for early-stage thymoma: video-assisted thoracoscopic thymectomy versus transsternal thymectomy. J Thorac Cardiovasc Surg 2014;147:1599-603.
6. Ashton RC Jr, McGinnis KM, Connery CP, et al. Totally endoscopic robotic thymectomy for myasthenia gravis. Ann Thorac Surg 2003;75:569-71.
7. Friedant AJ, Handorf EA, Su S, et al. Minimally Invasive versus Open Thymectomy for Thymic Malignancies: Systematic Review and Meta-Analysis. J Thorac Oncol 2016;11:30-8.
8. Ruffini E, Filosso PL, Guerrera F, et al. Optimal surgical approach to thymic malignancies: New trends challenging old dogmas. Lung Cancer 2018;118:161-70.
9. Batirel HF. Minimally invasive techniques in thymic surgery: a worldwide perspective. J Vis Surg 2018;4:7.
10. Toker A, Sonett J, Zielinski M, et al. Standard terms, definitions, and policies for minimally invasive resection of thymoma. J Thorac Oncol 2011;6:S1739-42.
11. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol 2009;62:e1-34.
12. Suda T, Kaneda S, Hachimaru A, et al. Thymectomy via a subxiphoid approach: single-port and robot-assisted. J Thorac Dis 2016;8:S265-71.
13. Fang W, Yao X, Antonicelli A, et al. Comparison of surgical approach and extent of resection for Masaoka-Koga Stage I and II thymic tumours in Europe, North America and Asia: an International Thymic Malignancy Interest Group retrospective database analysis. Eur J Cardiothorac Surg 2017;52:26-32.
14. Jurado J, Javidfar J, Newmark A, et al. Minimally invasive thymectomy and open thymectomy: outcome analysis of 263 patients. Ann Thorac Surg 2012;94:974-81; discussion 981-2.
15. Orsini B, Santelmo N, Pages PB, et al. Comparative study for surgical management of thymectomy for non-thymomatous myasthenia gravis from the French

national database EPITHOR. Eur J Cardiothorac Surg 2016;50:418-22.

16. Tomulescu V. Thoracoscopic approach for tumors of anterior mediastinum. Acta Endocrinologica 2008;4:309-19.

17. Tomulescu V. Minimally invasive approach for tumoral and non tumoral myasthenia gravis; from thoracoscopic to robotic surgery. Ann Fundeni Hosp 2011;16:78-86.

18. Youssef SJ, Louie BE, Farivar AS, et al. Comparison of open and minimally invasive thymectomies at a single institution. Am J Surg 2010;199:589-93.

19. Turchetti G, Pierotti F, Palla I, et al. Comparative health technology assessment of robotic-assisted, direct manual laparoscopic and open surgery: a prospective study. Surg Endosc 2017;31:543-51.

20. Barbash GI, Glied SA. New technology and health care costs--the case of robot-assisted surgery. N Engl J Med 2010;363:701-4.

21. Kajiwara N, Barron JP, Kato Y, et al. Cost-benefit performance of robotic surgery compared with video-assisted thoracoscopic surgery under the Japanese national health insurance system. Ann Thorac Cardiovasc Surg 2015;21:95-101.

22. Balduyck B, Hendriks JM, Lauwers P, et al. Quality of life after anterior mediastinal mass resection: a prospective study comparing open with robotic-assisted thoracoscopic resection. Eur J Cardiothorac Surg 2011;39:543-8.

23. Burt BM, Yao X, Shrager J, et al. Determinants of Complete Resection of Thymoma by Minimally Invasive and Open Thymectomy: Analysis of an International Registry. J Thorac Oncol 2017;12:129-36.

24. Cakar F, Werner P, Augustin F, et al. A comparison of outcomes after robotic open extended thymectomy for myasthenia gravis. Eur J Cardiothorac Surg 2007;31:501-4; discussion 504-5.

25. Casiraghi M, Galetta D, Borri A, et al. Robotic-assisted thymectomy for early-stage thymoma: a propensity-score matched analysis. J Robot Surg 2018;12:719-24.

26. Jun Y, Hao L, Demin L, et al. Da Vinci robot-assisted system for thymectomy: experience of 55 patients in China. Int J Med Robot 2014;10:294-9.

27. Kamel MK, Rahouma M, Stiles BM, et al. Robotic Thymectomy: Learning Curve and Associated Perioperative Outcomes. J Laparoendosc Adv Surg Tech A 2017;27:685-90.

28. Kang CH, Hwang Y, Lee HJ, et al. Robotic Thymectomy in Anterior Mediastinal Mass: Propensity Score Matching Study With Transsternal Thymectomy. Ann Thorac Surg 2016;102:895-901.

29. Kneuertz PJ, Kamel MK, Stiles BM, et al. Robotic Thymectomy Is Feasible for Large Thymomas: A Propensity-Matched Comparison. Ann Thorac Surg 2017;104:1673-8.

30. Marulli G, Comacchio GM, Schiavon M, et al. Comparing robotic and trans-sternal thymectomy for early-stage thymoma: a propensity score-matching study. Eur J Cardiothorac Surg 2018;54:579-84.

