PERGAMON

European Journal of Cancer 37 (2001) 2154–2160

European Journal of Cancer

www.ejconline.com

Review

Health-related quality of life and cost-effectiveness studies in the European randomised study of screening for prostate cancer and the US Prostate, Lung, Colon and Ovary trial

A.B. Miller[a,*], J.B. Madalinska[b], T. Church[c], D. Crawford[d], M.L. Essink-Bot[b], V. Goel[e], H.J. de Koning[b], L. Määttänen[f], T. Pentikäinen[g]

[a]*Division of Clinical Epidemiology, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany*
[b]*Department of Public Health, Erasmus University Rotterdam, PO box 1738, 3000 DR Rotterdam, The Netherlands*
[c]*School of Public Health, University of Minnesota, MCC 807 Mayo, 420 Delaware Street, SE. Minneapolis, MN 55455, USA*
[d]*Division of Urologic Oncology, University of Colorado Health Sciences Center, 4200 East Ninth Avenue, Box C-324, Denver, CO 80262, USA*
[e]*Department of Health Administration, University of Toronto, Toronto, ON, Canada M5S 1A8*
[f]*Finnish Cancer Registry, Liisankatu 21B, 00170 Helsinki, Finland*
[g]*VTT, Group for Technology Studies, PO Box 1002, FIN-02044 VTT, Finland*

Received 26 October 2000; received in revised form 11 June 2001; accepted 15 June 2001

Abstract

Decisions on policies for screening for prostate cancer require that information upon health-related quality of life (HRQL) and cost-effectiveness (CE) be available, as the lead time for some of the cases detected by screening will be very long and detriments in quality of life could have a major impact on the subjects remaining life-span. A framework within which both HRQL and cost-effectiveness of prostate cancer screening can be assessed is presented. Studies of both are ongoing in the European Randomised Study of screening for prostate cancer and the US Prostate, Lung, Colon and Ovary trial. Preliminary information confirms that it is important to study screened subjects and controls, and not to assume that inferences derived from study of prostate cancer outside screening trials can be extrapolated to the trials. However, it will require prolonged study to enable the overall effects on quality of life, and on cost-effectiveness to be determined. Such studies are ongoing for the two trials.

Keywords: Cost effectiveness; Quality of life; Prostate cancer; Screening

1. Introduction

The extent that prostate cancer screening improves or impairs overall health-related quality of life (HRQL), as well as the acceptability of its cost to the individual and the community, is an important evaluation measure [1]. Deciding the healthcare policy is only possible if information is available on HRQL and the health costs of screened and unscreened participants as well as the mortality reduction from screening. Modelling suggests that an 'optimistic' estimate of screening effectiveness is required in order for screening to be cost-effective [2,3]. HRQL and CE studies have been initiated within the European Randomised Study of Screening for Prostate Cancer (ERSPC) and the US Prostate, Lung, Colon and Ovary (PLCO) trials. Collaboration between the trials will facilitate resolving the complex issues concerning HRQL and cost effectiveness.

The principal endpoint for the trials is a reduction in mortality from prostate cancer. The only valid surrogate for mortality is believed to be a reduction in clinically advanced or metastatic cancer [4]. However, surveillance may bring forward the time of diagnosis of such disease among men found to have prostate cancer by prostate-specific antigen (PSA) screening, resulting in an excess of advanced disease in a screened group compared with an unscreened one. Thus, basing cost effectiveness on cancer detection, especially if they are small stage 1 tumours, or even all cancers irrespective of stage, would be wrong, as each of these are expected to be influenced by screening but are affected by lead-time, length, selection and over-diagnosis biases.

* Corresponding author. Tel.: +49-6221-42-2219; fax: +49-6221-42-2203.
E-mail address: a.miller@dkfz-heidelberg.de (A.B. Miller).

0959-8049/01/$ - see front matter
PII: S0959-8049(01)00288-X

There is a similar problem related to the time cost-effectiveness and HRQL events occur. Many cancers will be diagnosed earlier in the screening arm, and thus at a younger age than in the control arm. Given that the costs of the screening tests, and the costs and adverse HRQL associated with false-positives and from treating the cancers that occur relatively early, it could be concluded that the HRQL issues are overwhelming [5]. It will require a prolonged follow-up before the detrimental effects on HRQL associated with advanced cancer late in life, which may be prevented in the screened group, appear in the control group. Therefore, long-term follow-up of participants in the trials will be required to determine the late quality of life effects.

