

DETAILED COMMENTS ON SECTION 7 OF THE ScHARR PROSTATE CANCER REPORT AND RELATED APPENDICES

SUMMARY

One of the comments that the PC Support Federation raised on the ScHARR report was that it had been inadequately verified. These notes are intended to respond to that shortfall by looking in detail at Section 7 of the report and the related appendices.

It is shown that there are two clear examples of mathematical errors in the ScHARR results (which are presumably due to programming errors).

- According to the ScHARR model, all of the screening Policies lead to a very large increase in the number of hormone therapy treatments. This is counterintuitive. I did a spreadsheet calculation using the ScHARR data and modelling assumptions to check the ScHARR results and this shows that there should be a large decrease in the number of hormone therapy treatments due to screening.
- Let us consider the group of men who would have been diagnosed clinically with metastatic cancer but are instead screened and diagnosed with a local cancer. According to the ScHARR results, these men should die on average 1.5 years younger because of screening! However, this is precisely the group of men who should gain most from early diagnosis by screening.

In addition, there are a number of examples of ‘inadvertent’ bias introduced into the modelling by the selection of input data. As examples,

- There is an assumption that treatments would be offered in the same proportion for screened patients as prior to screening. The modelling shows that screening would find a large number of dormant cancers. It seems unreasonable to assume that there would be a major increase in the number of radical treatments offered in the UK in order to treat these dormant cancers. It is more likely that there would be a substantial increase in the number of men on Active Monitoring.
- More generally, the ScHARR report is backward looking in the range of treatments that are offered. For example, it does not take account of HDR brachytherapy which improves the outcomes of men with locally advanced cancers. The report should consider the range of treatments that will be available if screening were introduced rather than those that were available some years ago.
- There is an assumption that men who are screened at 50 would receive the treatments in the same proportion as currently received by men in their late 60’s. As a consequence, the modelling exaggerates the number of 50 year old men with local cancers who would only be offered hormone therapy.

INTRODUCTION

In these notes I am using spreadsheet calculations to check the accuracy of the computer results given in the SchARR report. These checks are targeted on results which appear to be anomalous.

I have tended to concentrate on Policy 1 (screening once at 50) simply because this is more suitable for spreadsheet calculations than the repeated screening policies. This policy also has the advantage of separating out issues associated with age, since 50 year old men are clearly not elderly and, but for the prostate cancer, there is a high probability that they would live for many years.

CALCULATION OF THE NUMBER OF ADDITIONAL PATIENT INTERVENTIONS FOR POLICY 1 (I.E A SINGLE SCREENING AT 50)

For Policy 1, there are 2280.5 cancers detected at screening at 50 that would otherwise have been detected clinically (Table 19). There are an additional $(2773.1 - 2280.5 =)$ 492.6 cancers detected because of the policy of screening at 50, and this corresponds to the 18% overdetection for this policy (data from Tables 17, 18, 19).

Appendix J implies that, for Policy 1, there is an increase of 578 in the number of patients receiving hormone therapy. This is counterintuitive since the men screened with dormant cancers do not require HT and the men with active cancers are treated earlier than otherwise, because of the screening, and so are less likely to need HT.

The report gives sufficient information to be able to calculate the number of additional interventions for screening compared to ‘no screening’ for Policy 1.¹ The following table is a comparison of my spreadsheet calculation with the results given in Appendix J for the “Additional patient interventions for screening compared to no screening”.

| | SchARR report | Verification calculation | Difference |
|--------------------------------------|---------------|--------------------------|------------|
| Radical prostatectomy | 377 | 320 | 57 |
| Radical radiotherapy | 218 | 193 | 25 |
| Radical radiotherapy + HT | 1 | -24 | 25 |
| Hormone therapy | 578 | -683 | 1261 |
| Active Monitoring / Watchful waiting | 298 | 327 | -29 |

