Comparative efficacy and safety of treatments for localised prostate cancer: an application of network meta-analysis

Tengbin Xiong, Rebecca M Turner, Yinghui Wei, David E Neal, Georgios Lyratzopoulos, Julian P T Higgins

ABSTRACT

Context: There is ongoing uncertainty about the optimal management of patients with localised prostate cancer.

Objective: To evaluate the comparative efficacy and safety of different treatments for patients with localised prostate cancer.

Design: Systematic review with Bayesian network meta-analysis to estimate comparative ORs, and a score (0–100%) that, for a given outcome, reflects average rank order of superiority of each treatment compared against all others, using the Surface Under the Cumulative RAnking curve (SUCRA) statistic.

Data sources: Electronic searches of MEDLINE without language restriction.

Study selection: Randomised trials comparing the efficacy and safety of different primary treatments (48 papers from 21 randomised trials included 7350 men).

Data extraction: 2 reviewers independently extracted data and assessed risk of bias.

Results: Comparative efficacy and safety evidence was available for prostatectomy, external beam radiotherapy (different types and regimens), observational management and cryotherapy, but not high-intensity focused ultrasound. There was no evidence of superiority for any of the compared treatments in respect of all-cause mortality after 5 years. Cryotherapy was associated with less gastrointestinal and genitourinary toxicity than radiotherapy (SUCRA: 99% and 77% for gastrointestinal and genitourinary toxicity, respectively).

Conclusions: The limited available evidence suggests that different treatments may be optimal for different efficacy and safety outcomes. These findings highlight the importance of informed patient choice and shared decision-making about treatment modality and acceptable trade-offs between different outcomes. More trial evidence is required to reduce uncertainty. Network meta-analysis may be useful to optimise the power of evidence synthesis studies once data from new randomised controlled studies in this field are published in the future.

BACKGROUND

Prostate cancer is a worldwide major public health issue. Nearly 75% of diagnosed cases, however, occur in developed countries, where it is typically the most common cancer in men. In the UK, about 40 000 men are diagnosed with prostate cancer and 10 000 men die from it every year. In the USA, there are 240 000 new diagnoses of prostate cancer, with 34 000 associated deaths every year. Most patients with prostate cancer are diagnosed at an early stage, and many diagnoses are made in asymptomatic men.

The main treatment options for localised prostate cancer include radical prostatectomy, external beam radiotherapy and observational management (ie, regular testing of clinical, biochemical or radiological markers or as prompted by occurrence of symptoms). As some of these treatments are associated with substantial risk of side effects, it is important to try to resolve the current uncertainty about the optimal treatment options.

Some randomised trials have compared the efficacy and safety of two or three treatments. For example, the SPCC-4 trial in Europe and the PIVOT study in the USA compared radical prostatectomy with observational management. The UK Prostate Testing for Cancer and Treatment (ProtecT)
trial is evaluating treatment effectiveness of active monitoring, radical prostatectomy and external beam radiotherapy for clinically localised prostate cancer in men aged 50–69 years identified through population-based prostate-specific antigen (PSA) testing. The recruitment phase for the ProtecT trial, which began in 1999, has been completed, but outcomes will not be available until a minimum follow-up period has been accrued.

It is unlikely that any single trial will compare all available treatment options. We therefore performed a network meta-analysis based on a systematic review of completed randomised trials comparing different interventions for patients with localised prostate cancer. The network meta-analysis allowed us to integrate evidence from direct comparisons (treatments compared head to head within a randomised trial) and indirect comparisons (treatments compared by combining the results of randomised trials with common comparators). Our objective was to apply the established methodology used in network meta-analysis to an area of clinical practice where no such previous studies existed. In doing so, our aims were to summarise existing evidence; ‘map out’ current gaps in comparative evidence to help motivate the design and conduct of future comparative studies and develop an approach ‘primed’ for subsequent updating and incorporation of future trial evidence.

