ELSEVIER

Journal of Clinical Epidemiology 58 (2005) 639–644

Journal of Clinical Epidemiology

BRIEF REPORTS

A reliable coding system to define screening prostate-specific antigen tests was developed in a case–control study

Peter S. Bunting[a,b,c,*], Jan Neuman[d,1], Vivek Goel[e], Neil Fleshner[f], Jacek Kopec[g,h], Peter Levers[i], Padraig Warde[j,k]

[a]*Division of Biochemistry, Department of Pathology and Laboratory Medicine, The Ottawa Hospital and University of Ottawa, 1053 Carling Avenue, Ottawa, Ontario K1Y 4E9, Canada*
[b]*Department of Laboratory Medicine, Sunnybrook and Women's College Health Sciences Centre, University of Toronto, Toronto, Ontario, Canada*
[c]*Gamma-Dynacare Medical Laboratories, Brampton, Ontario, Canada*
[d]*Department of Public Health, University of Toronto, Toronto, Ontario, Canada*
[e]*Departments of Health Policy, Management and Evaluation and Public Health Sciences, University of Toronto, Toronto, Ontario, Canada*
[f]*Division of Urology, Sunnybrook and Women's College Health Sciences Centre, University of Toronto, Toronto, Ontario, Canada*
[g]*Arthritis Research Centre of Canada, 895 West 10th Avenue, Vancouver, British Columbia, Canada*
[h]*Department of Health Care and Epidemiology, University of British Columbia, Vancouver, British Columbia, Canada*
[i]*Division of Urology, Department of Surgery, Southlake Regional Health Centre and Markham Stoufville Hospital, Markham, Ontario, Canada*
[j]*Department of Radiation Medicine, Princess Margaret Hospital, 610 University Avenue, Toronto, Ontario, Canada*
[k]*Department of Radiation Oncology, University of Toronto, Toronto, Ontario, Canada*

Accepted 25 October 2004

Abstract

Objective: To establish the reliability of a coding system for screening and diagnostic prostate-specific antigen (PSA) testing from patient charts.

Study Design and Setting: Two investigators reviewed 448 chart abstractions from a population-based case–control study of PSA screening in the Toronto area. The tests evaluated for reliability were transrectal ultrasound (TRUS), digital rectal examination (DRE), and prostate-specific antigen (PSA).

Results: DRE results were found in 87%, PSA results in 65%, and TRUS results in 12% of the 749 charts. Interobserver agreement was 94% for DRE texture ($\kappa = .885$), 95% for DRE asymmetry ($\kappa = .868$), 85% for DRE physician interpretation ($\kappa = .698$), 97% for final DRE result ($\kappa = .856$), and 87% for TRUS ($\kappa = .769$). Physician interpretation modified the final result in only 6.2% of DREs. Interobserver agreement for PSA coding was 91% ($\kappa = .787$). Of PSA results, pure PSA screening with no symptoms of obstructive urination was found in 19%, symptomatic PSA screening in 46%, and diagnostic PSA testing in 35%.

Conclusion: We have developed a practical and reliable coding system for TRUS, DRE, and PSA in the context of a case–control study of PSA screening.

Keywords: Prostate-specific antigen; Prostatic neoplasm; Screening; Reliability; Digital rectal examination; Case–control; Chart review

1. Introduction

Screening for prostate cancer with prostate-specific antigen (PSA) remains a highly controversial topic [1]. Randomized controlled trials of PSA screening are under way in Europe and the United States, but are several years from reaching their conclusion. Several approaches have tried to assess the mortality benefits of PSA screening. Many of these use health care databases as sources of information. A recent population-based study in Canada, using prostate cancer incidence after the introduction of PSA testing as a surrogate for screening, failed to show an association between intensity of PSA screening and decreases in prostate cancer mortality over 15 years [2]. Another population-based study in the United States, comparing areas with intensive PSA screening (Seattle) and low screening (Connecticut), drew the same conclusion, after 11 years of follow-up [3]. On the other hand, a study in Austria found markedly reduced mortality in the heavily screened state of Tyrol, compared with other states in which screening was much less used [4].

* Corresponding author. Tel.: 613-798-5555 ext. 16106; fax: 613-761-5401.
E-mail address: psbunting@ottawahospital.on.ca (P.S. Bunting).
[1] Present address: Leadership Sinai Centre for Diabetes, Mt Sinai Hospital, Toronto, Ontario, Canada.

0895-4356/05/$ – see front matter
doi: 10.1016/j.jclinepi.2004.10.014

These approaches continue to provide conflicting conclusions about whether there is mortality benefit from PSA screening.