31. Qian L, Chen X, Huang J, et al. A comparison of three approaches for the treatment of early-stage thymomas: robot-assisted thoracic surgery, video-assisted thoracic surgery, and median sternotomy. J Thorac Dis 2017;9:1997-2005.

32. Renaud S, Santelmo N, Renaud M, et al. Robotic-assisted thymectomy with Da Vinci II versus sternotomy in the surgical treatment of non-thymomatous myasthenia gravis: early results. Rev Neurol (Paris) 2013;169:30-6.

33. Rowse PG, Roden AC, Corl FM, et al. Minimally invasive thymectomy: the Mayo Clinic experience. Ann Cardiothorac Surg 2015;4:519-26.

34. Rückert JC, Swierzy M, Ismail M. Comparison of robotic and nonrobotic thoracoscopic thymectomy: A cohort study. J Thorac Cardiovasc Surg 2011;141:673-7.

35. Seong YW, Kang CH, Choi JW, et al. Early clinical outcomes of robot-assisted surgery for anterior mediastinal mass: its superiority over a conventional sternotomy approach evaluated by propensity score matching. Eur J Cardiothorac Surg 2014;45:e68-73; discussion e73.

36. Weksler B, Tavares J, Newhook TE, et al. Robot-assisted thymectomy is superior to transsternal thymectomy. Surg Endosc 2012;26:261-6.

37. Wilshire CL, Vallieres E, Shultz D, et al. Robotic Resection of 3 cm and Larger Thymomas Is Associated With Low Perioperative Morbidity and Mortality. Innovations (Phila) 2016;11:321-6.

38. Ye B, Li W, Ge XX, et al. Surgical treatment of early-stage thymomas: robot-assisted thoracoscopic surgery versus transsternal thymectomy. Surg Endosc 2014;28:122-6.

39. Ye B, Tantai JC, Li W, et al. Video-assisted thoracoscopic surgery versus robotic-assisted thoracoscopic surgery in the surgical treatment of Masaoka stage I thymoma. World J Surg Oncol 2013;11:157.

40. Buentzel J, Heinz J, Hinterthaner M, et al. Robotic versus thoracoscopic thymectomy: The current evidence. Int J Med Robot 2017;13.

41. Fok M, Bashir M, Harky A, et al. Video-Assisted Thoracoscopic Versus Robotic-Assisted Thoracoscopic Thymectomy: Systematic Review and Meta-analysis.

Innovations (Phila) 2017;12:259-64.

42. Gioutsos K, Kocher GJ, Schmid RA. Robotics in pulmonology and thoracic surgery: what, why and when? Panminerva Med 2016;58:318-28.

43. Buentzel J, Straube C, Heinz J, et al. Thymectomy via open surgery or robotic video assisted thoracic surgery: Can a recommendation already be made? Medicine (Baltimore) 2017;96:e7161.

44. Agrusa A, Romano G, Navarra G, et al. Innovation in endocrine surgery: robotic versus laparoscopic adrenalectomy. Meta-analysis and systematic literature review. Oncotarget 2017;8:102392-400.

45. Cerfolio RJ, Bryant AS, Minnich DJ. Starting a robotic program in general thoracic surgery: why, how, and lessons learned. Ann Thorac Surg 2011;91:1729-36; discussion 1736-7.

46. Ro CY, Derose JJ Jr, Connery CP, et al. Three-year experience with totally endoscopic robotic thymectomy. Innovations (Phila) 2006;1:111-4.

47. Ismail M, Swierzy M, Ruckert JC. State of the art of robotic thymectomy. World J Surg 2013;37:2740-6.

48. Reser D, Caliskan E, Tolboom H, et al. Median sternotomy. Multimed Man Cardiothorac Surg 2015;2015. doi: 10.1093/mmcts/mmv017
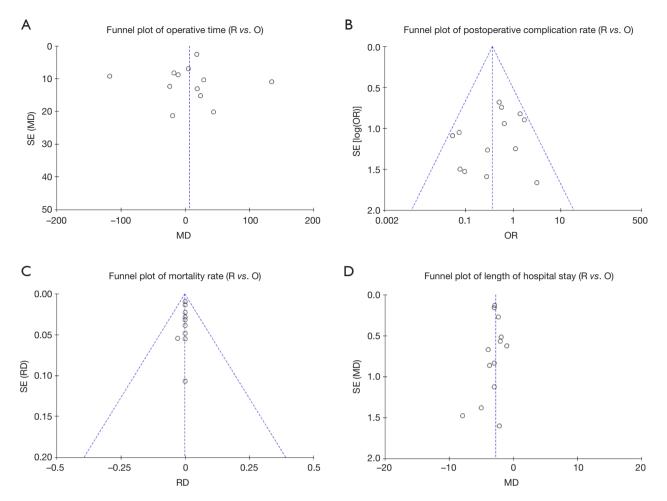
Figure S1 Funnel plots. Graphs showing funnel plots for any outcome with at least ten studies, including operative time for robotic *vs.* open (A), postoperative complication rate for robotic *vs.* open (B), mortality rate for robotic *vs.* open (C), and length of hospital stay for robotic *vs.* open (D).