METHODS
Eligibility criteria
We sought completed randomised trials in men with localised prostate cancer that had compared two or more of the following interventions (as primary treatment, with or without the same adjuvant therapy in all arms): prostatectomy, radiotherapy including brachytherapy, cryotherapy, high-intensity focused ultrasound (HIFU) and observational management. Observational management is characterised by testing of clinical, biochemical or radiological markers of disease progression at regular intervals (typically every 6 months) or as prompted by the occurrence of new symptoms, possibly leading to either radical or palliative treatment. We opted to use the term ‘observational management’ in preference to active surveillance or active monitoring because the latter terms typically aim to keep men in a window of curability so that only those who require it undergo radical treatment.

Eligible trials had to have reported any of the following efficacy and safety outcomes: all-cause mortality, prostate cancer mortality and gastrointestinal (GI) or genitourinary (GU) toxicity. Studies comparing treatment combinations or sequences (eg, per protocol management by surgery with subsequent radiotherapy) were excluded.

Identification of studies
We adopted the search strategy of a systematic review that supported the development of clinical guidelines on the diagnosis and treatment of prostate cancer by the UK’s National Institute for Health and Clinical Excellence (NICE) in 2008. Studies had been identified by searching MEDLINE (in 2006) and scanning reference lists of papers. We retrieved all relevant randomised trials identified in the NICE guidelines and implemented the same search strategies to update the collection of trials. We restricted the search to the period from January 2005 to September 2012. No language limits were placed on the searches (see online supplementary appendix I for full search strategies).

Data extraction
Two reviewers (TX and RMT) independently screened all the titles and abstracts of the studies retrieved by the searches for potentially eligible trials, and then independently assessed the full articles of those trials to confirm whether they met the eligibility criteria. The results were checked and discussed by TX and RMT to agree on a final list of included studies. Using a structured and piloted data collection form, all relevant data in each included paper were extracted by two reviewers independently (TX and RMT/YW). The data extracted were cross-checked and unresolved discrepancies were referred to a third reviewer; where necessary, problems were discussed in a panel meeting (TX, RMT, YW, JPTH and GL) while DEN acted as a clinical expert advisor.

For each included study, we extracted characteristics of participants and interventions, outcomes reported and collected, sample size (randomised and analysed) in each arm, numerical results, losses to follow-up and details of patients excluded from the analyses. To inform the appropriateness of including studies in the meta-analysis and facilitate assessment of the strength of the evidence we assessed the risk of bias in each included study using the Cochrane Collaboration’s Risk of Bias tool. Two reviewers (TX and either RMT or YW) completed this independently and agreed on final assessments. The tool assesses risk of bias arising from inadequacies in processes of generation of the random allocation sequence, concealment of the allocation sequence and blinding and from incomplete outcome data and selective outcome reporting.

Outcomes
We analysed all-cause mortality and cancer-related mortality at 5 years, late GI and late GU toxicity at 3 years. The choice of these follow-up times was pragmatic, as they were the ones most frequently reported in the included trials. Once these time points had been chosen, we extracted the outcome data from the time nearest to these targeted measurement times. Late GI and late GU toxicity were defined as scores ≥2 measured by the Radiation Therapy Oncology Group (RTOG) questionnaire scale at 3 years follow-up. We have not encompassed biochemical or clinical failure as operational definitions of either of those outcomes tend to be specific to different radical treatment modalities.
Statistical analyses

Initially, we compared each pair of treatments using direct evidence alone, for each outcome. Separate meta-analyses were performed for each pair-wise comparison of interventions: a random-effects model was fitted within each comparison,\(^23\) with a common between-study heterogeneity variance assumed across comparisons to allow for heterogeneity even when only a single study was available. Results are reported as ORs with 95% CIs, for every comparison evaluated directly in one or more studies.

