Systematic literature review of cost-effectiveness analyses of robotic-assisted radical prostatectomy for localised prostate cancer

Chao Song, Lucia Cheng, Yanli Li, Usha Kreaden, Susan R Snyder

ABSTRACT

Objectives Review and assess cost-effectiveness studies of robotic-assisted radical prostatectomy (RARP) for localised prostate cancer compared with open radical prostatectomy (ORP) and laparoscopic radical prostatectomy (LRP).

Design Systematic review.

Setting PubMed, Embase, Scopus, International HTA database, the Centre for Reviews and Dissemination database and various HTA websites were searched (January 2005 to March 2021) to identify the eligible cost-effectiveness studies.

Participants Cost-effectiveness, cost-utility, or cost-minimization analyses examining RARP versus ORP or LRP were included in this systematic review.

Interventions Different surgical approaches to treat localised prostate cancer: RARP compared with ORP and LRP.

Primary and secondary outcome measures A structured narrative synthesis was developed to summarise results of cost, effectiveness, and cost-effectiveness results (eg, incremental cost-effectiveness ratio [ICER]). Study quality was assessed using the Consensus on Health Economic Criteria Extended checklist. Application of medical device features were evaluated.

Results Twelve studies met inclusion criteria, 11 of which were cost-utility analyses. Higher quality-adjusted life-years and higher costs were observed with RARP compared with ORP or LRP in 11 studies (91%). Among four studies comparing RARP with LRP, three reported RARP was dominant or cost-effective. Among ten studies comparing RARP with ORP, RARP was more cost-effective in five, not cost-effective in two, and inconclusive in three studies. Studies with longer time horizons tended to report favorable cost-effectiveness results for RARP. Nine studies (75%) were rated of moderate or good quality. Recommended medical device features were addressed to varying degrees within the literature as follows: capital investment included in most studies, dynamic pricing considered in about half, and learning curve and incremental innovation were poorly addressed.

Conclusions Despite study heterogeneity, RARP was more costly and effective compared with ORP and LRP in most studies and likely to be more cost-effective, particularly over a multiple year or lifetime time horizon. Further cost-effectiveness analyses for RARP that more thoroughly consider medical device features and use an appropriate time horizon are needed.

PROSPERO registration number CRD42021246811.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This review provided a comprehensive, systematic and transparent literature search strategy covering multiple databases as well as the grey literature for health technology assessment (HTA) reports. However, private or confidential HTAs containing cost-effectiveness analysis might still be missing.

⇒ The review conducted bias assessment using the Consensus on Health Economic Criteria Extended checklist, suitable for economic evaluation studies.

⇒ Four additional criteria unique to medical devices (organisational impact, learning curve, incremental innovation and dynamic pricing) were assessed for each included study.

⇒ Internal validity of the systematic review synthesis depends on the quality of the limited number of primary studies included.

INTRODUCTION

Prostate cancer is the second most frequent malignancy (after lung cancer) in men worldwide.1 For men diagnosed with clinically localised prostate cancer, radical prostatectomy is one of the primary treatment options. Radical prostatectomy can be performed with open radical prostatectomy (ORP) or minimally invasive techniques, including laparoscopic radical prostatectomy (LRP) and robotic-assisted radical prostatectomy (RARP). Despite being a less invasive approach, application of LRP is low,2,3 possibly due to its technical difficulty in performing complex procedures (eg, bilateral nerve-sparing dissection and construction of watertight urethrovaginal anastomosis), steep learning curve, and limitations in dexterity and ergonomics.4,5,6 Robotic-assisted surgery (RAS) using the da Vinci surgical system (Intuitive Surgical Operations, Sunnyvale, California, USA) overcomes
the technical challenges encountered by LRP by allowing for additional wrist movements and three-dimensional visualisation of the operative field.\textsuperscript{10} Surgeons perform RARP using a surgical system that translates the surgeon’s hand movement from the console in real time. This is achieved through the instrument’s 7 degrees of motion, with precision and tremor filtration. Globally, use of RARP has been increasing, and has become a common surgical approach in many countries.\textsuperscript{11}

The clinical effectiveness of RARP has been well documented in literature. Compared with ORP, RARP has been shown to reduce postoperative complications (eg, blood loss and transfusion rate), reduce hospital length of stay and enable faster recovery.\textsuperscript{12,13} Compared with conventional LRP, RARP offers technical advantages (eg, the fully wristed dexterity, highly magnified three-dimensional high-resolution video) to overcome challenges from the complexity of radical prostatectomy and enables more patients to benefit from minimally invasive techniques.\textsuperscript{14,15}

