Decision models for distinguishing between clinically insignificant and significant tumors in prostate cancer biopsies: an application of Bayes' Theorem to reduce costs and improve outcomes

Arthur J. Swersey [1] · John Colberg [2] · Ronald Evans [3] · Michael W. Kattan [4] · Johannes Ledolter [5] · Rodney Parker [6]

Abstract

Prostate cancer is the second leading cause of death from cancer, behind lung cancer, for men in the U. S, with nearly 30,000 deaths per year. A key problem is the difficulty in distinguishing, after biopsy, between significant cancers that should be treated immediately and clinically insignificant tumors that should be monitored by active surveillance. Prostate cancer has been over-treated; a recent European randomized screening trial shows overtreatment rates of 40%. Overtreatment of insignificant tumors reduces quality of life, while delayed treatment of significant cancers increases the incidence of metastatic disease and death. We develop a decision analysis approach based on simulation and probability modeling. For a given prostate volume and number of biopsy needles, our rule is to treat if total length of cancer in needle cores exceeds c, the cutoff value, with active surveillance otherwise, provided pathology is favorable. We determine the optimal cutoff value, c^*. There are two misclassification costs: treating a minimal tumor and not treating a small or medium tumor (large tumors were never misclassified in our simulations). Bayes' Theorem is used to predict the probabilities of minimal, small, medium, and large cancers given the total length of cancer found in biopsy cores. A 20 needle biopsy in conjunction with our new decision analysis approach significantly reduces the expected loss associated with a patient in our target population about to undergo a biopsy. Longer needles reduce expected loss. Increasing the number of biopsy cores from the current norm of 10–12 to about 20, in conjunction with our new decision model, should substantially improve the ability to distinguish minimal from significant prostate cancer by minimizing the expected loss from over-treating minimal tumors and delaying treatment of significant cancers.

Keywords Prostate biopsy · Simulation · Decision analysis · Bayes' theorem

1 Introduction

Prostate cancer is the second leading cause of death from cancer, behind lung cancer, for men in the United States, with nearly 30,000 deaths per year (www.cancer.net/cancer-types/ prostate-cancer/statistics). Whether to screen men for prostate cancer is controversial, provoking heated arguments on both sides of the issue. Screening for the presence of prostate cancer consists of a blood test to measure prostate specific antigen (PSA), which is a marker for prostate cancer, and typically a digital rectal examination in which the physician looks for lumps or other prostate abnormalities. Under screening, an elevated PSA, usually defined as more than 4 ng/mL or an abnormal digital exam, leads to a prostate biopsy. In a biopsy, a number of needles, guided by ultrasound, referred to as trans-rectal ultrasonography (TRUS), are inserted into the prostate, with each needle extracting a core of tissue later examined by a radiologist for the presence of cancer. If cancer is detected, the patient is either treated by surgery (prostatectomy) or radiation. or not immediately treated but monitored over time by further PSA blood tests, a process referred to as "active surveillance."

Two major randomized trials, one in the U.S. [1], and the other in Europe [2], addressed the benefits of screening and

✉ Arthur J. Swersey
  arthur.swersey@yale.edu

1   Yale School of Management, Box 208200, New Haven, CT 06511, USA

2   Department of Urology, Yale Medical School, New Haven, CT, USA

3   Prognos, New York, NY, USA

4   Cleveland Clinic, Cleveland, OH, USA

5   Henry B. Tippie College of Business, University of Iowa, Iowa, IA, USA

6   Indiana University, Bloomington, IN, USA

yielded conflicting results. In the European trial, screening reduced the death rate from prostate cancer by an estimated 21% at a 13-year follow-up, while the U.S. study showed no significant reduction in the death rate for those that were screened. The argument in favor of screening is that doing so will mean detecting and treating significant prostate cancers early and thereby saving lives. The argument against screening is that prostate cancer is a disease that has frequently been over-diagnosed and over-treated [3, 4]. The European trial [2] showed that about 40% of surgeries in the screened group were unnecessary, meaning that the volume of cancer and its pathology were such that not treating would have resulted in minimal risk to the patient. The cost of misclassifying and treating such tumors that are called minimal or indolent includes the monetary cost of the treatment (surgery or radiation) plus its associated mortality and morbidity, as well as the common treatment side effects of impotence and incontinence. The European study also showed that screening resulted in a 40% reduction in metastatic disease at diagnosis for patients in the screened group [5]. Metastatic disease means the cancer has spread beyond the prostate and consequently the patient typically must undergo repeated hormone therapy with its associated negative side effects and faces diminished life expectancy.

In 1992, the American Urological Association (AUA) and the American Cancer Society (ACS) recommended annual PSA screening for men 50 years of age and older [6, 7]. In 2008, the US Preventive Services Task Force (USPSTF) recommended that men 75 years of age and older not be screened [8], and in a draft public statement in October 2011 followed by a final document in May 2012, the organization recommended against screening for all men, citing that the benefits did not outweigh the negative effects [9]. Drazer et al. [10] reported a reduction in screening rates after the USPSTF draft and final recommendations. The authors estimated that screening declined in the period from 2010 to 2013 from 33.3% to 24.8% for men aged 50–59, and from 51.2% to 46.3% for men aged 60–74. Jemal et al. [11] reported that for men 50 years of age and older, prostate cancer incidence per 100,000 men declined by 16% from 2011 to 2012, which the authors attributed to a reduction in the incidence of screening. Weiner et al. [12] estimated that the number of men with metastatic (advanced) prostate cancer increased by 72% from 2004 to 2013. While the authors concluded that this increase may in part be the result of the USPSTF recommendations against screening, they also recognized that the increase in metastatic disease began before the release of these recommendations.

A number of authors [13–15] have developed models that aim to optimize the screening policy. A screening policy specifies the age at which to start screening, the PSA level that triggers a biopsy, which typically is age dependent, the interval between PSA tests, and the age at which screening stops.

These authors use simulation or Markov models of cancer spread over time and typically aim to maximize expected quality-adjusted life years (QALYs). Bertsimas et al. [16] recognized that alternative screening models may lead to conflicting decisions. The authors' optimization framework compared three screening models [13, 15, 17] that estimate quality-adjusted life expectancy (QALE). From these models they identified 64 screening strategies (policies) that form an efficient frontier based on the tradeoff between the *average* of the three models' assessments of the change in QALE compared to no screening, and the *minimum* of the three models' assessments of the change in QALE compared to no screening.

In 1989, Hodge et al. [18] standardized the placement of needles in the biopsy procedure by developing the sextant method in which six needles were placed in specified positions. Because the sextant method missed many significant tumors, researchers including Presti et al. [19] and Epstein et al. [20] developed procedures in which additional needles were placed more laterally, that is, toward the edges of the prostate. In the next section we show the placement of needles in the sextant method as well as the locations of additional needles placed more laterally.

The output of a prostate biopsy typically shows the identity and number of needle cores that are positive for cancer, the length of cancer in each core, the fraction of core length containing cancer, and the pathology of the cancer detected as measured by the Gleason score. A higher Gleason score identifies a more aggressive tumor. The Gleason score consists of two integers or grades, each between 2 and 10. The first grade represents the score for the primary pattern of the tumor (more than 50% of the total pattern viewed), while the second grade represents the score of the secondary pattern. These two scores are added to obtain the total Gleason score. A total Gleason score of 6 or less would be classified as low grade cancer. Scores above 6 indicating higher grade tumors would generally lead to immediate treatment regardless of the amount of cancer found on biopsy.

The forgoing discussion highlights the screening dilemma. Screening will identify and then immediately treat some important cancers. But at the same time treatment (surgery or radiation) in response to positive biopsies has led to a high proportion of unnecessary treatments with negative consequences due to the risks of surgery or radiation, and the associated side effects of impotence and incontinence. A key problem in prostate cancer research and practice that we address in this paper is the difficulty in distinguishing between significant tumors that should be treated immediately and clinically insignificant tumors to be monitored by active surveillance. Boccon-Gibod aptly described this as "identifying the tigers from the pussycats." [21]. To address this problem, researchers using statistical regression analysis developed criteria and nomograms to predict cancers they call either

minimal or indolent based on biopsy results, prostate specific antigen (PSA), and other factors [22–29]. The well-known Epstein criteria [22, 23] that were developed using sextant biopsies predicts the presence of minimal cancer if 1) PSA density (PSA/prostate volume) < 0.15 ng/mL; 2) Gleason score ≤ 6; 3) cancer in two or fewer biopsy cores; and 4) no more than 50% cancer in any single core. The authors of nomograms [24–29] evaluated decision rules that specified no surgery (active surveillance) if the probability of minimal cancer exceeded varying thresholds. These studies have accuracies of 80% or more based on the area under the receiver operating characteristic curve (AUC), but the decision rules have performed much less well in correctly classifying tumors found on biopsy as either minimal or significant. Steyerberg et al. [26] reported that the rule no surgery if the probability of indolent cancer >30% would have misclassified only 7% of indolent tumors (calling for surgery and overtreatment) but 65% of important cancers (calling for no surgery). The rule no surgery if the probability of indolent cancer >60% would have misclassified only 15% of important cancers but 54% of indolent cancers. In a validation study of several nomograms, the Kattan et al. [24] and Steyerberg et al. [26] nomograms performed best, while all were more accurate in identifying significant cancers than they were at identifying indolent ones [30]. Nguyen and Kattan discuss the strengths and limitations of nomograms [31].