Next, we fitted a network meta-analysis model for each outcome separately,\(^22\) combining direct evidence for each comparison (eg, from studies comparing interventions A with B) with indirect evidence (eg, from studies comparing A with C and studies comparing B with C), for all pair-wise comparisons simultaneously. The model accounts explicitly for the binary nature of each outcome using a binomial likelihood function; allows for heterogeneity of treatment effects between trials of the same comparison (assuming the same amount of heterogeneity for each comparison, irrespective of how many trials address it) and enforces an underlying relationship between direct and indirect evidence for a particular comparison, assuming these are consistent between the two sources. For each ‘loop’ of treatment comparisons from three or more independent sources and for each outcome, we computed the difference between estimates from direct and indirect evidence on the log OR scale.\(^23\)

This provides a measure of inconsistency between the different sources. We did not implement more sophisticated methods for testing or adjusting for inconsistency due to the small number of loops in the network.

Results are reported as ORs with 95% credible intervals for all pair-wise comparisons of interventions. All analyses were performed within a Bayesian framework, using Markov chain Monte Carlo methods in WinBUGS (MRC Biostatistics Unit, Cambridge, UK).\(^24\) Informative prior distributions were used for the heterogeneity variance, from a published set of distributions for heterogeneity expected in meta-analyses examining particular intervention and outcome types.\(^25\) since heterogeneity is imprecisely estimated when the number of studies is small. For all-cause mortality, a log-normal (−3.93, 1.51) distribution was used. For GI and GU toxicity, a log-normal (−2.01, 1.64) distribution was used. For cancer-related mortality, a log-normal (−2.89, 1.91) distribution was used. Vague N (0, 10\(^4\)) priors were used for all other model parameters. Results were based on 100 000 iterations, following a burn-in of 20 000 iterations.

For each outcome, we estimated the probability that each intervention is superior to all others, the second best, the third best and so on, from the rank orderings of the treatments at each iteration of the Markov chain. These ranking probabilities were used to calculate a summary numerical value: the Surface Under the Cumulative RAnking curve (SUCRA).\(^26\) SUCRA values are expressed as percentages; if an intervention is certainly the best, its SUCRA value would be 100%, and if an intervention is certainly the worst, its SUCRA value would be 0%. If all interventions are equivalent, we would expect all SUCRA values to be near 50%. We also report the median ranks and 95% credible intervals for each intervention.

RESULTS

Included studies and interventions

The NICE systematic review\(^8\) had identified 20 reports relating to 14 randomised trials\(^27–46\) Our updated searches retrieved 1740 studies and identified 39 reports of relevant randomised trials, of which 30 had not been included in the NICE review (figure 1).\(^47–76\) One of these reports was the sole report of a trial providing data only on acute toxicity,\(^41\) one paper reported only clinical failure\(^39\) and one paper reported biochemical failure, biochemical disease-free survival and quality of life;\(^57\) these three studies were then excluded since they did not report the outcomes of interest to us. In addition to the remaining 47 full papers from peer-reviewed journals, we identified and included in the analysis data from a conference abstract, describing a randomised trial comparing external beam radiotherapy versus watchful waiting,\(^77\) and reporting data on long-term mortality not previously reported in full-text-related publications.\(^78–79\)

Our searches also identified 16 relevant systematic reviews,\(^80–95\) We scrutinised the reference lists of all these as well as any further systematic reviews identified by the NICE review, and found no further relevant randomised trials.

The 48 identified reports described 21 randomised trials comparing the effectiveness of different treatments for localised prostate cancer.\(^27–38\) 40 42–56 58–77 Seventeen trials reported all-cause mortality, 16 trials reported cancer-related mortality, 16 trials reported GI toxicity, 15 trials reported GU toxicity. The characteristics of included studies are summarised in online supplementary appendix 2.

The risk of bias assessments for the included trials is illustrated in figure 2. Most of the evidence was of moderate-to-good quality. About half of the studies did not report adequate information about allocation sequence generation and allocation sequence concealment. Unblinded designs were used in all trials included; we judged this unlikely to cause bias for objectively measured outcomes such as mortality, but generate bias in the reporting and assessment of patient-reported toxicity outcomes. The small number of studies precluded the investigation of potential reporting biases across studies (eg, using funnel plots). Our searches were appropriate, but the possibility of publication bias cannot be excluded. It is unclear, however, whether reporting biases would tend to favour any particular treatment (see online supplementary appendix 3 for details of bias assessments for included trials).