Factors that are detrimental for HRQL and that are related to therapy can be estimated in non-trial participants, as can costs. The quality of life of patients with advanced prostate cancer has already been measured in several studies [6–13]. However, the spectrum and distribution of disease identified as a result of screening is not the same as in the absence of screening, it is therefore necessary to measure HRQL and determine the costs directly from samples of subjects in the trials to permit an accurate modelling of the late effects and their consequences.

Thus, the ERSPC and PLCO trials are being conducted with the intent of evaluating the comprehensive value of screening.

2. A framework for HRQL and cost-effectiveness studies

HRQL and cost-effectiveness studies are imbedded in a framework such as Fig. 1. The framework helps to facilitate decisions on the measures and timing that may be required. Each numbered node indicates a point in the screening, diagnosis, treatment, follow-up, and final endpoint (death) process when HRQL changes and cost expenditures occur. In non-compliant participants allocated to the screened group, no screening costs are incurred; likewise, some control group participants will incur screening costs because of contamination. Because of this self-selection, the analysis must primarily make an intention-to-treat comparison of the allocated screened and control groups.

For simplicity, the different nodes are described below in relation to HRQL and cost-effectiveness studies separately, although it is recognised that they are closely integrated, since HRQL is often incorporated into cost-effective studies, usually as preference-based measures.

2.1. HRQL studies

(1) Eligible participants have a baseline quality of life that should be estimated from representative samples. In several ERSPC or PLCO HRQL studies, measurements are being performed on a sampling basis within strata of age, race, centre and previous screening history.

(2) There appears to be an immediate, short-lived, decrement in quality of life following screening. It is important to measure the HRQL effects of the tests, including pain, discomfort and anxiety [14], before the results of the tests are available.

(3) Those participants allocated usual care (UC) in the volunteer-based trials may be disappointed, and some may seek PSA testing to substitute for the lack of screening. Studies to estimate the frequency of such contamination and assess the impact on HRQL of randomising to UC are being done.

(4) Participants with positive screening tests experience anxiety [14]. It would be preferable to measure this in advance of the diagnostic tests that follow, but in most instances it is only possible to measure such effects retrospectively.

(5) Participants with negative screening tests are reassured [14]. The majority of those with negative results will be true-negatives. The false-negatives are not initially identifiable; some of them will appear later as interval or screen-detected cancers.

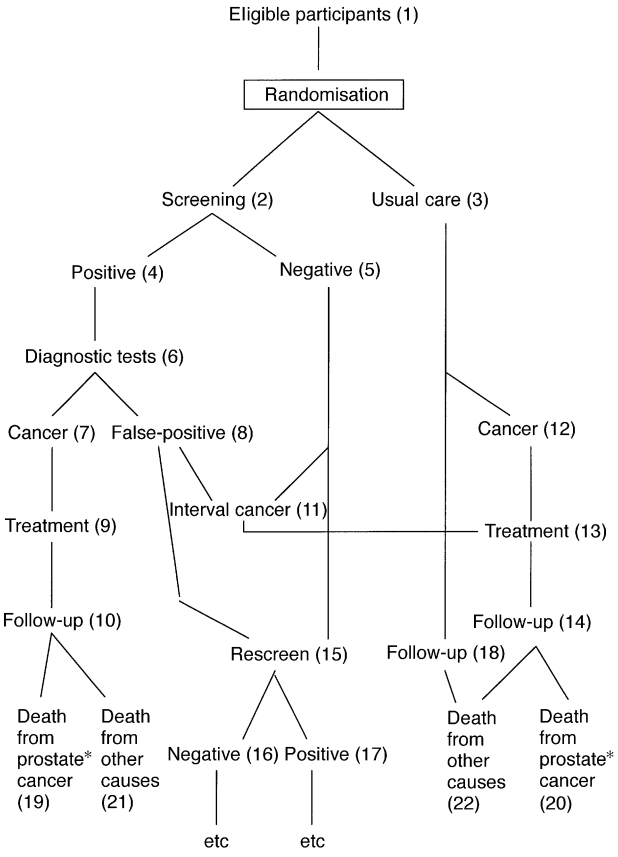

Fig. 1. Framework for cost-effectiveness and health-related quality of life (HRQL) measurements.