¹ My calculation uses the information in the SchARR report Tables 18 and 19 for the distribution of screen and clinically detected cancers. The distribution of patient interventions used was from the first table on page 115. The note, on page 115, assumes that 15% of patients who chose Active Monitoring go on to have radical treatment; I included this using the same distribution between the radical treatments as given in the first table of page 115. For the locally advanced cancers I used the treatment distribution given in Table 10. For the metastatic cancers I assumed that each man would have HT. The distribution of Gleason scores are given in Table 18 for the screen detected cancers. For the clinically detected cancers, I used the distribution for clinical detection for men aged 65 given in the first sub-table of Table 41 in Appendix I (this allows for a 15 year lead time between screen and clinical detection, as suggested in Table 17). According to the SchARR modelling (Table 20) the relative rate of death from PCa is identical with and without screening at 50, thus there would be the same number of patients receiving HT because of cancer progression after radical interventions.

| | | | |
|-----------------------|-----|-----|----|
| Other local treatment | 465 | 409 | 56 |
|-----------------------|-----|-----|----|

Table 1

It is not surprising to find small differences arising from detailed differences in the assumptions that I made compared to those used in the ScHARR model. However, there is a gross difference in the number of patients who receive hormone therapy. This is clearly an error in the ScHARR modelling.

This large reduction in the number of men having hormone therapy as a result of screening (using my figures) arises because 823.9 men (Table 19) are diagnosed with local cancers at screening that would otherwise been detected clinically at a later stage. Thus, the large increase in radical treatments, shown in the table, effectively displaces an equivalent number of hormone treatments that would have been given to the same men if their cancers had not been diagnosed by screening.

Appendix J, similarly suggests there are very large additional numbers of patients receiving hormone therapy for the other Policies. It follows from the above discussion that these are also erroneous and most probably all arise from the same ‘common cause’, such as an error in the computer modelling.

In the above calculation, I used the same data as in the ScHARR modelling. However, it is worth emphasising that this does not give best-estimate results.

For Policy 1, 18% of the patients, diagnosed through screening, would be ‘over detected’ with lead times of 12-15 years between the screened diagnosis and a clinical diagnosis. If clinicians were given this information then they would clearly increase the levels of Active Monitoring amongst men with local cancers and lower Gleason scores. In the ScHARR report, the BAUS data from 2008 is used to determine the distribution of treatments offered. In practice, this would not be directly applicable in the changed circumstances if screening were introduced. In these circumstances, the direct use of the BAUS 2008 data leads to a substantial exaggeration of the numbers of men who would receive radical treatments and thus introduces a bias into the report.

A more realistic estimate of the distribution of treatments can be achieved by scaling the BAUS 2008 results to make allowance for the ScHARR estimates of the numbers of dormant cancers detected by screening. For example, if Policy 1 was introduced, 492.6 of the 2582 local cancers would be ‘over detected’. Within the example that follows, the emphasis is on 50-year old men and I am making the assumption that the men with dormant cancers would have Gleason scores <7.² The modified distribution of treatments for patients screened under Policy 1 is given in the following table.

| 50-year old men with Gleason <7 | BAUS 2008 | Modified distribution for screened patients following introduction of Policy 1 |
|---------------------------------|-----------|--|
| Radical prostatectomy | 22.22% | 17.98% |
| Radical radiotherapy | 13.71% | 11.09% |

² The report by Hardie *et al* (2005), also included a small number of patients on Active Monitoring who had a Gleason score of 7. It would readily be possible to take this into account and make an analogous adjustment to the treatment distribution for such patients.

| | | |
|--------------------------------------|--------|--------|
| Radical radiotherapy + HT | 0.00% | 0.00% |
| Hormone therapy | 0.00% | 0.00% |
| Active Monitoring / Watchful waiting | 30.67% | 43.90% |
| Other/unknown | 33.40% | 27.03% |

Table 2

This is an example and, of course, there is scope for discussion around which precise set of numbers are used. However, the principle is that the distribution of treatments would change, following the introduction of screening, because of the higher numbers of ‘over detected’ cancers.

The ScHARR report assumes that 15% of all Active Monitoring patients go on to have radical treatments. This is based on the study by Hardie *et al* (2005) where 11 out of 75 (who survived the trial period) had radical treatments. However, 9 of these 11 had radical treatments because of the rate of rise of PSA levels and only 2 because of patient preference. If screening were introduced, there would be a large increase in the number of men with dormant cancers on Active Monitoring and these men would, of course, not show a significant rate of rise in PSA levels. Therefore, if there was screening, the assumption that 15% of men would go on to have radical treatments would represent a substantial overestimate and therefore introduces a bias into the calculation.