We categorised the interventions into the following eight categories: observational management,
prostatectomy, conventional radiotherapy (refers to two-dimensional external beam radiation therapy), conventional radiotherapy hypofractionated (refers to less than 20 fractions), conformal low dose (LD) radiotherapy (refers to less than 68 Gy), conformal high dose (HD) radiotherapy (refers to more than 74 Gy), conformal LD radiotherapy hypofractionated and cryotherapy. Twenty trials had two intervention arms. One trial compared three interventions; since two of the three interventions were very similar and both met our definition of conformal LD radiotherapy hypofractionated, we combined the data from these two arms and regarded the trial as a two-treatment comparison (conformal LD radiotherapy hypofractionated vs conformal HD radiotherapy). None of the reviewed studies assessed brachytherapy and HIFU. Figure 3 illustrates the full network of comparisons. There were two closed loops of comparisons, one connecting prostatectomy, observational management and radiotherapy modalities; and the other connecting different radiotherapy modalities. No inconsistency was detected in our estimates of the difference between direct and indirect evidence; however, precision was very low. Cryotherapy only had a single link to the network.

**All-cause mortality**

All-cause mortality was reported in 17 trials, covering all the eight interventions of interest. There is no evidence of superiority of any treatment for all-cause mortality. For each pair-wise comparison of interventions, the 95% intervals for ORs were wide and included 1. The lower-left triangle of results in table 1 presents ORs estimated from direct evidence alone, while the upper-right triangle of results presents ORs estimated from the network meta-analysis. The intervals are slightly narrower when based on indirect as well as direct evidence rather than direct evidence alone. The SUCRA values presented in table 2 summarise the ranking information for all interventions. With respect to all-cause mortality, the highest SUCRA values are 69% for conformal LD radiotherapy hypofractionated and 63% for conformal HD radiotherapy, indicating that these are most likely to be among the best treatments for this outcome. However, there is very
high uncertainty in the rankings of the interventions, as indicated by wide 95% credible intervals.

**Cancer-related mortality**

Cancer-related mortality was reported in 16 trials, covering eight of the interventions. This was a rare outcome in most treatment groups, as expected for patients with localised prostate cancer with a 5-year end point. OR estimates had wide 95% credible intervals, particularly in comparison for which only indirect evidence was available, and there was no evidence of superiority for any of the comparator treatments (table 3). Based on direct comparisons alone, conformal HD radiotherapy was superior to conventional radiotherapy (OR 0.21 (95% interval 0.03 to 0.97)) and prostatectomy was superior to observational management (OR 0.60 (95% interval 0.37 to 0.98)).

**GI and GU toxicity**

Late GI toxicity was reported in 16 trials and late GU toxicity was reported in 15 trials. There was evidence...
### Table 1
All-cause mortality: ORs (posterior mean with 95% intervals) for each pair-wise comparison of interventions, based on direct evidence alone (lower-left triangle) or direct and indirect evidence (upper-right triangle)

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Observational management</th>
<th>Prostatectomy</th>
<th>Conventional radiotherapy</th>
<th>Conventional radiotherapy hypofractionated</th>
<th>Conformal LD radiotherapy</th>
<th>Conformal HD radiotherapy</th>
<th>Conformal LD radiotherapy hypofractionated</th>
<th>Cryotherapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observational management</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prostatectomy</td>
<td>0.80 (0.61,1.06)</td>
<td>0.79 (0.61,1.02)</td>
<td>0.84 (0.48,1.57)</td>
<td>0.72 (0.37,1.49)</td>
<td>0.70 (0.40,1.28)</td>
<td>0.70 (0.40,1.28)</td>
<td>0.70 (0.40,1.28)</td>
<td></td>
</tr>
<tr>
<td>Conventional radiotherapy</td>
<td></td>
<td>1.06 (0.60,2.02)</td>
<td>0.91 (0.46,1.90)</td>
<td>0.85 (0.59,1.24)</td>
<td>0.86 (0.55,1.35)</td>
<td>0.82 (0.51,1.35)</td>
<td>0.62 (0.11,3.21)</td>
<td></td>
</tr>
<tr>
<td>Conventional radiotherapy hypofractionated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conformal LD radiotherapy</td>
<td>0.66 (0.35,1.21)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conformal HD radiotherapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conformal LD radiotherapy hypofractionated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cryotherapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Lower-left results compare row-defining interventions against column-defining interventions. Upper-right results compare column-defining interventions against row-defining interventions.