(6) Diagnostic tests heighten anxiety, and also affect HRQL through their interference with normal life [14]. Ideally, HRQL should be measured before the outcome is known, as there is a risk of recall bias if measurements are attempted later.
(7) Reaction to screen-detected cancers will vary in relation to whether their 'earlier' detection is perceived as a benefit derived from the trial. Such reactions should be captured before therapy is started, as in the Rotterdam HRQL study [15].
(8) A non-cancer outcome to positive diagnostic tests (false-positive) is likely to be reassuring, with a rapid reduction in anxiety. Measurement of HRQL at several points after the non-cancer outcome is feasible. There is some evidence that compliance with subsequent screen-related events is higher than for those with negative test results [16].
(9) HRQL should be measured soon after completion of therapy for screen-detected cancers, and during it also if treatment is prolonged, to detect adverse consequences such as impotence, incontinence, impaired bowel functioning, etc. [17].
(10) Measurements during follow-up should capture long-term increments and decrements of HRQL, as well as interference with life events caused by diagnostic testing for cancer recurrence. If there is recurrence, there will be further decrements of HRQL.
(11) Interval cancers will probably be similar to clinically detected cancers in the UC group (12). However, the fact that they occurred after a negative screening test may result in a different emotional reaction. Hence, they need study in their own right.
(12) Apart from those detected by opportunistic (spontaneous or self-selected) screening, the majority of cancers in the UC group will be symptomatic. They require careful study as they form the controls for (7).
(13) The HRQL decrements associated with treatment should also be measured. Measurement will be facilitated in the centres that provide diagnosis and therapy for UC participants.
(14) For comparison purposes, follow-up assessments should be scheduled at the same frequency as for (10).
(15) The routinely scheduled re-screens will induce changes in HRQL that will resemble those that follow the initial (prevalence) screen. Thus, there are similar measurement requirements as for (2).
(16)/(17) Negative and positive results from re-screening will induce measurement requirements similar to (4) and (5), to be followed by similar requirements to (6), etc.
(18) The follow-up requirements for the UC group are periodic, perhaps annual, for the duration of the follow-up in the trial.
(19)/(20) HRQL during the terminal illness will be assessed by proxy ratings, as obtaining these data directly from patients may be emotionally too burdensome for them or the screening centre may learn of the terminal illness after the death of the patient. Previous studies [18,19] showed that at the individual level, patient-proxy agreement was generally moderate to good. Although at the group level systematic differences between the patient and proxy mean scores were observed, with a tendency of relatives to report more impairments of patients' HRQL, the bias tended to be limited. Despite these limitations, when employing significant others as the proxy respondents of cancer patients' quality of life the proxy is a viable and acceptable method for obtaining HRQL data [18].
(21)/(22) Deaths from other causes will also have decrements associated with HRQL that contribute to the total HRQL burden for study subjects. If screening is effective there will eventually be more of them in the screen arm than in the UC group.

2.2. CE studies

(1) There is a cost associated with identifying subjects eligible for screening. However, the processes required for a trial usually differ from routine screening. This cost element will have to be acknowledged, but not necessarily evaluated in the trials.
(2) The costs associated with the screening tests are important, as they may be the major cost of the screening process. Some costs could be obtained from the budgets of the trials, but the costs incurred in routine practice, and the costs incurred by the participants in attending the screen, require special study.
(3) There is a cost associated with UC, including physician visits for symptoms associated with cancer, and any diagnostic tests.
(4)/(5) There are costs associated with notifying screen-test results.
(6)/(8) The costs of distinguishing true- from false-positives and managing false-positives require special study, as these may not be not under the control of the screening centres. Both insurance (HMO) and medicare costs have to be considered, since costs vary by insurance status as well as by age.
(7)/(9)/(10) The costs of treating true-positives will vary by stage. It cannot be assumed that the costs of treatment by stage for a screen-detected cancer are the same as for a non-screen-detected cancer.
(11)/(12)/(13)/(14) The costs of identifying, treating and managing interval and non-screen-detected cancers should be the same by stage, age and centre as for the general population. However, special study may be needed to obtain the detail required for trial purposes.

