Meta analysis of robot-assisted versus conventional laparoscopic fundoplication in children

Thomas P. Cundy a,b,⁎, Leanne Harling b, Hani J. Marcus a,b, Thanos Athanasiou b, Ara W. Darzi a,b

A R T I C L E  I N F O
Article history:
Received 16 September 2013
Received in revised form 30 November 2013
Accepted 6 December 2013

Key words:
Robotic
Robot-assisted
Laparoscopy
Gastro-esophageal reflux
Fundoplication
Children

A B S T R A C T

Background: Minimally invasive fundoplication may be performed using either a robot-assisted (RF) or conventional laparoscopic (LF) technique. Evidence comparing RF and LF in children remains unclear. This study aims to elucidate the comparative safety and efficacy of RF versus LF by systematic review and meta-analysis.

Methods: Comparative studies investigating RF versus LF in children were identified from multiple electronic literature databases. Meta-analysis was performed using random effects modeling. Safety parameters investigated were post-operative morbidity and intra-operative conversions. Efficacy outcomes of interest were operative success, re-operation, post-operative complications, length of hospital stay (LOS), total operating time (OT), analgesia requirement, and cost.

Results: Six observational studies met inclusion criteria, reporting outcomes of 297 children. No randomized controlled trials were identified. Pooled analysis determined no statistically significant differences between RF and LF for conversions, OT, LOS, and post-operative complications. There was no standardized follow up beyond the early post-operative period to enable data synthesis for remaining outcomes of interest. Limited evidence indicates higher costs with RF.

Conclusions: Safety and short-term efficacy seem comparable between RF and LF in children. There is insufficient evidence to assess comparative effectiveness for many important procedure specific outcome measures. Higher quality and longer follow-up studies are required.

© 2014 Elsevier Inc. All rights reserved.

Fundoplication is a high volume surgical procedure in the pediatric patient population. Indications are assorted and invariably related to symptoms or signs of gastro-esophageal reflux disease (GERD) [1]. Following the advent of endoscopic surgery, a minimally invasive approach to this anti-reflux procedure is now increasingly favored as standard of care [1].

Initial reports of laparoscopic anti-reflux surgery in infants and children were published in the early 1990’s [2,3]. Almost ten years thereafter, robot-assisted minimally invasive surgery was first described in these age groups [4,5], and fundoplication remains the most prevalently reported application of this technology in pediatric general surgery [6]. Robotic technology offers putative patient benefits through a range of features that are felt to enhance the surgeon’s ability to undertake minimally invasive surgery (MIS).

The role of robotic surgery in this setting remains unclear, generating a growing sentiment of polarized opinion amongst the surgical community, which is without a well-defined evidence base. The aim of this study is to critically appraise the literature comparing robot-assisted versus conventional laparoscopic MIS for fundoplication in order to further elucidate the comparative safety and efficacy of these techniques.

1. Methods

The study protocol was registered on the PROSPERO international prospective database of systematic reviews (CRD42013003971). Analysis was performed in accordance with recommendations outlined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement [7].

1.1. Search strategy

Systematic literature searches were undertaken of PubMed and EMBASE electronic databases using the following search strategy (“Surgery, Computer-Assisted”[MeSH] OR “robotics”[MeSH] OR “da Vinci” OR “telerobotic” OR “telesurgery” OR “robotic surgery”) AND (“Pediatrics”[MeSH] OR “Infant”[MeSH] OR “Child”[MeSH] OR “Adolescent”[MeSH]) AND (“Fundoplication”[MeSH] OR “Gastroesophageal reflux”[MeSH]). The search period was defined as June 2001 to June 2013 inclusively. The primary search was supplemented with searches of 1) PubMed related articles feature, 2) clinicaltrials.gov registry using the keyword “fundoplication”, and 3) abstracts of the
International Pediatric Endosurgery Group annual congress from 2002 to 2013.

1.2. Inclusion and exclusion criteria

All included studies satisfied the following criteria: 1) comparing robot-assisted (RF) versus conventional laparoscopic fundoplication (LF), 2) involving pediatric patients with mean or median study group ages < 18 years, 3) reporting ≥ 5 patients in each study group, and 4) investigating either objective clinical outcome measures or GERD symptoms via standardized questionnaires. No language restrictions were imposed. In the event that duplication of data was observed, more recent studies or those with larger sample sizes were preferentially considered, with subsequent exclusion of earlier, smaller studies.

