Author's response to reviews

Title: Factors predicting hospital length-of-stay after radical prostatectomy: a population-based study

Authors:

Maria Kelly (m.kelly@ncri.ie)
Linda Sharp (linda.sharp@ncri.ie)
Fiona Dwane (f.dwane@ncri.ie)
Tracy Kelleher (t.kelleher@ncri.ie)
Frances J. Drummond (f.drummond@ncri.ie)
Harry Comber (h.comber@ncri.ie)

Version: 2 Date: 26 March 2013

Author's response to reviews: see over
Dear Mr. Danrolf de Jesus

Factors predicting hospital length-of-stay after radical prostatectomy: a population-based study

Thank you for your letter. We are pleased to submit a revised manuscript addressing the issues raised by the reviewers.

In responding to the comments of the reviewers, we have made a series of substantive changes to the manuscript, including:

- removing all results relating to trends in radical prostatectomy, so that the paper focuses solely on length-of-stay;
- reducing the number and complexity of the tables to help guide the reader to the important results;
- adding detail to the methods so that the derivation of some of the figures and results is clearer; and
- correcting all inconsistencies and typographical errors.

As a consequence, the paper is now, in our view, substantially improved.

Please find below our detailed responses to each point raised by the reviewers.

We hope that you will find the revised paper acceptable and will be willing to consider it for publication in BMC Health Services Research.

Yours sincerely

Maria Kelly, on behalf of the authors
1. Abstract, Results section: This states that 27% had RP – this is 26% in the main results section.

We apologise for the inconsistency in this figure. In fact, 26.51% (2411/9096) of men diagnosed over the study period had a radical prostatectomy (RP). We have amended the abstract and main results section to read 26.5%

2. Abstract, Results section: The % increased from 20% to 23% - are these figures correct? How can it then be 27% (or 26%) overall?

In light of comments from reviewer 2 (see below), all results relating to trends in RP over time have been removed from the paper.

To provide clarity as regards the specific point raised by the reviewer, the figures 20% and 23% refer to the number of RPs expressed as a percentage of the number of cases diagnosed in a specific year. So, in 2002 942 men were diagnosed with prostate cancer and 189 men underwent RP in that year, which represents 20%. In 2008 1507 men were diagnosed with prostate cancer and 353 men had a RP in 2008, which represents 23%.

The figures 27% refer to the number of men who had a RP at any time between diagnosis and the end of follow-up/censoring date (31/12/2009) expressed as a percentage of all men who were diagnosed with prostate cancer in the period 2002-2008. Men diagnosed with prostate cancer often have repeated PSA tests or prostate biopsies as part of a ‘watchful waiting’ treatment plan and the decision to treat actively by radical prostatectomy can occur a long time after diagnosis.

We have clarified end of follow-up is 31/12/2009 on page 5, last paragraph.

3. The figures in the main results section say 27% of men diagnosed in 2002 and 22% in 2008. It’s not clear if these figures relate to different things.

As noted above, we have removed all results relating to trends in RP over time from the paper.

In terms of the specific point raised by the reviewer, the denominator for these figures are different (please see our response to point 2 above). Censoring occurred at 31/12/2009. Men diagnosed at the start of the study period (i.e. early in 2002) had almost seven years of follow-up to the censoring date at 31/12/2009, while men diagnosed at the end of the study period (i.e. late in 2008) only had just over one year follow-up. In 2002 942 men were diagnosed with prostate cancer 258 of whom had RP in the period up to censoring at 31/12/2009: this represents 27%. In 2008 1507 men were diagnosed with prostate cancer 337 of whom had RP in the period up to censoring at 31/12/2009: this represents 22%.

We apologise that the derivation of these figures was not clear in the original manuscript.

4. Methods, Statistical analysis, paragraph 1: Subjects with missing data were excluded from relevant analyses. Please elaborate more here. How many subjects were excluded?
To clarify, in the original analysis, we did include a missing category for each variable. The numbers categorized as missing were as follows: marital status (n=4), deprivation index (n=139), smoking status (n=328), patient status at discharge (n=112), grade (n=43) and stage (n=1128). With the exception of the variable for marital status the missing category was included in the multivariate analysis for all variables. We excluded the 4 cases with missing information on marital status from the analysis. This information is shown in table 1.