(15)/(16)/(17) Re-screening costs will be similar to the initial screening, although they involve costs associated with ensuring compliance.
(18) There are the study-associated costs of follow-up of the UC group. These will probably not require special documentation.
(19)/(20) The costs associated with terminal illness from fatal cancers may be incurred earlier in life in the UC than the study group. Therefore, costs associated with both (19) and (20) will have to be separately determined.
(21)/(22) The costs of caring for people dying of other causes will also require study. The time these events occur, and thus the influence of discounting, may be critical.

2.3. Difficulties in applying the framework

One of the major difficulties investigators will have in determining HRQL for many of the steps in the framework, is that they may learn of an event after much delay. This particularly affects items nos. (2), (4), (6), (7), (8), (9), (11), (12) and (13). However, the main concern in the trials has to be with long-term, persistent decrements of HRQL.

Costs related to the treatment of prostate cancer are available for the US [20]. However, they may differ from the costs of treating screen-detected or interval cancers. Although administrative data for costs may be available, it is nearly impossible to determine from routine medical records which costs are screen-related and which are not. Thus, the only unbiased way of comparing costs is by intention-to-treat, accruing costs to each allocated group and determining the difference. For a complete accounting, both indirect and intangible costs should be estimated, as well as direct costs.

For several of the CE measures, costs in the trials will not directly reflect future costs. Diagnosis and treatment will change in the future, and to guide policy in the future, the costs in the future will have to be included in the CE models. This can be partly overcome by ensuring that health care utilisation data are collected on all subjects. The unit costs for specific utilisation can then be indexed to a reference year when the CE analysis is done.

Several items, such as nos. (2), (4), (6), (7), (8), (9) and perhaps (19), require screening to be undertaken to determine the costs. Comparable costs in the UC group will have to be estimated over a similar time period.

3. Assessment of HRQL

HRQL is a multidimensional construct incorporating patients' functioning in physical, psychological and social domains. A clear distinction must be made between the description and evaluation of HRQL. Descriptive measures generate a profile of scores across different dimensions of HRQL and provide a detailed description of HRQL during different phases of screening and disease. Evaluative measurement yields a single summary index ('utility') that is obtained for each profile of HRQL scores (health state). Health state utilities are necessary for calculation of quality-adjusted life years (QALYs). A QALY is a composite health outcome measure, combining both duration and quality of life. Time lived with disease is made 'equivalent' to a shorter period in full health using a utility weight between 0 (death) and 1 (full health). QALYs are suited to the overall evaluation of a screening programme.

To assess HRQL effects, generic, disease-specific or domain-specific measures can be used. Generic questionnaires (e.g. Short Form-36 (SF36), Short Form-12) are comprehensive, non-specific HRQL measures. They allow for comparisons across diseases and between disease stages. Although generic measures are used mainly for descriptive purposes, some instruments provide a direct link to health state utilities. Measures with a link to utilities (EuroQoL-5D, Quality of Well-being Scale, Health Utility Index) provide a 'tariff' or scoring formula to transform descriptions of a patient's health status into a summary figure ('utility'). Preferences from the general public are commonly used to reflect the societal perspective in a decision-making context [21,22]. Recently, efforts have been made to derive utilities from the SF-36 [23,24].

Disease- and domain-specific measures are used to complement generic measures. The early ones (e.g. UCLA Prostate Cancer Index) assessed the extent of symptoms related to prostate cancer and its treatment (e.g. urinary incontinence, sexual dysfunction, gastrointestinal symptoms). Later ones (e.g. State-Trait Anxiety Inventory, Centre for Epidemiologic Studies Depression Scale) concentrate on the impact of the disease on a specific psycho-social domain of a patient's HRQL (e.g. anxiety, depression). Disease-specific instruments seem to be capable of detecting longitudinal differences in functioning of patients who undergo radical prostatectomy or primary radiotherapy [17], as well as differences between disease stages (localised versus metastatic prostate cancer). Whether post-treatment decrements in functional status have an impact on generic HRQL is unclear. Some studies could not detect significant changes between pre- and post-treatment SF-36 scores [25].

In ERSPC and PLCO, a commonly applied combination consists of descriptive generic, generic with link to utilities, disease-specific and domain-specific instruments. The studies explore the relationship between disease-specific and generic HRQL in prostate cancer patients. These efforts may result in the development of more sensitive instruments for capturing relevant HRQL changes in all phases of screening.