1.3. Outcomes of interest

Primary outcome measures of interest were intra-operative conversions, intra-operative complications, length of hospital stay (LOS), post-operative complications, operating time (OT), analgesia requirement and cost. Operating time was regarded as the ‘total’ time from first skin incision to skin closure. Secondary outcomes of interest were operative success, requirement for re-operation (i.e. due to wrap failure, post-operative symptoms related to surgery) and post-operative morbidity (defined as dysphagia, retching, belching). Operative success was regarded as improvement or resolution of surgical impairment, proportion of re-do cases, proportion of cases undergoing concomitant gastrostomy at the time of surgery.

Table 1
Characteristics of studies included in the meta-analysis.

<table>
<thead>
<tr>
<th>Study design</th>
<th>Study period</th>
<th>n</th>
<th>Fundoplication</th>
<th>Min months follow up [mean (range)]</th>
<th>Surgeons</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
<td>LF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ivascu 2004 [16]</td>
<td>RCC</td>
<td>2002–2003</td>
<td>17</td>
<td>34</td>
<td>Nissen</td>
</tr>
<tr>
<td>Lehner 2006 [14]</td>
<td>PC</td>
<td>2001–2003</td>
<td>10</td>
<td>10</td>
<td>Thal</td>
</tr>
<tr>
<td>Antao 2010 [15]</td>
<td>RC</td>
<td>1999–2009</td>
<td>19</td>
<td>33</td>
<td>Nissen</td>
</tr>
</tbody>
</table>

*“-” not reported. RCC = retrospective case-control study, PC = prospective cohort study, RC = retrospective cohort study.

Statistical analysis was conducted using Review Manager® Version 5.1.7 for Windows (The Cochrane Collaboration, Software Update, Oxford, UK) and STATA v.11 statistical analysis software (StataCorp LP, TX, USA). A weighted random-effects model was used for all analyses. Studies were weighted based on sample size and quality of study scoring. Pooled odds ratios (OR) were calculated as the summary statistic for dichotomous variables and weighted mean difference (WMD) calculated for continuous variables. Both OR and WMD are reported with 95% confidence intervals (CI). A P value < 0.05 was considered statistically significant. Determination of heterogeneity was undertaken using the χ² test (Cochran’s Q) and I² test; with I² ≥ 75% denoting high degree of statistically significant heterogeneity. Risk of publication bias was assessed by visual inspection of funnel plots in addition to statistical estimation with both Begg and Mazumdar’s test and Egger’s test for small study effects.

2. Results

Six studies met inclusion criteria, involving 135 robot-assisted and 162 conventional laparoscopic fundoplication procedures (Tables 1–2) [11–16]. A summary of the results of our search strategy is shown in Fig. 1.

2.1. Study characteristics and appraisal of quality of evidence

No randomized controlled trials were identified. Four of the included studies were cohort studies and the remaining two studies were case-controlled studies (Table 1). All reported RF cases were undertaken using the da Vinci Surgical System® (Intuitive Surgical, CA). All included studies reported utilization of the Nissen fundoplication technique, with the exception of the study by Lehner et al., that reported Thal fundoplications [14]. Only one included study was prospective in design, with patient allocation determined by parent preference following detailed explanation of both surgical techniques [14]. All other studies were retrospective and observational [11–13,15,16]; mostly with study group periods that were not synchronous and using historical LF controls [12,13,15]. All studies were single-institution in origin. No

Table 2
Distribution of patients with neurological impairment and also those requiring concomitant gastrostomy at the time of surgery.

<table>
<thead>
<tr>
<th></th>
<th>RF (%)</th>
<th>LF (%)</th>
<th>P value</th>
<th>Gastrostomy (%)</th>
<th>RF (%)</th>
<th>LF (%)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ivascu 2004 [16]</td>
<td>53%</td>
<td>53%</td>
<td>0.99</td>
<td>41%</td>
<td>46%</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>Lehner 2006 [14]</td>
<td>0%</td>
<td>0%</td>
<td>n/a</td>
<td>0%</td>
<td>0%</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>Copeland 2008 [13]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>34%</td>
<td>48%</td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>Al-Bassam 2009 [11]</td>
<td>64%</td>
<td>72%</td>
<td>0.55</td>
<td>60%</td>
<td>64%</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Anderberg 2009 [12]</td>
<td>79%</td>
<td>80%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Antao 2010 [15]</td>
<td>32%</td>
<td>45%</td>
<td>-</td>
<td>47%</td>
<td>58%</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

*“-” not reported. Ni = neurologic impairment.
studies reported pre-hoc power calculations to justify adequate sample size. The median mNOS score was 8 (range 5–14, Table 3).