Despite the increased worldwide adoption of RARP, its economic value compared with ORP and LRP remains controversial, drawing attention from policy makers, payers and health technology assessment (HTA) agencies. Multiple cost-effectiveness analyses\textsuperscript{16–19} have been performed in different healthcare settings using different methodologies over the past decade and have come to diverse conclusions. A systematic assessment of previously conducted cost-effectiveness analyses is critical for researchers and decision-makers to understand the value of RARP and to reach consensus on the appropriate methodology to quantify its cost-effectiveness.

Moreover, medical devices, such as those used to perform robotic surgery, are characterised by distinctive features that are less frequently found in pharmaceuticals\textsuperscript{20}, therefore, methods of conducting economic evaluations for medical devices need to consider additional attributes beyond those traditionally used in assessing drugs, such as the technology’s organisational impact, learning curve, incremental innovation and dynamic pricing.\textsuperscript{21,22} RAS is a good example to illustrate the need to evaluate medical device-specific features. First, substantial infrastructural investment in a robotic surgical systems triggers ‘organisational impact’. Second, a surgeon’s experience and proficiency with RAS impacts clinical outcomes and efficiency; in other words, a ‘learning curve’. Third, postlaunch innovation in robotic systems and instruments routinely occur over time and may be associated with changes in clinical outcomes, efficiency and cost. For example, since the first da Vinci surgery conducted in 2000, a total of four generations of da Vinci surgical systems have been launched with numerous instrument-level upgrades. Finally, the prices of medical devices typically decrease over time due to innovation, production scale, and market competition. This ‘dynamic pricing’ increases the level of uncertainty when assessing RAS. To our knowledge, no study in the existing cost-effectiveness literature for RARP has investigated these medical device features.

The aim of this systematic literature review is to assess the existing cost-effectiveness studies of RARP for localised prostate cancer, evaluate how medical device features have been considered in those studies, and provide insight for future cost-effectiveness studies in the field of robotic surgery.

METHODS

The protocol for this systematic review was developed in advance and registered with PROSPERO (registration number CRD4202146811) in May 2021.

Search strategy and study selection

Multiple databases were searched to retrieve studies published from 1 January 2005 (before the earliest cost-effectiveness literature on RARP) to 1 March 2021 (search date). Specifically, the databases included PubMed, Embase, Scopus, International HTA database and the Centre for Reviews and Dissemination database, which includes the NHS Economic Evaluation database, Database of Abstracts and Reviews of Effects, and the Health Technology Assessment (HTA) database. The detailed search strategy is outlined in online supplemental appendix A.

In addition to these five databases, we also performed a targeted grey literature search of various HTA websites (eg, UK National Institute for Health and Care Excellence, US Agency for Healthcare Research and Quality Evidence-Based Reports, US Institute for Clinical and Economic Review, Canadian Agency for Drugs and Technologies in Health and Tufts Cost-Effectiveness Analysis Registry) and Google. Keywords used included robotic surgery, robot assist, da Vinci, prostatectomy and prostate cancer.

Studies were included in this review if they were cost-effectiveness analyses, cost–utility analyses (CUA) or cost–minimisation analyses, and contained da Vinci-assisted radical prostatectomy as an intervention of interest. Studies were excluded if they were not in English, did not include prostatectomy or RAS, only had cost data without cost-effectiveness outcomes, were based on duplicate patient populations, or reviews that only included primary studies already captured in our review. The inclusion and exclusion criteria specifications are detailed in online supplemental appendix B.

Two reviewers independently screened the literature and any inconsistencies in the identification of potentially relevant studies were discussed to reach a consensus. The results are reported according to Preferred Reporting Items of Systematic Reviews and Meta-Analyses guidelines.\textsuperscript{23}

Data extraction

Data were extracted from included studies based on the study protocol. Extracted data elements included: study year, country, economic analysis type, comparator, perspective, time horizon, effectiveness measure

and outcome value, cost measure and outcome value, incremental costs and incremental effectiveness, cost-effectiveness value (eg, incremental cost-effectiveness ratio [ICER] if applicable), discount rate, sensitivity analysis parameters and authors’ conclusion.