4. Preliminary findings in the ERSPC and PLCO trials

4.1. On HRQL

In general, the screening process itself does not seem to result in appreciable differences between screened subjects and controls, nor between participants and non-participants, although participants with pre-existing anxiety tend to remain anxious [14]. Considerable attention is therefore being paid to the HRQL decrements associated with false-positive screening test results and to those with a positive screen who are found to have cancer. Those deemed to have a false-positive screen after a negative biopsy of the prostate are an important risk group for subsequent cancer diagnosis, as some may later be diagnosed with prostate cancer, whether as an interval finding or after a subsequent screen. In the group who come for re-screening, HRQL decrements could become more important as they age.

In the Rotterdam HRQL study, patients with screen-detected prostate cancer reported significantly better pre-treatment generic HRQL (physical aspects), compared with patients diagnosed in a clinical setting [15]. Nevertheless, HRQL scores of the latter group remained in the range of the population norm. No differences were found in patients' self-reported levels of urinary, bowel and sexual functioning. Pre-treatment comparison of patients scheduled either for prostatectomy or radiotherapy revealed that the radiotherapy patients were significantly older and had more co-morbidity. Problems with urinary, bowel and sexual functioning were uncommon; however, radiotherapy patients older than 65 years appeared to be less sexually active prior to the diagnosis. Radiotherapy patients also reported poorer levels of generic HRQL. These results indicate that patients with screen-detected prostate cancer come from a distinct, relatively healthy population, presumably due to some self-selection when responding to invitations to be screened.

4.2. On cost-effectivness

Costs are being determined at many steps in the framework, especially in the screened arm, in both trials. Cost implications of advanced prostate cancer have been determined from non-trial participants in Rotterdam [26]. In the Nordic countries, many of the required costs are readily available from the health care systems. In the US, with different healthcare organisations involved, costs vary, and many healthcare organisations either cannot, or are reluctant, to supply them. In one of the PLCO centres where the downstream costs of interventions after both positive and negative screens are being studied, preliminary estimates from one of the three healthcare organisations in that area have been derived. Additional estimates are needed from other PLCO centres with different healthcare organisations, especially those with a more minor participation.

5. Future HRQL and cost-effectiveness studies in ERSPC and PLCO

5.1. Specific aims

1. Collect serial HRQL data in intervention and UC subjects who remain free of prostate cancer, stratifying the intervention group according to whether the screening tests were negative or falsely-positive.
2. Measure the immediate and short-term HRQL effects from among those with positive and negative screening tests.
3. Collect serial HRQL data in intervention and UC subjects who develop a prostate cancer, including information about cancer-related side-effects and complications arising from treatment.
4. Determine the HRQL decrements from those activities that contribute to the indirect costs of screening (e.g. travel to the screening centre, time spent on screening, diagnostic tests, etc.).
5. Determine the differential in HRQL effects from the terminal illnesses of subjects who die in the intervention and UC arms separately for prostate cancer and other causes of death, and evaluate whether there are differences according to the age at which death occurs.
6. Track utilisation of health care associated with screening for prostate cancer for each country by centre and healthcare system.
7. Collect data on the cost of screen-related diagnostic and treatment procedures for suspected and confirmed prostate cancers. Compare these costs with corresponding costs in the UC group.
8. Collect data on the opportunity costs for attendance for screen-related diagnostic and treatment procedures.
9. Determine the differential between the costs of treatment and subsequent follow-up and terminal care for screen-detected and non-screen-detected cancers.
10. Determine a utility measure yearly in each arm for each trial within each country for the duration of the trials.
11. Develop methodology for adjusting comparisons in items 1–10 for underlying differences in the non-randomised comparison groups, based on data collected at enrolment or data available from medical records. Methods may be adapted from

those used to adjust for compliance in randomised trials [27–30], or to adjust for lead-time and length-bias in observational studies of screening [31].

5.2. Comments on future studies

A high priority is to decide on the instruments that should be used for the HRQL studies, as well as to determine the utility measure. There is a conflict between group (population) HRQL estimates which will be influenced by the healthy screenee effect, and individual (prostate cancer patient-based) estimates. In overall evaluations, the former could easily submerge the latter, yet it is the latter on which we wish to concentrate. The emerging ability to map generic HRQL measures such as the SF-36 to utilities, and the development of prostate-specific health status and utility measures such as the PORPUS [32], may facilitate collecting this full range of data while minimising the respondent burden.