Three studies explicitly described inclusion and exclusion criteria. The prospective cohort study by Lehnert et al. strictly limited their cohort to children with clinically relevant GERD, without co-existing disease (including neurological impairment) or recurrent GERD, not requiring gastrostomy and with no prior history of major congenital anomalies such as esophageal atresia, abdominal wall defects or diaphragmatic hernia [14]. Al-Bassam et al. and Anderberg et al. described inclusion criteria based on indications for fundoplication that consisted of GERD that was refractory to medical management and standardized diagnostic work-up findings [11,12]. Two studies reported intentional case matching. Ivascu et al. matched cases by age [16] and Copeland et al. reported case-matching with controls, however these factors were not clear [13]. Amongst the five studies that reported patient age data, children in the RF groups tended to be older than those in LF groups (overall mean age 8.44 years, RF versus 6.40 years, LF) and this approached overall statistical significance (Fig. 2a, WMD = 2.1 years; 95% CI −0.10 to 4.32; P = 0.06). Pooled analysis identified no overall difference in weight amongst the three studies that reported this data (P = 0.65, Fig. 2b). The youngest patients reported were 2 months and 3 months of age for RF and LF groups respectively [11]. The smallest patients reported were 3.5 kg and 3.0 kg for RF and LF groups respectively [11].

2.1.1. Primary outcomes: conversions

Conversions, length of hospital stay, post-operative complications, operating time, analgesia requirement and cost.

All 6 studies reported conversion data. Overall conversion rates were 3.0% (4/135) for RF and 6.2% (10/162) for LF, with combined estimate indicating 51% lower conversion rate in RF that was not statistically significant (Fig. 3, OR = 0.49; 95% CI 0.14–1.72; P = 0.27, Fig. 3). No statistical evidence of heterogeneity (I² = 0%) or publication bias (Egger’s Test = −1.87, P = 0.69; Begg and Mazumdar; P = 1.00) was detected.

In the RF group, reasons for conversions were due to intraoperative hypotension and arrhythmia (n = 1) [13], technical difficulty in a redo case (n = 1) [13], and unclear anatomy associated with a hiatus hernia (n = 1) [11]. In the LF group, reasons for conversion were due to gastric perforation (n = 1) [13], incidental discovery of a Morgagni hernia (n = 1) [12], “instrument problems” (n = 1) [12], and technical difficulty in re-do cases (n = 2) [11]. Although detailed reasons for conversions were not described by Antao et al., it was reported that most events in LF cases were due to technical difficulties related to spatially confined workspaces in smaller children [15].

2.2. Length of hospital stay

Duration of post-operative hospital stay was reported in 5 studies. No statistically significant difference was observed between RF and LF.

Table 3

<table>
<thead>
<tr>
<th>Study Reference</th>
<th>Group assignment</th>
<th>Representativeness</th>
<th>Age</th>
<th>Gender</th>
<th>Weight</th>
<th>Indication</th>
<th>GORD severity</th>
<th>Co-morbidities</th>
<th>Re-do cases</th>
<th>Other procedures</th>
<th>Technique</th>
<th>Surgeons</th>
<th>Defined outcomes</th>
<th>Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ivascu et al.16</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lehnert et al.14</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Copeland et al.13</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Al-Bassam et al.11</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anderberg et al.12</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antao et al.15</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

mNOS total /15: 6 10 6 14 9 5
groups (WMD = −0.02 days; 95% CI = −1.10 to 1.06; P = 0.98), with no heterogeneity detected (Fig. 4).

2.3. Post-operative complications

The overall frequency of post-operative complications was 12/135 (8.9%) and 13/162 (8.0%) for RF and LF groups respectively. No cases of mortality were identified. Reported complications are summarized in Table 4. Inadequate reporting detail prohibited categorization and analysis of complications according to the Clavien-Dindo classification [17]. There was no significant difference observed between RF and LF groups for complication frequency (Fig. 5, OR = 1.13; 95% CI = 0.47–2.70; P = 0.78), again without significant heterogeneity.