In the meta-analysis of direct comparisons (reported in lower-left triangle), between-trial heterogeneity $\tau^2$ was estimated as 0.009 (95% interval 0.001 to 0.08).

In the network meta-analysis (reported in upper-right triangle), between-trial heterogeneity $\tau^2$ was estimated as 0.009 (95% interval 0.001 to 0.07).

The small superscript numbers in the cells of lower-left triangle indicate the numbers of randomised trials which compared the two interventions directly.

HD, high dose; LD, low dose.

### Table 2
Ranking of interventions with respect to all-cause and cancer-related mortality, adverse gastrointestinal and genitourinary events: SUCRA values and median ranks (with 95% intervals)*

<table>
<thead>
<tr>
<th>Intervention</th>
<th>All-cause mortality</th>
<th>Cancer-related mortality</th>
<th>Adverse gastrointestinal events</th>
<th>Adverse genitourinary events</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SUCRA value (%)</td>
<td>Median rank (95% interval)</td>
<td>SUCRA value (%)</td>
<td>Median rank (95% interval)</td>
</tr>
<tr>
<td>Observational management</td>
<td>18</td>
<td>7 (2 to 8)</td>
<td>30</td>
<td>6 (3 to 8)</td>
</tr>
<tr>
<td>Prostatectomy</td>
<td>49</td>
<td>5 (1 to 7)</td>
<td>64</td>
<td>4 (1 to 7)</td>
</tr>
<tr>
<td>Conventional radiotherapy</td>
<td>35</td>
<td>6 (2 to 8)</td>
<td>16</td>
<td>7 (4 to 8)</td>
</tr>
<tr>
<td>Conventional radiotherapy hypofractionated</td>
<td>58</td>
<td>4 (1 to 8)</td>
<td>44</td>
<td>5 (1 to 8)</td>
</tr>
<tr>
<td>Conformal LD radiotherapy</td>
<td>57</td>
<td>4 (1 to 7)</td>
<td>61</td>
<td>4 (2 to 7)</td>
</tr>
<tr>
<td>Conformal HD radiotherapy</td>
<td>63</td>
<td>3 (1 to 7)</td>
<td>75</td>
<td>2 (1 to 6)</td>
</tr>
<tr>
<td>Conformal LD radiotherapy hypofractionated</td>
<td>69</td>
<td>1 (1 to 8)</td>
<td>85</td>
<td>1 (1 to 8)</td>
</tr>
<tr>
<td>Cryotherapy</td>
<td>50</td>
<td>4 (1 to 8)</td>
<td>24</td>
<td>7 (2 to 8)</td>
</tr>
</tbody>
</table>

The SUCRA value is a numerical summary of the estimated probabilities that each treatment is the best, second best, third best (and so on) for that particular outcome. Higher values indicate higher rankings compared with other treatments. For example, for cryotherapy, the high SUCRA value of 99% and median rank of 1 for adverse gastrointestinal events shows that cryotherapy is expected to be superior with respect to this outcome.

HD, high dose; LD, low dose; SUCRA, Surface Under the Cumulative RAnking curve.

*The SUCRA value is a numerical summary of the estimated probabilities that each treatment is the best, second best, third best (and so on) for that particular outcome. Higher values indicate higher rankings compared with other treatments. For example, for cryotherapy, the high SUCRA value of 99% and median rank of 1 for adverse gastrointestinal events shows that cryotherapy is expected to be superior with respect to this outcome.

