ELSEVIER

European Journal of Cancer 41 (2005) 411–415

European Journal of Cancer

www.ejconline.com

Determining overdiagnosis by screening with DRE/TRUS or PSA (Florence pilot studies, 1991–1994)

Stefano Ciatto [a,*], Ginetta Gervasi [a], Rita Bonardi [a], Paola Frullini [a], Paola Zendron [a], Claudio Lombardi [b], Emanuele Crocetti [a], Marco Zappa [a]

[a] *Centro per lo Studio e la Prevenzione Oncologica, Viale A. Volta 171, 50131 Florence, Italy*
[b] *Department of Urology, Torre Galli Hospital, Florence, Italy*

Received 23 June 2004; received in revised form 9 September 2004; accepted 28 September 2004
Available online 5 January 2005

Abstract

The rate of overdiagnosis of prostate carcinoma was assessed by following 6890 participants in pilot screening studies from 1991 to 1994. Observed/expected incidence and mortality were determined using data from the Cancer and Mortality Registry. The cancer detection rate (1.75%) and observed/expected ratio (12.5:1) were high at the first screening, and substantially lower at the second screening (0.65% or 4.10:1). According to the registry follow-up, prostate cancer occurred in 225 subjects in the whole study cohort, while 178.2 were expected with 50 652 men/years at risk. The standardised incidence rate was 1.66 in the screened (95%CI = 1.4–2.0), 0.97 in the non-responders (95%CI = 0.8–1.2) and 1.23 in subjects excluded from invitation due to previous cancer or major illness (95%CI = 0.8–1.5). A 66% excess incidence rate was observed in the screened subjects over a 9-year period, confirming previous estimates of overdiagnosis.

Keywords: Prostate cancer; Screening; Overdiagnosis

1. Introduction

Screening for prostate cancer is presently under evaluation in two large prospective randomised trials in the United States (US) (PCLO) and Europe (ERSPC) [1], both aim to assess the impact of screening on prostate cancer mortality. Thus far, no mortality data are available from these studies and only first screening round cancer detection rates are known from the ERSPC study [2], showing a high observed/detected ratio suggesting there is substantial overdiagnosis.

Although the efficacy of screening in reducing prostate cancer mortality is yet to be proven, opportunistic screening using prostate-specific antigen (PSA) levels is increasingly performed in Western countries, particularly in the US [3,4]. A sharp rise in prostate cancer incidence has been observed in the US as a consequence of such opportunistic screening [5]. This has not been followed by a decrease in mortality for several years, suggesting there may be diagnostic anticipation and overdiagnosis.

Detection of "latent" non-aggressive cancers that will not to become clinically evident (overdiagnosis) is an unavoidable consequence of screening, but its magnitude and the side-effects of overtreatment can be a major drawback, where screening might ultimately be more harmful than beneficial. This might be the case for screening for prostate cancer, for which the amount of overdiagnosis has been estimated to be high [6,7] and where radical treatment has major side-effects [8]. Thus, continuous monitoring and evaluation of the overdiagnosis rate using current data from existing screening programmes is necessary.

* Corresponding author. Tel.: +39 55 5012275; fax: +39 55 5001623.
E-mail address: s.ciatto@cspo.it (S. Ciatto).

0959-8049/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2004.09.030

A preliminary estimate of the overdiagnosis rate was published in 1998, based on the follow-up of two pilot studies of prostate cancer screening performed at the Centro per lo Studio e la Prevenzione Oncologica (CSPO) in Florence [9,10]. The present study analyses prostate cancer incidence in the same cohort after longer follow-up. The aim of the present study was to estimate the magnitude of overdiagnosis associated with screening.

2. Patients and methods

Two pilot studies assessing the feasibility of prostate cancer screening by digital rectal examination (DRE) + transrectal ultrasonography (TRUS) or by PSA were performed at the CSPO from 1991 to 1994, completing two biennial screening rounds (only first screening responders were invited to the second screening). Detailed features of these studies have already been reported in [11]. The screening protocol was not aggressive, as random sextant biopsy was limited to subjects with PSA values of 10 ng/ml or above, whereas directed biopsy was prompted by suspicious findings at DRE or TRUS. A random sample of National Health Service general practices (GP) were invited to join the study: as over 98% GPs accepted, the study population was assumed to be representative of the whole population in the District. Resident men aged 60–74 years registered at GPs and volunteering to be part of the studies were invited to screening. GPs were asked to exclude subjects with major, disabling illness, those unlikely to attend invitation, or those with known prostate cancer. As both screening studies showed almost the same prevalence/incidence ratio [11], with a comparable diagnostic anticipation, a pooled analysis of both cohorts was done with the purpose of estimating the overdiagnosis rate in a larger sample.