It clearly is not possible to make the assessments summarised above in relation to the framework on all study participants, nor is it necessary. However, there is a difficulty in sampling, as the chain of measurements desirable for a sequence of events, e.g. (2)–(4)–(6)–(8), would require different size samples to assess the state with precision, and provide the 'before' measurement for what could follow. The solution may be to combine a series of cross-sectional samples with repetitive resampling of a series of individuals. Cross-sectional samples may be optimal for (1), (2), (3), (5), (15), (16) and (22). The sampling fractions will require further consideration, but will need to be stratified by age, race, gender, study centre and calendar year, and could differ between sampling times. The ongoing pilot studies will provide guidance on the required sample sizes, instruments and the timing of their administration.

Another difficulty is that not all of the potential requirements are currently being subjected to study; therefore, empirical decisions may be necessary. Close to 100% samples might be desirable for cancer states, e.g. (7), (9), (10), (11), (12), (13), (14), (19) and (20). The remaining states could either require different cross-sectional sampling fractions for precision, e.g. for (4), (6), (8), (17), etc., or would be derived by following the same previously sampled individuals at their subsequent events, e.g. (18) would repetitively resample those sampled for (3), and (6) and (8) those sampled for (4) (less those in (7)).

In conclusion, assessment of quality of life and costs within a large screening trial is clearly not a simple exercise. There is potential for significant respondent burden which could adversely affect the main trial processes. At the same time, it is essential that feasible steps are taken to ensure that the best possible data are collected. Otherwise, we will be left with trying to assess quality of life and cost-effectiveness after the fact. Funding to enable the necessary studies to be completed is essential.[1]

Acknowledgements

The authors are grateful for the advice and input of Martin Brown, John Gohagan and Fritz Schroeder, during the preparation of this paper. We also thank John Nyman, Patricia McGovern, Cynthia Gross, Kathryn Taylor and David Johnson, for sharing with us their research plans and initial efforts.

References

1. The International Prostate Screening Trial Evaluation Group. Rationale for randomised trials of prostate cancer screening. *Eur J Cancer* 1999, **35**, 262–271.
2. Krahn MD, Mahoney JE, Eckman MH, *et al.* Screening for prostate cancer: a decision analytic view. *JAMA* 1994, **272**, 781–786.
3. Office of Technology Assessment, US Congress. *Assessment, Costs, and Effectiveness of Prostate Cancer Screening in Elderly Men*. Washington, DC, US Government Printing Office, May 1995.
4. Prorok PC, Chamberlain J, Day NE, Hakama M, Miller AB. UICC Workshop on the evaluation of screening programmes for cancer. *Int J Cancer* 1984, **34**, 1–4.
5. Adami HO, Baron JA, Rothman KJ. Ethics of a prostate cancer screening trial. *Lancet* 1994, **343**, 958–960.
6. Clark JA, Wray N, Brody B, *et al.* Dimensions of quality of life expressed by men treated for metastatic prostate cancer. *Soc Sci Med* 1997, **45**, 1299–1309.
7. Curran D, Fossa S, Aaronson N, *et al.* Baseline quality of life of patients with advanced prostate cancer. European Organization for Research and Treatment of Cancer (EORTC), Genito-Urinary Tract Cancer Cooperative Group (Gut-CCG). *Eur J Cancer* 1997, **33**, 1809–1814.
8. Litwin MS, Lubeck DP, Henning JM, *et al.* Differences in urologist and patient assessments of health-related quality of life in men with prostate cancer: results of the CaPSURE database. *J Urol* 1998, **159**, 1988–1992.
9. Lubeck DP, Litwin MS, Henning JM, *et al.* Changes in health-related quality of life in the first year after treatment for prostate cancer: results from CaPSURE. *Urology* 1999, **53**, 180–186.
10. Moinpour CM, Savage MJ, Troxel A, *et al.* Quality of life in advanced prostate cancer: results of a randomized therapeutic trial. *J Natl Cancer Inst* 1998, **90**, 1537–1544.
11. Rosendahl I, Kiebert GM, Curran D, *et al.* Quality-adjusted survival (Q-TWIST) analysis of EORTC trial 30853: comparing goserelin acetate and flutamide with bilateral orchiectomy in patients with metastatic prostate cancer. European Organization for Research and Treatment of Cancer. *Prostate* 1999, **38**, 100–109.
12. Sharp LK, Knight SJ, Nadler R, *et al.* Quality of life in low-income patients with metastatic prostate cancer: divergent and convergent validity of three instruments. *Qual Life Res* 1999, **8**, 461–470.