A shortcoming of the regression/nomogram approach is that it is based on a particular number of needle cores, and thus one cannot determine the effect of increasing that number and changing the length of needles, which we do in this paper. In most of the regression/nomogram approaches cited above biopsies mainly followed the sextant method [22–29]. In Section 5 (Discussion and conclusions), we identify and discuss three lesser-known nomograms based on 12-core biopsies, which perform better than the sextant-based regression/nomograms in terms of sensitivity and specificity [32–34], findings that are consistent with the predictions of our model.

According to the 2016 version of the European Association of Urology (EAU) Guidelines on prostate cancer, TRUS-guided systematic prostate biopsy is the standard of care with 10–12 cores taken [35]. In arriving at this recommended standard of care, the EAU guidelines paper discusses some encouraging studies of the benefits of multi-parametric magnetic resonance (MRI) imaging in detecting aggressive tumors, but report that two randomized trials to evaluate the benefits of MRI combined with systematic biopsies showed contradictory results [35].

The EAU guidelines paper notes that since the current U. S Preventative Task Force recommendation on screening there has been a significant number of men whose aggressive cancers have not been detected [35]. For screening to be useful, an effective decision rule is needed to decide, after biopsy, which patients should be treated immediately and which should be followed with active surveillance. In this paper we address this need. The overall aim of our new decision analysis approach, which is based on computer simulation and probability modeling, is to improve the ability to distinguish minimal (indolent) tumors from significant cancers, and in particular to determine how increasing the number of cores in TRUS-guided biopsy beyond the current norm of 10–12 would affect this ability to distinguish.

In a broader sense, our research relates to improving health care delivery. Health care researchers and clinicians have intensified their focus on improving patient experience and outcomes, and reducing costs. The Triple Aim methodology developed by the Institute for Healthcare Improvement (IHI) is a framework that identifies three aims: improving the health of populations, improving the experience of care (including quality and patient satisfaction), and reducing per capita costs of health care [36]. Better prostate cancer treatment decision making leads to improvements in health care delivery as it relates to the Triple Aim dimensions. Better decision models improve patient outcomes by not delaying the treatment of significant cancers, which will reduce the incidence of metastatic disease and prostate cancer deaths. It will also lower the number of unnecessary treatments that lead to diminished quality of life. Better models should improve patient satisfaction and the experience of care by giving the patient the opportunity to participate in the decision making process, providing the patient with the information needed to make more informed choices, including useful predictions of the severity of his disease, and incorporating patient preferences regarding the cost tradeoffs between treating an insignificant tumor and not immediately treating a small or medium cancer (which are inputs to the model developed in this paper). Better patient outcomes should also lead to increases in patient satisfaction. Finally, improvements in decision making will reduce the costs associated with unnecessary treatments as well as those related to managing the progression of disease in patients whose treatments were delayed.

2 Mathematical models

2.1 Simulation model

Our Excel computer simulation model is similar to that of Daneshgari et al. [37]. We represent the prostate as an ellipsoid and tumors as spheres, with the tumor centers randomly located in the posterior part of the prostate. We made use of the prostatectomy data set of Epstein et al. [22, 23] of 157 men with negative digital rectal exams who underwent radical prostatectomy. These authors reported after surgery tumor volume V (in cc) in six intervals: V < 0.2 (17% of cases), 0.2 ≤ V < 0.5 (17% of cases), 0.5 ≤ V < 1.0 (13% of cases), 1.0 ≤ V < 4.0 (42% of cases), 4 ≤ V < 10.0 (8% of cases), and V ≥ 10

(3% of cases). We combined the two smallest intervals and included the 3% of cases with $V \geq 10$ in the 4 cc to 10 cc interval. This gave us tumors of four different volumes: minimal, small, medium and large. We defined a minimal tumor as one that is less than 0.5 cc in volume [38], and represented it in the simulation with a tumor of 0.25 cc. We defined a small tumor as having a volume between 0.50 cc and 1.0 cc, and represented it by a tumor of 0.75 cc. A medium tumor had a volume between 1.0 cc and 4.0 cc, and was represented as 2.5 cc, and a large tumor had a volume between 4 cc and 10 cc, and was represented as 7 cc. Prostate tumors are generally multi-focal [39] and we assumed each tumor consisted of three foci, one major and two minor. The major tumors were 0.25 cc, 0.75 cc, 2.5 cc, and 7 cc, while each minor tumor was 1/6 the size of its corresponding major tumor. Following Daneshgari et al. [37] and based on their measurements of 159 radical prostatectomy specimens, our simulated prostates had ratios of length to height to width of 3: 2.7: 1.9. We overlaid three sets of lines onto the prostate as shown in the posterior (axial) view of the simulated prostate in Fig. 1a and b. The A line is used as a reference, the two B lines represent the potential positions of needles under the sextant method, while the two C lines represent the positions of laterally placed needles. The x axis represents prostate width, the y axis represents prostate length, and the z axis, which is perpendicular to the x-y plane, represents prostate depth. Figure 1a and b also show the horizontal centerline, which bisects the length of the prostate. We simulated the placement of 6 (sextant method), 14, and 20 needles in prostates of size 30 cc or 55 cc, with needles of length 1.5 cm inserted perpendicular to the x-y plane. We assumed each needle removed a 1.5 cm long core of tissue. The placements of 6 (sextant method) and 14 needles are shown in Fig. 1a and b. In the 6 needle case 3 needles were positioned on each B line, with each B line halfway between the A line and the far edge of the prostate. In the 14 needle case, additional needles were positioned laterally along the C lines. In this case (and the 20 needle case), each B line was moved to the left toward the center of the prostate to improve spacing. Each B line was 1/3 of the distance from the A line to the far edge of the prostate, and each C line was 2/3 of the distance from the A line to the far edge of the prostate. In the 14 needle case 5 needles were positioned on each B line and 2 needles were positioned on each C line. In the 20 needle case (figure not shown), 6 needles were equally spaced on each B line, while 4 needles were equally spaced on each C line.

The needles were not placed evenly across the prostate because tumor density is not uniform. Zeng et al. [40] showed that the density of tumor incidence is highest in the lateral areas of the prostate, which supports the practice of placing needles more laterally [19, 20]. Based on the Zeng et al. results [40], for each tumor size, we randomly placed the center of the tumor within the prostate in such a way that 1/3 of the tumor fell in each lateral area, from the C lines to the edges of the
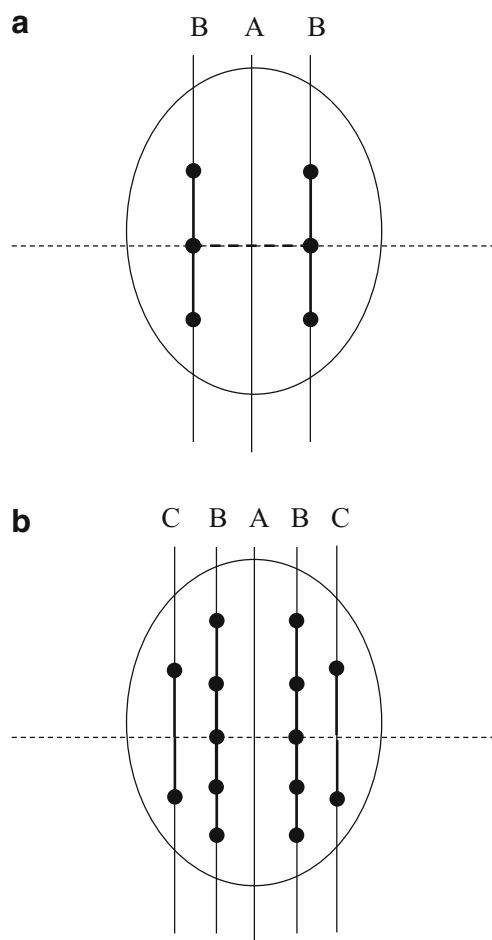
Fig. 1 **a** Positioning of Six Needles. The A line is used as a reference. 3 needles were placed on each B line as shown. Each B line was halfway between the A line and the far edge of the prostate. **b** Positioning of 14 Needles. The A line is used as a reference. Each B line was 1/3 of the distance from the A line to the far edge of the prostate, and each C line was 2/3 of the distance from the A line to the far edge of the prostate. As shown, 5 needles were placed on each B line, and 2 needles were placed on each C line.

prostate, with 1/3 in the remainder of the prostate. Because the lateral areas represent a relatively small volume of the prostate, placing 1/3 of the tumors in each lateral area reflects the higher tumor density in those areas. If a tumor sphere was partially outside the prostate, we added the volume outside to the volume inside. We did this by increasing the radius of the segment lying inside the prostate. Our decision rule to treat or not to treat is based on the total length of cancer found in biopsy cores. We determined from our simulation experiments that our results were not sensitive to the incidence distribution; in particular, we found that using an incidence of 40% of tumors in each lateral area produced the same optimal cutoff values for our decision rule as the tumor incidence of 1/3 in each lateral area.