**Data synthesis**

The approach used for data synthesis followed steps consistent with the International Society for Pharmacoeconomics and Outcomes Research Good Practices for cost and cost-effectiveness systematic review. A meta-analysis of the findings was planned if feasible and appropriate for the available data. However, if the published literature contained substantial variability in clinical outcomes, healthcare setting, methodology, effects, costs and willingness-to-pay thresholds, a structured narrative synthesis approach would be used.

Study characteristics, such as type of analysis, patient population, perspective and methodological choices, were summarised. We reported incremental costs, incremental effectiveness and ICERS. Cost-effectiveness results across studies were displayed using scatterplots by plotting incremental quality-adjusted life-years (QALYs) and incremental costs on the x-axis and y-axis, respectively. All cost data were converted and reported in 2021 US dollars using purchasing power parities along with the original cost data.

**Medical device features**

The application of four distinctive features recommended for economic evaluations of medical devices was evaluated in the included studies, and each one was categorised using three levels: ’adjustment made to model’, ’acknowledged but no model adjustment’ or ’not considered’. The study would be classified as ’adjustment made to model’ if adjustments were made to the base case or sensitivity analyses. If a study only mentioned the features in its writing but no adjustment was made in modelling, it would be considered as ’acknowledged but no model adjustment’.

**Critical appraisal of risk of bias**

Risk of bias of economic evaluations was assessed using the Consensus on Health Economic Criteria (CHEC)-Extended checklist. The CHEC-extended checklist contains guidelines for each criterion and scoring, and can be used to evaluate model-based and trial-based economic evaluations. A score of one point was assigned to each positive response, and zero to a negative response or for non-applicable items. The total score out of 20 items was converted to a score ranging from zero (low quality) to 100 (high quality). Included studies were categorised into four grades: low (≤40), moderate (41–50), good (51–75), moderate (76–95) and high quality (>95), respectively.

**Patient and public involvement**

The systematic review did not involve animal or human subjects. Patients and the public were not involved in the design and conduct of this systematic review since published studies were used to synthesise findings.

**RESULTS**

We identified 930 articles from the initial literature search. On reviewing, nine full-text studies from the database search and three articles from the targeted grey literature search met the inclusion criteria and were included in the final analysis (figure 1). Five studies were derived from HTA reports. Studies that were excluded tended to be costing-only studies or had mixed RARP with other surgical modalities in an intervention group.

**Characteristics of included studies**

Of the 12 studies included in this review, 11 are CUA reports. The characteristics of all studies are presented in table 1. High variability across clinical and healthcare settings was observed. Among those studies, three were conducted in Canada, two in Australia, one in USA, one in UK, two in Ireland and three in other countries. Ten studies compared RARP with ORP, four studies compared RARP with LRP and 1 study compared RARP versus routine care of mixed ORP and LRP. Cooperberg et al1 compared RARP with other surgical interventions as comparators.