We apologise that this was not made clear in the methods section. We have now clarified how missing covariate data was dealt with in the analysis, on a variable by variable basis, in the methods section (page 8, paragraph 1). The numbers of subjects with missing information are also shown in additional table.

5. Stage and smoking status have high levels of missing data. Should the missing group be included as a separate category? How do the results compare if this is done?

Please see our response to point 4. To clarify, in the original manuscript, we did not exclude subjects with missing information on stage and smoking status at diagnosis from the multivariate analysis. Instead, we included them as a separate level of the categorical variables. This was precisely because of the relative large percentages of subjects with missing information for these variables; to exclude them completely from the analysis would have been likely to have lead to biased results.

We apologise that this was not clear in the original manuscript. The revisions that we have made in response to point 4 should have clarified this issue.

6. Results, paragraph 1: The figures here do not tally with the abstract. Figure 2 needs a little more explanation here.

We apologize for the inconsistencies in the figures quoted in the abstract and the main text. As indicated above, we have now removed all results relating to trends in RP over time. Thus, “Figure 2” from the original manuscript is no longer in the paper.

7. What is the difference between the two bottom lines?

We assume this comment refers to Figure 2, which is no longer in the paper.

In terms of the specific point raised by the reviewer, the green line represented the number of RPs conducted by year, the red line represented the number of men diagnosed by year who subsequently went on to have an RP at any time up to censoring at 31/12/2009.

8. Table 1: This is too big and confusing with too many footnotes. Please condense it down to focus on the most important information (i.e. the ORs). I’m not sure there is any value in presenting a median LOS for every variable.

We have radically simplified table 1. This now contains the results from the final, multivariate, analysis and is limited to the variables which were statistically significant and included in that final model. For those readers who are interested in seeing the detailed univariate and multivariate data for every variable, we have provided this in a supplemental table intended for web publication only (additional table 2).
9. Table 2: This contains important information, but again could be presented in a better way (it is a large table for just a few key numbers).

We have reorganised the information in table 2 so the overall table size is reduced. We have removed blank header rows and adjusted column width. The information it contains has not changed.

10. Figure 2: I have mentioned this above. It’s not clear how the two bottom lines differ. It may be better to have the number of cancers and RPs and show the % by year rather than the numbers.

As indicted above we have now removed all results relating to trends in RP over time. Thus, “Figure 2” from the original manuscript is no longer in the paper.

11. Results, paragraph 3: The last sentence does not make sense. How can median be a range? Please re-word.

We have revised the sentence to improve clarity. It now reads “Median RP volume per surgeon varied from 1 to 17 per year among lower-volume surgeons (annual median \( \leq 17 \) RPs) and from 18 to 36 per year among higher-volume surgeons (annual median \( >17 \) RPs).” (Results section, 2nd paragraph, page 9).

12. Discussion, Factors associated with prolonged LOS, paragraph 2: The associations... has been previously reported (should be have).

We have made this minor drafting revision. The sentence now reads “The observed associations between LOS and provider volume have been previously reported.” (page 15, paragraph 2)

13. Discussion, Factors associated with prolonged LOS, paragraph 3: Have the authors considered analysing the data using multilevel modelling? This would enable them to look specifically at the amount of variation by surgeon and by hospital.

Yes, we did consider using multilevel modelling. However we were concerned we did not have a sufficient sample size for accurate estimation. We would have had small numbers at the group level (10 public hospitals and /or 40 surgeons working in public hospitals). There were also large variations in numbers of individuals within a group (that is, patients clustered by surgeon or patients clustered by hospital). Maas et al report a small sample size at level two (meaning a sample of 50 or less) leads to biased estimates of the second-level standard errors (1).

In addition, the matter was further complicated by the fact that in the Irish health system surgeons can (and often do) work in more than one public hospital. Thus some surgeons would have clustered under multiple hospitals. On reflection, therefore, we decided not the use a multilevel approach.