HD, high dose; LD, low dose; SUCRA, Surface Under the Cumulative RAnking curve.
that cryotherapy resulted in fewer adverse GI events than radiotherapy treatments (estimated ORs comparing cryotherapy against the five radiotherapy options ranged from 0.12 to 0.24, while all but one of the respective 95% credible intervals excluded 1). The SUCRA value of 99% for cryotherapy and the median rank of 1 (95% interval 1, 2) suggest that cryotherapy is almost certainly superior among the six treatments included in the network meta-analysis in relation to adverse GI events (tables 2 and 4). There was also evidence that GI toxicity was more likely with conformal LD radiotherapy than with conformal HD radiotherapy. Interpretation of such findings for toxicity should be more cautious than for the other outcomes, due to a concern that lack of blinding could have led to a risk of detection bias. For GU toxicity, there was no evidence favouring one intervention over another (table 5), although cryotherapy tended to receive better rankings than the five radiotherapy treatments (table 2), and the OR estimates favour cryotherapy, but the 95% intervals all included 1.

**DISCUSSION**

Using network meta-analysis, we were able to combine simultaneously all relevant evidence on treating patients with localised prostate cancer, even in the absence of direct comparative evidence for some treatment pairs, encompassing four efficacy and safety outcomes. Based on data from 21 trials including 7350 patients randomly assigned among eight different intervention regimes for localised prostate cancer, we found substantial uncertainty about the relative efficacy and safety of different interventions in respect of the studied outcomes.

Assumptions of consistency between direct and indirect evidence were tested to justify the joint synthesis of all studies; however, these tests had little power due to the relatively small number of trials available in most direct comparisons. Instead we must rely on judgements about the similarity of studies included in the analysis in aspects such as patient groups, outcome measures and study methodology. Although we defined the population of interest as patients with localised prostate cancer, there was heterogeneity between individual study populations in terms of the severity of disease. Some of the trials were conducted several decades ago, when surgery and radiology techniques may have been different, and we observed that stage migration has occurred in men diagnosed with prostate cancer due to emerging biomarker and image technologies. Furthermore, some of the trials used adjuvant therapy, although this was applied in all the arms within the trial.

Two further limitations warrant mention. Literature searches were completed in September of 2012. However, the results of one of the most important randomised trials—ProtecT study—has not been published so far, and to our knowledge no other new relevant randomised controlled trials (RCTs) have been
### Table 4  Adverse gastrointestinal events: ORs (posterior mean with 95% intervals) for each pair-wise comparison of interventions, based on direct evidence alone (lower-left triangle) or direct and indirect evidence (upper-right triangle)

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Observational management</th>
<th>Prostatectomy</th>
<th>Conventional radiotherapy</th>
<th>Conventional radiotherapy hypofractionated</th>
<th>Conformal LD radiotherapy</th>
<th>Conformal HD radiotherapy</th>
<th>Conformal LD radiotherapy hypofractionated</th>
<th>Cryotherapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observational management</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Prostatectomy</td>
<td>–</td>
<td>0.84 (0.33,1.88)</td>
<td></td>
<td>1.01 (0.19,5.29)</td>
<td>0.72 (0.29,1.59)</td>
<td>1.42 (0.57,3.39)</td>
<td>1.26 (0.35,4.30)</td>
<td>–</td>
</tr>
<tr>
<td>Conventional radiotherapy</td>
<td>–</td>
<td>–</td>
<td>1.01 (0.22,4.56)</td>
<td>0.70 (0.11,4.37)</td>
<td>1.40 (0.21,9.02)</td>
<td>1.24 (0.15,9.76)</td>
<td>0.17 (0.04,0.51)</td>
<td>–</td>
</tr>
<tr>
<td>Conventional radiotherapy hypofractionated</td>
<td>–</td>
<td>–</td>
<td>1.00 (0.17,1.16)</td>
<td>0.46 (0.07,2.95)</td>
<td>1.19 (0.21,9.02)</td>
<td>1.24 (0.15,9.76)</td>
<td>0.17 (0.04,0.51)</td>
<td>–</td>
</tr>
<tr>
<td>Conformal LD radiotherapy</td>
<td>–</td>
<td>–</td>
<td>2.66 (0.85,8.62)</td>
<td>1.73 (1.07,2.97)</td>
<td>0.89 (0.36,2.15)</td>
<td>0.12 (0.02,0.48)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Conformal HD radiotherapy</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.89 (0.39,1.96)</td>
<td>0.14 (0.02,0.48)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Conformal LD radiotherapy hypofractionated</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.89 (0.39,1.96)</td>
<td>0.14 (0.02,0.48)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Cryotherapy</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Lower-left results compare row-defining interventions against column-defining interventions. Upper-right results compare column-defining interventions against row-defining interventions.