Most (8 out of 12) of the studies were conducted from the payer’s perspective, with 2 studies taking a societal perspective. 1 study was conducted from a healthcare system perspective and 1 study from a hospital perspective. Regarding the time horizon, lifetime horizon was considered in two studies, 5–10 years was used in seven studies, 5–30 years was used in two studies and one study did not report the time horizon. Three studies are observational-based.
Table 1  Characteristics of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Type of literature</th>
<th>Population</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Perspective</th>
<th>Time horizon</th>
<th>Type of analysis</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSAC 2006</td>
<td>AUS</td>
<td>HTA report</td>
<td>localised PC</td>
<td>RARP</td>
<td>ORP</td>
<td>Societal</td>
<td>10 years</td>
<td>CUA</td>
<td>Decision Tree</td>
</tr>
<tr>
<td>O’Malley 2007</td>
<td>AUS</td>
<td>Journal article</td>
<td>PC</td>
<td>RARP</td>
<td>ORP</td>
<td>Payer*</td>
<td>*</td>
<td>CUA</td>
<td>Cohort based</td>
</tr>
<tr>
<td>Hohwü 2011</td>
<td>DEN</td>
<td>Journal article</td>
<td>localised PC</td>
<td>RARP</td>
<td>ORP</td>
<td>Societal</td>
<td>1 year</td>
<td>CEA</td>
<td>Cohort based</td>
</tr>
<tr>
<td>HIQA 2011</td>
<td>IRE</td>
<td>HTA report</td>
<td>PC T2-T3</td>
<td>RARP</td>
<td>Mix of ORP and LRP</td>
<td>Payer</td>
<td>5 years</td>
<td>CUA</td>
<td>Markov Model</td>
</tr>
<tr>
<td>Close 2013 /Ramsay 2012</td>
<td>UK /IRE</td>
<td>Journal article /HTA report</td>
<td>localised PC</td>
<td>RARP</td>
<td>LRP</td>
<td>Payer</td>
<td>10 years</td>
<td>CUA</td>
<td>Discrete event simulation</td>
</tr>
<tr>
<td>Cooperberg 2013</td>
<td>USA</td>
<td>Journal article</td>
<td>localised PC</td>
<td>RARP</td>
<td>ORP, LRP+other†</td>
<td>Payer</td>
<td>Lifetime</td>
<td>CUA</td>
<td>Markov Model</td>
</tr>
<tr>
<td>Teljeur 2014</td>
<td>IRE</td>
<td>Journal article</td>
<td>RP</td>
<td>RARP</td>
<td>ORP, LRP</td>
<td>Payer</td>
<td>Lifetime*</td>
<td>CUA</td>
<td>*</td>
</tr>
<tr>
<td>Ratchanon 2015</td>
<td>THA</td>
<td>Journal article</td>
<td>localised PC</td>
<td>RARP</td>
<td>LRP</td>
<td>Health system</td>
<td>10 years</td>
<td>CUA</td>
<td>Decision Tree</td>
</tr>
<tr>
<td>AHT 2017</td>
<td>CAN</td>
<td>HTA report</td>
<td>localised PC</td>
<td>RARP</td>
<td>ORP+others‡</td>
<td>Payer</td>
<td>9 years</td>
<td>CUA</td>
<td>Markov Model</td>
</tr>
<tr>
<td>HQO 2017</td>
<td>CAN</td>
<td>HTA report</td>
<td>localised PC</td>
<td>RARP</td>
<td>ORP</td>
<td>Payer</td>
<td>1 year</td>
<td>CUA</td>
<td>Markov Model</td>
</tr>
<tr>
<td>Parackal 2020</td>
<td>CAN</td>
<td>Journal article</td>
<td>localised PC</td>
<td>RARP</td>
<td>ORP</td>
<td>Payer</td>
<td>10 years</td>
<td>CUA</td>
<td>Markov Model</td>
</tr>
<tr>
<td>de Oliveira 2021</td>
<td>BRA</td>
<td>Journal article</td>
<td>localised PC</td>
<td>RARP</td>
<td>ORP</td>
<td>Hospital</td>
<td>5 years</td>
<td>CUA</td>
<td>Cohort based</td>
</tr>
</tbody>
</table>

*Not clearly stated
†IMRT, BT, 3DCRT, EBRT+BT.
‡Beam radiotherapy, brachytherapy and cryoablation.
AUS, Australia; BRA, Brazil; BT, brachytherapy; CAN, Canada; CEA, cost-effectiveness analysis; CUA, cost-utility analysis; 3DCRT, three-dimensional conformal radiation therapy; DEN, Denmark; EBRT, external beam radiation therapy; IMRT, Intensity-modulated radiation therapy; IRE, Ireland; LRP, local radical prostatectomy; ORP, open radical prostatectomy; PC, prostate cancer; RARP, robotic-assisted radical prostatectomy; RP, radical prostatectomy; THA, Thailand.
modelling studies, and the rest are simulation-based analyses. Among the nine simulation-based studies, eight studies clearly reported the model methods. Among these eight studies, Markov modelling was used in five studies,17 18 33 37 38 simple decision tree in two studies34 35 and discrete event simulation in one study.16

**Narrative synthesis of study results**

Based on the included studies, RARP was generally associated with higher effectiveness and higher cost. All studies showed RARP had higher QALYs than ORP or LRP across various time horizons, with one exception; Hohwü et al.19 reported RARP had lower QALYs than ORP with a 1-year time horizon. The range of incremental QALYs gained for RARP varied from 0.05 to 0.1 when compared with LRP, and 0.001 to 0.41 when compared with ORP. In 11 of the 12 studies, RARP had higher costs relative to the comparators. The one exception was an analysis from the US payer perspective by Cooperberg et al.17 which demonstrated lower cost for RARP compared with ORP and LRP; however, capital cost was not considered in the analysis. Results from the CUAs, in the form of incremental costs (standardised to 2021 USD) and incremental QALYs, were plotted and presented in figure 2. The corresponding summary is presented in table 2.