With a plane parallel to the x-y plane, we divided the prostate into two parts—posterior and anterior, and randomly generated tumor centers in the posterior part. The distance from

the dividing plane to the posterior apex was 0.4 times the prostate height. Rather than dividing the prostate into equal volume posterior and anterior parts, this approach made distant tumors more accessible to the needles. The simulation results we report are based on this approach. Sensitivity analysis showed that dividing the prostate into equal volume posterior and anterior parts did not change the optimal cutoff values of our decision rule. Each simulation experiment of 10,000 runs specified the prostate volume (30 cc or 55 cc), the volume of the tumor, and the number of needles. The output of each experiment was the probability of detection and the probability distribution of the total length of cancer found in the needle cores.

2.2 Estimating the prevalence of minimal, small, medium, and large tumors

We consider patients with negative digital rectal exam and PSA between 4 ng/mL and 10 ng/mL These patients are similar in that their PSA levels are elevated but not extremely so. Tanaka et al. refer to this range of PSA values as "the grey zone." [41]. We did not have access to individual PSA score values in the data set we used [22, 23], but the data did allow us to focus on patients in this important grey zone range of PSA values. Our decision rule is treat if total length (TL) of cancer in biopsy cores > c, the cutoff, and active surveillance otherwise, provided pathology is favorable (combined Gleason score less than 7). If combined Gleason score is 7 or more, decision is always to treat. Kajikawa et al. found that total length of cancer in cores is the optimal measure for predicting the presence of minimal cancer [33]. The determination of the optimal cutoffs, which we carry out in Section 3.2, required estimates of the prior probabilities of tumors of various sizes. Epstein et al. reported on 157 patients with negative digital rectal exams, median gland volume of 55.5 cc, and median PSA of 8.1 ng/mL, who underwent radical prostatectomy from 1988 through 1992 [22, 23]. We assumed most biopsies were performed using the sextant method (the standard at the time), and that all positive biopsies led to surgeries. We used their reported after-surgery tumor volume percentages for patients with PSA between 4 ng/mL and 10 ng/mL as posterior probabilities of minimal (Min), small (S), medium (Med), and large (L) tumors, given a positive sextant biopsy. Let P(Min), P(S), P(Med), and P(L) represent the prior (before biopsy) probabilities of minimal, small, medium, and large tumors respectively. Let $P(Min|+)$ be the probability of a minimal tumor given a positive sextant biopsy, let $P(+|Min)$ be the probability of a positive biopsy given a minimal tumor, and let P(+) be the probability of obtaining a positive biopsy for a patient in

our target population. Applying Bayes' Theorem, we backed out P(Min) in the following way. We have

$$P(Min|+) = \frac{P(Min)P(+|Min)}{P(+)},$$

Rearranging terms we have

$$P(Min) = \frac{P(Min|+)P(+)}{P(+|Min)}.$$

The three terms on the right hand side of this equation were determined as follows: $P(Min|+)$, the posterior probability of a minimal tumor, is 0.3, from the Epstein et al. data set [22, 23]; it is the fraction of patients identified after surgery as having minimal tumors (less than 0.5 cc), P(+) = 0.2. This value is based on data reported by Cooner et al. [42], Catalona et al. [43], and Schwartz et al. [44] for biopsies performed in periods 1987–1989, 1991–1992, and 1993–1997 respectively, when the sextant method was the norm, which showed that for patients with impalpable tumors and PSA between 4 and 10, about one in five biopsies were positive. $P(+|Min)$, the probability of detecting a minimal tumor in a 55 cc prostate (the median gland volume in the Epstein et al. data set [22, 23]) using the sextant method, was estimated from our simulations (it is 0.5103). Substituting these values into the above equation for P(Min) led to the prior (to biopsy) probability P(Min) = 0.118. Similar calculations led to the prior probabilities or prevalences of small, medium, and large tumors respectively. They were P(S) = 0.043, P(Med) = 0.086, and P(L) = 0.022. The probability of no cancer was 0.731. Bayes' Theorem is typically used to determine positive predictive values (posterior probabilities) given prior probabilities (prevalences). Here we reversed the procedure and worked backwards from the posterior probabilities to obtain the priors (prevalences).

3 Finding optimal cutoff values: expected loss equation

3.1 Two types of misclassification errors

As presented earlier, our decision rule is treat if total length (TL) of cancer in biopsy cores > c, the cutoff, and active surveillance otherwise, provided pathology is favorable (combined Gleason score less than 7). If combined Gleason score is 7 or more, the decision is always to treat. In applying this decision rule, there are two misclassifications: treating a minimal tumor, and not treating a small or medium tumor (In our simulations, we found for all reasonable cutoff values that large tumors were never misclassified). If the cancer is minimal it will be misclassified and treated if TL > c, the cutoff value. If the cancer is significant (small or medium), it will be

misclassified and not treated if TL ≤ c. The cost of misclassifying and treating a minimal tumor includes the monetary cost of the treatment (surgery or radiation) plus associated mortality and morbidity, as well as possible impotence and incontinence, while the costs of not treating larger tumors relate to disease progression. For a given number of needle cores and prostate volume, we determined the cutoff, c*, that minimizes the expected or average loss associated with a patient in our target population about to have a biopsy. The average loss is the sum of three terms. Each term is the prior probability of a tumor of specified size (minimal, small, or medium) times the probability of misclassification times the cost of misclassification. As the cutoff increases, the likelihood of treating a minimal tumor decreases while the likelihood of missing small and medium tumors increases. By varying the cutoff, we obtained the optimal cutoff value which minimized the expected loss.

3.2 Finding the optimal cutoff value

For each prostate volume (30 cc or 55 cc), we found by our simulation experiments the frequency distributions of total length (TL) of cancer in biopsy cores. For a specified decision rule cutoff c, a misclassification cost, C_{Min}, occurs if the cancer is minimal and treated (its $TL > c$); a misclassification cost, C_S, occurs if the tumor is small and not treated (its $TL \leq c$); and a misclassification cost, C_{Med}, occurs if the tumor is medium and not treated (its TL is $\leq c$). Rather than estimating these costs directly, which for C_S and C_{Med} would require a tumor spread model [14, 15], we employ sensitivity analysis varying the ratios of the cost of not detecting a small or medium tumor to the cost of treating a minimal tumor. We define these ratios as $k_S = C_S/C_{Min}$, and $k_{Med} = C_{Med}/C_{Min}$. Note that these ratios may vary with individual patient age and preferences.

The expected loss for a patient in our target population about to undergo a biopsy is given by

$$E(Loss) = C_{Min}P(TL > c|Min)P(Min) + C_S P(TL \leq c|S)P(S) + C_{Med}P(TL \leq c|Med)P(Med)$$
$$= C_{Min}P(TL > c|Min)P(Min) + k_s C_{Min}P(TL \leq c|S)P(S) + k_{Med}C_{Min}P(TL \leq c|Med)P(Med)$$
$$= C_{Min}\{P(TL > c|Min)P(Min) + k_s P(TL \leq c|S)P(S) + k_{Med}P(TL \leq c|Med)P(Med)\}$$

4 Results

Our simulation results show that increasing the number of needles increased the area under the receiver operating characteristic curve (ROC) (Fig. 2). The ROC curve is a graphical plot that shows the diagnostic ability of a binary classifier as its discrimination threshold is varied. In our case the discrimination threshold is the cutoff value, c, of our decision rule. Sensitivity (1 – type II error) is the probability of not treating a minimal tumor if it is present, while specificity (1 – type I error) is the probability of treating a small tumor if it is present. Each point on a particular curve represents a cutoff value, c. Moving from left to right on a curve shows increasing cutoff values in increments of 0.2 cm. As the cutoff value increases, the probability of not treating a minimal tumor (sensitivity) increases, while the probability of treating a small tumor (specificity) decreases and hence 1 – specificity increases. Ideally, one wants to achieve both high sensitivity and high specificity. In our case (Fig. 2), the AUC measures how well our model performs in distinguishing between minimal tumors that should not be treated and small tumors that should be. The diagnostic ability increases with increases in the area under the curve. As Fig. 2 shows, and as we discuss in more detail

below, sampling with 20 needles achieved both high levels of sensitivity and specificity.