Linkage of all study subjects (excluded, non-responders, examined) with regional registries was performed (deterministic linkage based on name and date of birth), and incident prostate cancers (population-based Tuscany Tumour Registry [12], updated to December 2000), as well as deaths (from all causes) (population based Regional Mortality Registry, updated to December 2001) were identified. Information on emigration from the screening area was available, showing that the emigration rate in the study cohort age group and in the study period was negligible. Prostate cancers occurring before the date of invitation or of exclusion were not considered. Trends of prostate cancer incidence were determined for the whole cohort, for different time periods, and for single subgroups (invited and screened, non-responders to invitation, excluded and not invited).

Overdiagnosis was determined as the ratio (standardised incidence ratio: SIR) of prostate cancers observed in the study period and the number of cancers expected according to age-specific incidence rates provided by the Tuscany Cancer registry and to men/years at risk. Standardised mortality ratio (SMR) and 95% confidence intervals (95%CI) were computed to compare the observed and expected mortality from all causes. The expected number of prostate cancer cases (and deaths) was calculated by multiplying the age-, period- and site-specific incidence (and mortality) rates by the Tuscany Cancer Registry (and Regional Mortality Registry) by corresponding person-years. Observed prostate cancers were compared with those expected according to SIR, through the observed/expected ratio. 95%CI were computed assuming a Poisson distribution for the observed cases.

3. Results

Table 1 shows the detection rates of the two pilot study screening rounds. The detection rates at first or second screening rounds were 1.75% and 0.65%, respectively. Corresponding values of observed to expected incidence were 12.5 or 4.10:1, respectively. The clinical stage of the screen-detected cancers was T1c in 3, T2a in 23, T2b in 14, T3 in 7 and Tx in one case at the first screening, whereas it was T2a in all cases at the second screening. The Gleason score was <7, 7, >7 or unknown in 52%, 25%, 20% or 2% of cases at the first screening, and in 61%, 23%, 0% or 15% at the second screening, respectively.

Table 2 reports incident prostate cancers observed in the study cohort. Overall, 225 cases were observed, while 178.2 were expected for 50 652 men/years at risk. SIR was 1.66 in the screened (95%CI = 1.4–2.0), 0.97 in the non-responders (95%CI = 0.8–1.2) and 1.23 in the excluded subjects (95%CI = 0.8–1.5), respectively. In the whole study cohort, a SIR of 1.26 (95%CI = 1.1–1.4) was observed.

Table 3 shows the observed SIR according to the different time periods (less than 5 years/more than 5 years after the invitation date): the only significant difference in observed compared with expected incidence occurred in invited and examined subjects in the first 5 years after the invitation (SIR = 2.11, 95%CI = 1.2–2.6).

Table 4 shows mortality from all causes in the study cohort. Mortality in the whole cohort was as expected (SMR = 0.72, 95%CI = 0.7–0.8). A significant excess in the SMR was evident for subjects excluded from invitation by their GP (SMR = 1.67, 95%CI = 1.5–1.9), and, to a much lower extent, and non-statistically significant, for invited, but not examined subjects (SMR = 1.05, 95%CI = 0.98–1.1), whereas it was significantly lower for the screened subjects (SMR = 0.72, 95%CI = 0.7–0.8).

Table 1
Prostate cancer detection rates at first and second screening rounds of our pilot study

Age (years)	Invited subjects	Examined subjects	Cancers detected	% Detection rate	Cancers expected	Observed/expected
First screening round						
60–64 years	2513	1144	8	0.69	0.59	13.5
65–69 years	2155	945	20	2.11	1.40	14.2
70–74 years	1525	651	20	3.07	1.83	10.9
Total	6193	2740	48	1.75	3.82	12.5
Second screening round						
60–64 years	687	614	2	0.32	0.31	6.45
65–69 years	1024	772	5	0.64	1.14	4.38
70–74 years	928	613	6	0.97	1.72	3.48
Total	2639	1999	13	0.65	3.17	4.10

Expected cases on the basis of underlying incidence and observed/expected ratio are reported [3].

Table 2
Distribution of observed and expected prostate cancers in the study cohort

Subjects	Subjects	Men/years at risk	Observed	Expected	SIR	95%CI
Excluded by GP, not invited	775	5108	23	18.7	1.23	0.8–1.5
Invited, non-responders	3451	25 756	88	91.0	0.97	0.8–1.2
Invited, screened	2664	19 788	114	68.5	1.66	1.4–2.0
Total	6890	50 652	225	178.2	1.26	1.1–1.4

The standardised incidence rate (SIR) of expected to observed cases and its 95% confidence interval (95%CI) is indicated.
GP, general practices.