Most study results comparing RARP with ORP demonstrate RARP being cost-effective, although there is considerable heterogeneity across studies. As such, it is not appropriate to pool the cost-effectiveness results together. Five studies17 33 34 36 37 showed RARP to be more cost-effective than ORP, while two other studies18 19 showed RARP having higher ICERs that exceeded the willingness to pay (WTP) threshold value. Three studies32 33 36 were inconclusive on the cost-effectiveness of RARP due to insufficient comparative effectiveness data or an unspecified WTP threshold. Interestingly, the study time horizon was observed to correlate with study conclusions (figure 3). The only two studies18 19 that showed RARP was not cost-effective compared with ORP used a short-term time horizon (1 year), while the other five studies17 33 34 36 37 evaluated cost-effectiveness over 5, 9 and 10 years, or lifetime, all showed RARP to be more cost-effective than ORP.

Cost-effectiveness results for RARP compared with LRP were inconclusive given the limited number of publications (four studies),16 17 31 35 but showed a tendency towards RARP being more cost-effective. One study17 showed RARP as the dominant surgical option (lower cost and more effective). Two studies,16 32 one from the UK and one from Ireland, demonstrated RARP to be cost-effective, while another study31 conducted in Thailand found the ICER of RARP to be much higher than the stated WTP threshold.

**Systematic review on inclusion of medical device features**

Distinctive medical device features were considered to various extents in the included studies. Capital investment, which is one aspect of organisation impact, was widely considered. Ten16 18 19 31–33 35–38 of 12 studies (83%) included capital investment, and one study17 (8.5%) justified why it was not included in the analysis. Capital equipment cost and procedure volume per system were considered as sensitive parameters in seven studies (58%).16 18 19 31 35 37 38 Dynamic pricing was reflected in five studies16 18 35 37 38 (42%) by evaluating the uncertainty of equipment or instrument prices in the sensitivity analyses. Although none of studies quantitatively evaluated the impact of surgeon experience on outcomes and efficiency, 7 of 12 studies16 18 19 31–33 35–38 (58%) mentioned ‘learning curve’. In terms of incremental innovation, two studies18 38 (17%) mentioned new generations of the surgical system and one study16 (8.5%) included different costs for the new generation system within the analysis.

**Systematic review on risk of bias**

Study quality for the 12 studies was evaluated using the CHEC-Extended checklist, and results are presented in online supplemental appendix C. Among the 12 studies, 1 study36 was classified as ‘excellent’, 8 studies16–19 31 34 37 38 as ‘good’, 1 study35 as ‘moderate’ and 2 studies32 36 as ‘low’ quality. The primary reason studies were classified as low or moderate quality is that they failed to appropriately measure, value or report cost and/or effectiveness. Among the 20 items on the checklist, the majority of the studies16–19 31 32 34–37 (10 out of 12) did not consider ethical and distributional issues, and 7 studies16–18 34 36–38 (58%) used utility data collected from different study populations.