14. Figure 1: This is quite straight forward and is explained in the text. This could be a supplementary file instead.

We think Figure 1 is important to understanding how the dataset was derived. Therefore, we would prefer to retain it within the main body of the paper. However, if the editor prefers, we would be happy to provide it as a supplementary figure instead.
1. **Reviewer 2** The paper addresses two seemingly separate issues: firstly changes in the incidence of prostate cancer and RP over time, and secondly factors predicting hospital LOS. This is confusing and dilutes the main message of the paper. I would strongly suggest that the two issues are totally separated and that this paper should focus on LOS only. The changes over time are best presented as a separate paper. I have not specifically commented on any subsequent issues relating to the incidence of prostate cancer over time in the rest of this review. I do accept that some issues relating to changes in RP over time that are relevant to the factors that predict LOS. Clearly these need to stay in.

On reflection, we agree with the reviewer and feel that the paper would be clearer if it focussed on a single issue – that is, length-of-stay. We have therefore, removed all results and discussion relating to trends in prostate cancer and RP over time from the paper. The focus and content of the paper is now to investigate factors predicting hospital LOS and readmissions in men who had RP following prostate cancer.

2. Cost effectiveness. Longer LOS may be more cost effective (e.g. if fewer readmissions). This needs exploring, and ideally a proper cost-effectiveness model built. In the last sentence of the third paragraph of the background, the authors suggest that LOS and readmissions are the likely driver of prostate cancer care. This is only partially true. What is the cost of a 10-day admission rather than a 9-day admission, and what is the difference in re-admission rates. It is a complex relationship that needs teasing out with modelling rather than conjecture. Without a full health economic model it is impossible reach the implications that the authors do.

We agree with the reviewer and did not intend to suggest a shorter LOS is more cost effective per se.

We have redrafted this part of the discussion. We have said that given the variation in volumes of RP carried out by hospitals and surgeons, centralisation might allow the implementation of clinically-based (e.g. standardized critical care pathways) and service-based (e.g. discharge, follow-up and elective readmission policies) initiatives to improve patient care that might simultaneously reduce costs. We have added a new heading called ‘Centralisation’ at the end of page 16 to make this point and referenced the need for a thorough evaluation of the cost and benefits of re-organising RP services (page17).

We have removed the sentence “This suggests that hospital length-of-stay following RP and any associated readmissions are likely to be a major driver of the costs of prostate cancer care” from the Introduction section.

3. The analysis would benefit by exploring the association between LOS and readmission.

We consider that it is important to distinguish between emergency and elective readmissions, especially since the latter varies so much by hospital, suggesting that overall rates of re-admission are driven (to some extent) by hospital policies (see below). For this reason, we do not consider it would be appropriate to develop a model to investigate associations between LOS and risk of readmission per se.

Emergency readmissions following RP are relatively uncommon. There were only 47 emergency readmissions with 28 days of discharge after the initial RP (3.1% of all men undergoing RP and only 5.5% of all readmissions); this is shown in table 3. In those men who had a prolonged LOS (>9 days) at the index episode 4.0% were readmitted as emergencies within 28 days (that is 15 emergency readmissions out of 375 men with prolonged LOS at the index episode). In those whose LOS was less than 10 days at the index episode 2.8% were readmitted as emergencies within 28 days (that is 32
emergency readmissions of 1160 men). These two percentages did not differ significantly (p=0.225). We have now provided this information on page 10, 3rd paragraph.

Even if we followed men for 90 days after discharge, only 59 were readmitted as emergencies. We, therefore, felt that the sample size was insufficient to permit us to build a multivariate model of the relationship between LOS and risk of emergency readmission.

Within 28 days of discharge, there were 807 elective readmissions (52.6% of all men undergoing RP and 94.5% of all readmissions), of which 503 (58.9% of all readmissions) were overnight readmissions and 304 (35.6% of all readmissions) were day cases, (table 3). The majority of overnight elective admissions were in higher volume hospitals while the majority of day case admissions were in lower volume hospitals. We believe this difference is a reflection of individual hospital policy on readmission. In multivariate analyses we would not be able to identify the true association between LOS and risk of elective readmission (if one exists); we, therefore, did not investigate this.

4. It is predictable that the results as they stand are largely self-evident and entirely predictable (i.e. longer LOS associated with unmarried, co-morbidity, and lower volume hospitals and surgeons).