Figure 3 (30 cc prostate) and Fig. 4 (55 cc prostate) show expected loss expressed in terms of C_{Min} versus cutoff for 6, 14, and 20 needles. For the 30 cc prostate with $k_S = C_S/C_{Min} = 2$ (that is $C_S = 2C_{Min}$) and $k_{Med} = C_{Med}/C_{Min} = 8$ ($C_{Med} = 8C_{Min}$), shown in Fig. 3, panel a, minimum expected losses for 6, 14, and 20 needles are $0.0466C_{Min}$, $0.0190C_{Min}$, and $0.0121C_{Min}$. For 14 needles, compared to the 6-needle base case, this is a reduction in expected loss of 59% at the optimal cutoffs, while 20 needles provide an additional reduction of 15% compared to the 6-needle base case. For the 30 cc prostate with $k_S = C_S/C_{Min} = 1$ (that is $C_S = C_{Min}$) and $k_{Med} = C_{Med}/C_{Min} = 4$ ($C_{Med} = 4C_{Min}$), shown in Fig. 3, panel b, minimum expected losses for 6, 14, and 20 needles are $0.0299C_{Min}$, $0.0121C_{Min}$, and $0.0072C_{Min}$. For 14 needles compared to the 6-needle base case, this reduction at optimal cutoffs is 60% for 14 needles, while 20 needles provide an additional reduction of 16% compared to the 6-needle base case. For the 55 cc prostate with $k_S = C_S/C_{Min} = 2$ (that is $C_S = 2C_{Min}$) and $k_{Med} = C_{Med}/C_{Min} = 8$ ($C_{Med} = 8C_{Min}$), shown in Fig. 4, panel a, minimum expected losses for 6, 14, and 20 needles are $0.0896C_{Min}$, $0.053C_{Min}$, and $0.0305C_{Min}$. For 14 needles compared to 6 needles, this is a reduction of 41%,

Fig. 2 ROC Curves for 6, 14, and 20 Needles: Sensitivity is the probability of not treating a minimal tumor if it is present, specificity is the probability of treating a small tumor if it is present; Panel **a** is 30 cc prostate, Panel **b** is 55 cc Prostate; Area under the curve increases with number of needles
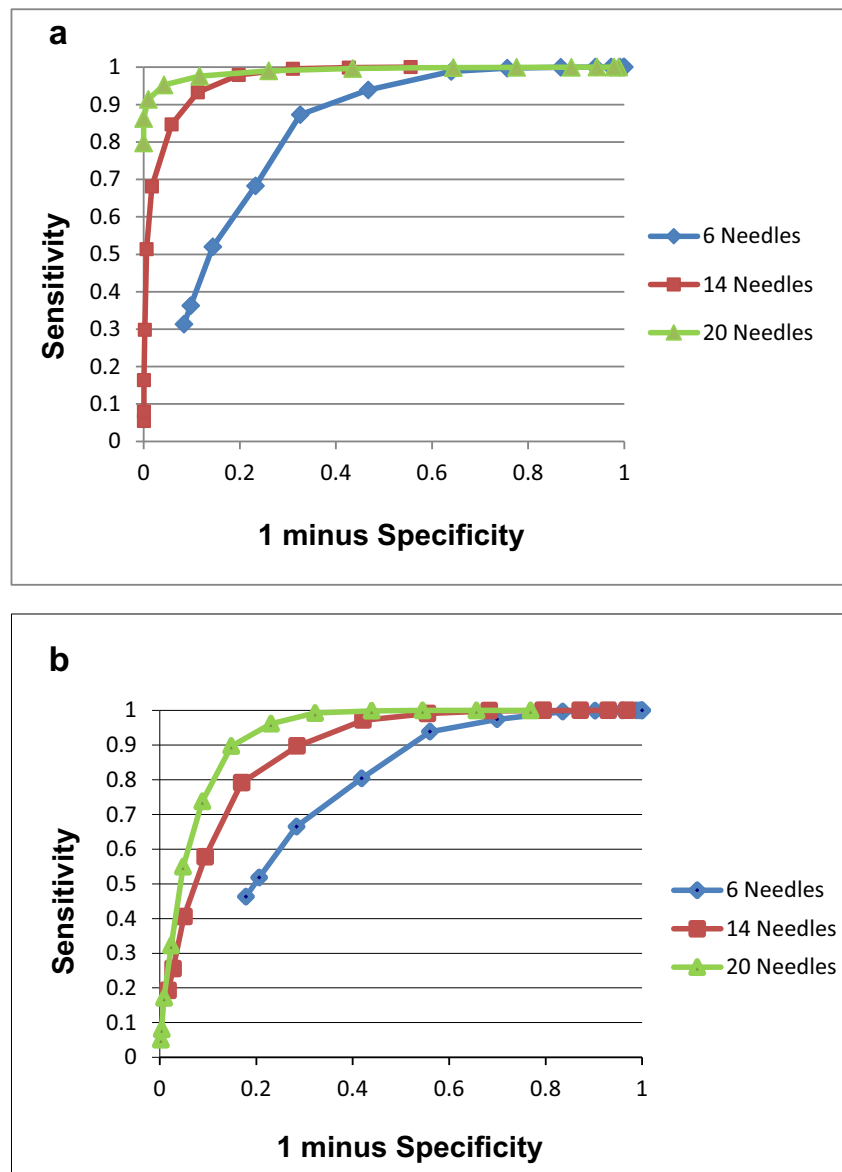

while 20 needles compared to the 6-needle base case, provides an additional reduction of 25%. For the 55 cc prostate with $k_S = C_S/C_{Min} = 1$ (that is $C_S = C_{Min}$) and $k_{Med} = C_{Med}/C_{Min} = 4$ ($C_{Med} = 4C_{Min}$), shown in Fig. 4, panel b, minimum expected losses for 6, 14, and 20 needles are $0.0636C_{Min}$, $0.0378C_{Min}$, and $0.0199C_{Min}$. For 14 needles compared to 6 needles, this is a reduction of 41%, while 20 needles compared to the 6-needle base case provides an additional reduction of 28%. Thus the benefit of 20 needles is significantly greater in the larger 55 cc prostate.

Table 1 shows optimal cutoffs as a function of prostate volume and number of needles. For each cutoff, the table lists the k_s and k_{Med} values, the percentages of minimal tumors hit (detected and treated) and the percentages of

small and medium tumors missed (not treated). Sensitivity is 1 – percent of minimal tumors hit and specificity is 1 – percent of small tumors missed. These percentages were obtained from our simulation experiments. Table 1 shows that increasing the number of needles (from 6 to 14 to 20) decreases both minimal tumors hit and small and medium tumors missed, hence increasing both sensitivity and specificity. The optimal cutoffs are largely insensitive to changes in the cost ratios. We observe in Table 1 that the optimal cutoff c* moved at most one interval of tumor length as the cost ratios, k_S and k_{Med} were varied. Also as shown in Figs. 3 and 4, each expected loss curve is very flat near its minimum cutoff value, and hence the difference in expected loss between adjacent cutoff values

Fig. 3 Expected Loss Versus
Cutoff for 6, 14, and 20 Needles,
30 cc Prostate; Panel **a** has cost of
missing small tumor 2 times cost
of hitting minimal tumor, cost of
missing medium tumor 8 times
cost of hitting minimal tumor;
Panel **b** has cost of missing small
tumor equal to cost of hitting
minimal tumor, cost of missing
medium tumor 4 times cost of
hitting minimal tumor
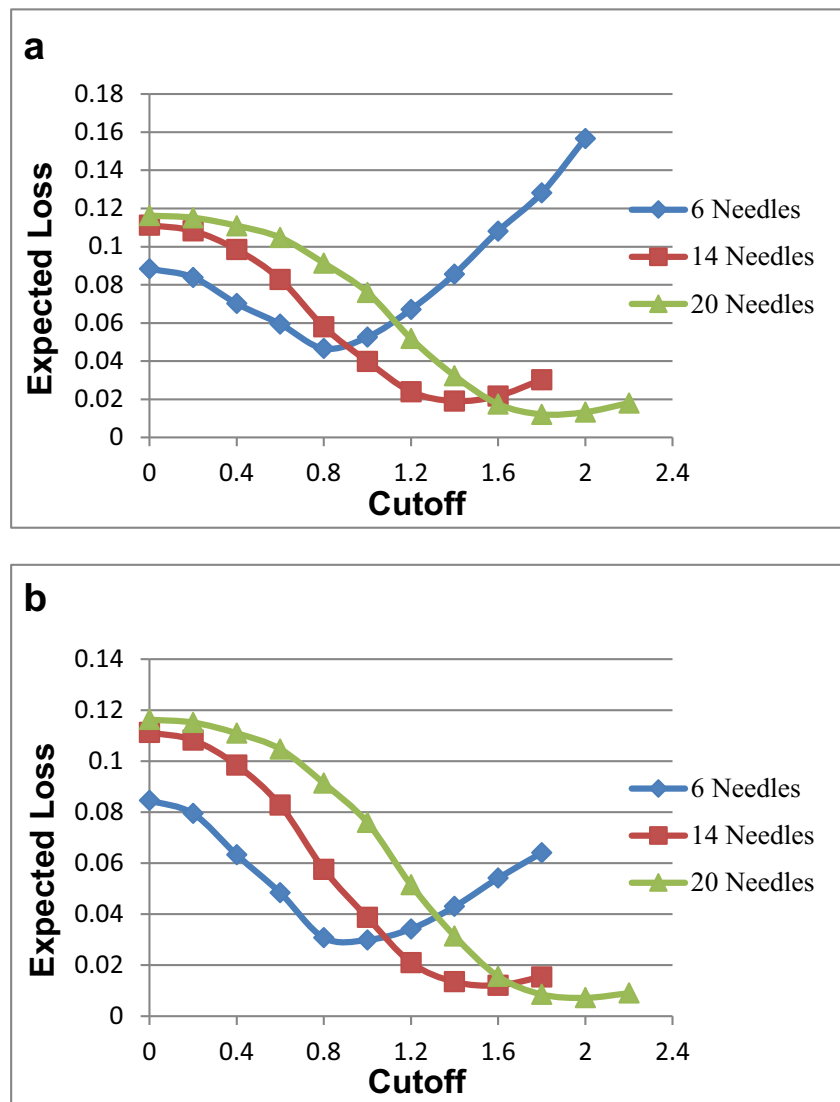
is very small. For example, as shown in Fig. 4, panel b, the optimal cutoff for 20 needles when $k_S = k_{Min}$ and $k_{Med} = 4k_{Min}$ is 1.4 cm with expected loss of $0.0199 C_{Min}$. At the adjacent cutoff of 1.2 cm on the same 20 needle curve of Fig. 4, panel b, which is the minimum cutoff for 20 needles in Fig. 4, panel a (based on $k_S = 2k_{Min}$ and $k_{Med} = 8k_{Min}$) the expected loss is $0.0206 C_{Min}$ which is only 3.5% higher than the loss of $0.0199 C_{Min}$ at the optimal cutoff. This supports our decision to use sensitivity analysis varying the cost ratios, rather than employing a tumor growth model to estimate the costs of missing small and medium tumors [14, 15].