**DISCUSSION**

Our systematic literature review identified 12 studies that evaluated the cost-effectiveness of RARP for localised prostate cancer patients. Three-quarters of studies were of excellent or good quality based on the CHEC-Extended
<table>
<thead>
<tr>
<th>Study</th>
<th>Incremental cost (in local unit)</th>
<th>Incremental cost (in 2021 USD)</th>
<th>Incremental QALY</th>
<th>ICER (Original)</th>
<th>ICER (in 2021 USD)</th>
<th>Willingness to pay</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RARP vs LRP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratchanon 2015</td>
<td>120359 baht</td>
<td>US$11,385</td>
<td>0.05</td>
<td>2407180 baht/QALY</td>
<td>US$227K/QALY</td>
<td>160K baht per QALY</td>
<td>RARP is not cost-effective</td>
</tr>
<tr>
<td>Close 2013/Ramsay 2012</td>
<td>GBPE1412</td>
<td>US$2464</td>
<td>0.08</td>
<td>GBP18 329/QALY</td>
<td>US$31K/QALY</td>
<td>£30K per QALY</td>
<td>RARP is cost-effective if volume &gt;150</td>
</tr>
<tr>
<td>Cooperberg 2013</td>
<td>Low risk: US$−591</td>
<td>Low risk: US$−732</td>
<td>Low risk: 0</td>
<td>ICER not calculated</td>
<td>ICER not calculated</td>
<td>N/A</td>
<td>No difference in effectiveness</td>
</tr>
<tr>
<td></td>
<td>High risk: US$−104</td>
<td>High risk: US$−129</td>
<td>High risk: 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teljeur 2014</td>
<td>No reported</td>
<td>No reported</td>
<td>€26 643/QALY</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td><strong>RARP vs ORP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSAC 2006</td>
<td>A$3742 for ED; A$4502 for UI</td>
<td>US$3613 for ED; US$4348 for UI</td>
<td>0.10 for ED; 0.01 for UI</td>
<td>A$37K/QALY for ED; A$450K/QALY for UI</td>
<td>US$36K/QALY for ED; US$435K/QALY for UI</td>
<td>Not reported</td>
<td>Lack of data</td>
</tr>
<tr>
<td>O’Malley 2007</td>
<td>A$2264</td>
<td>US$2049</td>
<td>0.09</td>
<td>A$24 457/QALY</td>
<td>US$40K/QALY</td>
<td>Not reported</td>
<td>RARP cost-effective</td>
</tr>
<tr>
<td>Hohwü 2011</td>
<td>€4506</td>
<td>US$7093</td>
<td>−0.07</td>
<td>€64 343/extra successful treatment</td>
<td>US$101K/extra successful treatment</td>
<td>N.A (main analysis CEA)</td>
<td>RARP not cost-effective</td>
</tr>
<tr>
<td>HIQA 2011</td>
<td>€2487</td>
<td>US$3730</td>
<td>0.09</td>
<td>€26 647/QALY</td>
<td>US$30K/QALY</td>
<td>No specified threshold</td>
<td>No specified threshold in Ireland</td>
</tr>
<tr>
<td>Cooperberg 2013</td>
<td>Low risk: $−344</td>
<td>Low risk: US$−425.87</td>
<td>Low risk: 0</td>
<td>ICER not calculated</td>
<td>ICER not calculated</td>
<td>N/A</td>
<td>No difference in effectiveness, RARP has lower cost</td>
</tr>
<tr>
<td></td>
<td>Intermediate risk: $−572</td>
<td>Intermediate risk: US$−708</td>
<td>Intermediate risk: 0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>High risk: $−1265</td>
<td>High risk: US$−1566</td>
<td>High risk: 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teljeur 2014</td>
<td>No reported</td>
<td>No reported</td>
<td>€26 920/QALY</td>
<td>US$40K/QALY</td>
<td>Not reported</td>
<td>Not reported</td>
<td>Not reported</td>
</tr>
<tr>
<td>AHT 2017</td>
<td>C$8541</td>
<td>US$7813</td>
<td>0.19</td>
<td>C$44 471/QALY</td>
<td>US$41K/QALY</td>
<td>C$50K per QALY</td>
<td>RARP cost-effective</td>
</tr>
<tr>
<td>HQO 2017</td>
<td>C$6234</td>
<td>US$5702</td>
<td>0.001</td>
<td>C$5.2M/QALY</td>
<td>US$55M/QALY</td>
<td>C$100K per QALY</td>
<td>RARP not cost-effective</td>
</tr>
<tr>
<td>Parackal 2020</td>
<td>C$1701</td>
<td>US$1457</td>
<td>0.07</td>
<td>C$25 704/QALY</td>
<td>US$22K/QALY</td>
<td>C$50K or 100K per QALY</td>
<td>RARP cost-effective</td>
</tr>
<tr>
<td>de Oliveira 2021</td>
<td>BRA R$9214</td>
<td>US$4368</td>
<td>0.41</td>
<td>R$22 690.83/QALY</td>
<td>US$11K/QALY</td>
<td>R$114 026.55 (3 times of GDP)</td>
<td>RARP cost-effective</td>
</tr>
</tbody>
</table>