We accept that aspects of our findings are predictable, but in itself, this is not an argument not to publish them. Our study is population-based and represents routine practice across an entire health system: there are few such series in the literature. The potential difficulty with series from a single institution is that they are likely to reflect local practice, include a quite selected group of patients, and thus to have limited generalisibility.

In terms of LOS, we provide further evidence to show that there is variation in LOS and the factors which are associated with this (notably hospital and surgeon volume, providing clear evidence of benefits of centralisation of services). Taken together, the accumulation of evidence can be of value in informing health service planners and providers where they may need to provide additional support and services if they wish to reduce LOS.

We also show – for the first time as far as we are aware - striking variations in elective readmissions across hospitals. For Ireland, these finding suggest that standardised follow-up care for all men having RP might be of benefit. In international terms, our results may stimulate investigation of patterns of readmission following RP in other healthcare settings.

5. The authors present medians and 75th centiles. These transpire to be only different by a day, so are largely meaningless clinically. Perhaps a better cut off maybe the 90th centile. Would these outliers tell us more about LOS? As it stands, the authors need to make a clear differentiation between statistical and clinical significance.

We accept that the 75th percentile is not particularly informative in reporting the descriptive data and now present medians with 10th and 90th percentiles (see additional file 2). We have retained the inter-quartile ranges so that it is clear to the reader the derivation of the cut-off used to define prolonged LOS (upper quartile) in the primary analysis.

The 90th centile was 13 days. To investigate whether it would be more clinically informative to focus on reducing LOS in the “more extreme” patients, we undertook a sensitivity analysis using the 90th centile to define prolonged LOS (>12 days); this represented a 5 day difference from the overall median LOS. Using this definition, 161 men had prolonged LOS. The factors which significantly predicted longer LOS were the same as in the primary analysis. The effects for all variables were in
the same direction as in the primary analysis. The risk estimates for marital status and surgeon volume were slightly further from unity than in the primary analysis. Men who were not married (OR=2.13, 95% CI 1.43-3.20), had co-morbidities (OR=1.64, 95% CI 1.13-2.38) or later stage III-IV (OR=1.98, 95% CI 1.12-3.50) were significantly more likely to have prolonged LOS. Those treated in higher volume hospitals (annual median > 49 RPs) or by higher volume surgeons (annual median > 17 RPs) were significantly less likely to have prolonged LOS (OR=0.31, 95% CI 0.21-0.45; OR=0.37, 95% CI 0.25-0.54 respectively).

We have described this sensitivity analysis in the Methods (page 8, paragraph 3) and the results are in the second last paragraph of the Results section, (page 11, 2nd paragraph).

6. Conclusion. The argument for centralization is tricky. It depends upon whose perspective this is from. Patients may accept longer stays closer to home, and there may be cost-benefits to some longer LOS. The real issue about centralisation is the quality of the surgery and LOS is only one of several proxy measures for this.

We agree with the reviewer that the issue of centralisation is not straightforward or trivial and that there are different perspectives that need to be considered. We did not intend to suggest that LOS is the only or most important metric to consider in this regard. Having said this, our study does provide evidence that adds to the existing evidence base on centralisation, and which might help convince service planners and providers (and also, perhaps, patients) about the benefits of centralisation.

Urologists agree that certain minimum volumes need to be maintained to maximise patient outcomes (2). In England the National Institute for Clinical Excellence (NICE) recommends that RP should be provided by specialised teams typically carrying out 50 or more operations per annum (3). In Ireland – as in other small countries - the only way these numbers of operations can be achieved is to centralise services to a few hospitals.

We have noted these points in the discussion (page 15, paragraph 2). In addition, as noted above (in response to point 2), we have redrafted this part of the discussion to emphasise the fact that different outcomes and perspectives need to be considered. We also state that there is a need for a thorough evaluation of the cost and benefits of centralisation of RP services (page 17).

7. The analysis needs to take into account the different surgical techniques and how these have changed over time

Unfortunately we do not have any information on different surgical techniques for RP. We discuss this in the last paragraph of the section “Comparisons in LOS between countries.”, page 13, last paragraph.