ANALYSIS OF THE ScHARR RESULTS TO DETERMINE THE AVERAGE NUMBER OF YEARS GAINED BY GROUPS OF MEN FOLLOWING SCREENING

The most striking finding of the ScHARR modelling is that it shows very poor outcomes following the screening. For Policy 1, the average number of years gained per person with a screened cancer that would have been diagnosed later with clinical PCa is only 0.5 years.³ This is surprisingly low and is worth investigating further. The aim here is to analyse these outcomes in more detail for the groups of men identified in Table 19 of the ScHARR report.

For convenience, I have copied the relevant part of Table 19 for Policy 1,

| | Once at 50 | Screen detected | | | Total |
|----------|------------|-----------------|------------------|------|--------|
| | | Local | Locally Advanced | Mets | |
| Clinical | Local | 1295.2 | | | 1295.2 |
| | LA | 352.5 | 51.0 | | 403.4 |
| | Mets | 471.4 | 106.2 | 4.2 | 581.8 |
| | Total | 2119.1 | 157.1 | 4.2 | 2280.5 |

Table 3

According to the modelling, there are 471.4 (21%) men who were found to have a local cancer at screening who would not have been diagnosed clinically until they had metastatic cancer. Similarly, a further 352.5 (15%) men who had a local cancer at screening, who would not have been diagnosed clinically until their cancer was

³ Based on data from Table 22, (=1127/2280).

locally advanced. The natural assumption is that these two sets of men would gain the most from the screening. If it is assumed that all of the 1127 years gained as a result of screening (Table 22) arise from these two sets of men then the average number of years gained is still only 1.4 years. This is still unreasonably low based on patient experience.

It is possible to analyse the results of the ScHARR report and investigate this further. Table 19 shows analogous data to those given above for four different policy options. In each case, there four different sets of possibilities,

- The screening detects a local cancer that would otherwise have been detected clinically as locally advanced.
- The screening detects a local cancer that would otherwise have been detected clinically as metastatic.
- The screening detects a locally advanced cancer that would otherwise have been detected clinically as metastatic.
- The screening detects a cancer at the same stage that it would have been if had been detected clinically.

The report does not say how much life is gained for each of the these possibilities. However, we can regard the average lengths of life gained as unknowns and since the ScHARR report gives us four sets results for the total number of life years gained (Table 22) we can solve these as a set of simultaneous equations.⁴ This shows that,

| | Groups of men where:- | Years gained |
|---|---|--------------|
| 1 | screening that detects a local cancer that would otherwise have been detected clinically as locally advanced | 5.0 |
| 2 | screening that detects a local cancer that would otherwise have been detected clinically as metastatic | -1.5 |
| 3 | screening that detects a locally advanced cancer that would otherwise have been detected clinically as metastatic | 0.6 |
| 4 | screening detects a cancer at the same stage that it would have been if had been detected clinically | 0.0 |

Table 4

In the case of group 4, this gives the reasonable expectation that there would be little or no benefit from screening for men who are detected at the same stage through screening as they would have been if the cancer had been detected clinically.

The most striking problem is for the men in group 2. These are men whose cancers are detected by screening at a local stage that would otherwise not have been detected clinically until they reached a metastatic stage. These men should see the greatest benefits of screening. Therefore, it is deeply puzzling why the ScHARR model implies that these men will die, on average, 1.5 years sooner as a result of the earlier diagnosis. This is clearly an error, though it is difficult to understand how such a major error can have crept into the analysis. The most obvious suggestion is that there

⁴ This is legitimate since the age bands for varying treatment are extremely coarse and the modelling assumes that all men who are less than 70 have the same treatment profiles (see Table 10 and Appendix D).

is an error in the computer algorithm were treatments are allocated to these men (such as giving these men only hormone therapy rather than the range of radical treatments).

Similarly, there is also a problem for the men in group 3. These men have locally advanced cancer. The distribution of treatments within the modelling is based on the BAUS data, which suggest that for men under 70, only 26% would be offered RT+HT and the remaining 74% would receive only HT. In other words, despite being diagnosed at an earlier stage, in the modelling 74% of these men would receive the same treatment as if they had been diagnosed with metastatic cancer.