A$, Australian Dollar; BRA, Brazil; C$, Canadian Dollar; CEA, cost-effectiveness analysis; ED, erectile dysfunction; GBP, British pound sterling; GDP, gross domestic product; ICER, incremental cost-effectiveness ratio; LRP, local radical prostatectomy; ORP, open radical prostatectomy; QALY, quality-adjusted life year; RARP, robotic-assisted radical prostatectomy; UI, urinary infection.
checklist. These studies vary widely in country/healthcare setting, comparators and time horizon. A majority of the studies found RARP to be more costly and more effective compared with ORP and LRP, although the cost-effectiveness conclusions (ie, the ICERs) varied and were dependent on the specific WTP thresholds used and influenced by time horizon. The four medical device features recommended to be included in economic evaluations were considered to varying degrees: capital investment in the surgical system (organisational impact) was widely considered, dynamic pricing was considered in about half of the studies, while learning curve and incremental innovation were poorly addressed in the included studies.

Cost, effectiveness and cost-effectiveness of RARP

Although the higher cost and effectiveness of RARP observed in this review were consistent with existing literature,11,39,40 the key question is whether the effectiveness gained is worth the increased cost. Although most of the CEAs comparing RARP with ORP or with LRP concluded RARP to be cost-effective in this systematic literature review, a definitive answer remains elusive given that cost-effectiveness thresholds varied across different countries and healthcare settings. Studies with longer time horizons tended to have more favourable cost-effectiveness results for RARP. Possible explanations for this correlation is that RARP incurs a high upfront capital, instrument and accessory cost in the perioperative time period, while improved patient outcomes observed after RARP including lower rates of positive surgical margins and better functional outcomes7,13,14,41 might lead to downstream healthcare cost savings11 and translate into better quality of life as well.

Medical device features in RARP

Despite recommendations in multiple authoritative methods publications,20-22 adoption of the four recommended special characteristics of medical devices within these cost-effectiveness evaluations was limited. Among most of the studies included in this review, the capital cost of acquiring a robotic system, one aspect of organisational impact, was considered in the cost calculation. The allocation of capital equipment cost per RARP case is challenging, given the complex financial allocation in different healthcare systems and sharing of the use of robotic systems across specialties. The capital cost calculation of robotic assisted surgery should reflect the actual cost allocation from the appropriate perspective and healthcare system. Only two included studies31,38 calculated the capital cost per RARP case by allocating the cost to multiple procedures across specialties to reflect real world practice. Additionally, if robotic capital cost is funded by a charitable donation, it is important to consider who actually paid for it and align cost calculations with the study perspective.

Learning curve, the second recommended medical device feature, is considered the most important characteristic associated with the use of a medical device.21 In this systematic review, while learning curve was mentioned among 58% of the studies, no action was taken to incorporate it within the analyses. Inclusion of a robotic surgical system’s learning curve could be considered on two fronts. First, surgeons need some practice to reach proficiency after adopting new technology, which could be accelerated by rigorous training. Second, higher surgical volumes may not only reduce cost per procedure (economies of scale), but also improve patient outcomes and reduce the operative time per procedure as the surgeons become more skilled and their proficiency increases.43,44 Scenario analysis could be considered to further understand the uncertainty related to surgical volume consistent with a learning curve effect.

Incremental innovation, the third recommended medical device feature, is common in RARP. For example, four generations of da Vinci RAS systems and instruments, with numerous product innovations, have been launched in the past 20 years. However, incremental innovation was considered in only one study in this review, and it focused exclusively on the differential costs by generations of systems without considering changes in effectiveness. This is likely due to lack of clinical studies that differentiate effectiveness among various generations of surgical systems. Postmarket observational studies for newer RAS generations or subgroup analysis for the different system/product generations are needed to address this gap.

Lastly, medical device pricing is considered more dynamic than drugs, and launching new generations of technology often influences the price of existing devices.21,22 Five studies in this review empirically tested varying equipment prices. In addition to using updated pricing information, researchers could consider estimating a threshold price at which RAS provides a minimally acceptable value, which decision makers could consider in future purchasing or leasing decisions. For a healthcare system, the threshold price for a new technology is at
the point of indifference between accepting and rejecting the technology, assuming all conditions for other options are equal. Analyses considering threshold price and technology generation were not found in this systematic review.