[1] Editor's note: On a related theme, please see the Special Issue on 'Economics and Cancer' [33].

13. Stanford JL, Feng Z, Hamilton AS, *et al*. Urinary and sexual function after radical prostatectomy for clinically localized prostate cancer: The Prostate Cancer Outcomes study. *JAMA* 2000, **283**, 354–360.
14. Essink-Bot M-L, de Koning HJ, Nijs HGT, *et al*. Short-term effects of population-based screening for prostate cancer on health-related quality of life. *J Natl Cancer Inst* 1998, **90**, 925–931.
15. Madalinska JB, Essink-Bot ML, de Koning HJ, *et al*. Health related quality of life in patients with screen-detected versus clinically diagnosed prostate cancer preceding primary treatment. *The Prostate* 2001, **46**, 87–97.
16. Burman ML, Taplin SH, Herta DF, Elmore JG. Effect of false-positive mammograms on interval breast cancer screening in a health maintenance organization. *Ann Intern Med* 1999, **131**, 60–62.
17. Madalinska JB, Essink-Bot ML, de Koning HJ, *et al*. Health related quality of life effects of radical prostatectomy and primary radiotherapy for (screen-detected or clinically diagnosed) prostate cancer. *J Clin Oncol* 2001, **19**, 1619–1628.
18. Sneeuw KCA, Aaronson NK, Sprangers MAG, *et al*. Comparison of patient and proxy EORTC QLQ-C30 ratings in assessing the quality of life of cancer patients. *J Clin Epidemiol* 1998, **51**, 617–631.
19. Hinton J. How reliable are relatives' retrospective reports of terminal illness? Patients' and relatives' accounts compared. *Soc Sci Med* 1996, **43**, 1229–1236.
20. Brown ML, Lipscomb J, Snyder C. The burden of illness of cancer: economic cost and quality of life. *Annual Review of Public Health* 2001, **22**, 91–113.
21. Brazier J, Deverall M, Green C. A review of the use of health status measures in economic evaluation. *J Health Serv Res Policy* 1999, **4**, 174–184.
22. Drummond MF, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*. Oxford, Oxford University Press, 1997.
23. Brazier J, Usherwood T, Harper R, *et al*. Deriving a preference-based single index from the UK SF-36 health survey. *J Clin Epidemiol* 1998, **51**, 1115–1128.
24. Meletiche DM, Doshi D, Lofland JH. Medical Outcomes Study Short Form 36: a possible source of utilities? *Clin Ther* 1999, **21**, 2016–2026.
25. Litwin MS, Shpall AI, Dorey F, *et al*. Quality-of-life outcomes in long-term survivors of advanced prostate cancer. *Am J Clin Oncol* 1998, **21**, 327–332.
26. Beemsterboer PMM, de Koning HJ, Birnie E, *et al*. Advanced prostate cancer: course, care and cost implications. *The Prostate* 1999, **40**, 97–104.
27. Cuzick J, Edwards R, Segnan N. Adjusting for non-compliance and contamination in randomized clinical trials. *Statistics in Medicine* 1997, **16**, 1017–1029.
28. Mark SD, Robins JM. A method for the analysis of randomized trials with compliance information: an application to MRFIT. *Controlled Clinical Trials* 1993, **14**, 79–97.
29. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983, **70**, 41–55.
30. Rubin DB. More powerful randomization-based p-values in double-blind trials with non-compliance. *Statistics in Medicine* 1998, **17**, 371–385.
31. Church TR. A novel form of ascertainment bias in case-control studies of cancer screening. *J Clin Epidemiol* 1999, **52**, 837–847.
32. Krahn M, Ritvo P, Irvine J, *et al*. Construction of the patient-oriented prostate utility scale (PORPUS): a multiattribute health state classification system for prostate cancer. *J Clin Epidemiol* 2000, **53**, 920–930.
33. Economics and cancer [special issue]. *Eur J Cancer* 2001, **37**.