8. Why were patients with previous cancer excluded? This may have been entirely curative and may have been years ago and of no consequence for this analysis.

To clarify, cases where the prostate tumour was a second, or subsequent, primary cancers (other than non-melanoma skin cancer) were excluded. In our study 794 of all men diagnosed with prostate cancer between 2002 and 2008 had a previous primary cancer; 110 (14%) of these men had RP surgery, 94 of them in a public hospital; these 94 men were excluded.

63 of these 94 men had a cancer in the 12 months prior to the prostate cancer diagnosis (56 within the previous 6 months and 7 between 6-12 months before the prostate cancer diagnosis). More than
95% of cancers that occurred in the year prior to the prostate cancer diagnosis were bladder cancers (C67).

The reason for excluding these 94 men was that the treatment of the previous cancer may have impacted on the treatment for the prostate cancer and hence on LOS. Thus our results are generalisable to the population of men undergoing RP for whom the prostate cancer is their first invasive cancer.

We have described numbers excluded in the “Methods” section (page 5, last paragraph). We describe the rationale for this exclusion and how this affects the generalisability of our study finding in the 3rd paragraph on page 12.

9. Why the cut-off at 70 rather than 75 or 80? Life expectancy still quite long in these age groups.

This cut-off was informed by clinical practice in Ireland. RP is very uncommon in men aged 65 or older. The frequency declines dramatically after 65 and almost no procedures are conducted in men aged 70 or over. In the years of our study, only 27 men aged 70 or more had RP. This suggests that RP is not routinely offered to those aged 70 or over and – because of this – we limited our study population to men aged less than 70 years. We have also inserted reference to a paper currently under review which clearly shows the pattern of RP by age group in Ireland; we expect that paper to be in press by the time the current paper has been reviewed.

10. It seems odd that a hospital only has to do one RP a week to be high-volume, and a surgeon only one every three weeks. What was the range / variation by hospital and by surgeon?

Our definition of “high” and “low” volume was driven by the low absolute numbers of RPs in Ireland and the distribution of the data. We defined “higher-volume” as >49 RPs per annum for hospitals and >17 RPs per annum for surgeons. For each hospital the annual volume of RPs performed and the median for the entire period was determined. Hospitals were classified as “lower” or “higher” volume by ordering hospitals on the period median volume and splitting so that approximately 50% of patients treated fell into each category. The same approach was used for surgeon volume. This information is in the “Methods” under the section headed “Data Sources”, last paragraph, page 6.

We do accept that our cut-off for “high volume” may not represent what would be considered high volume in absolute terms and/or in other healthcare systems. [We would note, however, that the cut-off for higher hospital volume is consistent with the recommendation of NICE in England that RP be provided in teams which conduct at least 50 procedures per annum]. To address this we have revised our terminology to refer to “higher-volume” and “lower-volume” throughout instead of high-volume and low-volume. We also note explicitly that compared to other healthcare settings, the number of RPs in Ireland is low (first paragraph under heading “Centralisation” on page 16).

In terms of range and variation, the number of hospitals performing RP ranged from a low of 16 in 2002 (8 public, 8 private) to a high of 20 in 2007 (10 public, 10 private). Half of all RPs carried out in public hospitals were done in two facilities; the remainder were carried out in eight institutions. In these eight hospitals the median volume varied from 1 to 43 RPs per year. In total 49 surgeons performed RP over the study period, 9 worked in private hospitals only, 9 worked in both public and private institutions and 31 worked in public hospitals only. The number of surgeons performing RP ranged from 20 in 2002 to 27 in 2008 and was highest (n=33) in 2007. Eight surgeons performed
50% of all public hospital RPs. Median volume per surgeon varied from 1 to 17 per year among lower-volume surgeons and from 18 to 36 RPs per year among higher-volume surgeons. This information is provided in the 2nd paragraph on page 9.