The problem here possibly lies in the coarseness of the age distribution within the BAUS data which are used to describe the distribution of treatments. In practice, 50-year old men would be more likely to receive radical treatments for locally advanced cancers than men who are almost 70 years old. In addition, the BAUS data reflect treatment patterns from 2008 before screening is introduced. These treatment patterns would change if the age of diagnosis of patients was systematically reduced by 12-15 years (based on the lead times quoted in Table 17). Furthermore, the ScHARR report effectively neglects high dose rate brachytherapy by assuming that it is equivalent to external beam radiotherapy. Åström *et al* (Radiotherapy & Oncology 74 (2005) 157-161) show that 61% of high risk patients who have HDR brachytherapy have no biochemical evidence of disease within 5 years.⁵ By not taking account of the shift in the age distribution and the newer treatments, the data used in the ScHARR report is effectively introducing a bias into the modelling.

According to the ScHARR modelling, Group 1 is the only group of men who would gain a significant benefit from early diagnosis and treatment. These are men that would have been clinically diagnosed as locally advanced if they had not been screened and diagnosed with local cancers. However, even in this case, the gain is fewer years than would be expected (because the modelling assumes that 74% of these men would only have received HT if they had been diagnosed clinically with locally advanced cancers). Instead all of these men would receive radical treatments that are potentially curative.⁶ For Policy 1, where men are screened once at 50 years old, a curative treatment could potentially extend their lives considerably beyond 5 years, which would lead to a large increase the average number of years of life gained for this group.

We can get some insight where the modelling is deficient if we consider the data in Appendix D. It is convenient here to continue to consider Option 1. Appendix D shows that the modelling assumes that 34-44% of all 50 year old men who are diagnosed with localised cancers and a Gleason score >7 are given hormone therapy with no radical treatment. It seems highly unlikely that such a high proportion of 50 year old men would be refused radical treatment for localised cancers. As an example,

⁵ I should declare a personal interest here. When I had HDR brachytherapy with external beam radiotherapy+HT for locally advanced PCa (aged 56, T3b, PSA=56.4, Gleason=7), my oncologist, Heather Payne said I had a 75% probability of no biochemical evidence of disease within 5-7 years. I am currently 4 years after treatment and have a PSA of 0.3 and have not had HT for the past 2 years.

⁶ Taken literally, the modelling would assume that a percentage of these men would have Active Monitoring. However, because all of these men on AM would show increases in PSA, it follows that they would all be offered radical treatments.

Åström *et al* (2005) show that with HDR brachytherapy, 82% of these men would survive for 5 years with no biochemical evidence of disease.

ADDITIONAL VERIFICATION POINTS

The data in Appendix O, the probabilistic sensitivity analysis, is not consistent with data in Section 7. For example, the mean relative rate of PCa death (50-74 years) is given as 0.87 in Appendix O, and 0.84 in table 20 of Section 7. There are a number of other examples which would suggest that these data are based on a different set of computer modelling assumptions.

In Appendix O, the units for the tables of PCa age specific incidence are not specified. These are small numbers given to two decimal places and so are clearly not counting numbers of men.

In Appendix I, the numbers in table 41 have no obvious relationship to those presented in Section 7. As an example, in Table 41, there are a total of 608 cancers detected in the screening at 50 (for the policy of screening once at 50), whereas in table 19, there are 2280.5 cancers detected for this group. I could find no link between the numbers in Table 41 and those used elsewhere in the paper. [It would be helpful if there was labelling describing the contents of the rows (at least for the first of the tables that together make up Table 41).]

It would be very useful if Appendix I contained a table equivalent to Table 19 giving the stage shift distribution for the full range of policy options considered (and not just the four options that are highlighted in the main text).

Following the discussion above, it would be extremely useful if the ScHARR report quoted the average life years gained for either the four groups of men highlighted in Table 4 above, or for the 6 groups of men identified in Table 3 above (based on Table 19 of the ScHARR report).

The paper is not at all clear which data have been used in Section 7 to model the outcome of treatments. Despite searching through the paper, I have no idea what data were used for the outcomes.

Section 7 does not consider the negative effects of radical treatments. However, if the results from Section 7 were corrected to remove mathematical errors and bias in the input data, then this would lead to a reduction in the levels of adverse side effects.