2.4. Operating time

Total operating time was reported in all 6 studies. No statistically significant difference was revealed between RF and LF groups, although the estimated effect favored RF (Fig. 6a, WMD 11.2 minutes; 95% CI = −14.85–37.28; P = 0.4). Significant heterogeneity was detected (I² = 88%), as well as small study effect publication bias (Egger's test = −3.88, P = 0.028). Sub-group analysis of high quality studies only (>8 median mNOS score) improved heterogeneity (I² = 83%), however with minor effect on overall estimate (Fig. 6b, WMD = 15.97 minutes; 95% CI = −18.55 to 50.49; P = 0.36). Statistical publication bias assessment of this subgroup was not possible due to the low number of high-quality studies.

In evaluating breakdown of operative times for defined procedural stages, Lehnert et al. found initial time from skin incision to start of hiatal region dissection to be significantly longer in the RF group (P = 0.002) [14]. Inversely, they also found that for the more technically difficult aspect of the procedure, hiatal dissection was significantly shorter in the RF group (P = 0.005) [14]. No significant time differences were found for hiatal repair and fundoplication stages. In the only study reporting comparative time data that examined set-up or robot ‘docking’ time, Antao et al. found total operating room time to be more contrasting between RF and LF, compared to operating time [15].

2.5. Analgesia requirement

Two studies reported comparative opiate analgesia requirements, although each used different methods of measurement that precluded quantitative data synthesis [11,12]. Al-Bassam et al. found no significant difference in the proportion of patients that required 1–2 doses of opiate analgesia post-operatively (36% RF versus 28% LF, P = 0.173) [11]. In this study, analgesia requirement was subjectively assessed in the context of vital sign measurements. Using a more objective and standardized Visual Analogue Scale (VAS) method to assess post-operative pain, Anderberg et al. found patients had shorter post-operative duration of opiate requirement (1.1 ± 0.9 days RF versus 1.6 ± 0.7 days LF) [12]. These findings are somewhat unexpected given that larger abdominal wall incisions are needed to accommodate 8 mm instruments in RF cases, compared to 3 mm-5 mm instruments used in LF cases.

2.6. Cost

Cost analysis by Anderberg et al. revealed 29% higher per case cost for RF compared to LF [12]. When indirect fixed costs such as capital outlay and maintenance service were excluded, RF remained more expensive on a per case basis (EUR 9,584 versus EUR 8,982). Of note, LF instrument costs were not included in these economic calculations as pre-existing reusable instruments were used during the study period.
2.6.1. Secondary Outcomes.

Operative success, re-operation rate and post-operative morbidity.

In general, length of follow up was short, with all studies focusing on peri-operative and short-term post-operative outcomes (Table 1). No studies clearly defined criteria or endpoints for operative success. As such, there was no standardized or protocol follow up beyond the early post-operative period to enable data synthesis for the secondary outcomes of interest outlined above.

Antao et al. reported 3 GERD recurrences, 2 in patients having undergone LF and 1 in a patient having undergone RF [15]. No other cases of recurrence were identified and no requirement for re-operation was reported in any of the included studies. During routine clinical follow up at undefined intervals within a median 14 month period post-operatively, Al-Bassam et al. observed complete symptom resolution in 64% of RF patients and 60% of LF patients [11]. Improvement in symptoms was noted in 36% and 40% of RF and LF patients in this same cohort respectively [11]. At 30-day clinical follow up, Copeland et al. observed comparable rates of transient symptoms (30% RF versus 28% LF) that included dysphagia, abdominal pain, feeding aversion and gas bloating [13].

2.6.2. Post-hoc sub-analysis.

Clinical outcomes are known to vary among the assorted laparoscopic fundoplication techniques (i.e. Nissen, Thal, Toupet, Dor). To explore the effect of fundoplication technique as a potential confounder in our results, we performed sub-analysis of studies involving Nissen fundoplication only. The only outcome measure affected in pooled meta-analysis.

To explore the effect of fundoplication technique as a potential confounder in our results, we performed sub-analysis of studies involving Nissen fundoplication only. The only outcome measure affected in pooled meta-analysis.