11. Might the casemix (i.e. potentially more complex cases) in the high volume hospitals be a confounder here?

The reviewer is correct – case-mix is indeed likely to be a confounder of the relationship between LOS and hospital volumes. We explicitly refer to this possibility in the last paragraph of the section ‘Factors associated with prolonged LOS’ on page 15, as follows:

“Although we found that both hospital and surgeon volume were statistically independent predictors of LOS, in practice these are likely to be related; higher volume centres are more likely to have more experienced surgeons who in turn may drive quality maintenance or improvements in perioperative care. However, higher-volume centres may also deal with more complex cases which cannot be adjusted for fully, using simple co-morbidity counts. Unfortunately we do not have information on other predictors of case complexity (e.g. frailty), and so we may have underestimated the strength of association between higher volume and lower LOS.”

12. Although the missing episodes were distributed across hospitals and years, they are non-random (in that there is always a cause for missing episodes) so it is probably incorrect to say that they are unlikely to be a cause for bias.

We have revised this sentence and it now reads, “The missing episodes were distributed across hospitals and years.” This can be found in the 2nd paragraph on page 12, last sentence.

13. The findings are really quite different when compared with data from the US (median 3 days). What are they doing in the US to achieve this, and what are their readmission rates (and other proxy measures for quality)?

The median 3 days from the US is post-operative stay while our results are based on combined pre-operative and post-operative stay. If we restricted consideration to post-operative stay only, then median LOS was 6 days with the upper quartile cut-off at 8 days. This information is shown in the 3rd paragraph of ‘Results’ section, top of page 10.

In the US, there has been a concerted effort to reduce post-surgery LOS following RP using collaborative care pathways. This, together with differences in casemix, type of procedure performed and public and private healthcare cultures, probably explains the higher LOS in Ireland than in the US. We refer to this in the in the 2nd paragraph of section ‘Comparison of LOS between countries’, page 13.

14. There is considerable over-discussion and over-interpretation of the results.

We have now focussed our discussion on the main results and how they compare with those from other healthcare settings.

15. Background – 1st sentence. The ‘most common’. Really?
We confirm that this statement is correct and is based on the best available international data. We derived this information from GLOBOCAN which is a programme of the International Agency for Research on Cancer. GLOBOCAN has compiled the best available cancer data internationally and used this to make estimates of the worldwide cancer burden. We provide the reference to GLOBOCAN (‘Ferlay J, Shin HR, Bray F, Forman D, Mathers C and Parkin DM. GLOBOCAN 2008, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10 [Internet]. Lyon, France: International Agency for Research on Cancer; 2010 Available from: http://globocan iarc fr, accessed on 01/09/2011’)

16. I would appreciate a greater explanation of the Elixhauser index, and in particular why this was used in preference to the more widely used Charlson index.

The Charlson comorbidity index was developed as a method for predicting risk of death in the coming year according to a score based on applying weights to series of comorbid conditions. It was originally developed in a cohort of medical patients (Charlson et al., 1987) and includes 17 comorbid conditions which were associated with 1-year mortality.

The Elixhauser method includes 31 comorbid conditions and was developed and validated using administrative data encompassing all adult acute-care inpatient hospitalizations that occurred in 1992 in California (4). The Elixhauser comorbidity measure was developed to predict a range of outcomes including in-hospital mortality, hospital charges and length-of-stay. We felt that it was, therefore, likely to be a more appropriate measure of comorbidity in the current study than the Charlson index. In the paper, we provide a reference to the Elixhauser index (page 6, reference 26 in paper).

17. Last para under ‘data sources’ best as an acknowledgement

We have moved the paragraph shown below to the ‘Acknowledgements’ section.

“The NCR has permission under the Health (Provision of Information) Act 1997 to collect and hold data on all persons diagnosed with cancer in Ireland. The use of that data for research is covered by the Statutory Instrument which established the Registry Board in 1991. The NCR were give permission to access anonymised HIPE data under the Health (Provision of Information) Act 1997. Permission was given by the Department of Health and Economic and Social Research Institute (ESRI) who are joint custodians of the data. All datasets were anonymised prior to analysis. “

18. ‘Gender’ appears in the model. I can only assume that this is a mistake to keep reviewers on their toes.

We apologise for this typographical error. It has been corrected.

19. Please clarify why ‘admission type’ in the model. There is no such thing as an emergency RP, surely?

We apologise for this typographical error; to clarify “admission type” was not considered in the main analysis. We have corrected the text to remove this error. In the analysis of readmissions we distinguish between elective and emergency readmission.

References