Figure 5a plots for a 55 cc prostate and 20 needles the frequency distributions of length of cancer in cores for a minimal tumor (lower panel) and a small tumor (upper panel); each distribution was found from 10,000 simulation runs. The misclassifications occurred because of the overlap in the two frequency distributions: a fraction of the

minimal tumor lengths fell above the cutoff and would mean treatment of the minimal tumor, while a fraction of the small tumor lengths fell below it and would mean missing the small tumor. The shaded area in the lower panel to the right of the optimal cutoff represents the fraction of minimal tumors treated (misclassifications), which is 0.090 ± 0.0029[1]. The shaded area in the upper panel to the left of the optimal cutoff represents the fraction of small tumors missed, which is 0.148 ± 0.0036[1] These misclassification percentages are also shown in the last line of Table 1. Figures for 6 and 14 needles (not shown) demonstrated that increasing the number of needles reduced the overlap of the two frequency distributions, which is reflected in the results in Table 1.

---

[1] We calculate the standard error of the proportion estimate from 10,000 independent simulation runs (trials) using the binomial distribution.

Fig. 4 Expected Loss Versus
Cutoff for 6, 14, and 20 needles,
55 cc Prostate; Panel **a** has cost of
missing small tumor twice cost of
hitting minimal tumor, cost of
missing medium tumor 8 times
cost of hitting minimal tumor;
Panel **b** has cost of missing small
tumor equal to cost of hitting
minimal tumor, cost of missing
medium tumor 4 times cost of
hitting minimal tumor
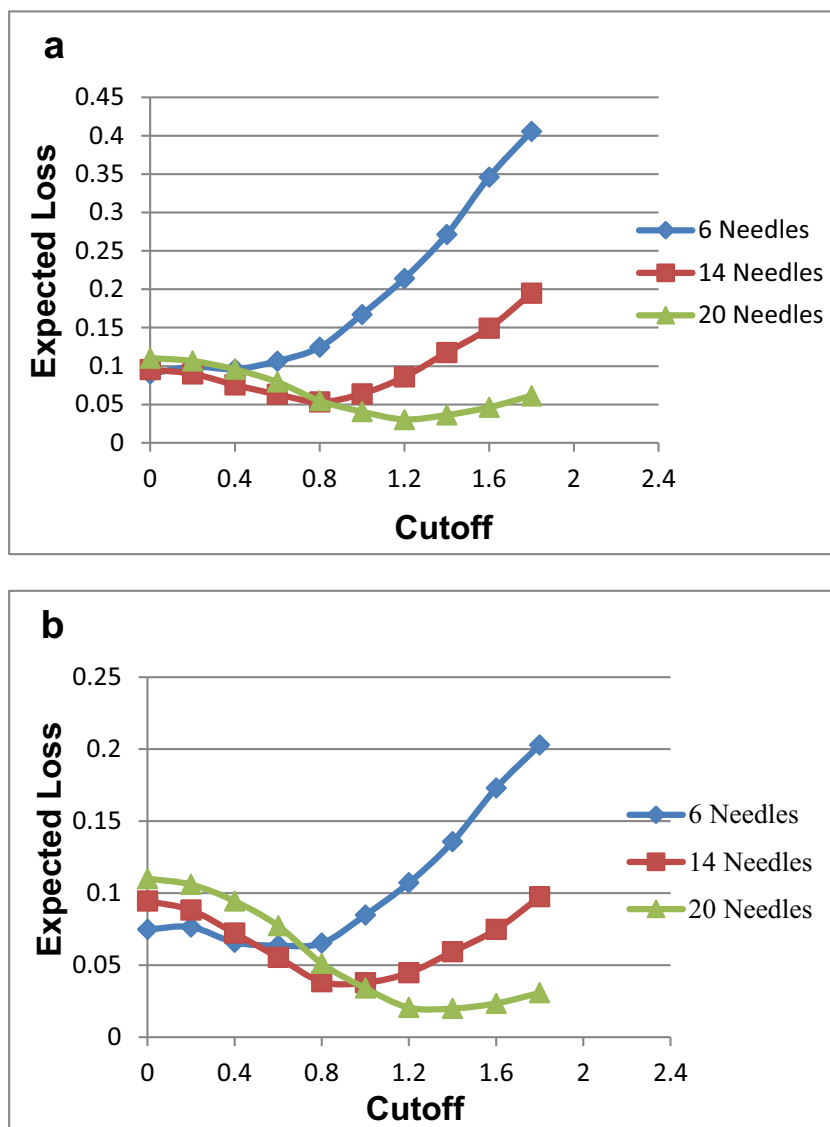
We experimented with different needle lengths and found that a shorter needle degraded performance greatly. Figure 5a and b compare for a 55 cc tumor and 20 needles, the fraction of small tumors missed and the fraction of minimal tumors treated under optimal cutoffs for needles (and cores) of length 1.5 cm. and 1.0 cm respectively. For the 1.5 cm needle and core, the average fraction of minimal tumors treated was a little smaller $(0.090 \pm 0.0029$ versus $0.103 \pm 0.00304)$, while the average fraction of small tumors missed was lower by 57% $(0.148 \pm 0.0036$ versus $0.343 \pm 0.00475)$. We found that for the 30 cc prostate and 14 needles, the 1.0 cm needle core compared to the 1.5 cm needle core increased the expected loss at the optimal cutoffs by 71% (Fig. 6), while for the 55 cc prostate, 20 needle case, the shorter 1.0 cm core increased the expected loss

at the optimal cutoffs by 55% (figure not shown). Several authors have shown that longer core lengths improved cancer detection and that core lengths obtained were often considerably shorter than the length of the needle [45, 46]. We are not suggesting that 1.0 cm needles are used in practice. But 1.5 cm needles are commonly used and result in a significant fraction of cores being 1.0 cm or less.

4.1 Posterior probabilities

An output of each 10,000 run simulation (gland size, tumor size, and number of needles) is the frequency distribution of the total cancer length in cores (TL) which we tabulated in 0.2 cm-wide intervals. Using these results, we apply Bayes' Theorem to estimate the posterior probabilities or positive

Table 1 Optimal cutoff values as a function of prostate volume and number of needles

Prostate volume	Number of needles	Cutoff c* (cm)	k_s	k_{Med}	Percent minimal hit	Percent small missed	Percent medium missed
30 cc	6	1.0	0.5	4	6.1	46.7	1.5
		1.0	1	4	6.1	46.7	1.5
		0.8	2	4	12.7	32.6	1.1
		0.8	2	8	12.7	32.6	1.1
	14	1.6	0.5	4	2.1	22.4	0.0
		1.6	1	4	2.1	22.4	0.0
		1.4	2	4	6.7	13.0	0.0
		1.4	2	8	6.7	13.0	0.0
	20	2.0	0.5	4	1.0	14.1	0.0
		2.0	1	4	1.0	14.1	0.0
		1.8	2	4	4.2	8.6	0.0
		1.8	2	8	4.2	8.6	0.0
55 cc	6	0.8	0.5	4	5.1	60.4	9.7
		0.6	1	4	17.7	46.8	6.6
		0.4	2	4	30.4	32.9	4.6
		0.4	2	8	30.4	32.9	1.7
	14	1.0	0.5	4	10.0	30.1	3.8
		1.0	1	4	10.0	30.1	3.8
		0.8	2	4	20.1	18.9	2.0
		0.8	2	8	20.1	18.9	2.0
	20	1.4	0.5	4	3.1	26.6	1.4
		1.4	1	4	3.1	26.6	1.4
		1.2	2	4	9.0	14.8	0.6
		1.2	2	8	9.0	14.8	0.6

The term k_s is the ratio of the cost of missing a small tumor to the cost of hitting and treating a minimal tumor, k_{Med} is the ratio of the cost of missing a medium tumor to the cost of hitting and treating a minimal tumor. The percent of minimal tumors hit, the percent of small tumors missed, and the percent of medium tumors missed all decrease with increases in the number of needles

predictive values of minimum, small, medium and large tumors. The posterior probability of the presence of a minimal tumor given that TL is greater than l centimeters and less than or equal to u centimeters, is.