Table 3
Prostate cancer incidence in the study cohort in different time periods

Period	Excluded by GP, not invited	Invited, not examined	Invited, examined	Total
<5 years after invitation				
Observed	13	41	76	130
Expected	10.1	46.7	36.1	92.9
SIR (95%CI)	1.29 (0.7–2.2)	0.88 (0.6–1.2)	2.11 (1.2–2.6)	1.40 (1.2–1.7)
>5 years after invitation				
Observed	10	47	38	95
Expected	8.6	44.3	32.4	85.3
SIR (95%CI)	1.16 (0.6–2.1)	1.06 (0.8–1.4)	1.17 (0.8–1.6)	1.1 (0.9–1.4)

The standardised incidence rate (SIR) of expected to observed cases and its 95% confidence interval (95%CI) is indicated.

Table 4
Mortality from all causes in the study cohort

	Subjects	Men/years at risk	Observed	Expected	SMR	95%CI
Excluded by GP, not invited	775	5544	320	191.4	1.67	1.5–1.9
Invited, not examined	3451	29344	955	910.5	1.05	0.98–1.1
Invited, examined	2664	23278	494	690.1	0.72	0.7–0.8
Total	6890	58166	1769	1792.0	0.99	0.94–1.03

The standardised mortality rate (SMR) of expected to observed cases and its 95% confidence interval (95%CI) is indicated.

4. Discussion

Overdiagnosis (i.e. the detection of a cancer which would not have been diagnosed without screening) and the subsequent overtreatment are often quoted as relevant negative effects of cancer screening. The magnitude of overdiagnosis depends on the magnitude of diagnostic anticipation (lead time) achieved by screening, and the importance of overtreatment depends on the morbidity and side-effects of treatment. Both this aspects are crucial in screening for prostate cancer, as the lead time likely exceeds 10 years [6] and major side-effects

of treatment are well known [8] (urinary incontinence, sexual impotence and bowel injury are commonly associated with radical prostatectomy or radiotherapy). In such a scenario, an estimate of the magnitude of overdiagnosis is absolutely fundamental.

The pilot screening study on which our present analysis is based is characterised by a limited diagnostic aggressiveness compared with the screening protocols adopted in the current ongoing screening trials [1,2], as random biopsy was performed in only a limited subset of cases, as is shown by the low detection rate of T1c stage cancers. A prevalence screening effect was evident at the first round (higher detection rate, higher observed/expected ratio, worse average stage and Gleason score compared with repeat screenings), and a lower detection rate was evident at the second screening round (observed/expected ratio of 4.10:1).

Some years ago, we reported an estimate of overdiagnosis, calculating the excess cumulative incidence after 5 biennial rounds of screening and after 14 years of follow-up [9]: assumptions for such a simulation were based on evidence from our pilot studies [11]. The overdiagnosis rate for subjects aged 60 or 65 years at entry was estimated to be 51% (95%CI 44–55) or 93% (95%CI 85–101) if incidence was assumed to be constant, and 25% or 65% if a 2% yearly increase in incidence was assumed in absence of screening in the general population.

In our present study, overdiagnosis was determined over a shorter observation period (9 years) and, during that period, a 4.5% yearly increment in prostate cancer incidence has been observed, that is mostly ascribed to the diffusion of opportunistic PSA screening in the Florence District [13]. The population used as a reference to estimate the underlying incidence included the study cohort, but the latter accounted for less than 3% of the former and therefore should not represent a significant contamination bias in the estimate of underlying incidence. Although an underestimation of overdiagnosis might occur due to these differences, the observed overdiagnosis is in line with that predicted from the previous study. In fact, a 1.66 SIR was observed after 9 years in the screened subjects, whereas the number of observed cancers among non-compliers were in the expected range (SIR = 0.97, 95%CI = 0.8–1.2). A higher cancer detection rate in responders to the screening invitation might be ascribed to a selection bias of subjects at higher risk of prostate cancer, but the excess detection rate persisted, although it was reduced, and was still significant when considering the whole study cohort (SIR = 1.26, 95%CI = 1.1–1.4).

It may be questioned as to what extent the study sample was representative of the general population. A selection bias associated with the study design (GP volunteer study) may be excluded as all GPs in a given area were invited to the study and almost all GPs volunteered. A screening effect is unlikely to have occurred in the whole study cohort, which showed an overall mortality equal to that expected and may thus be considered a representative sample of the resident population.