The comparison of prostatectomy with observational management was not linked to the rest of the network, so this evidence was included in the meta-analysis of direct comparisons only. In the meta-analysis of direct comparisons (reported in lower-left triangle), between-trial heterogeneity $\tau^2$ was estimated as 0.14 (95% interval 0.01 to 0.97). In the network meta-analysis (reported in upper-right triangle), between-trial heterogeneity $\tau^2$ was estimated as 0.24 (95% interval 0.02 to 1.23).

The small superscript numbers in the cells of lower-left triangle indicate the numbers of randomised trials which compared the two interventions directly.

### Table 5  Adverse genitourinary events: ORs (posterior mean with 95% intervals) for each pair-wise comparison of interventions, based on direct evidence alone (lower-left triangle) or direct and indirect evidence (upper-right triangle)

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Observational management</th>
<th>Prostatectomy</th>
<th>Conventional radiotherapy</th>
<th>Conventional radiotherapy hypofractionated</th>
<th>Conformal LD radiotherapy</th>
<th>Conformal HD radiotherapy</th>
<th>Conformal LD radiotherapy hypofractionated</th>
<th>Cryotherapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observational management</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Prostatectomy</td>
<td>2.27 (1.34,3.90)</td>
<td></td>
<td></td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.66 (0.22,2.00)</td>
</tr>
<tr>
<td>Conventional radiotherapy</td>
<td>–</td>
<td>–</td>
<td>1.00 (0.34,2.09)</td>
<td>0.91 (0.54,1.51)</td>
<td>1.19 (0.69,2.11)</td>
<td>1.41 (0.49,4.07)</td>
<td>0.66 (0.22,2.00)</td>
<td>–</td>
</tr>
<tr>
<td>Conventional radiotherapy hypofractionated</td>
<td>–</td>
<td>–</td>
<td>1.01 (0.34,3.00)</td>
<td>0.90 (0.27,2.97)</td>
<td>1.19 (0.36,4.04)</td>
<td>1.41 (0.31,6.44)</td>
<td>0.66 (0.14,3.04)</td>
<td>–</td>
</tr>
<tr>
<td>Conformal LD radiotherapy</td>
<td>–</td>
<td>–</td>
<td>0.80 (0.43,1.51)</td>
<td>0.90 (0.27,2.97)</td>
<td>1.32 (0.97,1.86)</td>
<td>1.56 (0.61,4.09)</td>
<td>0.73 (0.21,2.52)</td>
<td>–</td>
</tr>
<tr>
<td>Conformal HD radiotherapy</td>
<td>–</td>
<td>–</td>
<td>1.53 (0.62,3.82)</td>
<td>1.28 (0.93,1.86)</td>
<td>1.18 (0.48,2.92)</td>
<td>0.55 (0.16,1.91)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Conformal LD radiotherapy hypofractionated</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1.17 (0.48,2.91)</td>
<td>0.47 (0.10,2.15)</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Cryotherapy</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Lower-left results compare row-defining interventions against column-defining interventions. Upper-right results compare column-defining interventions against row-defining interventions.