$$P(Min|l < TL \leq u) = \frac{P(l < TL \leq u|Min)P(Min)}{P(l < TL \leq u)}$$
$$= \frac{P(l < TL \leq u|Min)P(Min)}{P(l < TL \leq u|Min)P(Min) + P(l < TL \leq u|S)P(S) + P(l < TL \leq u|Med)P(Med) + P(l < TL \leq u|L)P(L)}$$

This formula was used to find the posterior probabilities of a minimal (insignificant) tumor for each interval of tumor length. Replacing Min by each of the other tumor sizes (S, Med, L), yields the posterior probabilities of small, medium, and large tumors as a function of interval of tumor length, prostate volume, and number of needles. Table 2 (30 cc prostate) shows for each interval of cancer length, the posterior probabilities of an insignificant cancer. The probability of a significant cancer is 1 – the probability of an insignificant cancer. The table also shows the percent of positive biopsies falling in each interval, obtained from our simulations. Because our representations of the prostate and tumors are idealized, these posterior probabilities are not ready to use predictions for individual patients; but are encouraging with respect to what could be achieved using a more detailed model of the prostate and tumors. For the 20-needle biopsy a patient

Fig. 5 **a** Fraction of Small
Tumors Missed and Fraction of
Minimal Tumors Treated, 55 cc
Prostate, 20 Needles of Length
1.5 cm, Optimal Cutoff of 1.2 cm,
cost of missing small tumor is 2
times cost of hitting minimal
tumor, cost of hitting medium
tumor is 8 times cost of missing
minimal tumor. **b** Fraction of
Small Tumors Missed and
Fraction of Minimal Tumors
Treated, 55 cc Prostate, 20
Needles of Length 1.0 cm,
Optimal Cutoff of 1.0 cm, cost of
missing small tumor is 2 times
cost of hitting minimal tumor,
cost of missing medium tumor is
8 times cost of hitting minimal
tumor
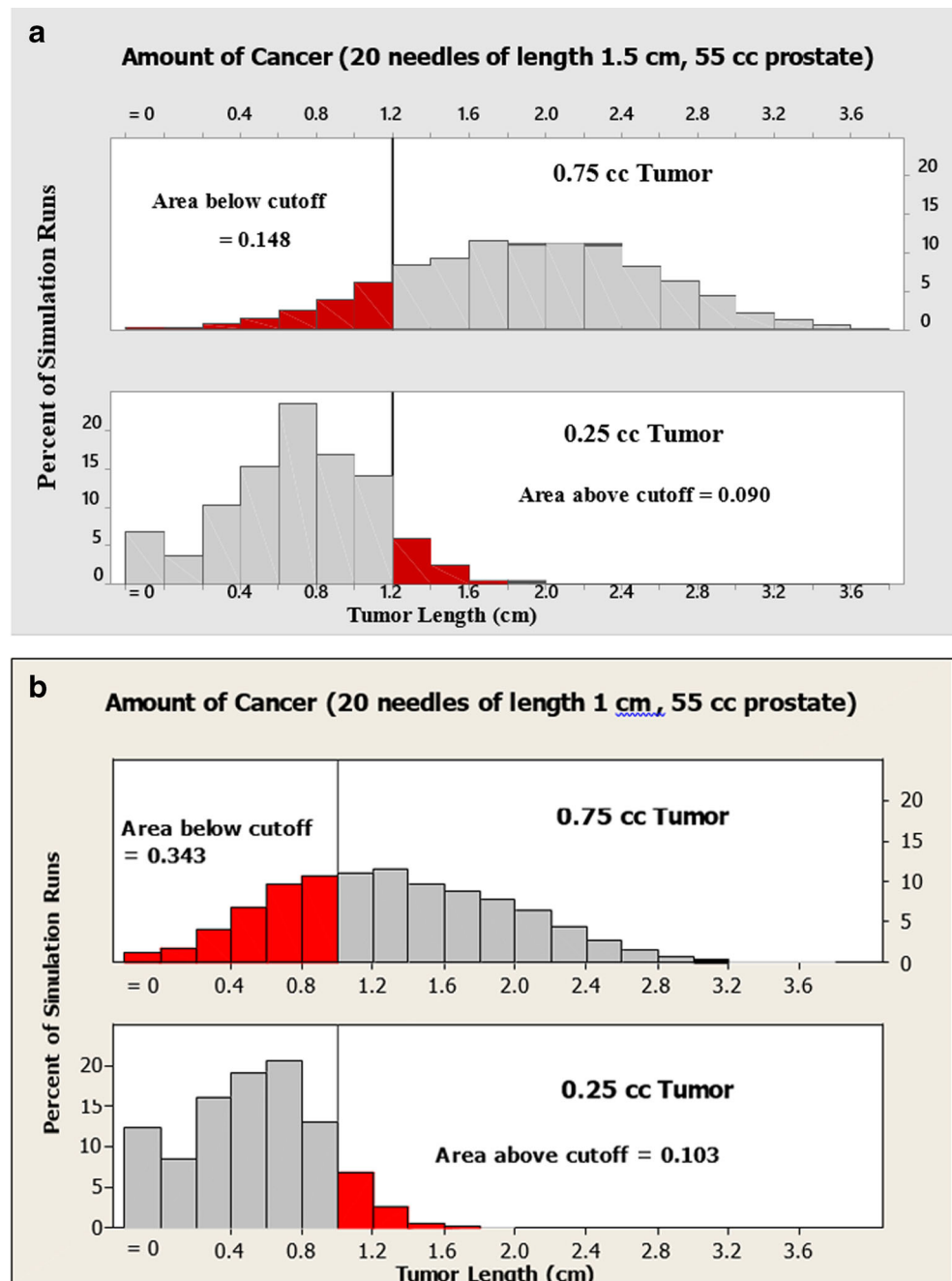
would have, in nearly all instances, a clear picture of whether his cancer is highly likely to be insignificant. Only the cancer length intervals of 1.6 - 1.8, 1.8 - 2.0, and 2.0 - 2.2 reflect some ambiguity. But for all other intervals of length of cancer (92.2% of cases) the predicted probabilities of insignificant cancer are either ≥0.94 or virtually 0, leaving little doubt about the likely outcome. For the 20 needle, 55 cc prostate case (table not shown), in 82.6% of cases the predicted probabilities of insignificant cancer are either ≥0.89 or virtually 0.

5 Discussion and conclusions

This research contributes to the literature by presenting a new approach for distinguishing minimal from significant cancers. Previous research has focused on how the probability of tumor detection changes with increases in the number of cores [47–49]. Bjurlin et al., in an extensive review that included 18 and 21 core biopsies, concluded that increasing the number of cores beyond 12 has only marginal benefits for cancer
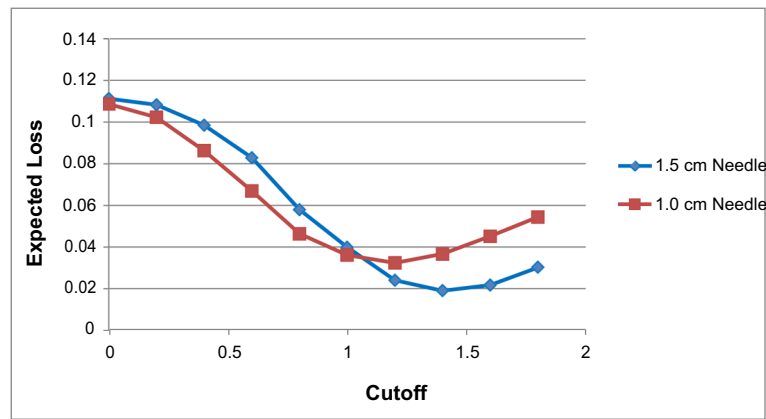
Fig. 6 Comparing Needle Lengths: 30 cc Prostate, 14 Needles; expected loss was 71% greater for the 1.0 cm needle compared to the 1.5 cm needle. Cost of missing small tumor 2 times cost of hitting minimal tumor, cost of missing medium tumor 8 times cost of hitting minimal tumor

detection [50]. Irani et al. compared a 20-core scheme to a 12-core approach and came to the same conclusion [49]. The simulation results reported in the current paper are consistent with the Irani et al. [49] findings, but show that the benefit of more needles (20 in this case) lies not in increased detection, but in the greater ability to distinguish minimal tumors from larger ones. This is a new finding not previously reported in the literature, and is consistent with statistical hypothesis testing, which shows that increasing the sample size reduces type II error. Classification errors occur because for an optimal cutoff c* on total cancer length in biopsy cores, some fraction of minimal tumor lengths will lie above c* and mean the minimal cancer will be treated, some fraction of small cancer

lengths will lie below c* and mean the small but significant tumor will not be treated, and some fraction of medium tumor lengths will lie below c* and mean the medium tumor will not be treated. Results confirmed the inadequacy of the sextant method and showed the large benefits of increasing to 14 needles, and the additional significant gains from 20 cores, especially in the case of the larger 55 cc prostate. These gains should be even greater for prostates larger than 55 cc. Increasing the number of needle cores on biopsy will also provide more accurate estimates of Gleason score. Results showed that needle core length matters in distinguishing minimal from larger tumors; the benefits of longer core lengths need to be explored further. This decision analysis took into