Suggestions for future RAS cost-effectiveness studies
Several opportunities to improve RAS economic evaluations were identified from this systematic literature review. First, the selection of time horizon should be long enough to capture the relevant differences in outcomes and costs to the various stakeholders. With emerging evidence demonstrating RAS’ long-term clinical benefits such as less positive surgical margin and better functional outcomes, researchers should consider applying appropriate time horizons consistent with the direct and, when relevant, indirect effects of the procedures on patient outcomes. Second, surgeon proficiency may affect patient outcomes and efficiency. Clinical studies might consider measuring and reporting the experience and proficiency of surgeons (eg, number of cases performed previously) when evaluating their surgical outcomes. Cost-effectiveness analyses could use clinical data from experienced surgeons who have passed the learning curve or consider stratified analyses by the performance of high-volume versus low-volume centres/surgeons to better examine the impact of surgeon proficiency. Moreover, with increasing numbers of new robotic surgical products and manufacturers, differentiation between products is critical for economic evaluations to inform decision-making, as it will be increasingly unlikely that all robotic surgery platforms are equivalent. Clinical studies that document clinical outcomes data by different brands and generations of devices are needed to enable the evaluation of incremental innovation of medical devices in cost-effectiveness analyses. Third, existing studies are primarily conducted from societal or payer perspectives. Future studies evaluating economic value of RAS should consider a healthcare systems perspective, given the purchasing decision is often made at this level. Researchers may need to carefully select cost and benefit components to align with the perspective of the study in the specific country. Fourth, the cost of infrastructure necessary to accommodate the device and any impact of the new device on procedure costs should be considered. This may include the cost of training, increase in surgical volume and conversion of procedure from inpatient to outpatient setting. Finally, the COVID-19 pandemic brings new challenges for constrained healthcare resources. The opportunity cost of using RAS to reduce downstream health resource use may be increasingly relevant in this environment.

Limitations of the systematic review
The current review is subject to several limitations. First, the literature search was limited to publicly available information. Private or confidential HTAs may contain cost-effectiveness analyses not included in this study, despite our effort to conduct a targeted grey literature search. Second, the internal validity of a systematic review synthesis depends on the quality of primary studies included. In our review, the general quality of included studies could be considered moderate to good, except for methods used to assess effectiveness. More than half of the studies did not use local utility data for prostate cancer. The lacking country-specific utility data increased the uncertainty of the published economic evaluations. Third, although studies with a cost-comparison design could provide insights on costs, they were excluded due to lack of effectiveness data. In addition, patient benefits that are not directly associated with clinical effectiveness measures, such as reduction of out-of-pocket cost and reduction on productivity loss, were not evaluated in most of the original studies in this review. Inclusion of these additional patient-focused outcomes may more accurately reflect value/cost. Finally, with the increasing use of non-surgical treatment for localised prostate cancer, such as high-intensity focused ultrasound, it is worthwhile to further investigate the cost-effectiveness across all treatment options.

CONCLUSIONS
To our knowledge, this is the first systematic literature review on the cost-effectiveness of RAS and evaluated the application of recommended medical device features. No conclusive cost-effectiveness result was identified in the literature due to study heterogeneity; however, RARP was found to be more costly and effective compared with ORP and LRP in most studies, providing a body of evidence supporting its cost-effectiveness. Analyses with longer time horizons showed more favourable cost-effectiveness results towards RARP. Further cost-effectiveness analyses for RARP that more thoroughly consider medical device features are needed to better understand and more appropriately estimate its economic value compared with other surgical and non-surgical treatments.

Acknowledgements We thank Sadaf Saaber, April Hebert, Ana Yankovsky and Ben Forest for support with literature screening and Rachael Mann, from Labcorp Drug Development, for editorial support.

Contributors CS, YL and UK were involved in the conception and design of the study. CS and LC were involved in the acquisition of data. All authors (CS, LC, UL, UK and SRS) made substantial contribution to the analysis and interpretation of the data, the drafting and critical revision of the manuscript. YL acts as the guarantor of this study.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests LC, YL and UK are full-time employees of Intuitive Surgical. CS is a former full-time employee of Intuitive Surgical and a current full-time employee of Union Chimique Belge (UCB).

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Ethics approval was not required since the systematic review aggregated results from published studies.

Provenance and peer review Not commissioned; externally peer reviewed.


35 MSAC. Laparoscopic remotely assisted radical prostatectomy, 2006.


38 HIQA. Health technology assessment of robot-assisted surgery in selected surgical procedures, 2011.


40 Ramsay C, Pickard R, Robertson C, et al. Systematic review and economic modelling of the relative clinical benefit and cost-