The comparison of prostatectomy with observational management was not linked to the rest of the network, so this evidence was included in the meta-analysis of direct comparisons only. In the meta-analysis of direct comparisons (reported in lower-left triangle), between-trial heterogeneity $\tau^2$ was estimated as 0.04 (95% interval 0.003 to 0.29). In the network meta-analysis (reported in upper-right triangle), between-trial heterogeneity $\tau^2$ was estimated as 0.04 (95% interval 0.002 to 0.26).

The small superscript numbers in the cells of lower-left triangle indicate the numbers of randomised trials which compared the two interventions directly.

HD, high dose; LD, low dose.
reported after this systematic review. Our choices of measurements may have favoured some treatments over others: for example, the RTOR scale had been used to define the late GI and late GU toxicity in the included studies, but it does not measure incontinence which could be the most common adverse event postprostatectomy.96

Methodologically, we used informative prior distributions based on external evidence for heterogeneity variances to increase precision in their estimation and improve estimation of treatment differences. Data-based informative priors have previously been considered by Lu and Ades,97 who used them for the between-study correlation structure. To our knowledge, our paper is the first application of network meta-analysis incorporating data-based informative priors for between-study heterogeneity.

Our findings have implications for research funding prioritisation and study design, and for clinical practice. The study identified particular ‘weak links’ in the network of comparative treatment options, which might be prioritised for future investment in RCTs. This is particularly the case for studies comparing HIFU (which currently is bereft of any comparative evidence) or brachytherapy against other treatment options, and also for trials examining the comparative efficacy and safety of prostatectomy versus conformal radiotherapy modalities. For clinicians, and for men diagnosed with prostate cancer, our findings highlight that the optimal treatment options may be different in respect of different outcomes: patients need to be given appropriate information about the uncertainty surrounding treatment choice currently, and be allowed to opt for “trade-offs” between efficacy and safety outcomes as they judge appropriately.98 Observational studies have consistently shown that radical prostatectomy has better cause-specific mortality outcomes compared with radiotherapy.99–103

In conclusion, clinically important information from high-quality randomised trials is still needed to inform decision-making regarding primary treatment options for men with localised prostate cancer. The findings of this study highlight the importance of informed patient choice and shared decision-making about treatment modality and acceptable trade-offs between multiple outcomes. The upcoming results of the ProteCt study,13 which is evaluating effectiveness of multiple therapies in men with PSA-detected localised prostate cancer, together with other treatment studies in progress, will hopefully contribute to the evidence base. It is, however, unlikely that evidential uncertainty about all relevant and important outcomes will be resolved by these trials, and an updated network meta-analysis incorporating new evidence may be useful to synthesise the new with the existing evidence. We demonstrate a high degree of uncertainty about treatment superiority in the management of localised prostate cancer. Clinicians and patients need to grapple with this uncertainty in the context of shared decision-making.

Author affiliations
1Department of Oncology, University of Cambridge, Addenbrooke’s Hospital, Cambridge, UK
2MRC Biostatistics Unit, Institute of Public Health, Forvie Site, Cambridge, UK
3MRC Clinical Trials Unit, London Hub for Trials Methodology Research, London, UK
4School of Computing and Mathematics, Plymouth University, Plymouth, UK
5Department of Public Health and Primary Care, Cambridge Centre for Health Services Research, University of Cambridge, Cambridge, UK
6Centre for Reviews and Dissemination, University of York, York, UK

Contributors TX, RMT, GL and JPTH conceived and designed the study. TX performed the literature searches. TX, RMT, YW, GL and JPTH performed the literature review and data extraction. TX, RMT, YW, GL and JPTH analysed the data. TX wrote the first draft of the manuscript. TX, RMT, YW, DEN, GL and JPTH contributed to the writing of the manuscript.

Funding TX was funded by the National Institute for Health Research Health Technology Assessment (NIHR HTA) programme (HTA 96/20/99). RMT was supported by Medical Research council grant U105285807. GL was supported by a Post-Doctoral Fellowship Award of the National Institute for Health Research (PDF-2011-04-047).

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/3.0/

REFERENCES