2.7.1. Evidence discrete roles of robot-assisted fundoplication

The impact of conventional laparoscopic fundoplication over the traditional open surgical technique was clearly disruptive, with well-described patient benefits justifying widespread adoption and progressive acceptance towards standard of care [19,20]. Given that all skin incisions, port sites, port numbers and operative steps are essentially indifferent between RF and LF approaches; it could be argued that RF should be considered as an adjunct to conventional laparoscopy, rather than a de novo surgical approach in itself. In this regard, any expected patient benefit associated with robot-assistance would more incremental when performed by expert surgeons, and therefore less discernable or even ‘hidden’ when evaluated using familiar discriminators in small comparative studies. Furthermore, some perceived benefits of robot-assistance might not be measurable by quantitative data obtained through clinical trials [21].

The enhanced features of robot-assistance are felt to offer a greater role in more difficult or high-risk cases such as those involving previous fundoplication or preservation of pre-existing gastrostomy [22,23]. Currently, there is no evidence to guide patient selection in these circumstances and this might be a focus of future research. Several authors also describe an attractively short and steep learning curve for RF in which proficiency can be attained after only 5–10 cases [11–13,24]. This would seem favourable against literature that reports the LF learning curve to be overcome after 25–50 cases [25–27]. These figures obviously need to be interpreted with caution as the majority of surgeons undertaking RF would have experience with LF and there would undoubtedly be a skill transfer effect. No studies identified in this review reported data indicating LF (or RF) experience prior to the study period, and this should be encouraged in future reporting such that any confounding effect can be estimated. Lastly, another attractive surgeon focused factor might include improved ergonomic set-up, limiting the burden of mental and physical fatigue that may be perceived as a potential threat to operative performance and therefore patient outcome [28].

3.2. Cost-benefit

From an exclusively evidence-based standpoint, it is challenging to robustly defend a cost-benefit argument for RF in the context of
presently available outcome and economic data. This scenario is certainly not unique to this particular indication and seems reflected across most adult and pediatric settings [29,30]. The high cost of robotic surgical technology is a major barrier to adoption, driven by a lack of market competition and expenses generally associated with early phase technology [30]. While affordability presently seems static at best, cost-benefit evaluations will obviously need to be periodically re-visited as pricings inevitably are driven down. In the meantime, costs of increasingly popular disposable laparoscopic instruments should not be ignored. On a per case cost basis, the expense of these single-use instruments approaches that of robotic instruments [31].

3.3. Limitations and future directions

The majority of comparative effectiveness literature is comprised of small retrospective studies with historical controls. The strength of this meta-analysis is therefore limited by inherent vulnerability of these study designs to multiple sources of bias and both known and unknown confounders. Many important outcome variables also remain unaccounted for in the literature. More high quality prospective studies and clinical trials are indisputably warranted to better inform the debate of RF versus LF for fundoplication in children.

Randomized controlled trials are considered the gold-standard method of comparative assessment for new surgical techniques and technology. Recognized difficulties to undertaking RCTs within surgical disciplines are heightened in pediatric surgery by additional ethical and scientific challenges [32]. It should be noted that amongst the small < 0.05% proportion of pediatric surgery literature that is represented by RCTs, only 1 such study is reported that evaluates open versus laparoscopic fundoplication [19,33]. In addition to generalizable barriers to undertaking clinical trials in surgery, pursuit of prospective comparative studies or trials in this field might be faced with local pressure from hospital management or other stakeholders to optimize utilization of the robot to justify the significant capital outlay investment.

Future studies should strive towards methodological quality that satisfies criteria of; 1) prospective design, 2) patient selection criteria accounting for many of the comparability variables outlined above (Table 3), 3) regular protocol follow up with clearly delineated outcome measures, 4) attention towards more procedure and pathology orientated outcomes, and 5) length of follow up that extends beyond the early post-operative period. Rigor of post-operative complication monitoring and recording might be improved with adherence to the Clavien-Dindo classification [17].

4. Conclusion

Meta-analysis of the current literature identifies comparable safety and short-term efficacy for robot-assisted laparoscopic fundoplication when compared to conventional laparoscopic fundoplication in children. However, there is insufficient evidence to assess comparative effectiveness for important procedure and pathology specific outcome measures such as success rate, post-operative surgery related morbidity and re-operation rates. Higher quality and longer-follow up studies are now required to more definitively and comprehensively determine the role of robotic-assisted techniques in this setting. Equipoise in short-term clinical outcomes emphasizes the need for further cost-effectiveness analysis to also be undertaken.

References


Fig. 5. Forest plot comparisons of post-operative complications for studies comparing RF versus LF in children.

Fig. 6. Forest plot comparisons of a) total operating time for studies comparing RF versus LF in children, and b) sub-group analysis of high quality studies only.