Table 2 Posterior probabilities of insignificant tumors, for 6, 14, and 20 needles, 30 cc prostate

Length of cancer in cm	6 needles		14 needles		20 needles	
	Percent positive biopsies	Probability insignificant	Percent positive biopsies	Probability insignificant	Percent positive biopsies	Probability insignificant
0–0.2	2.85	0.90	1.15	1.00	0.42	1.00
0.2–0.4	9.09	0.89	3.77	1.00	1.58	1.00
0.4–0.6	10.10	0.83	6.07	0.99	2.33	1.00
0.6–0.8	11.74	0.84	9.72	0.99	5.04	1.00
0.8–1.0	6.25	0.55	7.85	0.96	5.83	1.00
1.0–1.2	6.42	0.40	8.12	0.91	9.31	0.99
1.2–1.4	3.70	0.12	4.85	0.80	7.86	0.98
1.4–1.6	3.68	0.04	3.62	0.57	6.71	0.94
1.6–1.8	2.89	0.01	2.70	0.27	3.90	0.84
1.8–2.0	3.45	0.00	2.10	0.08	2.36	0.61
2.0–2.2	3.55	0.00	2.12	0.02	1.53	0.28
2.2–2.4	3.32	0.00	1.80	0.00	1.41	0.01
2.4–2.6	3.57	0.00	1.57	0.00	1.41	0.00
2.6–2.8	3.88	0.00	1.25	0.00	1.43	0.00
2.8–3.0	…	…	…	…	…	…
…	…	…	…	…	…	…

account the tradeoff between the cost of treating a minimal tumor and the costs of not treating larger tumors, in determining an optimal decision rule, which is not part of the nomogram modeling approach and hence heretofore has not been done. The research also showed how Bayes' Theorem can be used to estimate posterior probabilities of tumors of different sizes, i.e., to predict the severity of cancer, based on cancer length found on biopsy. As noted earlier, basing the decision rule on total length of cancer in biopsy cores is supported by the work of Kajikawa et al. [33] who found that total length of cancer in cores is the optimal measure for predicting the presence of small volume (minimal) cancer on radical prostatectomy, to identify candidates for active surveillance.

Table 1 shows that sensitivity and specificity both increase with increases in the number of needles. The modeling of the prostate and tumors in this paper is a simplified one, but the results concerning sensitivity and specificity are encouraging with respect to what might be achieved using this approach in conjunction with a more refined simulation such as the 3-D models of Noguchi et al. [51] that captures prostate anatomy, and Kanao et al. [47] that also model tumor anatomy.

Bul et al. [52] recognized that the Steyerberg et al. [26] nomogram that predicts the probability of indolent cancer and is used in the prostate cancer risk indicator (www.prostatecancer-riskcalculator.com) has limited applicability to extended biopsies of 12 or 18 needle cores. The Steyerberg et al. [26] nomogram predictions are based on the length of cancer found in sextant biopsies. Steyerberg et al. [26] developed adjustment factors on cancer length by performing extended biopsies on prostates obtained from autopsies. For 12- and 18-needle biopsies total cancer length would be divided by 2.03 and 2.72 respectively. This would provide estimates of what tumor lengths would have been if sextant biopsies had been performed. But it is important to emphasize that these adjustment factors are not substitutes for actually basing the nomogram on extended biopsies. Based on the results reported here, a regression-based model or nomogram to identify insignificant cancer based on a 12 or 18–20 needle protocol would improve both sensitivity and specificity with greatest gains for the 18–20 needle biopsy. Komai et al. [32], Kajikawa et al. [33], and Kim et al. [34] obtained results that are consistent with this prediction. Komai et al. [32] reported on a regression model based on extended biopsies with 12 or more needles. These authors developed a cutoff rule based on total cancer length in cores divided by number of positive cores and obtained an AUC of 91% and for a specific cutoff, sensitivity of 80% and specificity of 86%, results that to our knowledge are significantly better than results previously reported by others. Komai et al. [32] also applied the Epstein criteria [22, 23] to their patient population and found an AUC of only 81%. Kajikawa et al. [33] based on 12-core biopsies developed a nomogram with AUC of 86%

and sensitivity and specificity of 77% for a specific cutoff on total tumor length. The nomogram of Kim et al. [34] based on 12-core biopsies obtained an AUC of 87% and a sensitivity and specificity at a particular cutoff that from their Fig. 3 is about 80%. The Kattan et al. [24] nomogram and the Steyerberg et al. [26] nomogram are based on sextant biopsies.

The model developed in this paper shows the decision making benefit of increasing the number of biopsy cores from the current norm of 10–12 to about 20, as well as the benefit of using longer biopsy needles. These are important insights that follow from the analysis, but before this model can be used in practice it needs to be applied in conjunction with a computer simulation having more detailed prostate and tumor representations. In addition, future regression prediction models and nomograms based on this recommended increase in cores should significantly improve sensitivity and specificity.

The findings of this study have policy implications regarding the continuing debate about whether the benefits of screening for prostate cancer outweigh the costs. Distinguishing minimal from significant tumors will improve decision making and thereby reduce the costs and increase the benefits of screening. This research contributes to the Triple Aim objectives of improving the experience of care, reducing costs, and improving outcomes. Better decisions will lower the costs of unnecessary treatments, and improve outcomes by reducing prostate deaths, metastatic disease, and the overtreatment of insignificant tumors. It will enhance the experience of care by giving patients the opportunity to participate in decision-making while providing them with useful estimates of the likelihood of significant disease.

References

1. Andriole GL, Crawford ED, Grubb RL, Buys SS, Chia D, Church TR, Fouad MN, Gelmann EP, Kvale PA, Reding DJ, Weisfield JL, Yokochi LA, O'Brien B, Clapp JD, Rathmell JM, Riley TL, Hayes RB, Kramer BS, Izmirlian G, Miller AB, Pinsky PF, Prorok PC, Gohagan JK, Berg CD (2009) Mortality results from a randomized prostate cancer screening trial. N Engl J Med 360(13):1310–1319

2. Schroder FH, Hugosson J, Roobol MJ, Tammela TLJ, Zappa M, Nelen V, Kwiatkowski M, Lojan M, Maattdnen L, Lilja H, Denis LJ, Recker F, Paez A, Bangma CH, Carlsson S, Puliti D, Villers A, Rebillard X, Hakama M, Stenman U-H, Kojola P, Taari K, Aus G, Huber A, van der Kwast TH, van Schaik RHN, de Koning HJ, Moss SM, Auvinun A (2014) Screening and prostate cancer mortality: results of the European randomized study of screening for prostate cancer (ERSPC) at 13 years of follow-up. Lancet 384: 2027–2035

3. Bangma CH, Roemeling S, Schroder (2007) Overdiagnosis and overtreatment of early detected prostate cancer. World J Urol 25(1):3–9

4. Draisma G, Etzioni R, Tsodikov A, Mariotto A, Wever E, Gulati R, Feuer E, de Koning H (2009) Lead time and overdiagnosis in

prostate-specific antigen screening: importance of methods and context. J Natl Cancer Inst 101(6):374–383

5. Buzzoni C, Auvinen A, Roobol MJ, Carlsson S, Moss SM, Puliti D, de Koning HJ, Bangma CH, Denis LJ, Kwiatkowski M, Lujan M, Nelen V, Paez A, Randazzo M, Rebillard X, Tammela TLJ, Villers A, Hugosson J, Schroder FH, Zappa M (2015) Metastatic prostate cancer incidence and prostate-specific antigen testing: new insights from the European randomized study of screening for prostate cancer. Eur Urol 68:885–890

6. Wolf AM, Wender RC, Etzioni RB, Thompson IM, D'Amico AV, Volk RJ, Brooks DD, Dash C, Guessous I, Andrews K, DeSantis C, Smith RA (2010) American Cancer Society guideline for the early detection of prostate cancer: update 2010. CA Cancer J Clin 60:70–98

7. Greene KL, Albertsen PC, Babaian RJ, Carter HB, Gann PH, Han M, Kuban DA, Sartor AO, Stanford JL, Zietman A, Carroll P (2009) Prostate specific antigen best practice statement: 2009 update. J Urol 182(5):2232–2241

8. U.S. Preventive Services Task Force (2008) Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 149:185–191

9. Moyer VA (2012) Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 157(2):120–134

10. Drazer MW, Huo D, Eggener SC (2015) National prostate cancer screening rates after the 2012 U.S. Preventive Services Task Force recommendation discouraging prostate-specific antigen-based screening. J Clin Oncol 33(22):2416–2416

11. Jemal A, Fedewa SA, Ma J, Siegel R, Lin CC, Brawley O, Ward EM (2015) Prostate cancer incidence and PSA testing patterns in relation to USPSTF screening recommendations. JAMA 314(19):2054–2060

12. Weiner AB, Matulewicz RS, Eggener SE, Schaeffer EM (2016) Increasing incidence of metastatic prostate cancer in the United States (2004–2013). Prostate Cancer Prostatic Dis 19:395–397

13. Ross KS, Carter HB, Pearson JD, Guess HA (2000) Comparative efficiency of prostate-specific antigen screening strategies for prostate cancer detection. JAMA 284(11):1399–1405