Studies that have access to the original medical charts are at an advantage for capturing greater detail about PSA testing and the clinical circumstances associated with such testing. One approach using original medical charts is the case–control design; we have undertaken such a study. We know of only one other published study on the reliability of retrospective medical chart review of PSA screening [5]. These authors provided reasonably detailed descriptions of how they identified diagnostic and screening PSA tests. They reported a composite agreement of 79% with the attending physician's rating, among four evaluators reviewing 95 cases. Another case–control study based on chart review [6] reported a 30% reduction in mortality due to screening with PSA and/or digital rectal examination (DRE), but reliability information was not given.

Part of the strategy of PSA screening usually involves a DRE. This technique is notorious for its lack of reproducibility, as reviewed by Brawer [7]. Nonetheless, DREs remain a common tool in periodic health examinations, to detect both prostate and colorectal cancers. Procedural details for performing a DRE are rather loose in most studies, which led one group to define a procedure for DRE and to systematically assess interobserver variability [8]. That study, which involved DREs performed by both general practitioners and urologists, showed that the strongest interobserver agreement among physicians was found for prostate size, tenderness, the presence of midline furrow, symmetry, induration, and nodularity. Agreement, adjusted for chance, was moderately good (Cohen's κ ranged from .485 to .682, depending on the characteristic). An earlier study [9], comparing DRE performance among different urologists in a prostate cancer screening context, found only fair agreement adjusted for chance ($\kappa = .22$), probably because the process was not as carefully designed as was the case with, for example, Varenhorst et al. [8]. Other studies on DRE interobserver reliability are not really comparable, because they compared estimates of prostate size with DRE with those of imaging studies, such as that by Roehrborn et al. [10]. That is, the studies were diagnosing benign prostatic hypertrophy (BPH) rather than prostate cancer. Furthermore, all of these interobserver agreement studies were in the realm of clinical practice, not chart abstraction.

A few chart abstraction case–control studies involving use of DRE to screen for prostate cancer have been reported [11–13], but with little information on how charts were abstracted and interpreted, nor on interobserver variation in the information abstracted. A few reasons for rectal examination were found in common for two of these studies [12,13]: routine screening, symptoms of prostatism, gastrointestinal and rectal symptoms, lower back or pelvic pain, and other prostate lesions; however, no schema was reported for how these indicators were arrived at.

In our own population-based case–control study of PSA screening for prostate cancer, we assessed whether PSA measurements done on cases and controls were for purposes of diagnosis or screening, based on information abstracted from patient charts. In the present report, we describe the process that we used, with relevant interobserver reliability. This protocol may be of value in future studies of prostate cancer screening.

2. Methods

2.1. Patient information

Subjects were accrued from the area of metropolitan Toronto and five nearby counties in Ontario, Canada. Cases were identified over a 3-year period (August 1999–May 2002) through chart review for the two regional cancer centers responsible for delivering all radiation therapy to all cancer patients. Further cases were accrued by contacting 88 of the 90 urologists in Toronto and the counties monthly by telephone, mail, or e-mail (2 refused to participate). The urologists were sent a case report form on which they provided information such as patient name, date of diagnosis, date of metastases, and copies of bone scan reports or other imaging for patients with metastatic prostate cancer in their practice. For inclusion, cases (236 patients having prostate cancer with distant metastases, a surrogate for death) had to have been diagnosed with prostate cancer after 1 January 1990, the year PSA testing began in Ontario. Controls (462) were men matched for age, region of residence, and reference date for censoring (equal to the date of cancer diagnosis for cases). We reviewed charts from the 608 family physicians caring for these patients. Subject and physician consent was obtained to review the charts. The study was approved by the ethics review boards of the University of Toronto, Sunnybrook and Women's College Health Sciences Centre, and Princess Margaret Hospital–Ontario Cancer Institute.

Chart abstraction procedures were as follows. Words or phrases taken from the chart were recorded verbatim by trained and experienced chart abstractors, including one of us (J.N.), on a study chart review form (CRF); any interpretations based on those words or phrases were made by members of the study team. CRF interpretations relating to defining screening PSAs were made independently on cases and controls by P.S.B. and J.N., with the exception of very simple charts (e.g., with no PSA recorded or simply "rectal negative" with no other pertinent information), which were coded by J.N. alone. Any discrepancies were taken to an additional member of the study team (usually N.F. or V.G.). P.S.B. was blinded as to case or control identity (though this was sometimes obvious).

2.2. PSA coding

There is some controversy in the literature over whether PSA testing for patients who have symptoms of urinary