Differences in the underlying incidence of prostate cancer are evident across countries, independent of opportunistic PSA screening and probably depending on different prevalences of aetiological factors and genetic susceptibility. Although we do not have direct evidence, it is likely that differences in incidence are associated with differences in the prevalence of ''latent'' carcinomas: were this the case, a higher underlying incidence would be associated with a higher risk of latent carcinoma being detected at random biopsy and thus to a higher level of overdiagnosis. A recent increase in prostate cancer incidence was evident in Western countries, due to the diffusion of opportunistic screening in the general population: using recent incidence rates as a reference standard might bias overdiagnosis calculations towards underestimation. Such a bias should not affect our present study as opportunistic screening by PSA was almost absent in Florence District when the pilot studies were carried out.

An excess detection rate is not synonymous of overdiagnosis, as ''clinically relevant'' cancers, with surface symptomatically and may be detected several years in advance. However, our data show that the excess detection rate observed after screening was not followed by a decrease in incidence, as it would be expected with early detection of ''clinically relevant'' cancers, and persisted over 9 years of follow-up: no decline of incidence has been observed thus far, which suggests that true overdiagnosis is present.

Recently, overdiagnosis has been estimated by means of a simulation model [6], assumptions being based on data observed in the ERSPC study. Overdiagnosis resulted and was related to age at screening: for a single screening, ranging from 27% to 56% at 55 or 75 years, respectively. Similar values are reported in an another recent paper by Tornblom and colleagues [14]. Overdiagnosis values in our study are slightly higher: this might be at least partially explained by the relatively short follow-up (9 years) of our study, compared with a lifetime estimate in the quoted studies, as the effect of a limited screening period (e.g. two biennial rounds as in our study) is expected to decrease over time. Nevertheless, we would have expected the amount of overdiagnosis estimated by Draisma and colleagues [6] to be higher than estimates observed in our present analysis, as the screening protocol in our pilot study was less aggressive, and detection rates and observed/expected ratios were lower compared with ERSPC study standards [2].

In conclusion, the present analysis confirms a previous estimate of the magnitude of overdiagnosis due to screening for prostate cancer. High rates of overdiagno-

sis are confirmed for a screening experience adopting a non-aggressive protocol, using almost no routine random biopsy in the presence of elevated PSA levels.

Conflict of interest statement

None declared.

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. de Koning HJ, Auvinen A, Berenguer-Sanchez A, *et al*. Large-scale randomized prostate cancer screening trials; program performance in the ERSPC and PLCO trials. *Int J Cancer* 2002, **97**, 237–244.
3. Legler JM, Feuer EJ, Potosky AL, *et al*. The role of prostate specific antigen (PSA) testing pattern in the recent prostate cancer incidence decline in the United States. *Cancer Cause Control* 1998, **9**, 519–527.
4. Cooper GS, Yuan Z, Jethva RN, *et al*. Determination of county-level prostate carcinoma incidence and detection rates with Medicare claims data. *Cancer* 2001, **92**, 102–109.
5. Shibata A, Whittemore AS. Prostate cancer incidence and mortality in the United States and the United Kingdom. *J Natl Cancer Inst* 2001, **93**, 1109–1110.
6. Draisma G, Boer R, Otto SJ, *et al*. Lead times and overdetection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. *J Natl Cancer Inst* 2003, **95**, 8688–8778.
7. Etzioni R, Penson DF, Legler JM, *et al*. Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends. *J Natl Cancer Inst* 2002, **94**, 981–990.
8. 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.
9. Zappa M, Ciatto S, Bonardi R, *et al*. Overdiagnosis of prostate carcinoma by screening: an estimate based on the results of the Florence Screening Pilot Study. *Ann Oncol* 1998, **9**, 1297–1300.
10. Ciatto S, Zappa M, Bonardi R, *et al*. Prostate cancer screening: the problem of overdiagnosis and lessons to be learned from breast cancer screen. *Eur J Cancer* 2000, **36**, 1347–1350.
11. Ciatto S, Bonardi R, Mazzotta A, *et al*. Comparing two modalities of screening for prostate cancer: digital rectal examination + transrectal ultrasonography *vs*. prostate specific antigen. *Tumori* 1995, **81**, 225–229.
12. Paci E, Crocetti E, Miccinesi G, *et al*. Tuscany cancer registry. In Parkin DM, Whelan SL, Ferlay J, Teppo L, Thomas DB, eds. *Cancer incidence in five continents*, vol. VIII. Lyon, IARC Scientific Publications, 2002. pp. 362–363., [No. 155].
13. Crocetti E, Ciatto S, Zappa M. Prostate cancer: different incidence but not mortality. Trends within two areas of Tuscany, Italy. *J Natl Cancer Inst* 2001, **93**, 876–877.
14. Tornblom M, Eriksson H, Franzen S, *et al*. Lead time associated with screening for prostate cancer. *Int J Cancer* 2004, **108**, 122–129.