14. Zhang J, Denton BT, Balasubramanian H, Shah ND, Inman BA (2012) Optimization of PSA screening policies: a comparison of the patient and societal perspectives. Med Decis Mak 32(2):337–349

15. Underwood DJ, Zhang J, Denton BJ, Shah ND, Inman BA (2012) Simulation optimization of PSA-threshold based prostate cancer screening policies. Health Care Manag Sci 15(4):293–309

16. Bertsimas D, Silberholz J, Trikalinos T (2018) Optimal healthcare decision making under multiple mathematical models: application in prostate cancer screening. Health Care Manag Sci 21(1):105–118

17. Gulati R, Gore JL, Etzioni R (2013) Comparative effectiveness of alternative PSA-based prostate cancer screening strategies. Ann Intern Med 158(3):145–153

18. Hodge KK, McNeal JE, Terris MK, Stamey TA (1989) Random systematic versus directed ultrasound guided trans-rectal core biopsies of the prostate. J Urol 142:71–75

19. Presti JC Jr, O'Dowd GJ, Miller MC (2003) Extended peripheral zone biopsy schemes increase cancer detection rates and minimize variance in prostate specific antigen and age related cancer rates: results of a community multi-practice study. J Urol 169(1):125–129

20. Epstein JI, Walsh PC, Carter HB (2001) Importance of posterolateral needle biopsies in the detection of prostate cancer. Urology 57:1112–1116

21. Boccon-Gibod L (1996) Editorial: significant versus insignificant cancer-can we identify the tigers from the pussycats. J Urol 156:1069–1070

22. Epstein JI, Walsh PC, Brendler CB (1994a) Radical prostatectomy for impalpable prostate cancer: the Johns Hopkins experience with tumors found on transurethral resection (stages T1A and T1B) and on needle biopsy (stage T1C). J Urol 152:1721–1729

23. Epstein JI, Walsh PC, Carmichael M, Brendler CB (1994b) Pathologic and clinical findings to predict tumor extent of nonpalpable (stage T1c) prostate cancer. JAMA 271:368

24. Kattan MW, Eastham JA, Wheeler TM et al (2003) Counseling men with prostate cancer: a nomogram for predicting the presence of small, moderately differentiated, confined tumors. J Urol 170:1792–1797

25. Bastian PJ, Mangold LA, Epstein JI, Partin AW (2004) Characteristics of insignificant clinical T1c prostate tumors: a contemporary analysis. Cancer 101:2001–2005

26. Steyerberg EW, Roobol MJ, Kattan MW, van der Kwast TH, de Koning HJ, Schroder FH (2007) Prediction of indolent prostate cancer: validation and updating of a prognostic nomogram. J Urol 177:107–112

27. Jeldres C, Suardi N, Walz J, Hutterer GC, Ahyai S, Lattouf JB, Haese A, Graefen M, Erbersdobler A, Heinzer H, Huland H, Karakiewicz PI (2008) Validation of the contemporary Epstein criteria for insignificant prostate cancer in European men. Eur Urol 54:1306–1313

28. Chun FK, Haese A, Ahyai SA et al (2008) Critical assessment of tools to predict clinically insignificant prostate cancer at radical prostatectomy in contemporary men. Cancer 113:701–709

29. O'Brien BA, Cohen RJ, Ryan A, Senqupta S, Mills J (2011) A new preoperative nomogram to predict minimal prostate cancer: accuracy and error rates compared to other tools to select patients for active surveillance. J Urol 186:1811–1817

30. Iremashvili V, Soloway SW, Pelaez L, Rosenberg DL, Manoharan M (2013) Comparative validation of nomograms predicting clinically insignificant prostate cancer. J Urol 81:1202–1208

31. Nguyen CT, Kattan MW (2012) Formalized prediction of clinically significant cancer: is it possible? Asian J Androl 14:349–354

32. Komai Y, Kawakami S, Numao N, Fujii Y, Saito K, Kubo Y, Koga F, Kumagai J, Yamamoto S, Yonese J, Ishikawa Y, Fukui I, Kihara K (2012) Extended biopsy based on criteria incorporating cumulative cancer length for predicting insignificant prostate cancer. BJU Int 110:E564–E569

33. Kajikawa K, Kanao K, Kobayashi I, Nishikawa G, Yoshizawa T, Kato Y, Watanabe M, Zennami K, Nakamura K, Sumitomo M (2016) Optimal method for measuring tumor extent in needle biopsy specimens to identify small-volume prostate cancer. Int J Urol 23:62–68

34. Kim TH, Jeon HG, Seo IS et al (2017) Development of a new nomogram to predict insignificant prostate cancer in patients undergoing radical prostatectomy. Scand J Urol 51(1):27–32

35. Mottet N, Bellmunt J, Bolla M, Briers E, Cumberbatch MG, de Santis M, Fossati N, Gross T, Henry AM, Joniau S, Lam TB, Mason MD, Matveev VB, Moldovan PC, van den Bergh RCN, van den Broeck T, van der Poel HG, van der Kwast TH, Rouvière O, Schoots IG, Wiegel T, Cornford P (2017) EAU-ESTRO-SIOG guidelines on prostate cancer. Part 1: screening, diagnosis, and local treatment with curative intent. Eur Urol 71:618–629

36. Berwick DM, Nolan TW, Whittington J (2008) The triple aim: care, health, and cost. Health Aff 27(3):759–769

37. Daneshgari F, Taylor GD, Miller GJ, Crawford ED (1995) Computer simulation of the probability of detecting low volume carcinoma of the prostate with six random systematic core biopsies. Urology 45:604–609

38. Stamey TA, Freiha FS, McNeal JE, Redwine EA, Whttemore AS, Schmid H-P (1993) Localized prostate cancer: relationship of tumor volume to clinical significance for treatment of prostate cancer. Cancer 71(Suppl):933–938

39. Andreoiu MD, Cheng L (2010) Multifocal prostate cancer: biologic, prognostic, and therapeutic implications. Hum Pathol 41:781–793

40. Zeng J, Bauer JJ, Mun SK (2000) Modeling and mapping of prostate cancer. Comput Graph 24:683–694

41. Tanaka N, Shimada K, Nakagawa Y, Hirao S, Watanabe S, Miyake M, Anai S, Hirayama A, Konishi N, Fujimoto K (2015) The optimal number of initial prostate biopsy cores in daily practice: a prospective study using the Nara urological research and treatment group nomogram. BMC Res Notes 8: 689–697

42. Cooner WH, Mosley BR, Rutherford CL Jr et al (1990) Prostate cancer detection in a clinical urological practice by ultrasonography, digital rectal examination and prostate specific antigen. J Urol 143: 1146–1154

43. Catalona WJ, Richie JP, Ahmann FR, Hudson M'LA, Scardino PT, Flanigan RC, Dekernion JB, Ratliff TL, Kavoussi LR, Dalkin BL, Waters WB, Macfarlane MT, Southwick PC (1994) Comparison of digital rectal examination and serum prostate specific antigen in the early detection of prostate cancer: results of a multicenter clinical trial of 6,630 men. J Urol 151:1283–1290

44. Schwartz MJ, Hwang DH, Hung AJ, Han J, McClain JW, Shemtov MM, te AE, Sosa RE, Vaughan ED, Scherr DS (2008) Negative influence of changing biopsy practice patterns on the predictive value of prostate-specific antigen for cancer detection on prostate biopsy. Cancer 112:1718–1725

45. Iczkowski KA, Casella G, Seppala RJ, Jones GL, Mishler BA, Qian J, Bostwick DG (2002) Needle core length in sextant biopsy influences prostate cancer detection rate. Urology 59:698–703

46. Öbeka C, Doğanca T, Erdal S, Erdoğan DH (2012) Core length in prostate biopsy: size matters. J Urol 187:2051–2055

47. Kanao K, Eastham JA, Scardino P, Reuter VE, Fine SW (2013) Can transrectal needle biopsy be optimized to detect nearly all prostate cancer with a volume of ≥ 0.5 mL? A three dimensional analysis. BJU Int 112:898–904

48. Jones JS (2007) Saturation biopsy for detecting and characterizing prostate cancer. BJU Int 99:1340–1344

49. Irani J, Blanchet P, Salomon L, Coloby P, Hubert J, Malavaud B, Mottet N (2013) Is an extended 20-core prostate biopsy more efficient than the standard 12-core? A randomized multicenter trial. J Urol 190:77–83

50. Bjurlin MA, Carter HB, Schellhammer P, Cookson MS, Gomella LG, Troyer D, Wheeler TM, Schlossberg S, Penson DF, Taneja SS (2013) Optimization of initial prostate biopsy in clinical practice: sampling, labeling and specimen processing. J Urol 189:2039–2046

51. Noguchi M, Deguchi D, Toriwaki J et al (2006) Evaluation of a prostate biopsy strategy for cancer detection using a computer simulation system with virtual needle biopsy for three-dimensional prostate models. Int J Urol 13:1296–1303

52. Bul M, Delongchamps NB, Steyerberg EW et al (2011) Updating the prostate cancer risk indicator for contemporary biopsy schemes. Can J Urol 18(2):5625–5629

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.